



การพัฒนาแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย



โดย

นายจตุพงศ์ โมรา

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปรัชญาดุษฎีบัณฑิต
สาขาวิชาหลักสูตรและการสอน(การสอนภาษาอังกฤษ) แบบ 2.1 ปรัชญาดุษฎีบัณฑิต

ภาควิชาหลักสูตรและวิธีสอน

บัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร

ปีการศึกษา 2561

ลิขสิทธิ์ของบัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร

การพัฒนาแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย



วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปรัชญาดุษฎีบัณฑิต
สาขาวิชาหลักสูตรและการสอน(การสอนภาษาอังกฤษ) แบบ 2.1 ปรัชญาดุษฎีบัณฑิต
ภาควิชาหลักสูตรและวิธีสอน
บัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร
ปีการศึกษา 2561
ลิขสิทธิ์ของบัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร

DEVELOPMENT OF A UNIVERSITY ENGLISH PROFICIENCY TEST



By
MR. Jatupong MORA

A Thesis Submitted in Partial Fulfillment of the Requirements
for Doctor of Philosophy (CURRICULUM AND INSTRUCTION)

Department of Curriculum and Instruction
Graduate School, Silpakorn University

Academic Year 2018

Copyright of Graduate School, Silpakorn University

หัวข้อ	การพัฒนาแบบทดสอบวัดความสามารถภาษาอังกฤษระดับ มหาวิทยาลัย
โดย	จตุพงศ์ โมรา
สาขาวิชา	หลักสูตรและการสอน(การสอนภาษาอังกฤษ) แบบ 2.1 ปรัชญา ดุสิตบัณฑิต
อาจารย์ที่ปรึกษาหลัก	ผู้ช่วยศาสตราจารย์ ดร. บำรุง ไตรรัตน์

บัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร ได้รับพิจารณาอนุมัติให้เป็นส่วนหนึ่งของการศึกษา
ตามหลักสูตรปรัชญาดุสิตบัณฑิต

..... คณะบดีบัณฑิตวิทยาลัย
(รองศาสตราจารย์ ดร.จุไรรัตน์ นันทานิช)

พิจารณาเห็นชอบโดย

..... ประธานกรรมการ
(ดร. อุดมกฤษฏี ศรีนนท์)

..... อาจารย์ที่ปรึกษาหลัก
(ผู้ช่วยศาสตราจารย์ ดร. บำรุง ไตรรัตน์)

..... อาจารย์ที่ปรึกษาร่วม
(ผู้ช่วยศาสตราจารย์ ดร. เสงี่ยม ไตรรัตน์)

..... ผู้ทรงคุณวุฒิภายใน
(ผู้ช่วยศาสตราจารย์ ดร. ปราณี นิลกรณ)

..... ผู้ทรงคุณวุฒิภายนอก
(รองศาสตราจารย์ ดร. สุพัฒน์ สุขมลสันต์)

56254904 : หลักสูตรและการสอน(การสอนภาษาอังกฤษ) แบบ 2.1 ปรัชญาดุษฎีบัณฑิต

คำสำคัญ : การพัฒนาแบบทดสอบ, แบบทดสอบวัดความสามารถภาษาอังกฤษ

นาย จตุพงศ์ โมรา: การพัฒนาแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย
อาจารย์ที่ปรึกษาวิทยานิพนธ์ : ผู้ช่วยศาสตราจารย์ ดร. บำรุง โตรัตน์

มหาวิทยาลัยต่าง ๆ ในประเทศไทยเริ่มกำหนดให้นักศึกษาสอบวัดความสามารถภาษาอังกฤษก่อนจบการศึกษา หากมหาวิทยาลัยสร้างแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัยที่ได้มาตรฐานเพื่อใช้วัดความสามารถภาษาอังกฤษของนักศึกษา ผลการสอบจะมีความน่าเชื่อถือและยังช่วยประหยัดค่าใช้จ่ายในการสอบของนักศึกษาอีกด้วย การวิจัยนี้มีวัตถุประสงค์หลัก 2 ข้อ คือ 1) เพื่อสร้างและพัฒนาแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัยให้เป็นแบบทดสอบมาตรฐานที่มีลักษณะเฉพาะคือ คุณประโยชน์ของแบบทดสอบเป็นไปตามเกณฑ์ที่กำหนด คุณภาพรายข้อและความเชื่อมั่นเป็นไปตามเกณฑ์ที่กำหนด และ ความตรงร่วมสมัย (Concurrent Validity) สูง เมื่อเทียบกับแบบทดสอบมาตรฐาน และ 2) เพื่อศึกษาความพึงพอใจของผู้เกี่ยวข้องต่อแบบทดสอบที่สร้างขึ้น กลุ่มตัวอย่างได้แก่อาสาสมัครที่เป็นนิสิตระดับปริญญาตรี มหาวิทยาลัยทักษิณ วิทยาเขตสงขลา ซึ่งกำลังศึกษาอยู่ในภาคเรียนที่ 2 ปีการศึกษา 2561 จำนวน 45 คน วิเคราะห์ข้อมูลโดยใช้สถิติพื้นฐาน การวิเคราะห์คุณภาพรายข้อและความเชื่อมั่น และการหาความตรงร่วมสมัยโดยใช้ Pearson's Product-Moment Correlation Coefficients ผลการวิจัยพบว่า

1) แบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัยมีคุณประโยชน์โดยรวมอยู่ในระดับดี ($\bar{x} = 4.26$, S.D. = .175) และมีคุณประโยชน์รายด้าน 7 ด้าน เรียงตามลำดับ คือ ด้านความเป็นธรรม (Fairness) ($\bar{x} = 4.66$, S.D. = .491) ด้านความตรงเชิงโครงสร้าง (Construct Validity) ($\bar{x} = 4.24$, S.D. = .121) ด้านผลกระทบ (Impact) ($\bar{x} = 4.24$, S.D. = .164) ด้านความสมจริง (Authenticity) ($\bar{x} = 4.20$, S.D. = .084) ด้านความเชื่อมั่น (Reliability) ($\bar{x} = 4.19$, S.D. = .143) ด้านการมีปฏิสัมพันธ์ (Interactiveness) ($\bar{x} = 4.18$, S.D. = .109) และด้านความสะดวกใช้ (Practicality) ($\bar{x} = 4.14$, S.D. = .113)

2) แบบทดสอบปรนัย 150 ข้อ มีค่าความยากง่ายระดับปานกลาง ($P = .55$) ค่าอำนาจจำแนกอยู่ในระดับจำแนกได้ดี ($R = .49$) และค่าความเชื่อมั่น (KR20) เท่ากับ .83 ได้ข้อคำถามคุณภาพดี 122 ข้อ คิดเป็น 81.33 % และข้อคำถามที่ควรปรับปรุง 28 ข้อ คิดเป็น 18.67 % และแบบทดสอบอัตนัย 5 ข้อ ค่าความยากง่ายคุณภาพดีมาก ($P = .55$) ค่าอำนาจจำแนกอยู่ในระดับจำแนกได้ดี ($R = .38$) และค่าสัมประสิทธิ์ความเชื่อมั่นแอลฟาของครอนบาค (Coefficient Alpha) เท่ากับ .80 และเป็นข้อคำถามคุณภาพดีทั้ง 5 ข้อ คิดเป็น 100 %

3) ในภาพรวม แบบทดสอบทั้งฉบับมีค่าความตรงร่วมสมัย (Concurrent Validity) สูงกว่าเกณฑ์ ($r_{xy} = .90$) และมีความตรงร่วมสมัย (Concurrent Validity) ของแต่ละทักษะ เรียงตามลำดับ คือ ทักษะการเขียนและการพูด ($r_{xy} = .95$) การฟัง ($r_{xy} = .93$) การอ่าน ($r_{xy} = .87$) และการใช้ภาษา ($r_{xy} = .76$)

4) ในภาพรวม ผู้มีส่วนเกี่ยวข้องพึงพอใจแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัยอยู่ในระดับมาก ($\bar{x} = 4.27, S.D. = 0.770$) และมีความพึงพอใจด้านต่าง ๆ เรียงตามลำดับ คือ 1) ด้านประโยชน์ของแบบทดสอบ ($\bar{x} = 4.40, S.D. = 0.687$) 2) ด้านคำสั่งของแบบทดสอบ ($\bar{x} = 4.36, S.D. = 0.708$) 3) ด้านการจัดการสอบ ($\bar{x} = 4.32, S.D. = 0.842$) 4) ด้านการตรวจให้คะแนน ($\bar{x} = 4.25, S.D. = 0.784$) 5) ด้านรูปแบบและการจัดลำดับความ ($\bar{x} = 4.22, S.D. = 0.813$) และ 6) ด้านเนื้อหาสาระของแบบทดสอบ ($\bar{x} = 4.19, S.D. = 0.786$)

ข้อเสนอแนะสำหรับการวิจัยครั้งนี้ คือ ควรมีการศึกษาเพิ่มเติมเพื่อการปรับปรุงข้อคำถามบางส่วนให้มีคุณภาพตามเกณฑ์ และนำไปใช้กับกลุ่มตัวอย่างขนาดใหญ่ทั้งภายในมหาวิทยาลัยทักษิณและมหาวิทยาลัยอื่น ๆ



56254904 : Major (CURRICULUM AND INSTRUCTION)

Keyword : TEST DEVELOPMENT, ENGLISH PROFICIENCY TEST

MR. JATUPONG MORA : DEVELOPMENT OF A UNIVERSITY ENGLISH PROFICIENCY TEST
 THESIS ADVISOR : ASSISTANT PROFESSOR BAMRUNG TORUT, Ph.D.

Universities in Thailand currently require their students to take English proficiency test before graduation. If the universities construct a standardized university English proficiency test to assess their students' English proficiency, the test results will be reliable and this can also help save the students' expense in test-taking. Two main objectives of this research were 1) to construct and develop a university English proficiency test to be a standardized one with three specified characteristics: test usefulness was at a specified level, item analysis indexes and reliability were at a specified level, and there was high concurrent validity when compared the scored with a standardized test, and 2) to study the stake holders' satisfaction towards the test. The sample consisted of 45 Bachelor's degree students studying in the second semester of 2019 academic year at Thaksin University, Songkhla Campus. Basic statistics, item analysis and reliability coefficients, and Pearson's product moment correlation coefficients were computed to analyze the data. It was found that:

1) overall, test usefulness of UEPT was at a good level ($\bar{x} = 4.26$, S.D. = .175), and 7 individual qualities could be put in this order: Fairness ($\bar{x} = 4.66$, S.D. = .491), Construct Validity ($\bar{x} = 4.24$, S.D. = .121), Impact ($\bar{x} = 4.24$, S.D. = .164), Authenticity ($\bar{x} = 4.20$, S.D. = .084), Reliability ($\bar{x} = 4.19$, S.D. = .143), Interactiveness ($\bar{x} = 4.18$, S.D. = .109), and Practicality ($\bar{x} = 4.14$, S.D. = .113)

2) 150 objective test items had moderate difficulty index ($P = .55$), their discrimination was at a good level ($R = .49$), and the reliability coefficient (KR20) was at .83; 122 test items (81.33%) were collected for future use and 28 items (18.67%) needed improvement or reconstruction. 5 subjective test items also had moderate difficulty index ($P = .55$), their discrimination was at a good level ($R = .38$), and the reliability coefficient (Cronbach Alpha) was at .80 and all items were collected.

3) overall, the concurrent validity of UEPT was higher than the set value ($r_{xy} = .90$) and the concurrent validity of each test section could be put in this order: Writing and

Speaking ($r_{xy} = .95$), Listening ($r_{xy} = .93$), Reading ($r_{xy} = .87$), and Language Use ($r_{xy} = .76$).

Some recommendations for this research is that there should have a further study to improve some test items to have good qualities and the test should be tried out with a large group of sample within Thaksin University and in other universities.



กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จได้ด้วยความกรุณาอิงจาก อาจารย์ที่ปรึกษาหลัก ผศ.ดร. บำรุง โตรัตน์ อาจารย์ที่ปรึกษาร่วม ผศ.ดร. เสงี่ยม โตรัตน์ และ ผศ.ดร. ปราณี นิลกรณ์ ผู้วิจัยขอกราบขอบพระคุณอาจารย์ทั้ง 3 ท่านที่ได้เสียสละเวลาอันมีค่าในการให้คำปรึกษาและชี้แนะอย่างละเอียดในทุกขั้นตอนการดำเนินการ ด้วยความทุ่มเทตั้งแต่เริ่มต้นจนงานวิจัยชิ้นนี้สำเร็จลุล่วงไปได้ด้วยดี

ขอกราบขอบพระคุณ ดร. อุดมกฤษฎี ศรีนนท์ ที่กรุณาเป็นประธานกรรมการสอบให้กับผู้วิจัย และได้ให้คำแนะนำที่เป็นประโยชน์ในการปรับปรุงงานวิจัยให้ดียิ่งขึ้น

ขอกราบขอบพระคุณ รศ.ดร. สุพัฒน์ สุขมกลสันต์ ที่ให้ความกรุณาเป็นผู้ทรงคุณวุฒิ และให้คำชี้แนะเพื่อการปรับปรุงงานมาโดยตลอด และขอกราบขอบพระคุณ Mr.Neil C. Cuthbertson อาจารย์ ดร. ศิริรัตน์ สีนประจักษ์ผล และอาจารย์พรทิพย์ บดีพงศ์ ที่กรุณาเป็นผู้เชี่ยวชาญตรวจสอบเครื่องมือที่ใช้ในการวิจัย พร้อมแสดงความคิดเห็นและให้คำแนะนำที่เป็นประโยชน์ในการพัฒนาแบบทดสอบในครั้งนี้

ขอขอบคุณ Mr. Jeremy และ Mrs. Christina Smith ที่ช่วยเสียสละเวลา และให้ความกรุณาในการบันทึกเสียงแบบทดสอบการฟัง ขอขอบคุณ Dr. Irish Chan Seosan, Miss Nikolett Solymosi และ Mr. George Gyamfi ที่ช่วยตรวจสอบความเหมาะสมและเสนอความคิดเห็นในบทสนทนาของแบบทดสอบการฟัง

ขอขอบคุณ ผศ.ดร. เกียรติทราย วุฒิพงษ์ ผศ.มลฤดี ทองศรี และอาจารย์จรัสภรณ์ มะเลโลहित ที่ได้กรุณาเสียสละเวลาในการตรวจและให้คะแนนแบบทดสอบการเขียนตลอดการเก็บข้อมูล ขอขอบคุณ อาจารย์ ดร.ชูไฮลา อีเต อาจารย์ดิฐะพร วิสะมิตน์นัท และอาจารย์นวลทิพย์ มหามงคล ที่ได้ให้ความช่วยเหลือในการทดสอบและให้คะแนนการสอบการพูดของผู้เข้าสอบตลอดการวิจัย และขอขอบคุณอาจารย์ในสาขาวิชาภาษาตะวันตก คณะมนุษยศาสตร์ และสังคมศาสตร์ มหาวิทยาลัยทักษิณ วิทยาเขตสงขลาทุกท่านที่ให้ความสนใจ และให้ความร่วมมือในการตอบแบบประเมินคุณภาพแบบทดสอบ และให้คำแนะนำที่เป็นประโยชน์ในการพัฒนาแบบทดสอบในครั้งนี้

ขอขอบคุณนิสิตมหาวิทยาลัยทักษิณที่อาสาสมัครเป็นกลุ่มทดลองและกลุ่มตัวอย่างของการวิจัยในครั้งนี้ ทุกคนได้เสียสละเวลาในเข้าร่วมทำแบบทดสอบและให้ความร่วมมือตามขั้นตอน งานวิจัยสำเร็จลุล่วงเป็นอย่างดี

ขอขอบคุณนางสาวทิพยาพร จงพล ที่ให้ความช่วยเหลือในการพิมพ์งานและสูตรทางคณิตศาสตร์

ขอขอบคุณคณะผู้บริหารมหาวิทยาลัยทักษิณที่อนุมัติให้ลาศึกษาต่อแบบเต็มเวลา และให้ทุนสนับสนุนการศึกษาต่อระดับปริญญาเอกในครั้งนี้

ท้ายที่สุด ขอกราบขอบพระคุณคุณพ่อคุณแม่ที่เป็นแรงบันดาลใจให้มุ่งมั่น ตั้งใจ ในการทำหน้าที่ให้ลุล่วง และขอขอบคุณญาติพี่น้องทุกคนที่คอยให้กำลังใจ และช่วยเหลือให้ผู้วิจัยเล่าเรียนจนประสบความสำเร็จในที่สุด

สารบัญ

	หน้า
บทคัดย่อภาษาไทย.....	ง
บทคัดย่อภาษาอังกฤษ.....	ฉ
กิตติกรรมประกาศ.....	ช
สารบัญ.....	ฅ
สารบัญตาราง.....	จ
บทที่ 1	1
บทนำ.....	1
1.1 ความสำคัญ ปัญหา และความเป็นมาของการวิจัยการพัฒนาแบบทดสอบวัดความสามารถ ภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test).....	1
1.1.1 ความสำคัญของแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test)	1
1.1.1.1 ความสำคัญของแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test) ต่อนักศึกษา.....	1
1.1.1.2 แบบทดสอบวัดความสามารถภาษาอังกฤษโดยรวมที่ได้รับการยอมรับและมี การนำมาใช้ในระดับมหาวิทยาลัย.....	3
1.1.1.3 หลักการที่สำคัญในการสร้างแบบทดสอบให้ได้มาตรฐาน.....	5
1.1.1.4 สรุปความสำคัญของแบบทดสอบวัดความสามารถภาษาอังกฤษในระดับ มหาวิทยาลัยและการใช้หลักการพัฒนาแบบทดสอบให้ได้มาตรฐาน	6
1.1.2 สภาพปัญหาของประเทศไทยที่จำเป็นต้องมีแบบทดสอบวัดความสามารถภาษาอังกฤษ ระดับมหาวิทยาลัย	7
1.1.2.1 สภาพความสามารถของนักศึกษาที่จบปริญญาตรี	7
1.1.2.2 สภาพการวัดความสามารถภาษาอังกฤษในระดับมหาวิทยาลัยที่ไม่อิงเกณฑ์ มาตรฐาน.....	7

1.1.2.3	สรุปสภาพปัญหาการวัดความสามารถภาษาภาษาอังกฤษระดับมหาวิทยาลัย	8
1.1.3	ความเป็นมาของการวิจัยด้านการวัดความสามารถภาษาภาษาอังกฤษในระดับมหาวิทยาลัย	8
1.1.3.1	งานวิจัยด้านการพัฒนาแบบทดสอบวัดความสามารถภาษาภาษาอังกฤษในระดับมหาวิทยาลัยในประเทศไทย	8
1.1.3.2	งานวิจัยด้านการพัฒนาแบบทดสอบวัดความสามารถภาษาภาษาอังกฤษในระดับมหาวิทยาลัยในต่างประเทศ	9
1.1.3.3	แนวโน้มการวิจัยด้านความสามารถภาษาภาษาอังกฤษในระดับมหาวิทยาลัย	11
1.1.4	บริบทของการศึกษา	11
1.1.4.1	สภาพการจัดการเรียนการสอนภาษาอังกฤษพื้นฐานของมหาวิทยาลัยทักษิณ	11
1.2	กรอบแนวคิดวิจัย	13
1.3	ขอบเขตการวิจัย	14
1.3.1	ประชากร	14
1.3.2	กลุ่มตัวอย่าง	14
1.3.3	ตัวแปรที่ศึกษา	15
1.4	ข้อจำกัดในการวิจัย	15
1.5	นิยามศัพท์เฉพาะ	15
1.5.1	แบบทดสอบวัดความสามารถภาษาภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test)	16
1.5.2	แบบทดสอบวัดความสามารถภาษาภาษาอังกฤษระดับมาตรฐานสากล	16
1.5.3	ความตรงเชิงเนื้อหา (Content Validity)	16
1.5.4	ความตรงร่วมสมัย (Concurrent Validity)	17
1.5.5	ความตรงเชิงโครงสร้าง (Construct Validity)	17
1.5.6	ความลำเอียงของข้อคำถาม (Item Bias)	17

1.5.7	ดัชนีความสอดคล้องระหว่างข้อคำถามและวัตถุประสงค์ (Item-Objective Congruence Index)	18
1.5.8	คุณภาพแบบทดสอบ	18
1.5.9	ความพึงพอใจในการใช้แบบทดสอบ	18
1.5.10	ผู้เข้าสอบ.....	18
1.5.11	การประเมินคุณภาพแบบทดสอบ (Test Validation).....	19
1.5.12	ความเป็นธรรม (Fairness).....	19
1.6	ประโยชน์ที่คาดว่าจะได้รับ.....	19
1.6.1	ประโยชน์ต่อผู้เรียน.....	19
1.6.2	ประโยชน์ต่อผู้จัดการเรียนการสอน	20
1.6.3	ประโยชน์ต่อสถาบันและผู้มีส่วนได้ส่วนเสียต่อการผลิตบัณฑิต	20
บทที่ 2	22
วรรณกรรมที่เกี่ยวข้อง.....	22
2.1	แบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test).....	24
2.1.1	ความหมายและผลกระทบของแบบทดสอบวัดความสามารถภาษาอังกฤษระดับ มหาวิทยาลัย (University English Proficiency Test) ต่อผู้เกี่ยวข้อง	24
2.1.1.1	ความหมายของแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test).....	24
2.1.1.2	ผลกระทบของแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test) ต่อผู้เกี่ยวข้อง	25
2.1.2	เนื้อหาสาระที่เกี่ยวข้องกับการสร้างแบบทดสอบวัดความสามารถภาษาอังกฤษระดับ มหาวิทยาลัย (University English Proficiency Test).....	28
2.1.2.1	หลักสูตรภาษาอังกฤษระดับอุดมศึกษา (มหาวิทยาลัยทักษิณ, 2544).....	28
2.1.2.3	แบบทดสอบวัดความสามารถภาษาอังกฤษมาตรฐานสากล	35

2.2 การพัฒนาแบบทดสอบวัดความสามารถภาษาอังกฤษให้เป็นแบบทดสอบมาตรฐาน	133
2.2.1 องค์ประกอบที่เกี่ยวข้องกับการพัฒนาแบบทดสอบ	133
2.2.1.1 ขั้นตอนในการพัฒนาแบบทดสอบ	133
2.2.1.2 ตารางกำหนดลักษณะแบบทดสอบ (Table of Test Specifications)	141
2.2.2 การประเมินคุณภาพแบบทดสอบวัดความสามารถภาษาอังกฤษ	146
2.2.2.1 ความเชื่อมั่น (Reliability)	146
2.2.2.2 ความตรงและการประเมินความตรง (Validity and Validation)	151
2.2.2.4 การวิเคราะห์แบบทดสอบรายข้อ (Item Analysis)	176
2.2.2.5 การสอบถามความคิดเห็นผู้เกี่ยวข้องโดยใช้แบบสอบถาม	178
2.3 งานวิจัยที่เกี่ยวข้องกับการพัฒนาแบบทดสอบและการประเมินคุณภาพแบบทดสอบวัด	
ความสามารถภาษาอังกฤษ	186
บทที่ 3	224
วิธีดำเนินการวิจัย	224
3.1 ประชากรและกลุ่มตัวอย่าง	224
3.2 ตัวแปรที่ศึกษา	228
3.3 ระยะเวลา	228
3.4 รูปแบบการวิจัย	230
บทที่ 4	338
ผลการวิเคราะห์ข้อมูล	338
4.1 คุณประโยชน์ของแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University	
English Proficiency Test)	338
4.2 คุณภาพรายข้อและความเชื่อมั่นของแบบทดสอบวัดความสามารถภาษาอังกฤษระดับ	
มหาวิทยาลัย (University English Proficiency Test)	356
4.3 ความตรงร่วมสมัยของแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย	
(Concurrent Validity)	365

รายการอ้างอิง 390

ประวัติผู้เขียน 515



สารบัญตาราง

	หน้า
ตารางที่ 1 การสังเคราะห์สาระแบบทดสอบวัดทักษะการฟัง.....	67
ตารางที่ 2 การสังเคราะห์สาระแบบทดสอบวัดทักษะการพูด.....	79
ตารางที่ 3 การสังเคราะห์สาระแบบทดสอบวัดทักษะการอ่าน.....	99
ตารางที่ 4 การสังเคราะห์สาระแบบทดสอบวัดทักษะการเขียน.....	111
ตารางที่ 5 การสังเคราะห์สาระแบบทดสอบวัดทักษะการใช้ภาษา.....	124
ตารางที่ 6 จำนวนประชากรในการวิจัย.....	224
ตารางที่ 7 จำนวนกลุ่มทดลองและกลุ่มตัวอย่าง.....	227
ตารางที่ 8 Test Blueprint for Listening Test (TOEFL).....	251
ตารางที่ 9 Test Blueprint for Listening (University English Proficiency Test).....	253
ตารางที่ 10 Test Blueprint for Language Use (TOEFL).....	256
ตารางที่ 11 Test Blueprint for Language Use (University English Proficiency Test).....	260
ตารางที่ 12 Test Blueprint for Reading Test (TOEFL).....	265
ตารางที่ 13 Test Blueprint for Reading Test (University English Proficiency Test).....	266
ตารางที่ 14 Test Blueprint for Writing (IELTS).....	269
ตารางที่ 15 Test Blueprint for Writing (University English Proficiency Test).....	270
ตารางที่ 16 Test Blueprint for Speaking (IELTS).....	273
ตารางที่ 17 Test Specifications for Speaking (University English Proficiency Test).....	275
ตารางที่ 18 ตารางกำหนดกรอบเนื้อหาสาระของแบบประเมินคุณประโยชน์แบบทดสอบ.....	280
ตารางที่ 19 ตารางกำหนดกรอบเนื้อหาสาระของแบบประเมิน.....	282
ตารางที่ 20 รายงานผลคะแนนการสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย – การทดลองครั้งที่ 1.....	292

ตารางที่ 21 รายงานผลคะแนนการสอบการพูดโดยแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย – การทดลองครั้งที่ 1.....	295
ตารางที่ 22 รายงานผลคะแนนการสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัยโดยใช้แบบทดสอบมาตรฐาน- การทดลองครั้งที่ 1	298
ตารางที่ 23 รายงานผลคะแนนการสอบการพูดโดยแบบทดสอบวัดความสามารถภาษาอังกฤษมาตรฐาน- การทดลองครั้งที่ 1	301
ตารางที่ 24 สรุปคุณภาพแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัยหลังจากการทดลองใช้ครั้งที่ 1	310
ตารางที่ 25 สรุปความตรงร่วมสมัย (Concurrent Validity) ของแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (x) และแบบทดสอบมาตรฐาน (y) หลังจากการทดลองใช้ครั้งที่ 1	311
ตารางที่ 26 รายงานผลคะแนนการสอบพูดโดยแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย – การทดลองครั้งที่ 2.....	321
ตารางที่ 27 รายงานผลคะแนนการสอบพูดโดยแบบทดสอบวัดความสามารถภาษาอังกฤษมาตรฐาน- การทดลองครั้งที่ 2.....	328
ตารางที่ 28 สรุปคุณภาพแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัยหลังจากการทดลองใช้ครั้งที่ 2	333
ตารางที่ 29 สรุปความตรงร่วมสมัย (Concurrent Validity) ของแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (x) และแบบทดสอบมาตรฐาน (y) หลังจากการทดลองใช้ครั้งที่ 2	336
ตารางที่ 30 สรุปผลการวิเคราะห์ข้อมูลตามจุดประสงค์ของการวิจัยและสถิติที่ใช้	337
ตารางที่ 31 ผลการประเมินคุณประโยชน์ของแบบทดสอบด้านการฟัง.....	339
ตารางที่ 32 ผลการประเมินคุณประโยชน์ของแบบทดสอบด้านการใช้ภาษา	342
ตารางที่ 33 ผลการประเมินคุณประโยชน์ของแบบทดสอบด้านการอ่าน	345
ตารางที่ 34 ผลการประเมินคุณประโยชน์ของแบบทดสอบด้านการเขียน	348
ตารางที่ 35 ผลการประเมินคุณประโยชน์ของแบบทดสอบด้านการพูด	352
ตารางที่ 36 สรุประดับคุณประโยชน์ของแบบทดสอบวัดความสามารถภาษาอังกฤษระดับ.....	355
ตารางที่ 37 คุณภาพรายข้อและค่าความเชื่อมั่นของแบบทดสอบด้านการฟัง	357

ตารางที่ 38	คุณภาพรายข้อและค่าความเชื่อมั่นของแบบทดสอบด้านการใช้ภาษา.....	359
ตารางที่ 39	คุณภาพรายข้อและค่าความเชื่อมั่นของแบบทดสอบด้านการอ่าน.....	360
ตารางที่ 40	คุณภาพรายข้อและค่าความเชื่อมั่นของแบบทดสอบด้านการเขียน.....	362
ตารางที่ 41	คุณภาพรายข้อและค่าความเชื่อมั่นของแบบทดสอบด้านการพูด.....	363
ตารางที่ 42	สรุปคุณภาพรายข้อและความเชื่อมั่นของแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย.....	364
ตารางที่ 43	รายงานผลคะแนนการสอบวัดความสามารถภาษาอังกฤษโดยใช้แบบทดสอบ 366	
ตารางที่ 44	รายงานผลคะแนนการสอบวัดความสามารถภาษาอังกฤษโดยใช้แบบทดสอบ TOEFL (Test of English as a Foreign Language) ของกลุ่มตัวอย่าง.....	368
ตารางที่ 45	ความสัมพันธ์ระหว่างคะแนนแบบทดสอบปรนัยของแบบทดสอบวัด ความสามารถ.....	371
ตารางที่ 46	รายงานผลคะแนนการสอบวัดความสามารถภาษาอังกฤษโดยใช้แบบทดสอบวัด ความสามารถภาษาอังกฤษระดับมหาวิทยาลัย-การเขียน.....	374
ตารางที่ 47	รายงานผลคะแนนการสอบวัดความสามารถภาษาอังกฤษโดยใช้แบบทดสอบ IELTS (International English Language Testing System) -การเขียน.....	377
ตารางที่ 48	ความสัมพันธ์ของคะแนนแบบทดสอบอัตนัย (การเขียน) ระหว่างคะแนน..	379
ตารางที่ 49	รายงานผลคะแนนการสอบวัดความสามารถภาษาอังกฤษโดยใช้แบบทดสอบวัด ความสามารถภาษาอังกฤษระดับมหาวิทยาลัย-การพูด.....	381
ตารางที่ 50	รายงานผลคะแนนการสอบวัดความสามารถภาษาอังกฤษโดยใช้แบบทดสอบ IELTS (International English Language Testing System) -การพูด.....	384
ตารางที่ 51	แสดงความตรงของคะแนนแบบทดสอบอัตนัย (การพูด) ระหว่างคะแนน..	387
ตารางที่ 52	สรุปความสัมพันธ์ระหว่างคะแนนของแบบทดสอบวัดความสามารถภาษาอังกฤษระดับ มหาวิทยาลัยและแบบทดสอบมาตรฐานสากลในแต่ละด้าน.....	389

บทที่ 1

บทนำ

1.1 ความสำคัญ ปัญหา และความเป็นมาของการวิจัยการพัฒนาแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test)

1.1.1 ความสำคัญของแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test)

1.1.1.1 ความสำคัญของแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test) ต่อนักศึกษา

แบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัยมีความสำคัญต่อนักศึกษาด้วยเหตุผล 3 ประการ คือ 1) ช่วยยกระดับคุณภาพด้านภาษาของประชากรไทย 2) ส่งเสริมให้บัณฑิตไทยมีคุณสมบัติตามความต้องการของสังคม และ 3) เพิ่มโอกาสที่ดีให้แก่บัณฑิตไทย โดยมีรายละเอียดในแต่ละประการ คือ การให้นักศึกษาได้สอบวัดความสามารถภาษาอังกฤษ เป็นแนวทางในการช่วยยกระดับคุณภาพด้านภาษาของประชากรไทย เนื่องจากมีรายงานและผลการวิจัยต่าง ๆ ที่ระบุว่า ความสามารถด้านภาษาอังกฤษของคนไทยยังอยู่ในระดับที่ไม่เป็นที่พอใจ EF Education First (2015); (Jin, 2011) รัฐบาลจึงกำหนดแนวทางที่จะแก้ปัญหาดังกล่าว โดยนายกรัฐมนตรีได้มอบหมายให้กระทรวงศึกษาธิการไปพิจารณาหาแนวทางที่จะทำให้คนไทยทั่วไปสามารถใช้ภาษาอังกฤษในการสื่อสารได้บ้าง (ASTV ผู้จัดการออนไลน์, 2558) เพื่อแก้ปัญหาดังกล่าว รัฐมนตรีว่าการกระทรวงศึกษาธิการได้กล่าวข้อความตอนหนึ่งในการกล่าวปาฐกถา เรื่อง “ความร่วมมือของอุดมศึกษาไทยในการแก้ไขปัญหาของประเทศ” ในการประชุมวิชาการที่ประชุมอธิการบดีแห่งประเทศไทย ประจำปี 2558 ว่า ในปีการศึกษา 2559 จะให้นักศึกษาระดับปริญญาตรีทุกคน ทดสอบภาษาอังกฤษเพื่อการสื่อสารก่อนจบการศึกษา หรือ Exit Examination ต่อจากนั้นจะให้ผลสอบบันทึกไว้กับใบรับรองการจบการศึกษาหรือทรานสคริปต์ แต่จะไม่มีผลกับเกรด โดยเตรียมทำการทดสอบกับนักศึกษาชั้นปีที่ 1 ก่อน ทั้งนี้ เพื่อเป็นการเร่งรัดการแก้ไขปัญหาความสามารถทางภาษาอังกฤษของนักเรียน นักศึกษาไทย ที่ยังอยู่ในระดับที่ยังไม่น่าพึงพอใจนัก พร้อมทั้งเชื่อว่า หากมีนโยบายนี้จะทำให้ทุกคนตระหนัก และเร่งพัฒนาความสามารถทางภาษาของตนเอง (TNN Thailand 24 Hours, 2558)

ประการที่สอง การกำหนดให้มีการสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัยจะ ช่วยผลักดันให้บัณฑิตไทยมีคุณสมบัติตามความต้องการของสังคม โดยสอดคล้องกับข้อสรุปจากการ ระดมความคิดเห็นการประชุมผู้บริหารสถาบันอุดมศึกษา เรื่อง “ความสามารถในการแข่งขันของ สถาบันอุดมศึกษาไทย เมื่อเข้าสู่ประชาคมเศรษฐกิจอาเซียน (AEC) ในปี 2558 จัดทำโดย มูลนิธิเพื่อ สถาบันการศึกษาวิชาการจัดการแห่งประเทศไทย (ไอเมท) เมื่อวันที่ 24 พฤษภาคม พ.ศ. 2555 ณ โรงแรมสยามซิตี้ กรุงเทพฯ พบว่า ความต้องการบัณฑิตในประเทศ คือ มีความใฝ่รู้ รอบรู้ใน ด้านต่างๆ สู้งานหนัก มีความคิดริเริ่มสร้างสรรค์ รู้จักวิเคราะห์ รู้จักการแก้ปัญหา การนำเสนอ มี ทักษะด้านภาษา คือ ภาษาอังกฤษ ภาษาอาเซียน มีวิสัยทัศน์ที่เปิดกว้าง เพื่อสร้างมูลค่าเพิ่มให้ตัวเอง มี ความรอบรู้ด้านการศึกษารวมชนท้องถิ่น มีทักษะในการบริหารจัดการ มีทักษะด้านวัฒนธรรม และมี จิตสาธารณะ ส่วนความต้องการบัณฑิตในประชาคมเศรษฐกิจอาเซียน คือ มีความตระหนัก ต้องมี ความรอบรู้ในการเปลี่ยนแปลง ยอมรับความหลากหลาย มีทักษะในด้านความรู้ด้านการปฏิบัติงาน มี ทักษะด้านความเข้าใจวัฒนธรรมที่หลากหลาย มีทักษะด้านเทคโนโลยีสมัยใหม่ มีทักษะด้านการ บริหารความแตกต่าง มีทักษะในการทำงานที่หลากหลาย มีทักษะด้านมนุษยสัมพันธ์ และมีทักษะ ด้านภาษาอังกฤษและภาษาอาเซียน (IMET, 2555) จะเห็นได้ว่า ความสามารถภาษาอังกฤษเป็น คุณสมบัติหนึ่งของบัณฑิตที่พึงประสงค์ตามความต้องการของสังคมทั้งในประเทศและในประชาคม เศรษฐกิจอาเซียน

ประการที่สาม การทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัยจะทำให้ นักศึกษาได้รับโอกาสที่ดี 2 อย่าง คือ 1) ได้มีโอกาสในการพัฒนาการเรียนรู้ภาษาอังกฤษสำหรับ ตนเอง และ 2) มีโอกาสที่ดีในการใช้ชีวิตในอนาคต ในประเด็นแรก การทดสอบผลักดันให้นักศึกษา ต้องอ่านหนังสือ หรือบทกวี ตำรา ศึกษาค้นคว้าอย่างจริงจัง จึงทำให้มีโอกาสได้เรียนรู้มากขึ้น และ ถือว่าเป็นสิ่งที่นักศึกษาทำอยู่แล้วในสถานการณ์ปัจจุบัน ดังที่ (Saint John's for Education, 2555) กล่าวว่า ในการเตรียมความพร้อมเพื่อเข้าสู่ประชาคมเศรษฐกิจอาเซียนนั้น นักศึกษาต้องปรับตัวและ เตรียมการรับมือกับสถานการณ์ในอนาคตเช่นกัน สำหรับโอกาสในการใช้ชีวิตนี้ครอบคลุมการใช้ ชีวิตประจำวันและการทำงานในโลกอนาคต ปฏิเสธไม่ได้เลยว่าภาษาอังกฤษไม่จำเป็น ดังที่ (สมเกียรติ อ่อนวิมล, 2554) กล่าวว่า ความจำเป็นที่จะต้องใช้ภาษาอังกฤษเพื่อเป็นสิ่งที่หลีกเลี่ยง ไม่ได้ เพราะ กฎบัตรอาเซียนข้อ 34 บัญญัติว่า “The working language of ASEAN shall be English” “ภาษาที่ใช้ในการทำงานของอาเซียน คือ ภาษาอังกฤษ” ความหมายที่เป็นที่เข้าใจใน ขึ้นต้นเป็นเพียงเรื่องของทางราชการและภาคธุรกิจเอกชนเท่านั้น ซึ่งหากเป็นเพียงเท่านี้ก็เป็นเรื่อง ปกติธรรมดาของการทำงานในโลกปัจจุบันอยู่แล้ว แม้จะหมายความเพียงว่าเป็นการใช้ภาษาอังกฤษ ในการสื่อสารระหว่างกันในการทำงานร่วมกันของเจ้าหน้าที่รัฐบาล ตลอดจนองค์กรและหน่วยงาน ต่างๆที่เกี่ยวข้อง ทั้งภาครัฐและภาคเอกชน ทว่าความหมายของบทบัญญัติที่ให้ภาษาอังกฤษเป็น ภาษาของอาเซียนสำหรับการทำงานร่วมกันนั้นมีความหมายกว้างไกลไปถึงทุกส่วนของประชาคม อาเซียนด้วย หมายความว่าประชาชนพลเมืองใน 10 ประเทศอาเซียนจะต้องใช้ภาษาอังกฤษกันมาก ขึ้น นอกเหนือจากภาษาประจำชาติหรือภาษาประจำถิ่นของแต่ละชาติแต่ละชุมชนเอง เพราะไม่ เพียงแต่เจ้าหน้าที่รัฐเท่านั้นที่จะต้องไปมาหาสู่ร่วมประชุมปรึกษาหารือและสื่อสารกัน และไม่เฉพาะ นักธุรกิจและคนทำมาค้าขายระหว่างประเทศเท่านั้นที่จะต้องใช้ภาษาอังกฤษในการสื่อสารและการ

ติดต่อธุรกิจระหว่างกัน แต่ในเมื่อทุกคนที่อยู่ในอาเซียนล้วนแล้วแต่เป็นพลเมืองของอาเซียนด้วยกันทุกคน และทุกคนจะต้องไปมาหาสู่ เดินทางท่องเที่ยว ทำความรู้จักคุ้นเคยต่อกัน เรียนรู้ซึ่งกันและกัน และที่สำคัญทุกคนจะต้องเดินทางข้ามพรมแดนเพื่อหางานทำและแสวงหาโอกาสที่ดีกว่าให้กับชีวิต ดังนั้นภาษาอังกฤษจึงเป็นเครื่องมืออันดับหนึ่งสำหรับพลเมืองอาเซียนในการสื่อสารสร้างสัมพันธ์สู่โลกกว้างของภูมิภาคอาเซียน โลกแห่งมิตรไมตรีที่ขยายกว้างไร้พรมแดน โลกแห่งการแข่งขันไร้ขอบเขตภูมิศาสตร์และวัฒนธรรม

โอกาสที่ดียังครอบคลุมถึงการได้งานทำ เนื่องจากการสมัครงานทุกตำแหน่งต้องการผู้สมัครที่มีความสามารถภาษาอังกฤษ จะเห็นได้ว่าหลาย ๆ บริษัทให้ความสำคัญกับการใช้ภาษาอังกฤษในการทำงาน ในการลงประกาศรับสมัครงาน ภาษาอังกฤษจึงเป็นคุณสมบัติข้อหนึ่งที่คนทำงานต้องมี หากผู้สมัครคนใดไม่สามารถพูดคุยเป็นภาษาอังกฤษได้ หรือพูดได้เพียงเล็กน้อยหรือเขียนได้เพียงอย่างเดียว แต่ไม่มีความเข้าใจในการใช้ภาษาอังกฤษ จะมีโอกาสในการได้งานทำน้อยลง ความสามารถด้านภาษาอังกฤษจึงกลายเป็นอุปสรรคในการทำงานอย่างหลีกเลี่ยงไม่ได้ หากต้องการประสบความสำเร็จในการทำงาน ผู้สมัครจำเป็นต้องมีความสามารถทางภาษาอังกฤษด้วย (Jobsdb.com, 2015) นอกจากนี้ ผู้สมัครที่เก่งภาษาอังกฤษจะมีโอกาสได้งานที่เงินเดือนสูง เพราะบริษัทข้ามชาติหรือบริษัทที่มีการติดต่อชาวต่างชาติส่วนใหญ่จำเป็นต้องใช้ภาษาอังกฤษในการสื่อสารเป็นสำคัญ และมีกำลังในการจ่ายค่าตอบแทนสูงเช่นกัน เช่น การได้งานเป็นแอร์โฮสเตส สจ๊วต ที่มีค่าตอบแทนสูง แต่เกณฑ์สำคัญด่านแรกที่สายการบินรับคน ก็คือภาษาอังกฤษ โดยต้องยื่นคะแนน TOEIC (Test of English as International Communication) และต้องสัมภาษณ์เป็นภาษาอังกฤษ โอกาสได้งานของคนที่มีความสามารถด้านภาษาตั้งแต่ 2 ภาษาขึ้นไปมีสูง เช่น ไทย-อังกฤษ ไทย-จีน เป็นต้น และบางธุรกิจยินดีจ่ายค่าภาษาให้กับพนักงานเพิ่มเติมจากเงินเดือนอีกด้วย (British Council, 2015)

นอกจากโอกาสที่ดีในการประกอบอาชีพแล้ว ผู้ที่มีความสามารถภาษาอังกฤษจะมีโอกาสดีในการเดินทางไปต่างประเทศ เพราะการเดินทางหรือย้ายถิ่นฐานไปอยู่ต่างประเทศที่นอกเขตประเทศในเขตประชาคมเศรษฐกิจอาเซียนคงเป็นเรื่องธรรมดาที่จะต้องมีความรู้ในในอนาคต สิ่งที่สำคัญในการเดินทางไปศึกษาต่อ ทำงาน หรืออาศัยอยู่ถาวร คือ วิชา และในบางประเทศต้องการผลการสอบภาษาอังกฤษ เช่น ในการเข้าสหราชอาณาจักร เพื่อศึกษาต่อหรือทำงาน ต้องมีผลสอบ IELTS จึงจะดำเนินการขอวีซ่าได้ การได้มีโอกาสทดสอบวัดความสามารถภาษาอังกฤษไว้บ้าง จะทำให้นักศึกษามีประสบการณ์ในการสอบ และทำให้โอกาสสอบผ่านง่ายขึ้น และมีโอกาสดี ๆ ในชีวิตมากขึ้น (British Council, 2015)

1.1.1.2 แบบทดสอบวัดความสามารถภาษาอังกฤษโดยรวมที่ได้รับการยอมรับและมีการนำมาใช้ในระดับมหาวิทยาลัย

แบบทดสอบวัดความสามารถภาษาอังกฤษโดยรวมที่ได้รับการยอมรับมีอยู่หลายชนิด ได้แก่ ข้อสอบ TOEFL (Test of English as a Foreign Language) ข้อสอบ IELTS (International English Language Testing System) ข้อสอบ CET (College English Test) ของประเทศจีน และ

ข้อสอบ CU-TEP (Chulalongkorn University Test of English Proficiency) โดยแต่ละชนิดมีขอบข่ายการวัด คือ ข้อสอบ TOEFL (Test of English as a Foreign Language) วัดความสามารถของผู้เข้าสอบในการใช้และเข้าใจภาษาอังกฤษในระดับมหาวิทยาลัย และประเมินว่าผู้เข้าสอบสามารถใช้ทักษะการฟัง พูด อ่าน และเขียนในแวดวงวิชาการได้ดีในระดับใด (Educational Testing Service, 2010) 3) ข้อสอบ IELTS (International English Language Testing System) วัดความสามารถของผู้เข้าสอบใน 4 ทักษะ คือ การฟัง การอ่าน การเขียน และการพูด แบ่งออกเป็น 2 รูปแบบ คือ แบบ Academic และ General Training ข้อสอบในส่วนของการพูดและฟังจะเหมือนกันทั้ง 2 แบบ แต่ส่วนการอ่านและเขียนจะแตกต่างกัน ข้อสอบรูปแบบ Academic เหมาะสำหรับผู้ที่ต้องการเรียนในระดับต่ำกว่าปริญญาตรีหรือระดับสูงกว่า หรือผู้ที่กำลังหางานอาชีพที่ได้รับ การขึ้นทะเบียน เช่น แพทย์ และพยาบาล ส่วนรูปแบบ General Training เหมาะสำหรับผู้ที่ย้ายถิ่นฐานไปทำงานในประเทศที่ใช้ภาษาอังกฤษ เช่น ออสเตรเลีย แคนาดา นิวซีแลนด์ อังกฤษ หรือสำหรับผู้ที่ต้องการฝึกอบรมหรือเรียนในระดับที่ต่ำกว่าปริญญาตรี (IELTS, 2015) 4) ข้อสอบ CET (College English Test) ของประเทศจีนวัดความสามารถของนักศึกษาในการเข้าใจการสนทนาจากการพูด (Oral Conversation) หรือบทอ่าน (Passages) ตามมาตรฐานของการใช้ภาษาอังกฤษแบบ American English หรือ British English (Zheng Ying & Cheng Liying, 2008) และข้อสอบ CU-TEP (Chulalongkorn University Test of English Proficiency) วัดความสามารถทางภาษาอังกฤษใน 4 ทักษะ คือ ฟัง พูด อ่าน และเขียนของผู้ที่ต้องการศึกษาต่อในระดับปริญญาตรี หลักสูตรนานาชาติ หลักสูตรปริญญาโท-เอก ในสาขาวิชาต่าง ๆ ตามที่จุฬาลงกรณ์มหาวิทยาลัยได้กำหนดรับเข้าเรียน (สถาบันภาษา จุฬาลงกรณ์มหาวิทยาลัย, 2552)

ปัจจุบันหลายประเทศได้กำหนดให้นักศึกษาในระดับมหาวิทยาลัยต้องผ่านการสอบวัดความสามารถภาษาอังกฤษก่อนจบการศึกษา เพื่อเป็นการประกันคุณภาพของบัณฑิตและสร้างโอกาสในการได้งานทำ ตัวอย่างเช่น ประเทศจีนใช้ข้อสอบวัดความสามารถภาษาอังกฤษ CET (College English Test) สำหรับนักศึกษาที่ไม่ใช่เอกภาษาอังกฤษ และ TEM (Test for English Majors) สำหรับนักศึกษาวิชาเอกภาษาอังกฤษ ทั้งในระดับปริญญาตรีและบัณฑิตศึกษา (Zheng Ying & Cheng Liying, 2008) ประเทศมาเลเซียใช้ข้อสอบวัดความสามารถภาษาอังกฤษ MUET (Malaysia University English Test) สำหรับการสอบเข้าเรียนต่อในมหาวิทยาลัยในประเทศมาเลเซีย และประเทศสิงคโปร์ และยังใช้เป็นเงื่อนไขในการจบการศึกษาของนักศึกษาด้วย นอกจากนี้ ผู้ที่ประกอบอาชีพอื่น ๆ สามารถสอบ MUET ได้เช่นกัน เพื่อใช้ในการเลื่อนตำแหน่งและเพื่อให้มีคุณสมบัติด้านภาษา เช่น อาชีพแม่บ้าน ตำรวจ ครู และพนักงานบริษัท เป็นต้น (Malaysia Examination Council, 2006) และประเทศไต้หวันใช้ข้อสอบ GEPT (General English Proficiency Test) ในการสอบวัดความสามารถทางภาษาอังกฤษของนักเรียน โดยกำหนดระดับการสอบที่แตกต่างกันไปตามระดับการศึกษาของนักเรียน เช่น The GEPT Elementary level สำหรับนักเรียนระดับประถมถึงมัธยมต้น และ The GEPT Intermediate level สำหรับนักเรียนมัธยมปลายหรือนักศึกษาชั้นปีที่ 1 เป็นต้น (The Language Training Testing Center, 2014)

สำหรับในประเทศไทย มหาวิทยาลัยหลายแห่งได้กำหนดให้มีการสอบวัดความสามารถทางภาษาอังกฤษแล้ว สำหรับนักศึกษาระดับปริญญาตรีส่วนใหญ่จะเป็นการวัดความสามารถอย่างเดียว

ยังไม่ได้กำหนดให้เป็นเกณฑ์ในการจบการศึกษา สำหรับนักศึกษาระดับบัณฑิตศึกษามีการใช้เกณฑ์ความสามารถภาษาอังกฤษเป็นเกณฑ์ในการจบการศึกษาที่มีกฎระเบียบของแต่ละมหาวิทยาลัยอย่างชัดเจน (จุฬาลงกรณ์มหาวิทยาลัย., 2557; มหาวิทยาลัยธรรมศาสตร์, 2557; มหาวิทยาลัยศิลปากร, 2557) สำหรับแบบทดสอบที่ใช้กันอยู่ในขณะนี้มียู่ 2 ชนิด คือ แบบทดสอบวัดความสามารถภาษาอังกฤษมาตรฐานสากล เช่น ข้อสอบ TOEIC (Test of English for International Communication), ข้อสอบ TOEFL (Test of English as a Foreign Language) และ ข้อสอบ IELTS (International English Language Testing System) และแบบทดสอบวัดความสามารถภาษาอังกฤษที่แต่ละมหาวิทยาลัยจัดสร้างขึ้นเอง เช่น ข้อสอบ CU-TEP (Chulalongkorn University Test of English Proficiency) ของจุฬาลงกรณ์มหาวิทยาลัย, ข้อสอบ TU-GET (Thammasart University General English Test) ของมหาวิทยาลัยธรรมศาสตร์, ข้อสอบ SEPT (Silpakorn English Proficiency Test) ของมหาวิทยาลัยศิลปากร, ข้อสอบ PSU-TEP (Prince of Songkla University Test of English Proficiency) ของมหาวิทยาลัยสงขลานครินทร์, ข้อสอบ MU-TEST (Mahidol University Test) ของมหาวิทยาลัยมหิดล, ข้อสอบ STOU-EPT (Sukhothai Thammathirat Open University English Proficiency Test) ของมหาวิทยาลัยสุโขทัยธรรมราชา, ข้อสอบ RMUTSV-Test (Rajamangala University of Technology Srivijaya Test) ของมหาวิทยาลัย เทคโนโลยีราชมงคลศรีวิชัย, ข้อสอบ MFU-TEP (Mae Fah Luang Univeristy Test of English Proficiency) ของมหาวิทยาลัยแม่ฟ้าหลวง, และ ข้อสอบ KU-EPT (Kasetsart University English Proficiency Test) ของมหาวิทยาลัยเกษตรศาสตร์ การเลือกใช้ข้อสอบสำเร็จรูปจะมีความสะดวกมากกว่าการสร้างข้อสอบเอง และผู้สอบสามารถอ้างอิงผลการสอบเพื่อใช้ในระดับสากลได้อีกด้วย แต่ในความเป็นจริง ตามทฤษฎีการออกแบบและพัฒนาข้อสอบนั้น แบบทดสอบวัดความสามารถภาษาอังกฤษของผู้เรียนควรได้รับการออกแบบและสร้างขึ้นให้เหมาะสมและสอดคล้องกับผู้สอบเฉพาะกลุ่ม การนำผลการสอบมาใช้จึงจะน่าเชื่อถือและมีความเที่ยงตรง (Pan Yi-Ching, 2009a; The European Association for Quality Language Services, 2009; Urmston Alan, Raquel, Michelle, & Tsang Carrie, 2013) และเพื่อให้ผลการสอบสามารถอ้างอิงได้ในระดับสากล แบบทดสอบชุดดังกล่าว ต้องได้รับการเทียบเคียงคุณภาพกับมาตรฐานการอ้างอิงความสามารถด้านภาษาของสภายุโรป (CEFR) ซึ่งเป็นมาตรฐานที่ได้รับความนิยมกันในวงกว้างในปัจจุบัน (The European Association for Quality Language Services, 2009)

1.1.1.3 หลักการที่สำคัญในการสร้างแบบทดสอบให้ได้มาตรฐาน

การสร้างแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัยให้มีมาตรฐานระดับสากลสามารถทำได้ด้วยการปฏิบัติตามทฤษฎีการออกแบบและสร้างข้อสอบ 2 ขั้นตอน คือ ขั้นตอนแรกเกี่ยวกับการสร้างแบบทดสอบให้มีคุณลักษณะที่บ่งชี้ประโยชน์ของแบบทดสอบ 6 ประการ คือ 1) Reliability หมายถึง ความเที่ยงของผลการสอบ เชื่อถือได้เมื่อใช้สอบกับผู้เข้าสอบกลุ่มเดิม ไม่ว่าจะ

สอบก็ครั้ง และสอบเวลาใด หรือสถานที่ใด ผลการสอบที่ได้ในแต่ละครั้งไม่แตกต่างกัน 2) Construct Validity หมายถึง ความตรงเชิงโครงสร้าง แบบทดสอบที่สร้างขึ้นวัดโครงสร้างความสามารถตรงตามที่ต้องการจะวัด 3) Authenticity หมายถึง ความสมจริง ตรวจสอบได้จากการที่แบบทดสอบมีระดับความสอดคล้องของลักษณะภาระงานในตัวแบบทดสอบกับการใช้ภาษาในสถานการณ์จริงในอนาคตของผู้เข้าสอบ 4) Interactiveness หมายถึง ปฏิสัมพันธ์ระหว่างแบบทดสอบกับผู้เข้าสอบ คุณลักษณะเฉพาะของผู้เข้าสอบแต่ละคน เช่น ความรู้ทางภาษาในด้านต่าง ๆ กลยุทธ์การเรียนรู้ภาษา และความรู้ในเนื้อหาสาระต่าง ๆ มีส่วนช่วยในการทำภาระงานในแบบทดสอบให้บรรลุผลสำเร็จ 5) Impact หมายถึง ผลกระทบของแบบทดสอบ ต่อสังคม ต่อบุคคล ประเภทของการเตรียมตัวเข้าสอบ ประสบการณ์การเข้าสอบ ลักษณะของผลที่คาดหวัง และผลกระทบต่อการเรียนการสอน ภาษาอังกฤษ และ 6) Practicality หมายถึง การนำแบบทดสอบไปใช้ประโยชน์ได้จริง (Heaton, 1990; Bachman & Palmer, 1996; Alderson, Clapham, and Wall, 1995; Brown, 1996; McNamara, 2000) (Alderson, 1995; Bachman & Lyle F. & Palmer, 1996; McNamara Tim, 2000) นอกจากนี้ ก่อนที่จะนำแบบทดสอบไปใช้จริงกับกลุ่มตัวอย่าง ผู้วิจัยควรให้ผู้เชี่ยวชาญด้านการทดสอบและประเมินผลการสอนภาษาอังกฤษตรวจสอบคุณภาพแบบทดสอบก่อน ผู้เชี่ยวชาญประเมินคุณภาพแบบทดสอบตามแบบประเมินคุณภาพด้านต่าง ๆ พร้อมทั้งเขียนแสดงความคิดเห็นในตอนท้าย ข้อมูลที่ได้จากผู้เชี่ยวชาญสามารถใช้ปรับปรุงแบบทดสอบก่อนนำไปใช้จริงกับกลุ่มตัวอย่าง ขั้นตอนที่สองเกี่ยวกับการเทียบคุณภาพแบบทดสอบที่สร้างขึ้นกับแบบทดสอบมาตรฐานสากล การเทียบเคียงคุณภาพแบบทดสอบนั้นจะทำให้แบบทดสอบที่สร้างขึ้นมีมาตรฐานสามารถใช้อ้างอิงความสามารถของผู้สอบตามเกณฑ์ในระดับสากลได้ด้วย (Pan Yi-Ching, 2009b; The European Association for Quality Language Services, 2009; Urmston Alan et al., 2013)

1.1.1.4 สรุปความสำคัญของแบบทดสอบวัดความสามารถภาษาอังกฤษในระดับมหาวิทยาลัยและการใช้หลักการพัฒนาแบบทดสอบให้ได้มาตรฐาน

แบบทดสอบวัดความสามารถภาษาอังกฤษในระดับมหาวิทยาลัยมีประโยชน์ทั้งต่อผู้เรียน ผู้จัดการเรียนการสอน และผู้มีส่วนได้เสียต่อการผลิตบัณฑิต สำหรับผู้เรียน ช่วยให้ทราบระดับความสามารถภาษาอังกฤษของตนเอง และข้อบกพร่องทางภาษาที่ควรปรับปรุงเพื่อยกระดับความสามารถให้สูงขึ้น สำหรับผู้จัดการเรียนการสอน สามารถนำผลการสอบของผู้เรียนมาใช้วางแผนหรือปรับปรุงการสอนให้ดียิ่งขึ้น และสำหรับผู้รับผิดชอบหลักสูตรสามารถใช้ผลการสอบในการวางแผนหรือปรับปรุงหลักสูตร นอกจากนี้ผู้มีส่วนได้เสียต่อการผลิตบัณฑิต จะได้ข้อมูลเพื่อพัฒนาระบบการเรียนการสอนภาษาอังกฤษในระดับมหาวิทยาลัยที่ชัดเจนและตรงเป้าหมาย และสถานประกอบการมีโอกาสดำเนินงานที่มีความสามารถภาษาอังกฤษไว้ทำงานในองค์กร ดังนั้นการพัฒนาแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัยให้ถูกต้องตามหลักการและขั้นตอนในการพัฒนาแบบทดสอบให้ได้มาตรฐาน อีกทั้งให้ผู้เข้าสอบสามารถนำผลการสอบไปอ้างอิง

ได้เสมือนกับได้สอบวัดความสามารถด้วยแบบทดสอบมาตรฐานสากล จะทำให้ผลการสอบที่ได้มีความหมายและเป็นที่ยอมรับในระดับสากลด้วย

1.1.2 สภาพปัญหาของประเทศไทยที่จำเป็นต้องมีแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย

สภาพปัญหาที่ประเทศไทยกำลังประสบอยู่และเห็นว่าการมีแบบทดสอบวัดความสามารถภาษาอังกฤษอาจช่วยแก้ปัญหาดังกล่าวได้ ปัญหาเหล่านั้นเกี่ยวข้องกับความสามารถด้านภาษาของนักศึกษาระดับปริญญาตรี และการวัดความสามารถภาษาอังกฤษด้วยแบบทดสอบที่ไม่อิงเกณฑ์มาตรฐาน ดังรายละเอียดต่อไปนี้

1.1.2.1 สภาพความสามารถของนักศึกษาที่จบปริญญาตรี

การพัฒนาแบบทดสอบวัดความสามารถภาษาอังกฤษสำหรับนักศึกษาระดับมหาวิทยาลัยในประเทศไทย สามารถแก้ปัญหาสำคัญที่พบอยู่ในปัจจุบันนี้ คือ ปัญหาระดับความสามารถภาษาอังกฤษของบัณฑิตไทยอยู่ในระดับต่ำเมื่อเปรียบเทียบกับชาติอื่น มีการศึกษาหลายเรื่องที่มีการสำรวจความสามารถในการใช้ภาษาอังกฤษของบัณฑิตไทยและประชากรไทย และการสำรวจความพึงพอใจของสถานประกอบการต่อคุณภาพบัณฑิตไทยแล้วพบว่าในภาพรวมคนไทยมีความสามารถภาษาอังกฤษอยู่ในระดับต่ำและควรได้รับการปรับปรุง (EF Education First, 2015; นันทิตา เพชรภรณ์, 2553; สำนักพัฒนาคุณภาพการศึกษา มหาวิทยาลัยเชียงใหม่, 2555; สุบิน ยุระรัช, 2556; สุรัชชัย โภคิยะกุล, 2556)

สาเหตุของปัญหาที่คุณภาพการเรียนภาษาอังกฤษของคนไทยต่ำประการหนึ่งคือการไม่มีเครื่องมือวัดคุณภาพภาษาอังกฤษที่อยู่ในระดับสากล ทำให้การเรียนการสอนขาดเป้าหมายที่ชัดเจน ผลสัมฤทธิ์ที่ได้จึงไม่สอดคล้องกับความต้องการของผู้เกี่ยวข้อง หากมีการทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย นักศึกษาอาจจะมีความตั้งใจศึกษาเล่าเรียนอย่างจริงจัง และปลูกฝังนิสัยหรือพฤติกรรมการเรียนรู้ด้วยตนเองมากขึ้นด้วยการใช้เวลาหลังเลิกเรียนหรือนอกห้องเรียนในการศึกษาภาษาอังกฤษเพิ่มขึ้น เพื่อให้มีความสามารถภาษาอังกฤษในระดับที่จะนำไปใช้ประโยชน์ต่อได้ในอนาคต (Arune Wiryachitra, n.d.; Gao, 2010; Jin, 2011; Paltridge & Brian, 2007; Pan Yi-Ching, 2009a)

1.1.2.2 สภาพการวัดความสามารถภาษาอังกฤษในระดับมหาวิทยาลัยที่ไม่อิงเกณฑ์มาตรฐาน

ปัจจุบันนี้การใช้เครื่องมือวัดความสามารถด้านภาษาอังกฤษระดับมหาวิทยาลัยยังไม่อิงมาตรฐานสากล เพราะแบบทดสอบที่ใช้เป็นแบบทดสอบที่สร้างขึ้นเพื่อใช้ภายในแต่ละมหาวิทยาลัย แบบทดสอบบางชนิดวัดเพียงบางทักษะ ไม่สะท้อนความสามารถในการใช้ภาษาอังกฤษในชั้นการสื่อสาร แบบทดสอบที่มีอยู่ส่วนใหญ่ใช้เพื่อคัดเลือกบุคคลเข้าศึกษาต่อในระดับบัณฑิตศึกษา และเป็นแบบทดสอบวัดความสามารถก่อนจบการศึกษาสำหรับนักศึกษาระดับบัณฑิตศึกษามากกว่าที่จะให้นักศึกษาระดับปริญญาตรีสอบด้วย และที่สำคัญแบบทดสอบดังกล่าวยังไม่ได้เทียบคุณภาพกับแบบทดสอบมาตรฐานสากล ทำให้นักศึกษาไม่สามารถใช้ผลการสอบเพื่ออ้างอิงความสามารถภาษาอังกฤษในระดับสากลได้ สำหรับแบบทดสอบวัดความสามารถภาษาอังกฤษมาตรฐานสากลที่ใช้กันอยู่ เช่น TOEIC, TOEFL, และ IELTS นั้นได้มีการเทียบเคียงคุณภาพระหว่างกันแล้ว เพื่อให้สามารถอ้างผลการสอบแทนกันได้ และในปัจจุบันแบบทดสอบมาตรฐานทุกชนิดควรได้รับการเทียบคุณภาพกับกรอบอ้างอิงความสามารถด้านภาษาของสภายุโรป (CEFR) ซึ่งเป็นเกณฑ์ที่ได้รับการยอมรับและเป็นที่ยอมรับกันอย่างกว้างขวาง ดังนั้นหากมีการพัฒนาแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัยที่สามารถเทียบเคียงคุณภาพกับแบบทดสอบมาตรฐานสากลได้รับการเทียบเคียงคุณภาพกับกรอบอ้างอิงความสามารถด้านภาษาของสภายุโรป (CEFR) แล้ว จะทำให้นักศึกษาที่ผ่านการสอบดังกล่าวมีความสามารถภาษาอังกฤษเสมือนผ่านการสอบวัดความสามารถด้วยแบบทดสอบมาตรฐานสากลและความสามารถภาษาอังกฤษของบัณฑิตจะได้รับการยอมรับในระดับสากลอีกด้วย (Arkoudis & Sophie et.al, 2009; Arunee Wiryachitra, n.d.; Gao, 2010; Jin, 2011; Paltridge & Brian, 2007; Pan Yi-Ching, 2009a; The European Association for Quality Language Services, 2009)

1.1.2.3 สรุปสภาพปัญหาการวัดความสามารถภาษาภาษาอังกฤษระดับมหาวิทยาลัย

สรุปได้ว่า ปัญหาการวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย คือ บัณฑิตที่จบการศึกษาส่วนใหญ่มีความสามารถภาษาอังกฤษโดยรวมอยู่ในระดับที่ไม่น่าพอใจ อันเป็นผลมาจากการไม่มีการสอบวัดความสามารถภาษาอังกฤษอย่างเป็นทางการและไม่ได้สร้างความตระหนักให้ทุกฝ่ายเห็นความสำคัญ นอกจากนี้ การใช้แบบทดสอบที่แต่ละมหาวิทยาลัยจัดทำขึ้นใช้ภายใน ยังไม่ได้เทียบมาตรฐานกับแบบทดสอบวัดความสามารถภาษาอังกฤษในระดับสากล ทำให้ไม่สามารถอ้างอิงความสามารถของผู้เข้าสอบกับมาตรฐานสากลได้

1.1.3 ความเป็นมาของการวิจัยด้านการวัดความสามารถภาษาภาษาอังกฤษในระดับมหาวิทยาลัย

1.1.3.1 งานวิจัยด้านการพัฒนาแบบทดสอบวัดความสามารถภาษาภาษาอังกฤษในระดับมหาวิทยาลัยในประเทศไทย

อัจฉรา วงศ์โสธร (2532) ได้ทำการศึกษาวิจัยเพื่อพัฒนาแบบทดสอบอิงปริเฉทภาษาอังกฤษ สำหรับใช้กับนักศึกษาไทยในระดับต่าง ๆ เพื่อพัฒนาแบบทดสอบให้เป็นมาตรฐานที่มีคุณสมบัติตามเกณฑ์โดยกำหนดปริเฉทของการใช้ภาษาในทักษะมาลาทางเสียง และทางตัวอักษร ได้แก่ ฟัง พูด อ่าน เขียน ตามหลักสูตรของแต่ละระดับขึ้นเป็นแบบทดสอบอิงปริเฉท สำหรับมัธยมศึกษาตอนต้น ตอนปลาย และอุดมศึกษา ตลอดจนกำหนดระดับความสามารถในการใช้ภาษาอังกฤษอันเป็นเกณฑ์กลางของนักศึกษาในแต่ละระดับ และวิเคราะห์ปัจจัยความสามารถในการสื่อความหมายทางภาษาอังกฤษของนักศึกษาไทย พบว่าแบบทดสอบทุกระดับมีความเหมาะสมกับระดับความสามารถของนักศึกษาและสอดคล้องกับโครงสร้างของข้อสอบ (ICC) และมีดัชนีความง่ายกับดัชนีความไวในเกณฑ์มาตรฐานเป็นปริมาณร้อยละ 95 ของข้อสอบทั้งหมด ระดับความสามารถของนักศึกษาส่วนใหญ่อยู่ในระดับปานกลางต้องปรับปรุงแก้ไข

นอกจากนี้ อัจฉรา วงศ์โสธร และ ศิริพร พงษ์สุรพิพัฒน์ (2544) ได้ทำวิจัยด้านการสร้างและพัฒนาแบบทดสอบภาษาอังกฤษ เพื่อพัฒนาแบบทดสอบสมรรถภาพทั่วไปทางภาษาอังกฤษด้านการฟัง การอ่าน และการเขียน ได้แบบทดสอบปรนัยวัดทักษะการฟัง อ่าน และเขียน และแบบอัตนัยวัดทักษะการเขียน จำนวน 4 ฉบับ ที่มีความเที่ยงอยู่ในเกณฑ์ดี – ดีมาก และมีข้อเสนอแนะว่า ควรทำการวิจัยเปรียบเทียบระดับสมรรถภาพทางการใช้ภาษาอังกฤษของนิสิต นักศึกษาและบุคคลทั่วไปที่วัดโดยแบบทดสอบของสถาบันภาษาและและแบบทดสอบมาตรฐานที่มีการใช้อย่างแพร่หลาย เช่น แบบทดสอบ TOEFL (Test of English as a Foreign Language) ซึ่ง Educational Testing Service แห่ง Princeton, New Jersey พัฒนาขึ้นใช้ และแบบทดสอบ ELTS (English Language Testing Service) ซึ่ง British Council และ University of Cambridge Local Examinations Syndicate พัฒนาขึ้น

สรุปได้ว่างานวิจัยด้านการพัฒนาแบบทดสอบวัดความสามารถภาษาอังกฤษในระดับมหาวิทยาลัยในประเทศไทย เสนอแนะให้มีการทำวิจัยเพื่อเปรียบเทียบระดับความสามารถทางการใช้ภาษาอังกฤษของผู้เข้าสอบที่วัดโดยแบบทดสอบที่สถาบันสร้างเองและแบบทดสอบมาตรฐาน

1.1.3.2 งานวิจัยด้านการพัฒนาแบบทดสอบวัดความสามารถภาษาอังกฤษในระดับมหาวิทยาลัยในต่างประเทศ

งานวิจัยด้านการพัฒนาแบบทดสอบวัดความสามารถภาษาอังกฤษในระดับมหาวิทยาลัยในต่างประเทศสามารถดูพัฒนาการได้จากการพัฒนาแบบทดสอบ CET (College English Test) ของประเทศจีน แบบทดสอบ GEPT (General English Proficiency Test) ของประเทศไต้หวัน และแบบทดสอบ DELTA (The Diagnostic English Language Tracking Assessment) ของประเทศฮ่องกง โดยมีรายละเอียดดังต่อไปนี้

แบบทดสอบ CET (College English Test) ของประเทศจีน มีพัฒนาการจากแรกเริ่มจนถึงปัจจุบันดังนี้ 1) มีการทดสอบทักษะไวยากรณ์และคำศัพท์ โดยการใช้ข้อคำถามแบบเลือกตอบเพียงอย่างเดียว 2) สำหรับการปรับปรุงแบบทดสอบในยุคต่อมา มีการเพิ่มข้อสอบประเภท Compound Dictation, การให้คำตอบสั้น ๆ และการแปล 3) มีการเพิ่มการอ่านเร็ว (Fast Reading) โดยใช้เทคนิค Skimming and Scanning ในการปรับปรุงแบบทดสอบฉบับใหม่ 4) มีการพัฒนา

แบบทดสอบวัดทักษะการพูด หรือ CET Spoken English Test ในช่วงระหว่างปี 1990s โดยกำหนดให้มีการสอบสัมภาษณ์แบบเผชิญหน้ากับผู้ดำเนินการสอบ การสอบพูดแบบกลุ่ม 3 คน มีการทำภาระงานแบบพูดคนเดียว และพูดกับผู้เข้าสอบคนอื่น การนำเสนอรายบุคคล และการอภิปรายกลุ่ม และ 5) มีการจัดสอบผ่านระบบคอมพิวเตอร์หรืออินเทอร์เน็ต ซึ่งการทดลองครั้งแรกสำหรับแบบทดสอบ IB CET-4 เกิดขึ้นในเดือนธันวาคม ปี 2008 และสำหรับแบบทดสอบ IB CET-6 ในเดือนธันวาคม ปี 2009 (Jin, 2011)

แบบทดสอบ GEPT (General English Proficiency Test) เป็นแบบทดสอบวัดความสามารถทั่วไปภาษาอังกฤษที่ใช้กันอย่างเป็นทางการในประเทศไทยได้ทุกวัน โดยแบ่งแบบทดสอบออกเป็น 5 ระดับ ตามระดับความสามารถภาษาอังกฤษที่กำหนดไว้ในระบบการศึกษาของประเทศได้ทุกวัน คือ ผู้จบการศึกษาในระดับมัธยมต้น ควรมีความสามารถภาษาอังกฤษเทียบเท่าระดับต้น (The Elementary Level) ผู้จบการศึกษาในระดับมัธยม ปลายควรมีความสามารถภาษาอังกฤษเทียบเท่าระดับกลาง (The Intermediate Level) ผู้จบการศึกษาในระดับมหาวิทยาลัยที่ไม่ใช่เอกภาษาอังกฤษ ควรมีความสามารถภาษาอังกฤษเทียบเท่าระดับกลางสูง (The High-Intermediate Level) ผู้จบการศึกษาในระดับมหาวิทยาลัยเอกภาษาอังกฤษ ควรมีความสามารถภาษาอังกฤษเทียบเท่าระดับสูง (The Advanced Level) และระดับที่สูงที่สุด คือ ระดับสูงพิเศษ (The Superior Level) ซึ่งใกล้เคียงกับความสามารถของผู้พูดเจ้าของภาษาที่ได้รับการศึกษาดี การพัฒนาแบบทดสอบ GEPT ในยุคแรก ๆ เริ่มต้นด้วยการทดสอบเฉพาะการฟังและการพูด แต่ในปัจจุบันมีการทดสอบครั้งทั้ง 4 ทักษะคือ การฟัง การอ่าน การเขียน และการพูด และเนื่องจากภาษาอังกฤษเริ่มมีความสำคัญมากขึ้นเรื่อยๆ โดยเฉพาะการติดต่อสื่อสารกับนานาชาติ ทีมงานพัฒนาแบบทดสอบ GEPT จึงได้พยายามสร้างมาตรฐานให้กับแบบทดสอบ GEPT โดยการประเมินคุณภาพแบบทดสอบด้วยวิธีการ คือ 1) หาความเชื่อมั่นของแบบทดสอบ (Reliability) ส่วนใหญ่จะได้ค่าระดับสูง คือไม่ต่ำกว่า .80 และสามารถเปรียบเทียบความเชื่อมั่นได้กับแบบทดสอบที่ถูกให้ความสำคัญมากขึ้น ๆ ด้วย 2) มีการประเมินคุณภาพของแบบทดสอบ GEPT ในแง่มุมต่าง ๆ เช่น การวิเคราะห์ข้อมูลผลการสอบ และข้อมูลย้อนกลับจากผู้เข้าสอบ ความเชื่อมั่นในการให้คะแนน ความตรงคู่ขนาน ความตรงเชิงเนื้อหา ความตรงที่สัมพันธ์กับเกณฑ์ การวิเคราะห์คลังข้อมูลจากผลงานเขียน และการศึกษาผลกระทบจากการสอบ 3) การเทียบเคียงคุณภาพแบบทดสอบกับกรอบอ้างอิงความสามารถในการเรียนภาษาของสภายุโรป (CEFR) ซึ่งเป็นเกณฑ์ภายนอกที่ได้มาตรฐาน 4) การศึกษาความตรงภายใน (Internal Validity) เพื่อหาความสัมพันธ์ของแบบทดสอบแต่ละส่วน และ 5) สิ่งที่จะทำต่อไป คือ การสำรวจความตรงเชิงโครงสร้าง (Construct Validity) ความตรงเชิงผลลัพธ์ที่ตามมา (Consequential Validity) และผลกระทบต่อสังคมและระบบการศึกษา (Wu & Mei, 2012)

แบบทดสอบ DELTA (The Diagnostic English Language Tracking Assessment) ของประเทศฮ่องกง เป็นแบบทดสอบที่พัฒนาขึ้นโดยจัดสอบผ่านเว็บและใช้ตรวจสอบติดตามความสามารถของนักศึกษาในมหาวิทยาลัย 3 แห่งในประเทศฮ่องกง การพัฒนาแบบทดสอบ DELTA อาศัยโครงสร้างความสามารถทางภาษาของ Bachman and Lyle F. & Palmer (1996) และเป็นแบบทดสอบประเภทเลือกตอบอย่างเดียว วัดความสามารถทั้งด้านการฟัง คำศัพท์ การอ่าน และ

ไวยากรณ์ (โดยแบบทดสอบการเขียน และพูดอยู่ในระหว่างการพัฒนา) (Urmston Alan et al., 2013)

1.1.3.3 แนวโน้มการวิจัยด้านความสามารถภาษาอังกฤษในระดับมหาวิทยาลัย

สรุปได้ว่า งานวิจัยในอดีตด้านการทดสอบความสามารถภาษาอังกฤษระดับมหาวิทยาลัยยังมีช่องว่างที่ควรได้รับการเติมเต็ม 3 ประการ คือ 1) แบบทดสอบที่สร้างขึ้นวัดความสามารถของผู้เข้าสอบไม่ครบทุกทักษะ ไม่สะท้อนความสามารถในการใช้ภาษาเพื่อการสื่อสารได้อย่างครบถ้วน 2) การเลือกใช้แบบทดสอบวัดความสามารถภาษาอังกฤษในระดับสากลมาวัดความสามารถด้านภาษาอังกฤษของนักศึกษาไทย อาจเกิดความลำเอียงด้านวัฒนธรรมได้ เพราะแบบทดสอบในระดับสากลอาจใช้บริบทของเนื้อหาสาระข้อสอบที่ไม่เป็นที่คุ้นเคยของนักศึกษาไทย และ 3) ยังไม่มีการใช้กรอบอ้างอิงความสามารถในการเรียนภาษาของสภายุโรป (CEFR) เป็นฐานในการกำหนดเนื้อหาสาระของแบบทดสอบ ดังนั้น การวิจัยเพื่อพัฒนาแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัยครั้งนี้ จึงต้องการพัฒนาแบบทดสอบให้คุณสมบัติดังต่อไปนี้ 3 ประการ คือ 1) เป็นแบบทดสอบที่วัดความสามารถของผู้เข้าสอบครบทุกทักษะ สามารถสะท้อนความสามารถในการใช้ภาษาอังกฤษเพื่อการสื่อสารได้อย่างแท้จริง 2) เนื้อหาสาระของแบบทดสอบไม่มีความลำเอียงด้านวัฒนธรรมและเหมาะสมกับผู้เข้าสอบที่เป็นชาวไทย และ 3) ใช้กรอบอ้างอิงความสามารถการเรียนภาษาของสภายุโรป (CEFR) เป็นฐานในการกำหนดเนื้อหาสาระของแบบทดสอบ

1.1.4 บริบทของการศึกษา

1.1.4.1 สภาพการจัดการเรียนการสอนภาษาอังกฤษพื้นฐานของมหาวิทยาลัยทักษิณ

สภาพปัญหา

นิตินโยบายปริญญาตรีมหาวิทยาลัยทักษิณถูกกำหนดให้เรียนวิชาภาษาอังกฤษที่เป็นวิชาพื้นฐานทั่วไป 2 วิชา คือวิชาภาษาอังกฤษทั่วไป 1 และ 2 แต่สำหรับบางคนอาจจะได้เรียนเพิ่มอีก 1 วิชา คือ วิชาภาษาอังกฤษทั่วไป 3 ในกรณีที่นิตินโยบายเหล่านั้นเลือกเรียนวิชาภาษาอังกฤษทั่วไป 3 เป็นวิชาภาษาเลือก การลงเรียนวิชาภาษาอังกฤษส่วนมากแล้วจะลงตามแผนที่กำหนดในหลักสูตร กล่าวคือไม่เกินชั้นปีที่ 2 นิตินโยบายจะได้เรียนวิชาพื้นฐานทั้ง 3 วิชาเหล่านี้ ครบแล้ว หากเป็นนิตินโยบายที่ไม่ใช่เอกภาษาอังกฤษ หรือด้านบริหารธุรกิจ หรือในกลุ่มวิชาเอกที่มีวิชาภาษาอังกฤษสอดแทรกอยู่ นิตินโยบายดังกล่าวจะไม่ได้เรียนภาษาอังกฤษอีกเลยจนกระทั่งจบการศึกษา

จากการวิเคราะห์ผลการเรียนในรายวิชาภาษาอังกฤษทั่วไป 1-3 ประจำปีการศึกษา 2557 พบว่า ในภาพรวมนิตินโยบายมีความสามารถภาษาอังกฤษอยู่ในระดับอ่อนถึงปานกลางจำนวน 59.96 % ที่เหลือเป็นนิตินโยบายที่มีความสามารถอยู่ในระดับมากและมากที่สุด จำนวน 33.17 % และกลุ่มที่มีความสามารถระดับอ่อนมาก ไม่สามารถประเมินผลได้ และที่ต้องลงทะเบียนเรียนใหม่ จำนวน 6.88

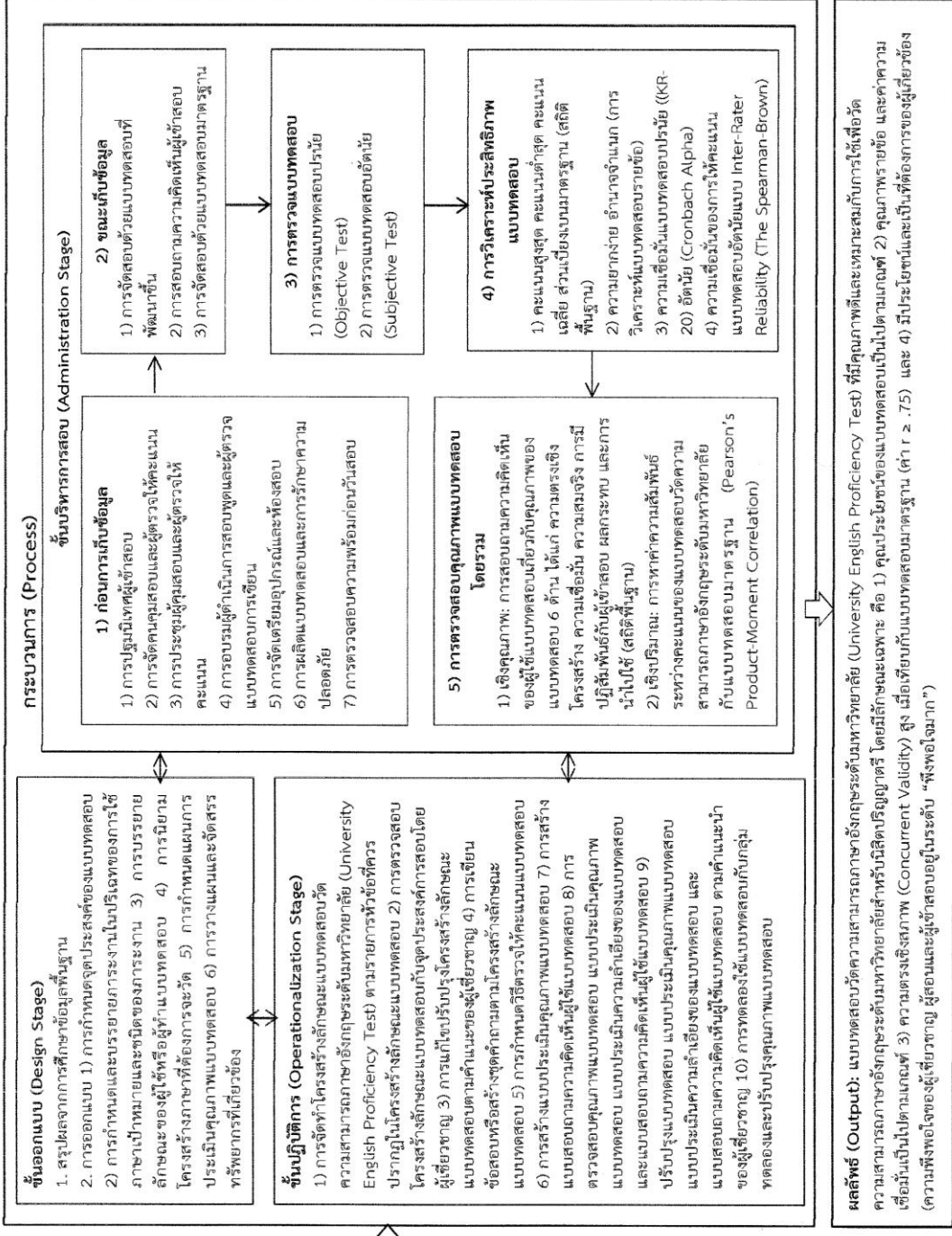
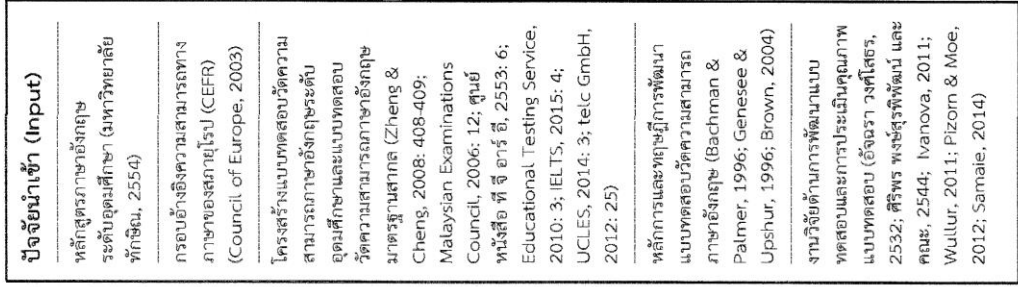
% (ฝ่ายวิชาการและประกันคุณภาพการศึกษา มหาวิทยาลัยทักษิณ, 2557) ผลสัมฤทธิ์ในการเรียน วิชาภาษาอังกฤษของนิสิตที่ไม่พึงประสงค์เช่นนี้ อาจมีสาเหตุมาจาก 2 ส่วน คือ จากนิสิต และจาก อาจารย์ผู้สอนและระบบบริหารจัดการ สาเหตุของผลสัมฤทธิ์ในการเรียนต่ำจากนิสิต อาจเป็นเพราะ นิสิตเองขาดการเอาใจใส่หรือกระตือรือร้นในการเรียน ขาดการศึกษาหาความรู้เพิ่มเติมนอกเวลาเรียน และไม่เห็นความสำคัญหรือไม่ชอบเรียนวิชาภาษาอังกฤษอยู่ก่อนแล้ว ส่วนสาเหตุที่อาจจะมีมาจาก อาจารย์ผู้สอนและระบบบริหารจัดการ คือ จำนวนนิสิตในแต่ละกลุ่มมีจำนวนมากเกินไป อาจารย์ไม่สามารถดูแลได้ทั่วถึง ส่งผลให้การจัดการเรียนการสอนไม่ได้เน้นให้นิสิตได้ฝึกการใช้ภาษาอังกฤษเพื่อ การสื่อสารเท่าที่ควร อีกทั้งระบบการประเมินความสามารถของนิสิตมีการสอบระหว่างภาคและปลาย ภาคที่เน้นแบบทดสอบเขียน วัดความรู้ด้านคำศัพท์ ไวยากรณ์ การใช้ภาษาในการสนทนา การอ่าน และการเขียนในระดับประโยค สำหรับการมอบหมายงานหรือการทดสอบย่อยด้านการพูดและเขียน ขึ้นอยู่กับการตกลงของอาจารย์และนิสิตเอง (ฝ่ายวิชาการและประกันคุณภาพการศึกษา มหาวิทยาลัยทักษิณ, 2557)

ความจำเป็นที่จะต้องมีการวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย

การมีแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test) จะช่วยผลักดันทั้งนิสิตและอาจารย์ให้เห็นความสำคัญและประโยชน์ของการเรียน การสอนภาษาอังกฤษมากขึ้น ในส่วนของนิสิตจะเห็นความสำคัญของการเรียนวิชาภาษาอังกฤษมากขึ้น เอาใจใส่ กระตือรือร้น ตั้งใจเรียนและหมั่นทบทวนบทเรียน ศึกษาหาความรู้เพิ่มเติม ฝึกหา โอกาสในการใช้ภาษาอังกฤษนอกห้องเรียนมากยิ่งขึ้น นอกจากนี้ การสอบวัดความรู้ยังช่วยให้นิสิต ทราบระดับความสามารถของตนเอง และจุดบกพร่องด้านภาษาของตน และสามารถหาแนวทางใน การปรับปรุงแก้ไขให้ดีขึ้นได้ (Bachman & Lyle F. & Palmer, 1996) ในกรณีที่มีการจัดสอบด้วย แบบทดสอบวัดความสามารถภาษาอังกฤษในระดับสากล นิสิตจะได้มีการผลการสอบวัดระดับ ความสามารถที่เป็นเครื่องยืนยันความสามารถในระดับสากล และอาจจะมีโอกาสมากในการได้รับการ คัดเลือกเข้าทำงาน (Jobsdb.com, 2015) ในส่วนของอาจารย์การทดสอบวัดความสามารถ ภาษาอังกฤษระดับมหาวิทยาลัยจะเป็นตัวผลักดันให้อาจารย์ปรับเปลี่ยนรูปแบบการสอนเพื่อเน้นให้ นิสิตได้ฝึกใช้ภาษาอังกฤษเพื่อการสื่อสารมากขึ้น ปรับระบบการวัดผลการเรียนที่เน้นการใช้ ภาษาอังกฤษเพื่อการสื่อสาร (Cohen & Andrew D, 1994; Heaton, 1975) ซึ่งจะสะท้อนรูปแบบ การใช้ภาษาเป้าหมายในสถานการณ์จริงของผู้เรียนมากยิ่งขึ้น นอกจากนี้สามารถนำข้อมูลย้อนกลับ จากการทดสอบวัดความรู้ไปใช้ในการปรับปรุงระบบการจัดการเรียนการสอนวิชาภาษาอังกฤษ พื้นฐานให้เกิดประโยชน์อย่างแท้จริง และช่วยกันยกระดับความสามารถภาษาอังกฤษของนิสิตให้ สูงขึ้น

1.2 กรอบแนวคิดวิจัย

1.2 กรอบแนวคิดการวิจัย



การวิจัยเรื่องการพัฒนาแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test) มีวัตถุประสงค์ 2 ข้อ คือ

1.) เพื่อสร้างและพัฒนาแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test) ให้เป็นแบบทดสอบมาตรฐานที่มีลักษณะเฉพาะ คือ

- คุณประโยชน์ของแบบทดสอบเป็นไปตามเกณฑ์ที่กำหนด
- คุณภาพรายข้อและความเชื่อมั่นเป็นไปตามเกณฑ์ที่กำหนด
- ความตรงร่วมสมัย (Concurrent Validity) สูง เมื่อเทียบกับแบบทดสอบมาตรฐาน

2.) เพื่อศึกษาความพึงพอใจของผู้เกี่ยวข้องต่อแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test)

1.3 ขอบเขตการวิจัย

ขอบเขตในการวิจัยครั้งนี้ครอบคลุมประชากร กลุ่มตัวอย่าง และตัวแปรที่ศึกษา อันมีรายละเอียดดังตารางต่อไปนี้

1.3.1 ประชากร

ประชากรในการศึกษาครั้งนี้ เป็นนิสิตระดับปริญญาตรี ภาคปกติ ชั้นปีที่ 1-4 ที่ลงทะเบียนและกำลังศึกษาอยู่ในภาคเรียนที่ 2 ปีการศึกษา 2561 จำนวน 8,167 คน ประกอบด้วย 5 คณะ คือ

- 1) คณะมนุษยศาสตร์และสังคมศาสตร์ 2) คณะศึกษาศาสตร์ 3) คณะศิลปกรรมศาสตร์ 4) คณะเศรษฐศาสตร์และบริหารธุรกิจ และ 5) คณะนิติศาสตร์

1.3.2 กลุ่มตัวอย่าง

กลุ่มตัวอย่างในการศึกษาครั้งนี้ ได้มาโดยการรับสมัครกลุ่มสนใจที่ต้องการสอบวัดความสามารถภาษาอังกฤษ โดยการทำหนังสือประชาสัมพันธ์การทดสอบไปยังคณะต่าง ๆ การประชาสัมพันธ์ในสื่อสังคมออนไลน์ และการเข้าพบนิสิตในห้องเรียน ในมหาวิทยาลัยทักษิณวิทยาเขตสงขลา ซึ่งกำหนดเป็นกรอบในการรับสมัครกลุ่มตัวอย่างไว้ดังนี้

ขั้นตอนที่ 1 ผู้วิจัยกำหนดขนาดกลุ่มทดลองและกลุ่มตัวอย่างที่สามารถบริหารจัดการได้ในการเก็บข้อมูลแต่ละครั้ง โดยการทดสอบข้อสอบปรนัย สามารถมีจำนวนผู้เข้าสอบสูงสุด คือ 300 คน โดยแบ่งออกเป็น 10 ห้องสอบ ห้องสอบละ 30 คน และใช้กรรมการคุมสอบ ห้องละ 2 คน รวมใช้กรรมการ 10 คน และการทดสอบการพูด สามารถมีจำนวนผู้เข้าสอบสูงสุด คือ 45 คน โดยเลือกมาจากผู้เข้าสอบที่ได้คะแนนในกลุ่มสูง ปานกลาง และต่ำ กลุ่มละ 15 คน และใช้กรรมการทดสอบจำนวน 3 คน

ขั้นตอนที่ 2 การจะนำผลคะแนนสอบมาหาความสัมพันธ์กันระหว่างแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test) และแบบทดสอบวัดความสามารถภาษาอังกฤษมาตรฐาน ผู้เข้าสอบจะต้องเข้าสอบครบ 2 ครั้ง และมีคะแนนที่สามารถนำมาเทียบเคียงได้ หากผู้เข้าสอบขาดสอบครั้งใดครั้งหนึ่ง จะไม่สามารถนำคะแนนมาหาค่าความสัมพันธ์ได้

1.3.3 ตัวแปรที่ศึกษา

กำหนดตัวแปรที่ใช้ในการวิจัยดังนี้

- 1) คะแนนที่ได้จากการทดสอบทักษะการฟัง การใช้ภาษา การอ่าน การเขียนและการพูดของแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test) และแบบทดสอบวัดความสามารถภาษาอังกฤษมาตรฐาน (แบบทดสอบปรนัยใช้คะแนนจากแบบทดสอบ TOEFL และแบบทดสอบอัตนัยใช้คะแนนจากแบบทดสอบ IELTS)
- 2) ความพึงพอใจของผู้เชี่ยวชาญ อาจารย์ผู้สอนภาษาอังกฤษ และผู้เข้าสอบต่อแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test)

1.4 ข้อยกจำกัดในการวิจัย

การวิจัยครั้งนี้มีข้อจำกัด คือ จำนวนของกลุ่มทดลองและกลุ่มตัวอย่างน้อยเกินไป โดยการทดลองครั้งที่ 1 มีกลุ่มทดลองจำนวน 28 คน การทดลองครั้งที่ 2 จำนวน 94 คน และการเก็บข้อมูลครั้งที่ 3 มีกลุ่มตัวอย่างจำนวน 45 คน กลุ่มทดลองและกลุ่มตัวอย่างได้มาโดยการรับสมัครนิสิตที่สนใจจะสอบวัดความสามารถภาษาอังกฤษ จำนวนดังกล่าวอาจไม่เพียงพอที่จะเป็นตัวแทนของประชากรในการวิจัยครั้งนี้

1.5 นิยามศัพท์เฉพาะ

คำศัพท์สำคัญที่ปรากฏในงานวิจัยควรได้รับการให้คำนิยามเพื่อให้เข้าใจถึงบริบทในการวิจัยครั้งนี้ คำศัพท์ที่สำคัญได้แก่ แบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test) แบบทดสอบวัดความสามารถภาษาอังกฤษระดับมาตรฐานสากล คุณภาพแบบทดสอบ ความพึงพอใจในการใช้แบบทดสอบ ผู้เข้าสอบ การประเมินคุณภาพแบบทดสอบ (Test Validation) และความเป็นธรรม (Fairness)

1.5.1 แบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test)

แบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test) คือ แบบทดสอบวัดความสามารถภาษาอังกฤษที่นักศึกษาในระดับปริญญาตรีทั่วไปควรสอบเพื่อวัดความสามารถภาษาอังกฤษในชีวิตประจำวันและภาษาอังกฤษเชิงวิชาการที่ผู้วิจัยสร้างขึ้นเพื่อวัดความสามารถภาษาอังกฤษของนักศึกษาในระดับมหาวิทยาลัย ใน 5 ทักษะ คือ 1) การฟัง 2) การใช้ภาษา 3) การอ่าน 4) การเขียน และ 5) การพูด มีเนื้อหาสาระที่ได้มาจากการสังเคราะห์เนื้อหาข้อสอบจาก 1) หลักสูตรภาษาอังกฤษระดับอุดมศึกษา 2) กรอบอ้างอิงความสามารถทางภาษาของสภายุโรป (CEFR) และ 3) โครงสร้างแบบทดสอบวัดความสามารถภาษาอังกฤษระดับอุดมศึกษาและแบบทดสอบวัดความสามารถภาษาอังกฤษมาตรฐานสากล เป็นแบบทดสอบที่สร้างขึ้นตามกระบวนการสร้างข้อสอบตามแนวทางของ (Bachman & Lyle F. & Palmer, 1996) และเป็นแบบทดสอบที่มีคุณภาพของแบบทดสอบทั้งฉบับสามารถเทียบเคียงกับแบบทดสอบวัดความสามารถภาษาอังกฤษมาตรฐานสากล

1.5.2 แบบทดสอบวัดความสามารถภาษาอังกฤษระดับมาตรฐานสากล

แบบทดสอบวัดความสามารถภาษาอังกฤษระดับมาตรฐานสากลที่นำมาใช้มี 2 ชนิด คือ 1) แบบทดสอบ TOEFL ITP (Test of English as a Foreign Language – Institutional Testing Program) ซึ่งใช้เทียบคะแนนของแบบทดสอบปรนัย คือ ทักษะการฟัง การใช้ภาษา และการอ่าน และ 2) แบบทดสอบ IELTS (International English Language Testing System) ซึ่งใช้เทียบคะแนนของแบบทดสอบอัตนัย คือ ทักษะการเขียน และการพูด

1.5.3 ความตรงเชิงเนื้อหา (Content Validity)

ความตรงเชิงเนื้อหา (Content Validity) หมายถึง ความสอดคล้องระหว่างข้อคำถามกับวัตถุประสงค์การสอบ ในการวิจัยครั้งนี้ มีการตรวจสอบความตรงเชิงเนื้อหาโดยให้ผู้เชี่ยวชาญด้านการสอนภาษาอังกฤษระดับมหาวิทยาลัยและด้านการวัดและประเมินผลการสอนภาษาอังกฤษ ตรวจสอบความสอดคล้องระหว่างตารางกำหนดโครงสร้างแบบทดสอบ (Test Blueprint) กับตัวข้อคำถามโดยใช้แบบประเมินค่า IOC ที่ผู้วิจัยสร้างขึ้น

1.5.4 ความตรงร่วมสมัย (Concurrent Validity)

ความตรงร่วมสมัย (Concurrent Validity) หมายถึง ความสัมพันธ์ระหว่างคะแนนของผู้เข้าสอบคนใดคนหนึ่งที่ได้จากการสอบแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัยและจากแบบทดสอบวัดความสามารถภาษาอังกฤษมาตรฐาน ตรวจสอบโดยการหาค่าความสัมพันธ์ระหว่างคะแนนของแบบทดสอบทั้ง 2 ชุด โดยใช้ Pearson's Product-Moment Correlation Coefficient โดยกำหนดค่าความสัมพันธ์ที่ยอมรับได้ คือ $r_{xy} \geq 0.75$

1.5.5 ความตรงเชิงโครงสร้าง (Construct Validity)

ความตรงเชิงโครงสร้าง (Construct Validity) หมายถึง ความสอดคล้องระหว่างตัวแบบทดสอบแต่ละส่วนกับองค์ประกอบที่แบบทดสอบที่ดีพึงมี ดำเนินการตรวจสอบโดยผู้เชี่ยวชาญและอาจารย์ผู้สอนภาษาอังกฤษระดับมหาวิทยาลัยประเมินแบบทดสอบโดยใช้แบบประเมินคุณประโยชน์ของแบบทดสอบตามแนวคิดของ (Bachman & Lyle F. & Palmer, 1996) โดยมีประเด็นที่ตรวจสอบ คือ 1) มีการระบุวัตถุประสงค์การสอบชัดเจนหรือไม่ 2) วัตถุประสงค์การสอบสอดคล้องกับจุดประสงค์หลักหรือไม่ 3) ภาระงานการสอบสอดคล้องกับจุดประสงค์หลักหรือไม่ 4) วิธีการตรวจให้คะแนนเหมาะสมหรือไม่ และคะแนนสอบสะท้อนความสามารถของผู้เข้าสอบหรือไม่ และ 5) คะแนนสอบสามารถทำนายความสามารถในอนาคตของผู้เข้าสอบได้หรือไม่ ในการวิจัยครั้งนี้ คาดว่าแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัยควรมีความตรงเชิงโครงสร้างอยู่ในระดับมาก ($\bar{X} \geq 3.5$)

1.5.6 ความลำเอียงของข้อคำถาม (Item Bias)

ความลำเอียงของข้อคำถาม (Item Bias) หมายถึง ความได้เปรียบหรือเสียเปรียบในการทำแบบทดสอบได้ระหว่างผู้เข้าสอบสายวิทย์และสายศิลป์ ข้อคำถามที่มีความลำเอียงคือ ข้อที่สนับสนุนให้ผู้เข้าสอบฝ่ายหนึ่งฝ่ายใดได้คะแนนมากกว่าอีกฝ่าย คำนวณหาค่าความลำเอียงของข้อคำถามแต่ละข้อโดยใช้วิธีแปลงค่าความยาก ตามแนวคิดของ Angoff and Ford (1973) และข้อคำถามที่มีความลำเอียง คือ ข้อที่มีค่าความลำเอียงมากกว่า +0.75 และน้อยกว่า -0.75

1.5.7 ดัชนีความสอดคล้องระหว่างข้อคำถามและวัตถุประสงค์ (Item-Objective Congruence Index)

ดัชนีความสอดคล้องระหว่างข้อคำถามและวัตถุประสงค์ (Item-Objective Congruence Index) หมายถึง ค่าที่ได้จากการตรวจสอบความตรงเชิงเนื้อหาของแบบทดสอบโดยใช้แบบประเมิน IOC ของผู้เชี่ยวชาญ 3 คน เพื่อกำหนดค่าความสอดคล้องระหว่างข้อคำถามและวัตถุประสงค์ ตามแนวคิดของ Rovinelli and Hambleton (1976) โดยข้อคำถามที่มีความสอดคล้องระหว่างข้อคำถามและวัตถุประสงค์ควรมีค่า $IOC \geq 0.5$

1.5.8 คุณภาพแบบทดสอบ

คุณภาพแบบทดสอบ หมายถึง แบบทดสอบที่ดำเนินการสร้างมีค่าคุณภาพตามเกณฑ์คุณลักษณะของแบบทดสอบที่ดี โดยมี 1) คุณภาพโดยรวมและรายด้านอยู่ในระดับ “คุณภาพดี” ($\bar{X} \geq 3.5$) 2) ค่าความยากง่าย (P) อยู่ระหว่าง 0.20 - 0.80 3) ค่าอำนาจจำแนก (R) ตั้งแต่ 0.20 ขึ้นไป 4) ค่าความเชื่อมั่น (Reliability) 0.75 ขึ้นไป 5) ค่าความเชื่อมั่นของผู้ตรวจให้คะแนน (Inter-rater Reliability) 0.75 ขึ้นไป และ 6) ค่าความสัมพันธ์ระหว่างคะแนนจากแบบทดสอบที่สร้างขึ้นกับแบบทดสอบมาตรฐานสากล (Pearson's Product-Moment Correlation) 0.75 ขึ้นไป

1.5.9 ความพึงพอใจในการใช้แบบทดสอบ

ความพึงพอใจในการใช้แบบทดสอบ หมายถึง คะแนนที่ได้แบบประเมินความพึงพอใจในการใช้แบบทดสอบ ที่มีผู้เชี่ยวชาญ อาจารย์ผู้สอนภาษาอังกฤษ และผู้เข้าสอบเป็นผู้ประเมินแบบทดสอบใน 6 ด้าน คือ 1) คำสั่งของแบบทดสอบ 2) เนื้อหาสาระของแบบทดสอบ 3) รูปแบบและการจัดลำดับความ 4) การตรวจให้คะแนน 5) ประโยชน์ของแบบทดสอบ และ 6) การจัดการสอบ และมีความพึงพอใจอยู่ในระดับ “พึงพอใจมาก” ($\bar{X} \geq 3.5$)

1.5.10 ผู้เข้าสอบ

ผู้เข้าสอบ หมายถึง นิสิตชั้นปีที่ 1-4 สาขาวิชาต่าง ๆ ในคณะต่างๆ ของมหาวิทยาลัยทักษิณ วิทยาเขตสงขลา และมีความประสงค์ที่จะทดสอบวัดความสามารถภาษาอังกฤษของตนเอง และลงทะเบียนเลือกรอบเข้าสอบด้วยความสมัครใจ

1.5.11 การประเมินคุณภาพแบบทดสอบ (Test Validation)

การประเมินคุณภาพแบบทดสอบ (Test Validation) หมายถึง การประเมินคุณภาพประโยชน์ของแบบทดสอบที่สร้างขึ้น 7 ด้าน ได้แก่ 1) ความตรงเชิงโครงสร้าง (Construct Validity) 2) ความเชื่อมั่น (Reliability) 3) ความสมจริง (Authenticity) 4) การมีปฏิสัมพันธ์ (Interactiveness) 5) ผลกระทบ (Impact) 6) ความสะดวกใช้ (Practicality) และ 7) ความเป็นธรรม (Fairness) โดยผู้เชี่ยวชาญด้านการสอนภาษาและการทดสอบภาษา การประเมินคุณภาพแบบทดสอบรายชื่อ ค่าความเชื่อมั่นให้เป็นไปตามเกณฑ์ และการเทียบเคียงคะแนนผลการสอบของตอนต่าง ๆ ของแบบทดสอบที่สร้างขึ้นกับแบบทดสอบมาตรฐาน

1.5.12 ความเป็นธรรม (Fairness)

ความเป็นธรรม (Fairness) หมายถึง คุณลักษณะที่แบบทดสอบพึงมี โดยเนื้อหาสาระของแบบทดสอบเหมาะสมกับผู้เข้าสอบทุกคน และเป็นหัวเรื่องที่ปราศจากความลำเอียง โดยไม่มีการกล่าวถึงการแบ่งแยกเพศ เชื้อชาติ สถานทางเศรษฐกิจและสังคม และไม่นำไปสู่ความรุนแรงใด ๆ

1.6 ประโยชน์ที่คาดว่าจะได้รับ

การวิจัยครั้งนี้ให้คุณประโยชน์แก่ผู้เกี่ยวข้องหลายฝ่ายด้วยกัน ได้แก่ ผู้เรียน ผู้จัดการเรียนการสอน และสถาบันและผู้มีส่วนได้ส่วนเสียต่อการผลิตบัณฑิต

1.6.1 ประโยชน์ต่อผู้เรียน

การได้สอบวัดความสามารถภาษาอังกฤษด้วยแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย มีประโยชน์ต่อผู้เรียนหลายประการ เช่น ช่วยให้ผู้เรียน

1. ประหยัดค่าใช้จ่ายมากกว่าการสอบวัดความสามารถภาษาอังกฤษด้วยแบบสอบมาตรฐานสากล
2. ทราบระดับความสามารถภาษาอังกฤษของตนเอง และข้อบกพร่องทางภาษาที่ควรปรับปรุงเพื่อยกระดับความสามารถของตนเอง
3. มีผลการรับรองความสามารถด้านภาษาอังกฤษที่สามารถอ้างอิงกับมาตรฐานสากล เพื่อใช้ในการสมัครงานและสมัครเรียนต่อในประเทศต่าง ๆ ได้
4. สามารถนำความรู้ที่ได้จากการสอบไปใช้ในการประกอบอาชีพและการดำเนินชีวิตประจำวันของตนเอง

1.6.2 ประโยชน์ต่อผู้จัดการเรียนการสอน

นอกจากแบบทดสอบวัดความสามารถภาษาอังกฤษก่อนจบการศึกษาระดับมหาวิทยาลัยจะมีคุณประโยชน์แก่ผู้เรียนแล้ว ผู้จัดการเรียนการสอน อันได้แก่ ผู้สอนและผู้รับผิดชอบหลักสูตรได้รับผลกระทบที่ดีจากการใช้แบบทดสอบฉบับนี้ด้วย โดยอาจารย์ผู้สอนสามารถ

1. ใช้ผลการสอบของนักเรียนที่ตนเองสอน มาใช้ในการวางแผนหรือปรับปรุงการสอนให้สอดคล้องกับผู้เรียนและสถานการณ์รอบด้าน
 2. จัดการเรียนการสอนอย่างเต็มศักยภาพ และให้สอดคล้องกับเกณฑ์มาตรฐานการสอนภาษาของมหาวิทยาลัยและเกณฑ์มาตรฐานการเรียนรู้ภาษาอังกฤษในระดับสากล
 3. เป็นผู้กำหนดวิวัฒนาการการสอนภาษาอังกฤษ และพร้อมที่จะปรับเปลี่ยนแนวการสอนให้ทันสมัยอยู่เสมอ
- และสำหรับผู้รับผิดชอบหลักสูตรสามารถ

1. ใช้ผลการสอบของนักเรียน มาใช้ในการวางแผนหรือปรับปรุงหลักสูตรให้สอดคล้องกับมาตรฐานในระดับมหาวิทยาลัยและมาตรฐานสากล
2. จัดกิจกรรมนอกชั้นเรียนเพื่อเสริมสร้างทักษะและความสามารถทางภาษาอังกฤษของผู้เรียนมากยิ่งขึ้น
3. กำกับให้อาจารย์ผู้สอนจัดการเรียนการสอนภาษาอังกฤษโดยคำนึงมาตรฐานที่ผู้เรียนจะต้องปฏิบัติได้ทั้งที่เป็นมาตรฐานของมหาวิทยาลัยและมาตรฐานสากล

1.6.3 ประโยชน์ต่อสถาบันและผู้มีส่วนได้ส่วนเสียต่อการผลิตบัณฑิต

แบบทดสอบวัดความสามารถภาษาอังกฤษก่อนจบการศึกษาได้ส่งผลประโยชน์ต่อสถาบันและผู้มีส่วนได้ส่วนเสียในการผลิตบัณฑิตอีกด้วย ในด้านประโยชน์ต่อสถาบันนั้น แบบทดสอบวัดความสามารถภาษาอังกฤษก่อนจบการศึกษาระดับมหาวิทยาลัยช่วยให้ผู้บริหารสามารถ

1. ได้ข้อมูลเพื่อพัฒนาระบบการเรียนการสอนภาษาอังกฤษในระดับมหาวิทยาลัยที่ชัดเจนและตรงเป้าหมาย
2. ผลิตบัณฑิตที่ได้คุณภาพ เป็นที่ยอมรับและต้องการของสถานประกอบการ และเสริมสร้างชื่อเสียงให้แก่มหาวิทยาลัย
3. ฝ่ายวิชาการ มหาวิทยาลัยทักษิณ มีแบบทดสอบมาตรฐานที่ใช้วัดความสามารถภาษาอังกฤษของผู้เรียนก่อนจบการศึกษา

สำหรับประโยชน์ของแบบทดสอบวัดความสามารถภาษาอังกฤษต่อผู้มีส่วนได้ส่วนเสียต่อการผลิตบัณฑิตดูได้จากผลประโยชน์ของแบบทดสอบต่อสถานประกอบการ เพราะแบบทดสอบฉบับนี้จะช่วยให้สถานประกอบการ

1. ได้พนักงานที่มีความสามารถภาษาอังกฤษเทียบเท่ามาตรฐานสากล
2. ประหยัดงบประมาณในการจัดคอร์สเพื่อฝึกอบรมภาษาอังกฤษเพิ่มเติมให้แก่พนักงานประจำปี
3. สามารถขยายเครือข่ายหรือทำธุรกิจกับต่างชาติได้ง่ายขึ้น และทำให้องค์กรเป็นที่รู้จักในระดับสากล



บทที่ 2

วรรณกรรมที่เกี่ยวข้อง

ในการวิจัยและพัฒนาแบบทดสอบเพื่อใช้วัดความสามารถภาษาอังกฤษของผู้เรียนในระดับมหาวิทยาลัย ผู้วิจัยได้ดำเนินการศึกษาวรรณกรรมที่เกี่ยวข้อง แนวคิด หลักการต่างๆ ได้แก่ 1) แบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test) 2) การพัฒนาแบบทดสอบวัดความสามารถภาษาอังกฤษ และ 3) งานวิจัยที่เกี่ยวข้องกับการพัฒนาแบบทดสอบและการประเมินคุณภาพแบบทดสอบวัดความสามารถภาษาอังกฤษ เพื่อนำข้อมูลมาเป็นแนวทางในการพัฒนาเครื่องมือในบทที่ 3 โดยมีลำดับการนำเสนอ ดังนี้

2.1 แบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test)

- 2.1.1 ความหมายและผลกระทบของแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test) ต่อผู้เกี่ยวข้อง
 - 2.1.1.1 ความหมายของแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test)
 - 2.1.1.2 ผลกระทบของแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test) ต่อผู้เกี่ยวข้อง
- 2.1.2 เนื้อหาสาระที่เกี่ยวข้องกับการสร้างแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test)
 - 2.1.2.1 หลักสูตรภาษาอังกฤษระดับมหาวิทยาลัย (มหาวิทยาลัยทักษิณ, 2544)
 - 2.1.2.2 กรอบมาตรฐานความสามารถในการเรียนภาษาอังกฤษของสภายุโรประดับ (Council of Europe, 2003)
 - 2.1.2.3 แบบทดสอบวัดความสามารถภาษาอังกฤษมาตรฐานสากล
 - 2.1.2.4 แบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test)
 - 2.1.2.5 สรุปเนื้อหาสาระสำหรับการสร้างแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test)

2.2 การพัฒนาแบบทดสอบวัดความสามารถภาษาอังกฤษให้เป็นแบบทดสอบมาตรฐาน

- 2.2.1 องค์ประกอบที่เกี่ยวข้องกับการพัฒนาแบบทดสอบ
 - 2.2.1.1 ขั้นตอนในการพัฒนาแบบทดสอบ
 - 2.2.1.2 ตารางกำหนดลักษณะแบบทดสอบ (Table of Test Specifications)
 - 2.2.1.3 หลักการเขียนข้อคำถามในแบบทดสอบภาษาให้มีคุณภาพดี

2.2.2 การประเมินคุณภาพแบบทดสอบวัดความสามารถภาษาอังกฤษ

2.2.2.1 ความเชื่อมั่น (Reliability)

2.2.2.2 ความตรงและการประเมินความตรง (Validity and Validation)

2.2.2.3 ความลำเอียงและความยุติธรรมในการทดสอบ

2.2.2.4 การวิเคราะห์แบบทดสอบรายข้อ (Item Analysis)

2.2.2.5 การสอบถามความคิดเห็นผู้เกี่ยวข้องโดยใช้แบบสอบถาม

2.3 งานวิจัยที่เกี่ยวข้องกับการพัฒนาแบบทดสอบและการประเมินคุณภาพแบบทดสอบวัดความสามารถภาษาอังกฤษ

2.3.1 งานวิจัยที่เกี่ยวข้องกับการพัฒนาแบบทดสอบวัดความสามารถภาษาอังกฤษ

2.3.2 งานวิจัยด้านการประเมินคุณภาพแบบทดสอบวัดความสามารถภาษาอังกฤษ



2.1 แบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test)

2.1.1 ความหมายและผลกระทบของแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test) ต่อผู้เกี่ยวข้อง

2.1.1.1 ความหมายของแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test)

ความหมายของแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test) ได้มาจากการศึกษาและวิเคราะห์ความหมายของแบบทดสอบวัดความสามารถภาษาอังกฤษโดยทั่วไป ที่ผู้เชี่ยวชาญด้านการทดสอบภาษาอังกฤษได้กล่าวไว้ ดังนี้ Alderson (1995) แบบทดสอบที่วัดโดยไม่ต้องอาศัยพื้นฐานของหลักสูตรใด ๆ และถูกออกแบบเพื่อใช้ทดสอบความสามารถของนักเรียนที่มีภูมิหลังด้านการเรียนภาษาที่แตกต่างกัน สำคัญของแบบทดสอบมักจะอาศัยการวิเคราะห์ความต้องการตามวัตถุประสงค์ของการสอบแต่ละประเภท ตัวอย่างเช่น แบบทดสอบวัดความสามารถสำหรับผู้ควบคุมการจราจรทางอากาศ ควรเกี่ยวข้องกับทักษะภาษาที่จำเป็นในหอคอยควบคุมการจราจรทางอากาศ เป็นต้น

Hughes (1989) แบบทดสอบวัดความสามารถที่ถูกออกแบบเพื่อวัดความสามารถด้านภาษาของบุคคล โดยไม่คำนึงว่าบุคคลเหล่านั้นจะต้องมีความรู้พื้นฐานใดมาก่อน ดังนั้น สารของแบบทดสอบวัดความสามารถจึงไม่เกี่ยวข้องกับพื้นฐานหรือจุดประสงค์ของรายวิชาด้านภาษาที่ผู้สอบอาจจะต้องเรียนมา แต่แบบทดสอบจะทดสอบตามตารางกำหนดลักษณะแบบทดสอบ ที่แสดงว่าผู้เข้าสอบต้องสามารถใช้ภาษาเพื่อทำอะไรได้ถึงจะได้รับการพิจารณาว่าเป็นผู้มีความสามารถทางภาษา

Heaton (1975) แบบทดสอบวัดความสามารถไม่มีความเกี่ยวข้องใด ๆ กับหลักสูตรหรือการสอน และสามารถใช้นักเรียนที่มาจากโรงเรียนและประเทศต่าง ๆ หรือที่มีภูมิหลังทางภาษาที่แตกต่างกัน แบบทดสอบจะเกี่ยวข้องกับการวัดการควบคุมการใช้ภาษาของนักเรียน และวัดในสิ่งที่ถูกคาดหวังว่านักเรียนจะต้องทำหรือได้ปฏิบัติในอนาคต ตัวอย่างเช่น นักเรียนมีความรู้ภาษาอังกฤษเพียงพอที่จะเรียนรู้รายวิชาต่าง ๆ ในระดับมหาวิทยาลัยหรือวิทยาลัยบางแห่งที่ใช้ภาษาอังกฤษเป็นสื่อกลางในการสอนหรือไม่ นักเรียนมีความรู้ภาษาอังกฤษเพียงพอที่จะปฏิบัติหน้าที่อย่างมีประสิทธิภาพในตำแหน่งงานที่ได้สมัครหรือไม่ ดังนั้น แบบทดสอบวัดความสามารถจึงไม่ได้วัดทักษะในภาพกว้าง แต่วัดทักษะเฉพาะอย่างของทักษะภาษาที่จำเป็นต้องใช้ในการศึกษาหรือการทำงานในอนาคตของนักเรียน

สรุปได้ว่า แบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test) หมายถึง แบบทดสอบที่ไม่ได้วัดตามพื้นฐานของหลักสูตรหรือการสอน

ใดๆ ที่นักศึกษาได้เรียนมา แต่เน้นการวัดความสามารถทางภาษาที่ผู้เรียนต้องใช้ในการสื่อสารภาษาอังกฤษทั่วไป การศึกษาหรือการทำงานตามระดับความรู้ของผู้เรียน

2.1.1.2 ผลกระทบของแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test) ต่อผู้เกี่ยวข้อง

แบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test) มีผลกระทบที่ดีต่อผู้ใช้แบบทดสอบเหมือนกับแบบทดสอบทั่วไป ตามความคิดเห็นที่นักวิชาการด้านการทดสอบภาษาอังกฤษได้กล่าวไว้ ดังนี้

Henning (1987) กล่าวว่า แบบทดสอบเป็นตัวบ่งชี้จุดแข็งและจุดอ่อนในความสามารถด้านภาษาของนักเรียน เช่น ในการทดสอบอาจพบว่านักเรียนมีความสามารถในการออกเสียงดีมาก แต่ความสามารถในการอ่านเพื่อความเข้าใจอยู่ในระดับต่ำ และในการสอบครั้งต่อไป อาจจะพบอีกว่าความเชี่ยวชาญของนักเรียนในการใช้คำศัพท์มากหรือน้อย เป็นปัจจัยสำคัญที่สัมพันธ์กับการมีความสามารถในการอ่านเพื่อความเข้าใจต่ำของนักเรียนคนเดิม เป็นต้น การใช้แบบทดสอบในลักษณะนี้เป็นการใช้แบบทดสอบเพื่อการวินิจฉัยปัญหาระหว่างเรียน ซึ่งมีประโยชน์ในการให้ข้อมูลสำคัญสำหรับนักเรียน ครู และผู้บริหาร เพื่อปรับปรุงกระบวนการเรียนการสอนให้มีประสิทธิภาพมากขึ้น ถ้าขาดข้อมูลเหล่านี้ ผู้สอนอาจจะสอนไปโดยเน้นเฉพาะการออกเสียงแก่นักเรียนคนดังกล่าว แต่ไม่มีโอกาสทราบถึงความบกพร่องในด้านคำศัพท์ของนักเรียน

Bachman and Lyle F. & Palmer (1996) กล่าวว่า ครูสามารถใช้แบบทดสอบประกอบการตัดสินใจเกี่ยวกับผู้เรียน เช่น ช่วยในการวินิจฉัยปัญหา ข้อมูลที่ได้จากการทดสอบสามารถใช้สำหรับการวินิจฉัยเนื้อหาหรือประเด็นที่เป็นจุดแข็งและจุดอ่อน เพื่อที่จะค้นหาชนิดและระดับของกิจกรรมการเรียนการสอนที่เหมาะสมสำหรับผู้เรียน โดยทั่วไปแล้ว ข้อทดสอบทางภาษาทุกชนิดมีศักยภาพที่จะทำหน้าที่ให้ข้อมูลสำหรับการวินิจฉัยปัญหา แบบทดสอบสำหรับการวินิจฉัยปัญหาสามารถเป็นได้ทั้งแบบทดสอบวัดผลสัมฤทธิ์ในการเรียน (Achievement Tests) หรือแบบทดสอบวัดความสามารถทั่วไป (General Proficiency Tests) นอกจากนี้ แบบทดสอบสามารถเป็นตัวช่วยบอกความก้าวหน้าและตัดสินผลการเรียน ข้อมูลย้อนกลับต่อประสิทธิภาพในการเรียนรู้ของผู้เรียน ข้อมูลที่อยู่ในความสนใจของทั้งครูและนักเรียน ข้อมูลเหล่านี้มีประโยชน์ต่อครูสำหรับการประเมินปฏิบัติการ (Formative Evaluation) โดยครูสามารถใช้ข้อมูล เพื่อการตัดสินใจปรับเปลี่ยนกระบวนการเรียนการสอนและกิจกรรมการเรียนรู้อื่นๆ ให้เหมาะสม สำหรับผู้เรียน แบบทดสอบถูกนำมาใช้เพื่อการประเมินสรุป (Summative Evaluation) ในความก้าวหน้าของผู้เรียนในรูปแบบของเกรดหรือคะแนนในตอนท้ายของการเรียนการสอนหรือเมื่อจบวิชา

Madsen (1980: 3-5) กล่าวว่า แบบทดสอบที่ได้รับการสร้างและออกแบบอย่างดี สามารถช่วยนักเรียนได้อย่างน้อย 2 ประการ คือ 1) ช่วยสร้างทัศนคติเชิงบวกต่อการเรียน โดยแบบทดสอบที่มีระดับความยากที่เหมาะสม มีข้อความบางส่วนที่ถามในสิ่งที่สูงกว่าระดับความสามารถของผู้เรียน และครอบคลุมทักษะที่กำหนดไว้ว่าจะประเมิน สามารถทำให้นักเรียนมีทัศนคติที่ดีซึ่งมีผลมาจากการ

รู้สึกว่าการทดสอบมีความยุติธรรมและวัดได้ตรงจุดประสงค์ และ 2) แบบทดสอบช่วยให้ผู้เรียนมีความรู้ภาษาเพิ่มขึ้น ความรู้ภาษาจะเกิดขึ้นในขณะที่ผู้เรียนอ่านหนังสือเพื่อการสอบ และเมื่อผลการสอบออกมา มีการนำเรื่องราวที่สอบมาพูดคุยกัน เมื่อมีการสอบหลาย ๆ ครั้ง การเรียนรู้สามารถเกิดขึ้นได้ในตัวผู้เรียน โดยที่ผู้เรียนมีความตระหนักรู้ในวัตถุประสงค์การเรียน และประเด็นที่ควรเน้นในรายวิชามากขึ้น นอกจากนี้ แบบทดสอบสามารถผลักดันให้เกิดการเรียนรู้ ที่มาจากคุณลักษณะของแบบทดสอบที่ช่วยวินิจฉัยปัญหา โดยนอกจากแบบทดสอบช่วยยืนยันว่าผู้เรียนแต่ละคนได้มีความรู้ในเรื่องใดเรื่องหนึ่งแล้ว ยังชี้ให้เห็นถึงประเด็นภาษาที่จำเป็นต้องให้ความสำคัญต่อไป

แบบทดสอบยังมีความสำคัญต่อผู้สอนอีกด้วย โดยช่วยตอบคำถามสำคัญที่ผู้สอนใช้ถามตัวเองว่า ฉันสอนได้มีประสิทธิภาพหรือไม่ หรือจะกล่าวได้ว่า ผู้สอนสามารถใช้ข้อทดสอบเพื่อวินิจฉัยความพยายามของตนเอง เช่นเดียวกับความพยายามของนักเรียน และในขณะที่ผู้สอนกำลังบันทึกคะแนนจากการสอบ อาจถามตัวเองด้วยคำถามว่า บทเรียนของฉันอยู่ในระดับที่เหมาะสมหรือไม่ ฉันกำหนดเป้าหมายการเรียนการสอนของฉันสูงหรือต่ำไปหรือไม่ ด้านใดบ้าง หรือเรื่องอะไรบ้างที่เราต้องเน้นหรือให้ความสนใจมากขึ้น เรื่องไหนจำเป็นต้องทบทวน และควรหรือไม่ที่ฉันต้องใช้เวลามากหรือน้อยในสื่อชนิดนี้กับนักเรียนที่จะเข้ามาเรียนปีหน้า และประการสุดท้าย ข้อทดสอบสามารถให้เกิดมุมมองในการปรับกระบวนการสอบได้เช่นเดียวกัน เช่น คำสั่งในแบบทดสอบชัดเจนหรือไม่ ทุกคนสามารถทำแบบทดสอบเสร็จทันภายในเวลาที่กำหนดให้หรือไม่ แบบทดสอบได้ทำให้เกิดความกังวลโดยไม่จำเป็นหรือไม่ และผลการสอบสะท้อนได้อย่างถูกต้องถึงสิ่งที่นักเรียนได้เรียนในห้องเรียนและในงานที่ได้รับมอบหมายหรือไม่ ดังนั้น ข้อทดสอบสามารถให้ประโยชน์ต่อนักเรียน ครู และแม้กระทั่งผู้บริหาร โดยการยืนยันจุดประสงค์ที่ได้ถูกกำหนดไว้และแสดงให้เห็นว่าควรจะปรับความพยายามในอนาคตไปในทิศทางใด นอกจากนี้ แบบทดสอบที่ดียังทำให้ความรู้สึกที่ดีต่อการเรียนและห้องเรียน และช่วยส่งเสริมการเรียนรู้

Heaton (1975) กล่าวถึงเหตุผลที่ว่าทำไมต้องมีการสอบไว้ 4 ประการ คือ 1) ช่วยประเมินประสิทธิภาพการสอนของครู โดยให้ครูได้ปรับการสอนของตนเองให้เหมาะสมกับนักเรียนเฉพาะกลุ่ม และให้นักเรียนได้รับผลจากการสอนของครูมากยิ่งขึ้น และให้ทุกคนทุกกลุ่มได้รับประโยชน์จากการสอนของครูอย่างเท่าเทียมกัน 2) ผลการสอบจะทำให้ครูรู้ว่าความยากเฉพาะจุด หรือเนื้อหาใดที่เป็นเรื่องยากสำหรับนักเรียน เหมือนกับคุณหมอมองต้องตรวจสอบคนไข้ก่อนว่าเป็นโรคอะไร แล้วจึงจะทำการรักษาได้ ครูก็เช่นเดียวกัน เมื่อรู้ปัญหาการเรียนรู้ของนักเรียน ก็จะสามารถให้การช่วยเหลือได้ถูกวิธีโดยผ่านการให้งานเสริมและทำแบบฝึกหัดเพิ่มเติม 3) แบบทดสอบที่ดีจะทำให้ครูรู้ว่า ส่วนใดของหลักสูตรทางภาษาที่ถูกรับว่าเป็นปัญหาของนักเรียนทั้งหมด เพื่อให้ครูสามารถประเมินประสิทธิภาพของหลักสูตร เช่นเดียวกับวิธีการ และสื่อ วัสดุที่ครูใช้สอนด้วย และ 4) แบบทดสอบที่วัดความสามารถนักเรียนอย่างยุติธรรมหรือไม่ลำเอียง สามารถช่วยสร้างความมั่นใจนักเรียนได้ แบบทดสอบที่ถูกสร้างขึ้นอย่างดี จะช่วยให้นักเรียนมีโอกาสที่จะแสดงความสามารถในการรับสารและส่งสารได้อย่างถูกต้อง โดยมีข้อแม้ว่านักเรียนควรได้ทราบผลการสอบทันทีที่สอบเสร็จเท่าที่จะเป็นไปได้ และมีโอกาสได้เรียนรู้จากความผิดพลาดของตนเองและเชื่อมโยงกับสิ่งที่ได้เรียนมา ในกรณีเช่นนี้ แบบทดสอบที่ดีสามารถใช้เป็นเครื่องมือการสอนที่มีคุณค่ามาก

J. D. Brown (2005) กล่าวว่า บางครั้งผู้บริหารจำเป็นต้องทำการตัดสินใจทางการศึกษาโดยอาศัยระดับความสามารถทางภาษาโดยทั่วของนักเรียน ประเด็นหลักของการตัดสินใจดังกล่าวมักเกี่ยวกับความรู้หรือทักษะโดยทั่วไปที่นักเรียนต้องมีมาก่อนเข้าเรียนหรือก่อนที่จะออกไปจากสถาบันการศึกษาบางแห่ง เช่น มหาวิทยาลัยในอเมริกาใช้การตัดสินใจด้านความสามารถดังกล่าว ในการกำหนดมาตรฐานการรับนักศึกษาเข้าเรียนและคุณภาพของนักศึกษาที่ก่อนจบออกไปในหลักสูตรใด หลักสูตรหนึ่ง การปรับระดับของวัตถุประสงค์ต่าง ๆ ในหลักสูตรให้เหมาะสมกับความสามารถของนักเรียน หรือในการเปรียบเทียบความสามารถของผู้เรียนระหว่างหลักสูตร ดังนั้น การตัดสินใจด้านความสามารถมักจะเกิดขึ้นเมื่อหลักสูตรต้องเกี่ยวข้องกับโลกภายนอก เช่น ในกรณีของการรับนักศึกษาเข้ามาเรียน นักศึกษาเหล่านั้นจะเหมาะสมกับหลักสูตรได้อย่างไร และในกรณีที่นักศึกษาจะจบออกไป ระดับความสามารถของนักศึกษาสูงพอเพื่อให้พวกเขาประสบความสำเร็จในการใช้ภาษาในสถาบันอื่น ๆ หรือไม่

Cohen and Andrew D (1994) กล่าวว่า แบบทดสอบมีความสำคัญทั้งต่อผู้เรียน และผู้สอน โดยต่อผู้เรียนนั้น แบบทดสอบมีส่วนส่งเสริมการมีส่วนร่วมอย่างมีความหมายของผู้เรียนต่อสื่อวัสดุที่เกี่ยวข้องกับวัตถุประสงค์การสอนประจำรายวิชา เพื่อให้ความเกี่ยวข้องอย่างมีความหมายนี้เกิดขึ้น เป้าหมายของภาระงานที่ใช้ในการประเมินต้องสะท้อนเป้าหมายของรายวิชา และเป้าหมายนี้จะต้องทำให้ผู้เรียนชัดเจนก่อนการวัดด้วย ถ้าผู้เรียนรับรู้ว่าการสอบสำคัญต่อความจำเป็นในรายวิชานั้น ๆ ผู้เรียนจะมีแรงบันดาลใจในการทำแบบทดสอบให้บรรลุความต้องการหรือเป้าหมายของตนเอง นอกจากนี้ การประเมินผลการเรียนอย่างสม่ำเสมอสามารถทำให้ผู้เรียนได้รับข้อมูลป้อนกลับเกี่ยวกับความสามารถทางภาษาของตนเอง ขณะที่การวัดกำลังเกิดขึ้น ผู้เรียนจะได้รับข้อมูลย้อนกลับว่าพวกเขาสามารถทำภาระงานในการสอบนั้นได้ดีเพียงใด เมื่อได้รับผลการสอบ ทราบคะแนน ผลการประเมิน นักเรียนจะรู้ว่าพวกเขาทำในสิ่งถูกประเมินได้มากน้อยเพียงใด ขึ้นอยู่กับคุณภาพของข้อมูลย้อนกลับ และความใส่ใจในการสอบของผู้เรียน พวกเขาอาจจะได้ทราบบางสิ่งบางอย่างเกี่ยวกับเรื่องที่เป็นจุดแข็งและเรื่องหรือเนื้อหาที่เป็นจุดอ่อนของตนเอง ทำให้เกิดการเรียนรู้ต่อไปหรือมีความตั้งใจที่จะทบทวน

ในส่วนของความสำคัญของแบบทดสอบต่อผู้สอนนั้น แบบทดสอบช่วยเป็นแรงบันดาลใจให้ครูตัดสินใจเกี่ยวกับเป้าหมายของการสอน โดยคำนึงถึงเนื้อหาวิชา ทักษะ และระดับของการบรรลุผลที่ต้องการ ในทางทฤษฎี ถือว่าครูต้องมีความชัดเจนในสิ่งเหล่านั้นตั้งแต่ต้น แต่ในทางปฏิบัติ การเตรียมการสำหรับการวัดประเมินผล ทำให้ครูต้องถามตัวเองว่า เป้าหมายที่แท้จริงของตนเองคืออะไร และเมื่อผลการสอบออกมา ครูจะดูได้ว่า ผู้เรียนทำในสิ่งที่ได้รับการประเมินได้ดีมากน้อยเพียงใด และตรวจสอบความแตกต่างระหว่างความคาดหวังกับความสามารถที่ทำได้จริง ข้อมูลดังกล่าวอาจจะบอกได้ว่านักเรียนรู้ได้ดีมากน้อยเพียงใด หรือถ้าพวกเขาบรรลุผลสิ่งที่ได้เรียนไป ครูได้จัดการสอนอย่างไร มีการเรียนข้อทดสอบได้ถูกเขียนได้ดีเพียงใด ข้อมูลป้อนกลับไปยังครูจะเสนอแนะวิธีการสอน สำหรับการทบทวน หรือการปรับปรุงการประเมินในอนาคต ข้อมูลดังกล่าวข้างต้นไม่ได้ถูกส่งผ่านไปยังผู้สอนได้ง่าย ๆ นั่นคือ ครูต้องใช้เวลาในการวิเคราะห์แบบทดสอบลงลึกในรายข้อ เพื่อให้ทราบคุณภาพของแบบทดสอบแต่ละข้อ ว่ามีความยากง่าย เหมาะสมหรือไม่ และสามารถที่แยกแยะกลุ่มผู้เรียนออกจากกันได้หรือไม่ ข้อใดบ้างที่ควรได้รับการปรับปรุงในครั้งต่อไป

สรุปได้ว่า แบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test) มีความสำคัญทั้งสำหรับผู้เรียน ผู้สอน และผู้บริหาร ความสำคัญสำหรับผู้เรียน ได้แก่ 1) ผลการสอบช่วยให้ผู้เรียนได้รับข้อมูลย้อนกลับที่เป็นจุดอ่อนหรือจุดแข็งในความสามารถด้านภาษาของตนเอง 2) การสอบช่วยสร้างทัศนคติเชิงบวกและแรงจูงใจในการเรียนภาษาอังกฤษ และ 3) การสอบช่วยผลักดันให้มีความรู้ด้านภาษาอังกฤษเพิ่มมากขึ้น ความสำคัญสำหรับผู้สอน คือ อาจารย์สามารถใช้ผลการสอบเพื่อการปรับปรุงประสิทธิภาพการจัดการเรียนการสอนภาษาอังกฤษของตนเอง และความสำคัญสำหรับผู้บริหาร คือ ผู้บริหารหลักสูตรสามารถใช้ข้อมูลจากผลการสอบในการปรับปรุงหลักสูตร วิจัยปัญหา และเปรียบเทียบความสามารถของผู้เรียนกับมาตรฐานภายนอกหรือความสามารถในการนำภาษาไปใช้ในอนาคต

2.1.2 เนื้อหาสาระที่เกี่ยวข้องกับการสร้างแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test)

เนื้อหาสาระที่เกี่ยวข้องกับการสร้างแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test) ครอบคลุมจุดประสงค์ (Purpose) สาระภาษา (Language Knowledge/Subskills) 27 หัวเรื่อง (Topical Knowledge) ตัวกระตุ้น (Input Text) ผลผลิตที่ได้ (Output Text) ชนิดของเนื้อความ (Text Type) รูปแบบภาระงาน (Task Type) รูปแบบการทดสอบ (Test Type) และการตรวจให้คะแนน (Scoring Technique) ซึ่งได้มาจากการศึกษาและวิเคราะห์เนื้อหาสาระจาก 4 แหล่งข้อมูล คือ 1) หลักสูตรภาษาอังกฤษระดับอุดมศึกษา 2) กรอบอ้างอิงความสามารถทางภาษาของสภายุโรป (CEFR) 3) แบบทดสอบวัดความสามารถภาษาอังกฤษมาตรฐานสากล และ 4) แบบทดสอบวัดความสามารถภาษาอังกฤษระดับอุดมศึกษา โดยมีรายละเอียดดังต่อไปนี้

2.1.2.1 หลักสูตรภาษาอังกฤษระดับอุดมศึกษา (มหาวิทยาลัยทักษิณ, 2544)

หลักสูตรภาษาอังกฤษระดับอุดมศึกษาในการวิจัยครั้งนี้ คือ หลักสูตรระดับปริญญาตรี หมวดวิชาศึกษาทั่วไป หลักสูตรปรับปรุง พ.ศ. 2554 ของมหาวิทยาลัยทักษิณ (2554) ซึ่งประกอบด้วยรายวิชาภาษาอังกฤษพื้นฐาน 3 รายวิชา คือ วิชา 0000121 ภาษาอังกฤษทั่วไป 1 วิชา 0000122 ภาษาอังกฤษทั่วไป 2 และวิชา 0000221 ภาษาอังกฤษทั่วไป 3 โดยแต่ละวิชามีคำอธิบายรายวิชาดังนี้

คำอธิบายรายวิชาของวิชา 0000121 ภาษาอังกฤษทั่วไป มีความว่า “ศึกษาและฝึกทักษะทางภาษาด้านการฟัง อ่าน พูด และเขียน เพื่อการสื่อสาร”

คำอธิบายรายวิชาของ วิชา 0000122 ภาษาอังกฤษทั่วไป 2 มีความว่า “ศึกษาและฝึกทักษะทางภาษาด้านการฟัง อ่าน พูด และเขียน เพื่อการสื่อสารในระดับที่สูงขึ้นกว่าภาษาอังกฤษทั่วไป 1”

คำอธิบายรายวิชาของวิชา 0000221 ภาษาอังกฤษทั่วไป 3 มีความว่า “พัฒนาทักษะการฟัง อ่าน พูด และเขียนเพื่อการสื่อสาร สามารถลำดับความคิดเป็นประเด็นที่มีเอกภาพ สัมพันธภาพและสารัตถภาพ โดยใช้ภาษาในระดับที่สูงขึ้น”

สรุปได้ว่าสาระและทักษะภาษาที่นิสิตระดับปริญญาตรีควรได้รับจากการเรียนวิชา ภาษาอังกฤษพื้นฐาน คือ สาระที่เกี่ยวข้องกับการใช้ภาษาอังกฤษเพื่อการสื่อสาร และทักษะความสามารถในการใช้ภาษาอังกฤษในการฟัง การพูด การอ่าน และการเขียน โดยสามารถลำดับความคิดเป็นประเด็นที่มีเอกภาพ สัมพันธภาพและสารัตถภาพ

2.1.2.2 กรอบมาตรฐานความสามารถในการเรียนภาษาอังกฤษของสภายุโรป (CEFR) (Council of Europe, 2003)

จากการศึกษาและวิเคราะห์กรอบมาตรฐานความสามารถในการเรียนภาษาอังกฤษของสภายุโรป (CEFR) (Council of Europe, 2003) ได้เนื้อหาสาระแบบทดสอบที่จะนำไปใช้ออกแบบแบบทดสอบในแต่ละส่วน 3 ประการ คือ 1) จุดประสงค์ (Purpose) 2) สาระภาษา (Language Knowledge/Subskills) และ 3) 28 หัวเรื่อง (Topical Knowledge) โดยมีรายละเอียดในแต่ละทักษะ ดังนี้

การทดสอบทักษะการฟัง

จากการศึกษาและวิเคราะห์กรอบมาตรฐานความสามารถในการเรียนภาษาอังกฤษของสภายุโรป (CEFR) (Council of Europe, 2003) ที่เกี่ยวกับทักษะการฟัง ได้ จุดประสงค์ (Purpose) สาระภาษา (Language Knowledge/Subskills) และ 28 หัวเรื่อง (Topical Knowledge) สำหรับการวัดทักษะการฟังดังนี้

จุดประสงค์ (Purpose) ในการทดสอบทักษะการฟัง คือ เพื่อวัดความสามารถของผู้เข้าสอบในการฟังเพื่อจับใจความสำคัญและเข้าใจรายละเอียดของเสียงพูดที่ได้บันทึกไว้

สาระภาษา (Language Knowledge/Subskills) ในการทดสอบทักษะการฟัง ประกอบด้วย

- 1) ความเข้าใจบทสนทนาระหว่างเจ้าของภาษา
- 2) การฟังในฐานะสมาชิกคนหนึ่งของกลุ่มของผู้ฟัง
- 3) การฟังประกาศและการให้คำแนะนำ และ
- 4) การฟังสื่อทางเสียงและสิ่งที่ได้บันทึกไว้ โดยผู้เข้าสอบควรมีความสามารถในการ
 - 1) ติดตามประเด็นสำคัญของการพูดแบบต่อเนื่องรอบ ๆ ตัว โดยที่บทพูดมีการใช้สำเนียงมาตรฐานและออกเสียงชัดเจน;
 - 2) จับใจความสำคัญการสนทนาของภาพแอนิเมชันระหว่างผู้พูดเจ้าของภาษา;
 - 3) จับใจความสำคัญของบทพูดรอบ ๆ ตัวได้เป็นส่วนใหญ่;
 - 4) ติดตามการบรรยาย หรือการพูดในสาขาของตนเอง;
 - 5) ติดตามการพูดสั้น ๆ ที่มีโครงสร้างชัดเจนในหัวข้อที่คุ้นเคย และนำเสนอด้วยสำเนียงพูดมาตรฐานและออกเสียงชัดเจน;
 - 6) ติดตามประเด็นสำคัญๆ ของการบรรยาย การพูด และการรายงาน และรูปแบบอื่น ๆ ของการนำเสนอทางวิชาการและวิชาชีพ ซึ่งซับซ้อนทั้งทางความหมายและทางภาษา;
 - 7) เข้าใจข้อมูลด้านเทคนิคง่าย ๆ เช่น คำแนะนำการใช้อุปกรณ์ในชีวิตประจำวัน เป็นต้น;
 - 8) ติดตามการบอกทิศทางที่มีรายละเอียดชัดเจน;
 - และ 9) เข้าใจการประกาศและข่าวสารในหัวข้อเกี่ยวกับรูปธรรมและนามธรรมที่พูดด้วยสำเนียง

มาตรฐานและที่ความเร็วปกติ; 10) เข้าใจสาระสำคัญโดยทั่วไปของสื่อระบบเสียงที่บันทึกไว้และถ่ายทอดสด; 11) เข้าใจประเด็นสำคัญของบทสรุปข่าวทางวิทยุและสื่อบันทึกง่าย ๆ; 12) เข้าใจสิ่งที่ได้บันทึกไว้ด้วยสำเนียงมาตรฐานมักจะพบเจอได้ในการดำเนินชีวิตในสังคม การงานอาชีพ หรือแวดวงวิชาการ และแสดงความคิดเห็น บอกทัศนคติ และสาระสำคัญของสารได้; และ 13) เข้าใจสารคดีทางวิทยุและสื่อระบบเสียงที่บันทึกไว้และถ่ายทอดสดอื่น ๆ ที่ส่งผ่านโดยสำเนียงมาตรฐานได้เป็นส่วนใหญ่ และสามารถบอกอารมณ์ น้ำเสียง ฯลฯ ของผู้พูดได้

29 หัวเรื่อง (Topical Knowledge) ที่สามารถนำมาใช้ทดสอบการฟังครอบคลุม 1) ความเข้าใจบทสนทนาระหว่างเจ้าของภาษา 2) การฟังในฐานะสมาชิกคนหนึ่งของการเป็นผู้ชมผู้ฟังสด 3) การฟังประกาศและการให้คำแนะนำ และ 4) การฟังสื่อทางเสียงและสิ่งที่ได้บันทึกไว้ โดยในแต่ละด้านมีรายละเอียดสาระแบบทดสอบ คือ ในด้านความเข้าใจบทสนทนาเจ้าของภาษา มีสาระเกี่ยวกับ 1) การพูดคุยแบบต่อเนื่องรอบ ๆ ตัว ซึ่งเป็นบทพูดที่มีการออกเสียงชัดเจนและใช้สำเนียงมาตรฐาน; 2) การสนทนาของภาพแอนิเมชันระหว่างผู้พูดเจ้าของภาษา ในด้านการฟังในฐานะสมาชิกคนหนึ่งของการเป็นผู้ชมผู้ฟังสด มีสาระเกี่ยวกับ 3) การบรรยายที่มีเนื้อหาสาระเป็นที่คุ้นเคยและนำเสนอโดยใช้ภาษาที่ไม่ซับซ้อนและจัดวางโครงสร้างอย่างชัดเจน; 4) การพูดสั้น ๆ ที่มีการวางโครงสร้างชัดเจนในหัวข้อที่คุ้นเคย; 5) การบรรยาย การพูด และการรายงาน และรูปแบบอื่น ๆ ของการนำเสนอทางวิชาการและวิชาชีพซึ่งมีความซับซ้อนทั้งด้านความหมายและภาษา ในด้านการฟังประกาศและการให้คำแนะนำ มีสาระเกี่ยวกับ 6) คำแนะนำการใช้อุปกรณ์ในชีวิตประจำวัน; 7) การบอกทิศทาง; 8) การประกาศและข่าวสารในหัวข้อที่เกี่ยวกับรูปธรรมและนามธรรมที่พูดด้วยสำเนียงมาตรฐานและความเร็วปกติ และในด้านการฟังสื่อทางเสียงและสิ่งที่ได้บันทึกไว้ มีสาระเกี่ยวกับ 9) หัวข้อที่เป็นความสนใจส่วนตัวและนำเสนอด้วยเสียงพูดมาตรฐาน ออกเสียงชัดเจน; 10) บทสรุปข่าวทางวิทยุและสื่อบันทึกง่าย ๆ ในหัวข้อที่คุ้นเคยและนำเสนออย่างค่อนข้างซ้ำและชัดเจน; 11) สิ่งที่ได้บันทึกไว้ด้วยสำเนียงมาตรฐานและมักจะพบเจอได้ในการดำเนินชีวิตในสังคม การงานอาชีพ หรือแวดวงวิชาการ; 12) สารคดีทางวิทยุ; และ 13) สื่อระบบเสียงที่บันทึกไว้และถ่ายทอดสดอื่น ๆ ที่นำเสนอด้วยสำเนียงมาตรฐาน

การทดสอบทักษะการพูด

จากการศึกษาและวิเคราะห์กรอบมาตรฐานความสามารถในการเรียนภาษาอังกฤษของสภายุโรป (CEFR) (Council of Europe, 2003) ที่เกี่ยวกับทักษะการพูด ได้ จุดประสงค์ (Purpose) สาระภาษา (Language Knowledge/Subskills) และ หัวเรื่อง (Topical Knowledge) สำหรับการวัดทักษะการพูดดังนี้

จุดประสงค์ (Purpose) ในการทดสอบทักษะการพูด คือ เพื่อวัดความสามารถของผู้เข้าสอบในการใช้ภาษาพูดเพื่อการสื่อสาร

สาระภาษา (Language Knowledge/Subskills) ในการทดสอบทักษะการพูดครอบคลุม 1) การสนทนา 2) การพูดคุยแบบไม่เป็นทางการ (กับเพื่อน) 3) การแลกเปลี่ยนข้อมูล 4) บทพูดคนเดียว (การบรรยายประสบการณ์) และ 5) การพูดกับผู้ชมหรือผู้ฟัง โดยผู้เข้าสอบควรมีความสามารถใน 1) เข้าถึงการสนทนาในหัวข้อที่คุ้นเคยได้โดยไม่ต้องเตรียมตัว ; 2) ติดตามการพูดที่ออกเสียงชัดเจนในการ

สนทนาประจำวัน แม้ว่าบางครั้งจะมีการขอให้กล่าวซ้ำสำหรับบางคำหรือวลี; 3) ดำเนินการสนทนาหรือพูดคุยให้ต่อเนื่องได้ แม้บางครั้งจะจับความสิ่งที่คู่สนทนาต้องการสื่อได้ยาก; 4) แสดงออกและตอบสนองความรู้สึกของตนเอง เช่น ความประหลาดใจ ความสุข ความเศร้า ความสนใจ และความไม่แตกต่าง; 5) ร่วมการสนทนาต่อเนื่องในหัวข้อต่างๆไป ได้เป็นส่วนใหญ่ ในลักษณะการมีส่วนร่วมได้อย่างลงตัว แม้จะอยู่ภาวะแวดล้อมที่เสียงดัง; 6) รักษาความสัมพันธ์กับผู้พูดเจ้าของภาษาโดยไม่ได้ขำขันหรือล้อเลียนพวกเขาอย่างไม่ตั้งใจ หรือขอให้พวกเขาแสดงพฤติกรรมที่นอกเหนือจากที่เขาจะทำได้กับเจ้าของภาษาด้วยกัน; 7) สื่อระดับอารมณ์และแสดงจุดเด่นของความสำเร็จของเหตุการณ์และประสบการณ์ส่วนตัว; 8) ติดตามสิ่งที่พูดคุยกันในทุกหัวข้อ ใดๆ ไปได้เป็นส่วนใหญ่ โดยที่คู่สนทนาใช้ภาษาที่เข้าใจง่ายและออกเสียงโดยชัดเจน; 9) แสดงความคิดของตนเกี่ยวกับหัวข้อที่เป็นนามธรรมหรือเกี่ยวกับวัฒนธรรม เช่น ดนตรี ภาพยนตร์ และสามารถอธิบายว่าเหตุใดบางอย่างถึงเป็นปัญหา; 10) ให้ข้อเสนอแนะสั้นๆ ต่อความคิดเห็นของบุคคลอื่น; 11) เปรียบเทียบและชี้ให้เห็นความแตกต่างของทางเลือก อภิปรายสิ่งที่ต้องทำ สถานที่ที่จะไป ใครหรือสิ่งไหนที่ต้องเลือก เป็นต้น; 12) ติดตามประเด็นสำคัญในการพูดคุยแบบไม่เป็นทางการกับเพื่อนได้เป็นส่วนใหญ่ โดยคู่สนทนาออกเสียงที่ชัดเจนด้วยสำเนียงมาตรฐาน; 13) ให้หรือถามหาแนวคิดหรือความคิดเห็นส่วนตัวในการอภิปรายหัวข้อที่สนใจ; 14) ทำให้ความคิดเห็นของตนและการตอบสนองเป็นที่เข้าใจเพื่อเป็นข้อเสนอแนะในการแก้ปัญหา; 15) แสดงความเชื่อ ความคิดเห็น การเห็นด้วยหรือไม่เห็นด้วย อย่างสุภาพ; 16) พูดได้ในอัตราเดียวกันกับการพูดคุยอย่างมีรสชาติระหว่างผู้พูดเจ้าของภาษา; 17) แสดงความคิดหรือความคิดเห็นของตนที่มีความชัดเจน นำเสนอและตอบสนองต่อข้อโต้แย้งที่ซับซ้อนได้อย่างน่าเชื่อถือ; 18) มีส่วนร่วมอย่างกระฉับกระเฉงในการพูดคุยแบบไม่เป็นทางการในบริบทที่คุ้นเคย; 19) แสดงความคิดเห็น ให้ประเด็นความเห็นได้อย่างชัดเจน ประเมินข้อเสนอที่เป็นทางเลือก พร้อมทั้งสร้างและตอบสมมติฐาน; 20) จับประเด็นที่พูดคุยกันได้เป็นส่วนมากโดยต้องใช้ความพยายามอยู่บ้าง แต่อาจจะพบว่ามีส่วนร่วมอย่างมีประสิทธิภาพในการพูดคุยกับผู้พูดเจ้าของภาษาจำนวนมาก ซึ่งไม่ปรับแต่งภาษาของพวกเขาเลย; 21) ออกและยืนยันความเห็นของตนในการพูดคุย โดยการให้คำอธิบายที่เป็นประโยชน์ ข้อโต้แย้ง และข้อเสนอแนะ ;ในด้านการแลกเปลี่ยนข้อมูล ผู้เข้าสอบควรมีความสามารถในการ; 22) แลกเปลี่ยน ตรวจสอบ และยืนยันข้อมูลที่เป็นความจริงที่ในหัวเรื่องประจำวัน ที่คุ้นเคย และเกี่ยวกับสาขาของตนด้วยความมั่นใจ ; 23) บรรยายวิธีการทำอะไรบางอย่าง พร้อมทั้งให้คำแนะนำอย่างละเอียด; 24) ค้นหาและถ่ายทอดข้อมูลที่เป็นข้อเท็จจริงอย่างตรงไปตรงมา; 25) สอบถามข้อมูลและไปตามการบอกทิศทางที่ละเอียดได้; 26) รับข้อมูลที่มีความละเอียดได้มากขึ้น; 27) เข้าใจและแลกเปลี่ยนข้อมูลที่ซับซ้อนและคำแนะนำในเรื่องต่าง ๆ ที่สัมพันธ์กับบทบาททางอาชีพของตน; 28) ส่งผ่านข้อมูลที่มีรายละเอียดได้อย่างน่าเชื่อถือ; 29) ให้คำอธิบายที่ละเอียดชัดเจนของการดำเนินการตามขั้นตอน; 30) สังเคราะห์และรายงานข้อมูลและข้อโต้แย้งจากแหล่งข้อมูลหลายแหล่ง; 31) ให้คำอธิบายหัวข้อต่าง ๆ ที่คุ้นเคยภายในสาขาความสนใจของตนได้อย่างตรงไปตรงมา; 32) เชื่อมโยงความสัมพันธ์ของการเล่าเรื่องอย่างตรงไปตรงมาหรือการอธิบายตามลำดับประเด็นได้อย่างคล่องแคล่วและสมเหตุสมผล สามารถเล่าประสบการณ์ได้อย่างละเอียด พร้อมทั้งอธิบายความรู้สึกและปฏิกิริยาตอบสนอง; 33) เชื่อมโยงรายละเอียดของเหตุการณ์ที่ไม่สามารถทำนายได้ เช่น อุบัติเหตุ เป็นต้น; 34) เชื่อมโยงโครงเรื่องของหนังสือหรือภาพยนตร์และบรรยายปฏิกิริยา

ตอบสนองของตน; 35) บรรยายความฝัน ความหวัง และความปรารถนาอันสูงสุด; 36) สามารถบรรยายเหตุการณ์จริงหรือในจินตนาการ; 37) สามารถเล่าเรื่องได้; 38) ให้คำอธิบายที่ละเอียดชัดเจน ในหัวข้อต่าง ๆ อย่างหลากหลายที่สัมพันธ์กับสาขาความสนใจของตน; 39) ทำการนำเสนออย่างตรงไปตรงมาโดยที่ต้องเตรียมมาก่อนในหัวข้อที่คุ้นเคยภายในสาขาของตน มีความชัดเจนพอที่จะติดตามได้โดยไม่ต้องรีบร้อนเกินไป และประเด็นสำคัญได้รับการอธิบายด้วยการตัดสินใจที่สมเหตุสมผล; 40) จัดการกับคำถามที่สอบถามข้อมูลเพิ่มเติม แต่อาจจะขอให้มีการกล่าวซ้ำถ้าการพูดนั้นเร็วเกินไป; 41) ทำการนำเสนอที่ได้รับการพัฒนาอย่างเป็นระบบ ชัดเจน โดยมีการนำเสนอประเด็นสำคัญและรายละเอียดสนับสนุนที่สำคัญ; 42) ละสายตาจากเอกสารที่เตรียมมาได้โดยทันทีและติดตามประเด็นที่น่าสนใจที่หยิบยกโดยสมาชิกที่เป็นผู้ฟัง มักจะแสดงให้เห็นถึงความคล่องแคล่วและสะดวกที่จะแสดงความเห็นอย่างสังเกตได้; 43) นำเสนอที่ได้รับการเตรียมตัวชัดเจนให้เหตุผลในการสนับสนุนหรือขัดแย้งบางประเด็นความคิดเห็น และแสดงให้เห็นถึงประโยชน์และข้อดีของทางเลือกหลายอย่าง และ 44) จัดการกับชุดคำถามที่เกิดขึ้นด้วยระดับความคล่องแคล่วและความทันทีทันใดซึ่งไม่ได้สร้างความตึงเครียดให้กับตนเองหรือผู้ฟัง

หัวเรื่อง (Topical Knowledge) ที่สามารถนำมาใช้ทดสอบการพูดครอบคลุม 1) การสนทนา 2) การพูดคุยแบบไม่เป็นทางการ (กับเพื่อน) 3) การแลกเปลี่ยนข้อมูล 4) บทพูดคนเดียว (การบรรยายประสบการณ์) และ 5) การพูดกับผู้ชมหรือผู้ฟัง โดยในแต่ละด้านมีรายละเอียดสาระแบบทดสอบ คือ ในด้านการสนทนา มีสาระเกี่ยวกับ 1) หัวข้อที่คุ้นเคยและเรื่องราวประจำวัน; 2) การสนทนาต่อเนื่องในหัวข้อทั่วไป ในด้านการพูดคุยแบบไม่เป็นทางการ (กับเพื่อน) มีสาระเกี่ยวกับ 3) หัวข้อที่เป็นนามธรรมหรือเกี่ยวกับวัฒนธรรม เช่น ดนตรี ภาพยนตร์ เป็นต้น; 4) การพูดคุยแบบไม่เป็นทางการกับเพื่อน; 5) การอภิปรายในหัวข้อที่สนใจ ในด้านการแลกเปลี่ยนข้อมูล มีสาระเกี่ยวกับ 6) การสนทนาในชีวิตประจำวัน ; 7) วิธีการทำอะไรบางอย่าง ; 8) การสืบค้นข้อมูลและส่งต่อข้อมูล; 9) การสอบถามทิศทาง; 10) ข้อมูลที่ซับซ้อนและคำแนะนำในเรื่องต่าง ๆ ที่สัมพันธ์กับบทบาททางอาชีพของตน ในด้านการพูดคนเดียว (การบรรยายประสบการณ์) มีสาระเกี่ยวกับ 11) หัวข้อที่หลากหลายของเรื่องที่คุ้นเคยภายในสาขาความสนใจของตน – B1; 12) หนังสือหรือภาพยนตร์; 13) ความฝัน ความหวัง และความปรารถนาอันสูงสุด; 14) เหตุการณ์จริงหรือในจินตนาการ; 15) หัวข้อต่าง ๆ อย่างหลากหลายที่สัมพันธ์กับสาขาความสนใจของตน และในด้านการพูดกับผู้ชมหรือผู้ฟัง มีสาระเกี่ยวกับ 16) หัวข้อที่คุ้นเคยภายในสาขาของตน

การทดสอบทักษะการอ่าน

จากการศึกษาและวิเคราะห์กรอบมาตรฐานความสามารถในการเรียนภาษาอังกฤษของสภายุโรป (CEFR) (Council of Europe, 2003) ที่เกี่ยวกับทักษะการอ่าน ได้ จุดประสงค์ (Purpose) สาระภาษา (Language Knowledge/Subskills) และ 31 หัวเรื่อง (Topical Knowledge) สำหรับการวัดทักษะการอ่านดังนี้

จุดประสงค์ (Purpose) ในการทดสอบทักษะการอ่าน คือ วัดความสามารถของผู้เข้าสอบในการอ่านเพื่อความเข้าใจบทอ่านชนิดต่าง ๆ ทั้งความเข้าใจในประเด็นสำคัญและรายละเอียดเฉพาะของบทอ่าน

สาระภาษา (Language Knowledge/Subskills) ในการทดสอบทักษะการอ่านครอบคลุม 1) การอ่านจดหมาย 2) การอ่านเพื่อจุดประสงค์เฉพาะ 3) การอ่านเพื่อหาข้อมูลและโต้แย้ง และ 4) การอ่านคำแนะนำการใช้ โดยผู้เข้าสอบควรมีความสามารถในการ 1) เข้าใจคำบรรยายเหตุการณ์ ความรู้สึก และความปรารถนาในจดหมายส่วนตัวได้เป็นอย่างดีพอที่จะสื่อสารกับเพื่อนทางจดหมายได้อย่างเป็นปกติ; 2) อ่านจดหมายที่เกี่ยวข้องกับสาขาความสนใจของตนเองและจับความหมายที่สำคัญได้ทันที; 3) กวาดสายตาอ่านบทอ่านที่ยาวขึ้นเพื่อหาข้อมูลที่ต้องการ และเก็บข้อมูลจากส่วนต่าง ๆ ของบทอ่าน หรือจากบทอ่านต่าง ๆ เพื่อมาเติมเต็มภาระงานเฉพาะอย่างของตน; 4) หาและเข้าใจข้อมูลสำคัญของสื่อวัสดุที่พบประจำวัน เช่น จดหมาย โบรชัวร์ และเอกสารทางราชการสั้นๆ; 5) กวาดสายตาอย่างรวดเร็วในการอ่านบทอ่านที่ยาวและซับซ้อน เพื่อหารายละเอียดที่สำคัญได้; 6) ระบุอย่างรวดเร็วในสาระและความสำคัญของประเด็นข่าว บทความ และรายงาน ในหัวข้อที่เกี่ยวข้องกับสาขาอาชีพที่หลากหลาย ตัดสินใจได้ว่าการศึกษายละเอียดคุ้มค่าหรือไม่; 7) ระบุข้อสรุปสำคัญของบทอ่านเชิงโต้แย้งที่มีการให้สัญลักษณ์ไว้อย่างชัดเจน; 8) ระบุบรรทัดของข้อโต้แย้งเพื่อแก้ปัญหาที่ถูกนำเสนอ โดยไม่จำเป็นต้องให้รายละเอียด; 9) บอกประเด็นสำคัญจากการอ่านบทความในหนังสือพิมพ์ที่นำเสนอหัวเรื่องที่คุ้นเคยและใช้ภาษาไม่ซับซ้อน; 10) รับข้อมูล ความคิด และความเห็นจากแหล่งเฉพาะด้านมาก ๆ ในสาขาของตน; 11) เข้าใจบทความพิเศษนอกสาขาของตนเอง โดยที่สามารถใช้พจนานุกรมอยู่เป็นบางครั้งเพื่อยืนยันการตีความในคำศัพท์เฉพาะของตนเอง; 12) เข้าใจบทความและรายงานเกี่ยวกับปัญหาในปัจจุบันที่ผู้เขียนเอาตัวอย่างหรือมุมมองมานำเสนอ; 13) เข้าใจคำแนะนำการใช้สำหรับอุปกรณ์อย่างใดอย่างหนึ่งที่เข้าใจได้ไม่ยาก และเขียนไว้อย่างชัดเจน และ 14) เข้าใจคำแนะนำการใช้ในสาขาของตนเองที่ซับซ้อนและมีความยาว รวมทั้งรายละเอียดในเรื่องอื่นและคำเตือน โดยที่ต้องอ่านซ้ำส่วนที่ยากอยู่บ้าง

หัวเรื่อง (Topical Knowledge) ที่สามารถนำมาใช้ทดสอบการอ่านครอบคลุม 1) การอ่านจดหมาย 2) การอ่านเพื่อจุดประสงค์เฉพาะ 3) การอ่านเพื่อหาข้อมูลและโต้แย้ง และ 4) การอ่านคำแนะนำการใช้ โดยในแต่ละด้านมีรายละเอียดสาระแบบทดสอบ คือ ในด้านการอ่านจดหมาย มีสาระเกี่ยวกับ 1) จดหมายส่วนตัว และ 2) จดหมายที่เกี่ยวข้องกับสาขาความสนใจของตนเอง ในด้านการอ่านเพื่อจุดประสงค์เฉพาะ มีสาระเกี่ยวกับ 3) บทอ่านต่าง ๆ ที่ยาวขึ้น; 4) สื่อวัสดุที่พบประจำวัน เช่น จดหมาย โบรชัวร์ และเอกสารทางราชการสั้น ๆ; 5) บทอ่านที่ยาวและซับซ้อน และ 6) ข่าว บทความ และรายงาน ในหัวข้อที่เกี่ยวข้องกับสาขาอาชีพที่หลากหลาย ในด้านการอ่านเพื่อหาข้อมูลและโต้แย้ง มีสาระเกี่ยวกับ 7) บทอ่านเชิงโต้แย้งที่มีการให้สัญลักษณ์ไว้อย่างชัดเจน; 8) บทความจากหนังสือพิมพ์ในหัวเรื่องที่คุ้นเคยที่นำเสนอด้วยภาษาที่ไม่ซับซ้อน; 9) ข้อมูลเฉพาะด้านอย่างแท้จริงในสาขาของตน; 10) บทความพิเศษนอกสาขาของตนเอง และ 11) บทความและรายงานเกี่ยวกับปัญหาในปัจจุบัน และในด้านการอ่านคำแนะนำการใช้ มีสาระเกี่ยวกับ 12) คำแนะนำการใช้สำหรับอุปกรณ์อย่างใดอย่างหนึ่งที่เข้าใจได้ง่ายและเขียนไว้อย่างชัดเจน และ 13) คำแนะนำการใช้ในสาขาของตนเองที่ยาวและซับซ้อน พร้อมทั้งมีรายละเอียดในเรื่องอื่นและคำเตือน

การทดสอบทักษะการเขียน

จากการศึกษาและวิเคราะห์กรอบมาตรฐานความสามารถในการเรียนภาษาอังกฤษของสภายุโรป (CEFR) (Council of Europe, 2003) ที่เกี่ยวกับทักษะการเขียน ได้ จุดประสงค์ (Purpose) สารระภาษา (Language Knowledge/Subskills) และ 32 หัวเรื่อง (Topical Knowledge) สำหรับการวัดทักษะการเขียนดังนี้

จุดประสงค์ (Purpose) ในการทดสอบทักษะการเขียน คือ วัดความสามารถของผู้เข้าสอบในการสื่อสารโดยใช้ภาษาเขียน

สารระภาษา (Language Knowledge/Subskills) ในการทดสอบทักษะการเขียนครอบคลุม 1) การเขียนจดหมาย และ 2) การเขียนเชิงสร้างสรรค์ โดยผู้เข้าสอบควรมีความสามารถในการ 1) เขียนจดหมายส่วนตัวเพื่อส่งข่าวและแสดงความคิดเห็นในหัวข้อที่เกี่ยวข้องกับนามธรรมหรือวัฒนธรรม; 2) เขียนจดหมายส่วนตัวอธิบายประสบการณ์ ความรู้สึก และเหตุการณ์ โดยให้รายละเอียดพอสมควร; 3) เขียนจดหมายเพื่อสื่อระดับของอารมณ์และเสนอการให้ความสำคัญส่วนตัวต่อเหตุการณ์และประสบการณ์และแสดงความคิดเห็นต่อสิ่งใหม่และความคิดเห็นในตัวสาร; 4) เขียนบรรยายที่มีรายละเอียดไม่ซับซ้อน ในหัวข้อต่าง ๆ ที่คุ้นเคย ภายในสาขาความสนใจของตน; 5) เขียนบอกเล่าประสบการณ์ บรรยายความรู้สึก และปฏิกิริยาตอบสนอง ด้วยข้อความที่เกี่ยวข้องกันแบบง่ายๆ; 6) เขียนบรรยายเหตุการณ์ การเดินทางที่เพิ่งผ่านมา ตามความจริงหรือจินตนาการ; 7) เขียนเล่าเรื่องได้; 8) เขียนบรรยายที่ละเอียด ชัดเจนของเหตุการณ์และประสบการณ์จริงหรือจินตนาการ โดยแสดงความสัมพันธ์ระหว่างความคิดในข้อความที่ต่อเนื่องกันชัดเจน และปฏิบัติตามข้อตกลงที่กำหนดไว้ตามประเภทการเขียนที่เกี่ยวข้อง; 9) เขียนบรรยายที่ละเอียด ชัดเจน ในหัวเรื่องที่หลากหลาย เกี่ยวข้องกับสาขาความสนใจของตน และ 10) เขียนวิจารณ์ภาพยนตร์ หนังสือ หรือละคร

หัวเรื่อง (Topical Knowledge) ที่สามารถนำมาใช้ทดสอบการเขียนครอบคลุม 1) การเขียนจดหมาย และ 2) การเขียนเชิงสร้างสรรค์ โดยในแต่ละด้านมีรายละเอียดสาระแบบทดสอบ คือ ในด้านการเขียนจดหมาย มีสาระเกี่ยวกับ 1) การเขียนจดหมายส่วนตัวเพื่อส่งข่าวและแสดงความคิดเห็นในหัวข้อที่เกี่ยวข้องกับนามธรรมหรือวัฒนธรรม; 2) การเขียนจดหมายส่วนตัวอธิบายประสบการณ์ ความรู้สึก และเหตุการณ์ โดยให้รายละเอียดพอสมควร และ 3) การเขียนจดหมายเพื่อสื่อระดับของอารมณ์และเสนอการให้ความสำคัญส่วนตัวต่อเหตุการณ์และประสบการณ์และแสดงความคิดเห็นต่อสิ่งใหม่และความคิดเห็นในตัวสาร และในด้านการเขียนเชิงสร้างสรรค์ มีสาระเกี่ยวกับ 4) การเขียนบรรยายเกี่ยวกับหัวข้อต่าง ๆ ที่คุ้นเคย ภายในสาขาความสนใจของตน; 5) การเขียนบอกเล่าประสบการณ์ บรรยายความรู้สึก และปฏิกิริยาตอบสนอง ด้วยข้อความที่เกี่ยวข้องกันแบบง่ายๆ; 6) การเขียนบรรยายเหตุการณ์ เช่น การเดินทางที่เพิ่งผ่านมา โดยอาจเป็นเหตุการณ์จริงหรือตามจินตนาการ; 7) การเขียนเล่าเรื่อง; 8) การเขียนบรรยายเหตุการณ์และประสบการณ์จริงหรือตามจินตนาการที่ละเอียดและชัดเจน โดยแสดงความสัมพันธ์ระหว่างความคิดในข้อความที่ต่อเนื่องกันได้ชัดเจน และปฏิบัติตามข้อตกลงที่กำหนดไว้ตามแต่ละประเภทการเขียนที่เกี่ยวข้อง; 9) การเขียน

บรรยายในหัวเรื่องที่หลากหลาย เกี่ยวข้องกับสาขาความสนใจของตนที่ละเอียดและ ชัดเจน และ 10) การเขียนวิจารณ์ภาพยนตร์ หนังสือ หรือละคร

การทดสอบทักษะการใช้ภาษา

จากการศึกษาและวิเคราะห์กรอบมาตรฐานความสามารถในการเรียนภาษาอังกฤษของสภายุโรป (CEFR) (Council of Europe, 2003) ที่เกี่ยวกับทักษะการใช้ภาษา ได้ จุดประสงค์ (Purpose) สารภาษา (Language Knowledge/Subskills) และ หัวเรื่อง (Topical Knowledge) สำหรับการวัดทักษะการใช้ภาษาดังนี้

จุดประสงค์ (Purpose) ในการทดสอบทักษะการใช้ภาษา คือ เพื่อวัดความสามารถของผู้เข้าสอบในการใช้คำศัพท์และโครงสร้างไวยากรณ์ในการสื่อสารได้อย่างถูกต้องและเหมาะสม

สารภาษา (Language Knowledge/Subskills) ในการทดสอบทักษะการเขียนครอบคลุม 1) ปริมาณคำศัพท์ และ 2) ความถูกต้องตามหลักไวยากรณ์ โดยผู้เข้าสอบควรมีความสามารถในการ 1) มีคำศัพท์เพียงพอที่จะสื่อสารได้เข้าได้ในหัวข้อต่าง ๆ ที่สัมพันธ์กับชีวิตประจำวัน เช่น ครอบครัว งานอดิเรก ความสนใจ งาน การเดินทาง และเหตุการณ์ปัจจุบัน; 2) เปลี่ยนรูปของคำศัพท์เพื่อหลีกเลี่ยงการใช้คำซ้ำ; 3) ความถูกต้องของคำศัพท์ส่วนใหญ่อยู่ในระดับสูง; 4) สื่อสารได้ถูกต้องอย่างสมเหตุผลในบริบทที่คุ้นเคยโดยทั่วไป มีการควบคุมการใช้ภาษาได้ดี มีอิทธิพลของการใช้ภาษาแม่และข้อผิดพลาดที่สังเกตได้บ้าง แต่สามารถเข้าใจได้ว่าผู้พูดกำลังสื่อสารอะไร; 5) ใช้ไวยากรณ์ได้สมเหตุสมผลอย่างถูกต้อง ในรูปแบบหรือโครงสร้างที่มักใช้เป็นประจำ ในชีวิตประจำวัน ที่สัมพันธ์กับสถานการณ์ที่สามารถทำนายได้; 6) ควบคุมการใช้ไวยากรณ์ได้ดี มีความผิดพลาดอยู่บ้างหรือเป็นความผิดพลาดที่ไม่เป็นระบบ และ 7) แสดงให้เห็นถึงระดับการควบคุมไวยากรณ์ที่ค่อนข้างสูง ไม่สร้างความผิดพลาดที่นำไปสู่การเข้าใจผิด

หัวเรื่อง (Topical Knowledge) ที่สามารถนำมาใช้ทดสอบการใช้ภาษาครอบคลุม 1) ปริมาณคำศัพท์ และ 2) ความถูกต้องตามหลักไวยากรณ์ โดยในแต่ละด้านมีรายละเอียดสาระแบบทดสอบ คือ ในด้านปริมาณคำศัพท์ มีสาระเกี่ยวกับ 1) คำศัพท์เพื่อการสื่อสารในหัวข้อที่สัมพันธ์กับชีวิตประจำวัน เช่น ครอบครัว งานอดิเรก และความสนใจ งาน การเดินทาง และเหตุการณ์ปัจจุบัน ; 2) คำศัพท์เพื่อกับการสื่อสารในเรื่องราวหรือเนื้อหาที่เกี่ยวข้องกับสาขาของตน; และ 3) คำศัพท์เพื่อการสื่อสารในหัวข้อโดยทั่วไป และในด้านความถูกต้องตามหลักไวยากรณ์ มีสาระเกี่ยวกับ 4) โครงสร้างที่มักใช้เป็นประจำ ในชีวิตประจำวัน ที่สัมพันธ์กับสถานการณ์ที่สามารถทำนายได้มากกว่า

2.1.2.3 แบบทดสอบวัดความสามารถภาษาอังกฤษมาตรฐานสากล

จากการศึกษาและวิเคราะห์แบบทดสอบวัดความสามารถภาษาอังกฤษมาตรฐานสากล 5 ชนิด ได้แก่ แบบทดสอบ TOEFL iBT (Test of English as a Foreign Language internet-Based

Test) (Educational Testing Service, 2010) แบบทดสอบ TOEFL ITP (Test of English as a Foreign Language – Institutional Testing Program) (Educational Testing Service, 2014) แบบทดสอบ IELTS (International English Language Testing System) (IELTS, 2015) แบบทดสอบ FCE (First Certificate in English) (UCLES, 2014) และแบบทดสอบ telc B1-B2 English (telc GmbH, 2012) ได้เนื้อหาสาระแบบทดสอบที่จะนำไปใช้ออกแบบทดสอบในแต่ละส่วน 7 ประการ คือ 1) จุดประสงค์ (Purpose) 2) สาระภาษา (Language Knowledge/Subskills) 3) หัวเรื่อง (Topical Knowledge) 4) ชนิดของเนื้อความ (Text Type) 5) รูปแบบภาระงาน (Task Type) 6) รูปแบบการทดสอบ (Test Type) และ 7) การตรวจให้คะแนน (Scoring Technique) โดยมีรายละเอียดในแต่ละชนิดแบบทดสอบ ดังต่อไปนี้

แบบทดสอบ TOEFL iBT (Test of English as a Foreign Language internet-Based Test)

แบบทดสอบ TOEFL iBT (Test of English as a Foreign Language internet-Based Test) (Educational Testing Service, 2010) มีเนื้อหาสาระแบบทดสอบในการทดสอบทักษะการฟัง การพูด การอ่าน และการเขียน ดังรายละเอียดต่อไปนี้

การทดสอบทักษะการฟัง

จากการศึกษาและวิเคราะห์แบบทดสอบ TOEFL iBT (Test of English as a Foreign Language internet-Based Test) (Educational Testing Service, 2010) ที่เกี่ยวกับทักษะการฟัง ได้เนื้อหาสาระแบบทดสอบ คือ 1) จุดประสงค์ (Purpose) 2) สาระภาษา (Language Knowledge/Subskills) 3) 35 หัวเรื่อง (Topical Knowledge) 4) รูปแบบการทดสอบ (Test Type) และ 5) การตรวจให้คะแนน (Scoring Technique) โดยมีรายละเอียด ดังนี้

จุดประสงค์ (Purpose) ในการทดสอบทักษะการฟัง คือ เพื่อวัดความสามารถของผู้เข้าสอบ ในการเข้าใจภาษาพูดในบริบทวิชาการที่ผู้ใช้ภาษาต้องสามารถฟังการบรรยายและบทสนทนาได้

สาระภาษา (Language Knowledge/Subskills) ในการทดสอบทักษะการฟังครอบคลุมความสามารถในการ 1) เข้าใจความคิดหลัก ประเด็นหลัก และรายละเอียดที่สำคัญ ที่สัมพันธ์กับความคิดหลัก 2) ตระหนักรู้ในทัศนคติของผู้พูดและระดับความ 3) ตระหนักรู้ในหน้าที่หรือความประสงค์ของผู้พูด 4) ตระหนักรู้ในการจัดลำดับของเนื้อความที่ 5) เข้าใจความสัมพันธ์ระหว่างความคิดที่นำเสนอ (ตัวอย่างเช่น ความเหมือน/ความต่าง เหตุ/ผล หรือ ขั้นตอนในกระบวนการ) 6) ทำนายหรือหาข้อสรุปโดยอาศัยสิ่งที่ถูกกล่าวเป็นนัย ๆ ในเรื่องที่ฟัง 7) ทำการเชื่อมโยงระหว่างส่วนต่าง ๆ ของข้อมูลในบทสนทนาหรือการบรรยาย และ 8) ตระหนักรู้ในการเปลี่ยนหัวเรื่องในการบรรยายและบทสนทนา และตระหนักรู้ในบทสนทนาและบทสรุปของการบรรยาย

35 หัวเรื่อง (Topical Knowledge) ของบทฟังเกี่ยวกับการบรรยายทางวิชาการและบทสนทนาที่สะท้อนถึงการฟัง-พูดที่เกิดขึ้นในห้องเรียน ทั้งที่เป็นการบรรยายของอาจารย์และการ

แลกเปลี่ยนความคิดเห็นระหว่างอาจารย์กับนักศึกษา บทสนทนาในบริบทวิชาการอาจเกิดขึ้นในระหว่างการประชุมในสำนักงานกับอาจารย์ หรือผู้ช่วยสอน หรือระหว่างการติดต่อที่เคาน์เตอร์บริการกับเจ้าหน้าที่ของมหาวิทยาลัย เนื้อหาของการสนทนาในสำนักงานมักจะเป็นเรื่องด้านวิชาการหรือสัมพันธ์กับข้อกำหนดในรายวิชา และที่เคาน์เตอร์บริการมักจะเป็นการสนทนาเกี่ยวข้องกับการจ่ายค่าที่พัก การลงทะเบียนเรียน และการสอบถามข้อมูลในห้องสมุด

รูปแบบการทดสอบ (Test Type) ในการวัดทักษะการฟังส่วนใหญ่เป็นแบบเลือกตอบแบบดั้งเดิม 4 ตัวเลือก และมีคำตอบถูกเพียง 1 ตัวเลือก และยังมีข้อคำถามแบบอื่นๆ อีก เช่น คำถามแบบเลือกตอบพร้อมด้วยคำตอบมากกว่า 1 ตัวเลือก (ตัวอย่างเช่น มีคำตอบที่ถูก 2 คำตอบ จากตัวเลือก 4 หรือมากกว่า) คำถามที่ให้ผู้เข้าสอบเรียงลำดับเหตุการณ์หรือขั้นตอนในกระบวนการ และคำถามที่ให้ผู้เข้าสอบจับคู่วัตถุ ข้อความกับประเภทในตาราง

การตรวจให้คะแนน (Scoring Technique) เป็นแบบตอบถูกได้ 1 หรือ ตอบผิดได้ 0

การทดสอบทักษะการพูด

จากการศึกษาและวิเคราะห์แบบทดสอบ TOEFL iBT (Test of English as a Foreign Language internet-Based Test) (Educational Testing Service, 2010) ที่เกี่ยวกับทักษะการพูดได้เนื้อหาสาระแบบทดสอบ คือ 1) จุดประสงค์ (Purpose) 2) สาระภาษา (Language Knowledge/Subskills) 3) ชนิด 36 ภาระงาน (Task Type) และ 4) การตรวจให้คะแนน (Scoring Technique) โดยมีรายละเอียด ดังนี้

จุดประสงค์ (Purpose) ในการทดสอบทักษะการพูด คือ เพื่อวัดความสามารถของผู้เข้าสอบในการพูดภาษาอังกฤษอย่างมีประสิทธิภาพในบริบททางวิชาการ ทั้งในระหว่างการเรียนและนอกห้องเรียน

สาระภาษา (Language Knowledge/Subskills) สำหรับการทดสอบทักษะการพูดครอบคลุมความสามารถในการ 1) ตอบคำถามและมีส่วนร่วมในการอภิปรายทางวิชาการ 2) สรุปสิ่งที่ได้อ่านหรือได้ฟังและแสดงความคิดเห็นในหัวข้อที่กำลังพูดคุยกัน 3) มีส่วนร่วมในการสนทนาตามปกติ และ 4) แสดงความคิดเห็นและสื่อสารกับบุคคลในสถานที่ต่าง ๆ เช่น ร้านหนังสือ ห้องสมุด ร้านอาหาร และสำนักงานที่พัก

ชนิด 36 ภาระงาน (Task Type) ในการทดสอบทักษะการพูดเลียนแบบสถานการณ์ในชีวิตจริงที่นักศึกษาอาจจะเจอในระหว่างการเรียน นักศึกษาได้รับการคาดหวังให้สามารถตอบคำถาม มีส่วนร่วมในการอภิปรายทางวิชาการ สรุปสิ่งที่ได้อ่านหรือได้ฟัง และแสดงความคิดเห็นในหัวข้อที่กำลังพูดคุยกัน ในสถานการณ์นอกห้องเรียน นักศึกษามีส่วนร่วมในการสนทนาตามปกติ แสดงความคิดเห็นและสื่อสารกับบุคคลในสถานที่ต่าง ๆ เช่น ร้านหนังสือ ห้องสมุด ร้านอาหาร และสำนักงานที่พัก ในขั้นการสอบ ผู้เข้าสอบจะได้รับการบอกให้พูดในหัวข้อที่หลากหลายที่มาจากประสบการณ์ส่วนบุคคล สถานการณ์ในมหาวิทยาลัยและเนื้อหาเชิงวิชาการ ส่วนการพูดจะใช้เวลา 20 นาที ประกอบด้วยคำถาม 6 ข้อ ในการตอบคำถาม 2 ข้อแรก เรียกว่าเป็นภาระงานการพูดโดยอิสระ (Independent Speaking Tasks) เพราะผู้เข้าสอบต้องใช้ความคิดตนเอง ความคิดเห็นและ

ประสบการณ์ในการตอบคำถาม อีก 4 คำถามที่เหลือเรียกว่าภาระงานการพูดเชิงบูรณาการ (Integrated Speaking Tasks) เพราะผู้เข้าสอบต้องบูรณาการทักษะภาษาอังกฤษของตนเอง เช่น ฟังแล้วพูด หรือ ฟัง อ่านแล้วพูด เหมือนกับที่ผู้เข้าสอบได้เจอในระหว่างการเรียนหรือนอกห้องเรียน (Educational Testing Service, 2010: 3)

การตรวจให้คะแนน (Scoring Technique) ของแบบทดสอบ TOEFL iBT (Test of English as a Foreign Language internet-Based Test) ใช้การให้คะแนนแบบองค์รวม (Holistic Criteria) ผู้ตรวจฟังการพูดเพื่อหาลักษณะต่าง ๆ ในภาษาพูด และให้คะแนนในทักษะพูดโดยรวมในคำตอบ แม้ว่าเกณฑ์ให้คะแนนจะแปรผันไปตามลักษณะคำถาม โดยภาพรวมแล้วผู้ให้คะแนนจะฟังเพื่อพิจารณาสิ่งต่อไปนี้ในคำตอบของผู้เข้าสอบ ความพรั่งพรู (Delivery) การพูดชัดเจนและสั้นไหลรวมทั้งออกเสียงได้ดี การเว้นจังหวะ และใช้เสียงสูงต่ำที่เป็นธรรมชาติ การใช้ภาษา (Language Use) ใช้หลักไวยากรณ์และคำศัพท์เพื่อนำเสนอความคิดได้อย่างมีประสิทธิภาพ การนำเสนอสอดคล้องกับหัวเรื่อง (Topic Development) ตอบตรงคำถามและนำเสนอความคิดได้อย่างต่อเนื่อง สัมพันธ์กัน การตอบที่ดีมักจะใช้เวลาที่ให้ได้ครบหรือส่วนใหญ่ และความสัมพันธ์ระหว่างความคิดและความต่อเนื่องของความคิดหนึ่งกับความคิดถัดไป ชัดเจน และติดตามได้ง่าย

การทดสอบทักษะการอ่าน

จากการศึกษาและวิเคราะห์แบบทดสอบ TOEFL iBT (Test of English as a Foreign Language internet-Based Test) (Educational Testing Service, 2010) ที่เกี่ยวกับทักษะการอ่านได้เนื้อหาสาระแบบทดสอบ คือ 1) จุดประสงค์ (Purpose) 2) สาระภาษา (Language Knowledge/Subskills) 3) 37 หัวเรื่อง (Topical Knowledge) 4) รูปแบบการทดสอบ (Test Type) และ 5) การตรวจให้คะแนน (Scoring Technique) โดยมีรายละเอียด ดังนี้

จุดประสงค์ (Purpose) ในการทดสอบทักษะการอ่าน คือ เพื่อวัดความสามารถของผู้เข้าสอบในการเข้าใจบทความวิชาการและบทความอื่น ๆ ในระดับมหาวิทยาลัย

สาระภาษา (Language Knowledge/Subskills) ในการทดสอบทักษะการอ่านครอบคลุมความสามารถในการ 1) อ่านบทความเพื่อค้นหาประเด็นหลักและข้อมูลที่สำคัญได้อย่างมีประสิทธิภาพ 2) เพิ่มความคล่องแคล่วและความเร็วในการอ่าน 3) เข้าใจหัวเรื่องโดยทั่วไปหรือความคิดหลักประเด็นสำคัญ ข้อเท็จจริงที่สำคัญและรายละเอียด คำศัพท์และบริบทและการใช้คำสรรพนาม 4) เชื่อมโยงความเข้าใจเกี่ยวกับสิ่งที่ถูกกล่าวโดยนัยในบทความ 5) จัดลำดับเนื้อความและจุดประสงค์ของบทความ 6) เข้าใจความใจความสัมพันธ์ระหว่างความคิด 7) จัดการข้อมูลเป็นชาร์ตเพื่อแบ่งประเภทหรือหาบทสรุป เพื่อให้จำประเด็นหลัก ๆ และรายละเอียดที่สำคัญได้ และ 8) เข้าใจการเชื่อมโยงความคิดต่าง ๆ กันตลอดบทความ

37 หัวเรื่อง (Topical Knowledge) ของบทความเป็นส่วนตัดต่อกับตำราเรียนระดับมหาวิทยาลัยและสื่อวิชาการอื่น ๆ ที่เขียนเป็นภาษาอังกฤษ ซึ่งครอบคลุมเนื้อหาในหลากหลายสาขาวิชา นอกจากนี้บทความมักจะนำเสนอข้อมูลเกี่ยวกับหัวเรื่องที่มีความสำคัญมากกว่าหนึ่ง

ความคิดหรือแง่มุม และจะมีอย่างน้อยหนึ่งคำถามที่ต้องการให้ผู้เข้าสอบแสดงความเข้าใจในการจัดลำดับเนื้อความของบทอ่าน

รูปแบบการทดสอบ (Test Type) ในการวัดทักษะการอ่านมีลักษณะข้อคำถาม 2 ชนิด คือ 1) คำถามพร้อมด้วยตัวเลือก 4 ตัวเลือก และมีคำตอบที่ถูกเพียงตัวเลือกเดียว ตามรูปแบบของแบบเลือกตอบดั้งเดิม ที่ให้ผู้เข้าสอบเติมประโยคในบริเวณที่เหมาะสมในตัวบทอ่าน และ 2) คำถามในส่วนการอ่านเพื่อการเรียนรู้ (Reading to learn) พร้อมด้วยตัวเลือกที่มากกว่า 4 ตัวเลือกและคำตอบที่ถูกมีได้มากกว่า 1 ตัวเลือก

การตรวจให้คะแนน (Scoring Technique) เป็นแบบตอบถูกได้ 1 หรือ ตอบผิดได้ 0 การทดสอบทักษะการเขียน

จากการศึกษาและวิเคราะห์แบบทดสอบ TOEFL iBT (Test of English as a Foreign Language internet-Based Test) (Educational Testing Service, 2010) ที่เกี่ยวกับทักษะการเขียนได้เนื้อหาสาระแบบทดสอบ คือ 1) จุดประสงค์ (Purpose) 2) สาระภาษา (Language Knowledge/Subskills) 3) ชนิด 38 ภาระงาน (Task Type) และ 4) การตรวจให้คะแนน (Scoring Technique) โดยมีรายละเอียด ดังนี้

จุดประสงค์ (Purpose) ในการทดสอบทักษะการเขียน คือ เพื่อวัดความสามารถของผู้เข้าสอบในการเขียนภาษาอังกฤษในบริบทวิชาการ

สาระภาษา (Language Knowledge/Subskills) สำหรับการทดสอบทักษะการเขียนครอบคลุมความสามารถในการ 1) จัดบันทึกสิ่งที่ได้ยินและอ่าน และใช้บันทึกเพื่อจัดระเบียบข้อมูลก่อนการเขียน 2) สรุป ตัดแปลงสำนวนภาษา และอ้างอิงข้อมูลให้ถูกต้องจากแหล่งข้อมูล 3) เขียนในทำนองที่บอกว่าสิ่งที่ผู้เข้าสอบได้ยินสัมพันธ์กับสิ่งที่ผู้เข้าสอบได้อ่าน และ 4) เขียนเรียงความที่แสดงและสนับสนุนความคิดเห็น

ชนิด 38 ภาระงาน (Task Type) ในการทดสอบทักษะการเขียนประกอบด้วย 2 ชนิด คือ ภาระงานเขียนเชิงบูรณาการ (Integrated Writing Task) และภาระงานเขียนอิสระ (Independent Writing Task) ภาระงานเขียนเชิงบูรณาการ (Integrated Writing Task) ต้องการให้บูรณาการข้อมูลที่นักศึกษาได้ฟังจากการบรรยาย กับสิ่งที่พวกเขาได้อ่านในตำราหรือสื่ออื่น ๆ โดยที่ผู้เข้าสอบต้องสามารถ จัดบันทึกสิ่งที่ได้ยินและอ่าน และใช้บันทึกในจัดระเบียบข้อมูลก่อนการเขียน สรุป ตัดแปลงสำนวนภาษา และอ้างอิงข้อมูลให้ถูกต้องจากแหล่งข้อมูล และเขียนในทำนองที่บอกว่าสิ่งที่ผู้เข้าสอบได้ยินสัมพันธ์กับสิ่งที่ผู้เข้าสอบได้อ่าน ภาระงานเขียนอิสระ (Independent Writing Task) ผู้เข้าสอบต้องสามารถเขียนเรียงความที่แสดงและสนับสนุนความคิดเห็นได้ด้วย ผู้เข้าสอบแสดงความความคิดเห็นและสนับสนุนความคิดเห็นด้วยความรู้และประสบการณ์ของตนเอง ตัวอย่างเช่น ผู้เข้าสอบอาจจะถูกกำหนดให้เขียนเรียงความเกี่ยวกับประเด็นที่อยู่ในความขัดแย้งกัน (a controversial issue) ผู้เข้าสอบอาจใช้ประสบการณ์ส่วนตัวในอดีตเพื่อสนับสนุนความคิดเห็นของตนเอง

การตรวจให้คะแนน (Scoring Technique) ของแบบทดสอบ TOEFL iBT (Test of English as a Foreign Language internet-Based Test) มีกระบวนการ คือ คำตอบสำหรับงานเขียนในทุกภาระงานจะถูกส่งไปยัง the ETS Online Scoring Network คำตอบถูกให้คะแนนโดย

ผู้ตรวจ 4 คน เป็นผู้ตรวจที่เป็นคน 2 คน สำหรับภาระงานบูรณาการ (integrated task) และผู้ตรวจเป็นบุคคล 1 คน และ e-rater 1 คน สำหรับงานเขียนอิสระ (independent task) (สำหรับการตรวจด้วย e-rater งานเขียนจะถูกตรวจด้วยคอมพิวเตอร์) งานเขียนจะได้รับการให้คะแนนตามเกณฑ์ 0 ถึง 5 ตามที่ระบุในเกณฑ์การให้คะแนนงานเขียน (the Writing Scoring Guides) คะแนนเฉลี่ยจาก 2 ภาระงาน จะถูกเปลี่ยนเป็นตัวเลขคะแนนตามเกณฑ์ 0 ถึง 30 สำหรับงานเขียนบูรณาการ จะได้รับการให้คะแนนตามคุณภาพของการเขียน (การจัดระบบความคิด organization, การใช้ไวยากรณ์และคำศัพท์ที่เหมาะสมและถูกต้อง precise use of grammar, and vocabulary) และ ความสมบูรณ์และถูกต้องของเนื้อหาสาระที่เขียน (the completeness and accuracy of the content) ส่วนการเขียนเรียงความอิสระ ได้รับการให้คะแนนตามคุณภาพงานเขียนโดยรวม ตามประเด็นของ การพัฒนาความคิด development จัดระบบความคิด organization, การใช้ไวยากรณ์และคำศัพท์ที่เหมาะสมและถูกต้อง precise use of grammar, and vocabulary ไม่จำเป็นว่าผู้เข้าสอบจะเห็นด้วยหรือไม่เห็นด้วยตามหัวข้อ ผู้ให้คะแนนได้รับการอบรมให้ยอมรับความหลากหลายทางความคิดทั้งหมดที่จะได้เจอ สำหรับงานเขียนทั้ง 2 ชนิด ผู้ให้คะแนนจะคิดว่าคำตอบที่คุณตอบเป็นฉบับร่าง พวกเขาไม่คาดหวังว่าผู้เข้าสอบต้องเขียนเรียงความได้อย่างยอดเยี่ยม ผู้เข้าสอบสามารถได้คะแนนสูงด้วยเรียงความที่มีข้อผิดพลาดอยู่บ้าง

แบบทดสอบ TOEFL ITP (Test of English as a Foreign Language – Institutional Testing Program) (Educational Testing Service, 2014)

แบบทดสอบ TOEFL ITP (Test of English as a Foreign Language – Institutional Testing Program) (Educational Testing Service, 2014) มีเนื้อหาสาระแบบทดสอบในการทดสอบทักษะการฟัง การใช้ภาษา และการอ่าน ดังรายละเอียดต่อไปนี้

การทดสอบทักษะการฟัง

จากการศึกษาและวิเคราะห์แบบทดสอบ TOEFL ITP (Test of English as a Foreign Language – Institutional Testing Program) (Educational Testing Service, 2014) ที่เกี่ยวกับทักษะการฟัง ได้เนื้อหาสาระแบบทดสอบ คือ 1) จุดประสงค์ (Purpose) 2) สาระภาษา (Language Knowledge/Subskills) 3) หัวเรื่อง (Topical Knowledge) 4) รูปแบบการทดสอบ (Test Type) และ 5) การตรวจให้คะแนน (Scoring Technique) โดยมีรายละเอียด ดังนี้

จุดประสงค์ (Purpose) ในการทดสอบทักษะการฟัง คือ เพื่อวัดความสามารถของผู้เข้าสอบในการฟังสาระที่เกี่ยวข้องกับชีวิตประจำวันและเป็นที่คุ้นเคย

สาระภาษา (Language Knowledge/Subskills) ที่เกี่ยวข้อง คือ ความสามารถในการ Identify 1) the meaning of what the second speaker said, 2) the synonyms of the difficult words, 3) words having similar sounds, 4) the passive forms, 5) Who and What with multiple nouns, 6) negative expressions, 7) double negative expressions, 8)

“almost negative” expressions, 9) negative with comparatives, expressions of agreement, 10) expressions of uncertainty and suggestion, 11) emphatic expressions of surprise, 12) untrue conditions, 13) two-and three- part words, 14) idioms, and 15) draw conclusions about who, what, where, 16) to anticipate the topics, 17) to anticipate the questions, 18) to determine the topic, และ 19) to draw conclusions about who, what, when, where.

หัวข้อที่ใช้ในการทดสอบ มักจะเกี่ยวข้องกับ เรื่องราวในชีวิตประจำวันและเป็นที่ยุคสมัย วิทยาศาสตร์ และประวัติศาสตร์ แหล่งที่มาของข้อมูล เป็น บทสนทนาสั้น ๆ บทสนทนายาว และการบรรยายหรืออธิบาย

รูปแบบการทดสอบ (Test Type) ในการวัดทักษะการฟังส่วนใหญ่เป็นแบบเลือกตอบแบบ ดั้งเดิม 4 ตัวเลือก และมีคำตอบถูกเพียง 1 ตัวเลือก

การตรวจให้คะแนน (Scoring Technique) เป็นแบบตอบถูกได้ 1 หรือ ตอบผิดได้ 0

การทดสอบทักษะการใช้ภาษา

จากการศึกษาและวิเคราะห์แบบทดสอบ TOEFL ITP (Test of English as a Foreign Language – Institutional Testing Program) (Educational Testing Service, 2014) ที่เกี่ยวกับ ทักษะการใช้ภาษา ได้เนื้อหาสาระแบบทดสอบ คือ 1) จุดประสงค์ (Purpose) 2) สาระภาษา (Language Knowledge/Subskills) 3) หัวเรื่อง (Topical Knowledge) 4) รูปแบบการทดสอบ (Test Type) และ 5) การตรวจให้คะแนน (Scoring Technique) โดยมีรายละเอียด ดังนี้

จุดประสงค์ (Purpose) ในการทดสอบทักษะการใช้ภาษา คือ เพื่อวัดความสามารถของผู้เข้า สอบในความเข้าใจภาษาเขียนมาตรฐานและความสามารถในการระบุข้อผิดพลาดในการใช้ภาษา

สาระภาษา (Language Knowledge/Subskills) ที่เกี่ยวข้อง คือ 1) Be sure the sentence has a subject and a verb, 2) Be careful of objects of prepositions, 3) Be careful of present participles, 4) Use reduced adjective clauses correctly, 5) Use adverb Time and Cause connectors correctly, 6) Use noun clause connector/subjects correctly, 7) Use other adverb connectors correctly, 8) Invert the subject and verb with place expressions, 9) Be careful of past participles, 10) Use noun clause connectors correctly, 11) Invert the subject and verb with conditionals, 12) Use reduced adverb clauses correctly, 13) Invert the subject and verb with negatives, 14) Use adjective clause connectors correctly, 15) Invert the subject and verb with comparisons, 16) After be, use the present participle or the past participle, 17) Form comparatives and superlatives correctly, 18) Make verbs agree after certain words, 19) After Will, Would, or other Modals, use the base form of the verb, 20) Use the correct tense with time expressions, 21) Distinguish subject and object pronouns, 22) Distinguish possessive adjectives and pronouns, 23) Use parallel structure with paired

conjunctions, 24) Make verb agree after prepositional phrases, 25) Use comparatives and superlatives correctly, 26) Distinguish A and An, 27) Use the correct tense with Will and Would, 28) Distinguish Like, Alike, Unlike, and Dislike, 29) Make inverted verbs agree, 30) Distinguish countable and uncountable nouns, 31) Distinguish specific and general ideas, 32) Position adjectives and adverbs correctly, 33) Recognize active and passive meanings, 34) Use articles with singular nouns, 35) Use predicate adjectives correctly, 36) Distinguish Make and Do, 37) Use the correct form of the passive, 38) Recognize when preposition have been omitted, 39) Distinguish the person from the thing, และ 40) Use adjectives after linking verbs

หัวเรื่องที่ใช้ในการทดสอบ เป็นเนื้อหาเชิงวิชาการ และข้อมูลจะเป็นประโยคสั้น ๆ

รูปแบบการทดสอบ (Test Type) ในการวัดทักษะการฟังส่วนใหญ่เป็นแบบเลือกตอบแบบ
ดั้งเดิม 4 ตัวเลือก และมีคำตอบถูกเพียง 1 ตัวเลือก

การตรวจให้คะแนน (Scoring Technique) เป็นแบบตอบถูกได้ 1 หรือ ตอบผิดได้ 0

การทดสอบทักษะการอ่าน

จากการศึกษาและวิเคราะห์แบบทดสอบ TOEFL ITP (Test of English as a Foreign Language – Institutional Testing Program) (Educational Testing Service, 2014) ที่เกี่ยวกับ
ทักษะการอ่าน ได้เนื้อหาสาระแบบทดสอบ คือ 1) จุดประสงค์ (Purpose) 2) สาระภาษา
(Language Knowledge/Subskills) 3) หัวเรื่อง (Topical Knowledge) 4) รูปแบบการทดสอบ
(Test Type) และ 5) การตรวจให้คะแนน (Scoring Technique) โดยมีรายละเอียด ดังนี้

จุดประสงค์ (Purpose) ในการทดสอบทักษะการอ่าน คือ เพื่อวัดความสามารถของผู้เข้า
สอบในการอ่านสาระเชิงวิชาการ

สาระภาษา (Language Knowledge/Subskills) ที่เกี่ยวข้อง คือ 1) To answer main
idea question correctly, 2) To use context to determine meanings of simple words, 3)
To answer transition questions correctly, 4) To answer stated detail questions
correctly, 5) To find pronoun referents, 6) To determine meanings from word parts, 7)
To answer implied detail question correctly และ 8) To determine where specific
information is found

หัวเรื่องที่ใช้ในการทดสอบ เกี่ยวข้องกับเทคโนโลยี ศิลปะ สัตว์ ธรรมชาติ และคน ข้อมูลจะ
เป็นบทอ่านสั้น ๆ จากตำรา หนังสือ เชิงวิชาการ

รูปแบบการทดสอบ (Test Type) ในการวัดทักษะการฟังส่วนใหญ่เป็นแบบเลือกตอบแบบ
ดั้งเดิม 4 ตัวเลือก และมีคำตอบถูกเพียง 1 ตัวเลือก

การตรวจให้คะแนน (Scoring Technique) เป็นแบบตอบถูกได้ 1 หรือ ตอบผิดได้ 0

แบบทดสอบ IELTS (International English Language Testing System)

แบบทดสอบ IELTS (International English Language Testing System) (IELTS, 2015) มีเนื้อหาสาระแบบทดสอบในการทดสอบทักษะการฟัง การพูด การอ่าน และการเขียน ดังรายละเอียดต่อไปนี้

การทดสอบทักษะการฟัง

จากการศึกษาและวิเคราะห์แบบทดสอบ IELTS (International English Language Testing System) (IELTS, 2015) ที่เกี่ยวกับทักษะการฟังได้เนื้อหาสาระแบบทดสอบ คือ 1) จุดประสงค์ (Purpose) 2) สาระภาษา (Language Knowledge/Subskills) 3) หัวเรื่อง (Topical Knowledge) 4) รูปแบบการทดสอบ (Test Type) และ 5) การตรวจให้คะแนน (Scoring Technique) โดยมีรายละเอียด ดังนี้

จุดประสงค์ (Purpose) ในการทดสอบทักษะการฟัง คือ เพื่อวัดความเข้าใจอย่างละเอียดในประเด็นเฉพาะต่าง ๆ หรือ ความเข้าใจโดยทั่วไป ในประเด็นหลัก ๆ ของเสียงบันทึก

สาระภาษา (Language Knowledge/Subskills) ในการทดสอบทักษะการฟังครอบคลุมความสามารถในการ 1) เข้าใจประเด็นเฉพาะและข้อมูลโดยทั่วไปของเสียงบันทึกอย่างถ่องแท้ 2) ฟังเพื่อหาข้อมูลรายละเอียด 3) ติดตามการสนทนาระหว่างคน 2 คน 4) ตระหนักรู้ในความสัมพันธ์ของข้อเท็จจริงในเสียงบันทึก 5) เข้าใจคำบรรยาย เช่น การบรรยายสถานที่ ควรบอกได้ว่าคำบรรยายนี้สัมพันธ์กับรูปภาพอย่างไร 6) เข้าใจคำอธิบายว่าสิ่งต่าง ๆ อยู่ที่ไหน และสามารถเดินตามเส้นทางเพื่อไปยังจุดหมาย (เช่น เดินตรงไป หรือ ผ่านประตูที่อยู่ไกลออกไป) 7) ระบุข้อมูลที่สำคัญในเสียงบันทึก อาจจะต้องเข้าใจความสัมพันธ์ระหว่างความคิด ข้อเท็จจริง เหตุการณ์ เช่น เหตุและผล เป็นต้น และ 8) ฟังเพื่อหาข้อเท็จจริง เช่น สถานที่ ราคา เวลา ที่ได้ยินในเสียงบันทึก

42 หัวเรื่อง (Topical Knowledge) ของบทฟังเกี่ยวกับสถานการณ์ในชีวิตประจำวันและในสังคม โดยมีทั้งบทสนทนาระหว่างผู้พูด 2 คน เช่น การสนทนาเกี่ยวกับการเตรียมการในการเดินทาง และบทพูดที่มีผู้พูดคนเดียว เช่น การกล่าวสุนทรพจน์เกี่ยวกับสิ่งอำนวยความสะดวกในท้องถิ่น นอกจากนี้ยังมีบทพูดที่เป็นสถานการณ์ทางการศึกษาและการฝึกอบรม ทั้งบทสนทนาระหว่างผู้พูดหลัก 2 คน เช่น นักศึกษามหาวิทยาลัย 2 คน พูดคุยกัน หรืออาจจะเป็นการแนะนำโดยตัวต่อตัว และบทพูดที่มีผู้พูดคนเดียว เช่น การพูดเกี่ยวกับเนื้อหาในรายวิชาทางวิชาการ

รูปแบบการทดสอบ (Test Type) ในการวัดทักษะการฟังมีการใช้ลักษณะคำถามที่หลากหลาย โดยเลือกจากแบบเลือกตอบ (Multiple Choice) แบบจับคู่ (Matching) การเติมความในแผนผัง/แผนที่/ไดอะแกรม (Plan/Map/Diagram Labeling) การเติมบทสรุปจากแบบฟอร์ม/โน้ต/ตาราง/ชาร์ต (Form/Note/Table/Flow chart/Summary Completion) การเติมประโยคให้สมบูรณ์ (Sentence Completion) และ คำถามให้ตอบสั้น ๆ (Short-Answer Questions)

การตรวจให้คะแนน (Scoring Technique) เป็นแบบตอบถูกได้ 1 หรือ ตอบผิดได้ 0

การทดสอบทักษะการพูด

จากการศึกษาและวิเคราะห์แบบทดสอบ IELTS (International English Language Testing System) (IELTS, 2015) ที่เกี่ยวกับทักษะการพูดได้เนื้อหาสาระแบบทดสอบ คือ 1) จุดประสงค์ (Purpose) 2) สาระภาษา (Language Knowledge/Subskills) 3) ชนิดภาระงาน (Task Type) และ 4) การตรวจให้คะแนน (Scoring Technique) โดยมีรายละเอียด ดังนี้

จุดประสงค์ (Purpose) ในการทดสอบทักษะการพูด คือ เพื่อทดสอบความสามารถในการพูด สัมภาษณ์แบบเผชิญหน้า

สาระภาษา (Language Knowledge/Subskills) สำหรับการทดสอบทักษะการพูด ครอบคลุมความสามารถในการ 1) พูดเพื่อตอบคำถามสั้น ๆ 2) พูดได้อย่างต่อเนื่องในหัวข้อที่คุ้นเคย และมีปฏิสัมพันธ์กับผู้ทดสอบ 3) แสดงความคิดเห็นและให้ข้อมูลเกี่ยวกับเรื่องราวในชีวิตประจำวัน และประสบการณ์หรือสถานการณ์ต่างๆ ไป 4) พูดได้อย่างต่อเนื่อง ๆ ในหัวข้อที่กำหนดให้ โดยใช้ภาษา และจัดระบบความคิดได้อย่างเหมาะสม และ 5) อธิบายความคิดเห็น วิเคราะห์ อภิปราย และพยากรณ์เกี่ยวกับประเด็นที่พูดถึง

ชนิด 43 ภาระงาน (Task Type) ในการทดสอบทักษะการพูด ประเมินผู้เข้าสอบในการพูด เพื่อตอบคำถามสั้น ๆ การพูดยาว ๆ ในหัวข้อที่คุ้นเคย และมีปฏิสัมพันธ์กับผู้ทดสอบด้วย การทดสอบพูดเป็นการสัมภาษณ์แบบเผชิญหน้าระหว่างผู้เข้าสอบและผู้ประเมิน การพูดจะถูกบันทึกไว้แบบทดสอบประกอบด้วย 3 ส่วน และในแต่ละส่วน ผู้เข้าสอบต้องทำตามภาระงานที่กำหนดให้ตามลำดับ ส่วนที่ 1 การแนะนำตัวและการสัมภาษณ์ (Introduction and Interview) ในส่วนนี้ ผู้ประเมินแนะนำตัวเอง และเช็คตัวตนของผู้เข้าสอบ หลังจากนั้นผู้ประเมินถามคำถามต่างๆ ไป ในหัวข้อที่คุ้นเคย เช่น บ้าน ครอบครัว งาน การเรียนหรือความสนใจ ใช้เวลาประมาณ 4-5 นาที ส่วนนี้วัดความสามารถของผู้เข้าสอบในการแสดงความคิดเห็นและให้ข้อมูลเกี่ยวกับเรื่องราวในชีวิตประจำวันและประสบการณ์หรือสถานการณ์ต่างๆ ไป โดยการตอบคำถามหลายข้อ ส่วนที่ 2 การพูดคนเดียวที่ยาวขึ้น ผู้ประเมินจะให้บัตรภาระงานแก่ผู้เข้าสอบเพื่อให้พูดเกี่ยวกับหัวข้อใดหัวข้อหนึ่ง ข้อความในบัตรภาระงานจะบอกว่าผู้เข้าสอบควรพูดประเด็นใดบ้าง และกำหนดให้ผู้เข้าสอบอธิบายเพียงแง่มุมเดียวของหัวข้อนั้น ในส่วนนี้ทดสอบความสามารถของผู้เข้าสอบในการพูดยาว ๆ ในหัวข้อที่กำหนดให้ การใช้ภาษาที่เหมาะสมและการจัดระบบความคิดได้อย่างเหมาะสม ผู้เข้าสอบจำเป็นต้องคิดเกี่ยวกับประสบการณ์ของตนเองเพื่อพูดส่วนนี้ให้สำเร็จ และส่วนที่ 3 ผู้เข้าสอบและผู้ประเมินอภิปรายประเด็นที่สัมพันธ์กับหัวข้อในส่วนที่ 2 โดยภาพกว้างๆ และในลักษณะที่เป็นนามธรรม และบางที่อาจคุยเจาะลึกลงไปได้ ใช้เวลาประมาณ 4-5 นาที ส่วนนี้ทดสอบความสามารถของผู้เข้าสอบในการอธิบายความคิดเห็นและวิเคราะห์ อภิปราย และพยากรณ์เกี่ยวกับประเด็นที่พูดถึง

การตรวจให้คะแนน (Scoring Technique) ของแบบทดสอบ IELTS (International English Language Testing System) ได้รับการตรวจโดยผู้ตรวจแบบทดสอบ IELTS ที่ได้รับการรับรองทำหน้าที่ประเมินความสามารถของผู้เข้าสอบตลอดช่วงสอบ มีเกณฑ์การประเมิน 4 ด้าน (สิ่งที่ผู้ประเมินต้องระลึกไว้เมื่อตัดสินใจที่จะให้คะแนนงานของผู้เข้าสอบ) คือ ความคล่องแคล่วและความต่อเนื่อง (Fluency and coherence) การใช้คำศัพท์ที่เหมาะสม (Lexical resource) ไวยากรณ์ที่ใช้

และความถูกต้อง (Grammatical range and accuracy) และการออกเสียง (Pronunciation) (IELTS, 2015: 5)

การทดสอบทักษะการอ่าน

จากการศึกษาและวิเคราะห์แบบทดสอบ IELTS (International English Language Testing System) (IELTS, 2015) ที่เกี่ยวกับทักษะการอ่านได้เนื้อหาสาระแบบทดสอบ คือ 1) จุดประสงค์ (Purpose) 2) สาระภาษา (Language Knowledge/Subskills) 3) หัวเรื่อง (Topical Knowledge) 4) รูปแบบการทดสอบ (Test Type) และ 5) การตรวจให้คะแนน (Scoring Technique) โดยมีรายละเอียด ดังนี้

จุดประสงค์ (Purpose) ในการทดสอบทักษะการอ่าน คือ เพื่อประเมินผู้เข้าสอบด้านความเข้าใจในรายละเอียดของประเด็นเฉพาะหรือความเข้าใจโดยทั่วไปเกี่ยวกับประเด็นหลัก ๆ ของเนื้อความ

สาระภาษา (Language Knowledge/Subskills) ในการทดสอบทักษะการอ่านครอบคลุมความสามารถในการ 1) เข้าใจในรายละเอียดของประเด็นเฉพาะหรือความเข้าใจโดยทั่วไปเกี่ยวกับประเด็นหลัก ๆ ของเนื้อความ 2) จำแนกข้อมูลจำเพาะที่ให้ไว้ในเนื้อความ 3) จำแนกความคิดเห็นหรือประเด็นความคิดต่าง ๆ 4) กวาดสายตาอ่านเนื้อความเพื่อที่จะหาข้อมูลเฉพาะ 5) ระบุมุมมองหลักของย่อหน้าหรือตอน 6) จำแนกความแตกต่างระหว่างความคิดหลักกับความคิดสนับสนุน 7) ระบุมุมมองความสัมพันธ์และการเชื่อมโยงระหว่างข้อเท็จจริงในเนื้อความ 8) ระบุมุมมอง และทฤษฎี และ 9) อ่านคร่าว ๆ และอ่านเจาะลึกเฉพาะที่ในเนื้อความเพื่อหาข้อมูลอย่างรวดเร็ว

44 หัวเรื่อง (Topical Knowledge) ในการวัดทักษะการอ่านประกอบด้วยบทอ่าน (Reading passages) จำนวน 3 บท มาจากหนังสือ วารสาร นิตยสาร หนังสือพิมพ์ และแหล่งทรัพยากรออนไลน์ ที่เขียนสำหรับผู้อ่านที่ไม่ใช่ผู้เชี่ยวชาญเฉพาะด้าน หัวข้อต่าง ๆ อยู่ในความสนใจโดยทั่วไปของนักศึกษาระดับปริญญาตรีหรือระดับที่สูงกว่า บทอ่านอาจจะมาจากหลาย ๆ รูปแบบตัวอย่างเช่น เชิงพรรณนา (Narrative) เชิงบรรยาย (Descriptive) หรือ แบบอ้อมค้อมหรือโต้แย้ง (Discursive/Argumentative) และจะมีอย่างน้อย 1 บทที่มีรูปแบบการเขียนเชิงโต้แย้งที่มีรายละเอียด บทอ่านอาจจะประกอบด้วยไดอะแกรม (Diagram) กราฟ (Graph) หรือภาพประกอบต่าง ๆ (Illustrations)

รูปแบบการทดสอบ (Test Type) ในการวัดทักษะการอ่านมีรูปแบบข้อคำถามในแบบทดสอบวัดทักษะการอ่านที่อาจจะสุ่มเลือกมาจากข้อคำถามแบบให้เลือกคำตอบ (Multiple Choice) การระบุข้อมูลว่าถูก/ผิด/หรือไม่มีข้อมูลให้ (Identifying Information (True/False/Not Given) การระบุความเห็นหรือการอ้างของผู้เขียนว่าถูก/ผิด/หรือไม่มีข้อมูลให้ (Identifying Writer's Views/Claims (Yes/No/Not given) การจับคู่ข้อมูล จับคู่หัวเรื่อง จับคู่ลักษณะ จับคู่ส่วนท้ายของประโยค (Matching Information, Matching Headings, Matching Features, Matching Sentence Endings) การเติมประโยคให้สมบูรณ์ (Sentence Completion) บทสรุป การเติมโน้ตตาราง ชาร์ตให้สมบูรณ์ (Summary/Note/Table/Flow Chart Completion) การเติมไดอะแกรม

ให้สมบูรณ์ (Diagram Label Completion) และคำถามให้ตอบสั้น ๆ (Short-Answer Questions) (IELTS, 2015)

การตรวจให้คะแนน (Scoring Technique) เป็นแบบตอบถูกได้ 1 หรือ ตอบผิดได้ 0 การทดสอบทักษะการเขียน

จากการศึกษาและวิเคราะห์แบบทดสอบ IELTS (International English Language Testing System) (IELTS, 2015) ที่เกี่ยวกับทักษะการเขียนได้นั้นเนื้อหาสาระแบบทดสอบ คือ 1) จุดประสงค์ (Purpose) 2) สาระภาษา (Language Knowledge/Subskills) 3) ชนิดภาระงาน (Task Type) และ 4) การตรวจให้คะแนน (Scoring Technique) โดยมีรายละเอียด ดังนี้

จุดประสงค์ (Purpose) ในการทดสอบทักษะการเขียน คือ เพื่อทดสอบความสามารถในการใช้ภาษาเขียนของผู้เข้าสอบ

สาระภาษา (Language Knowledge/Subskills) สำหรับการทดสอบทักษะการเขียนครอบคลุมความสามารถในการ 1) เขียนเล่าเรื่องจากการดูรูปภาพได้อย่างมีระบบ 2) ใช้ภาษาที่เหมาะสมกับหัวเรื่องและรูปแบบ จัดระบบ นำเสนอ และอาจจะเปรียบเทียบข้อมูล บรรยายขั้นตอนของกระบวนการหรือวิธีการ บรรยายวัตถุ เหตุการณ์หรือลำดับของเหตุการณ์ และอธิบายวิธีการทำงานของบางสิ่งอย่าง 3) เขียนเชิงโต้แย้งโดยจัดระบบข้อมูลได้ดี ชัดเจน มีประโยชน์ แสดงหลักฐานหรือตัวอย่างเพื่อสนับสนุนความคิดและใช้ภาษาได้อย่างถูกต้อง และ 4) นำเสนอวิธีแก้ปัญหาต่อปัญหาแต่ละอย่าง นำเสนอและปรับเปลี่ยนความคิดเห็น เปรียบเทียบความต่าง ความเหมือนของหลักฐาน ความคิดเห็นและข้อเสนอแนะ และประเมินและชี้แจงน้ำหนักความคิด หลักฐานหรือข้อโต้แย้ง

ชนิด 45 ภาระงาน (Task Type) ในการทดสอบทักษะการเขียนประเมินการใช้ภาษาเขียนของผู้เข้าสอบโดยภาระงาน 2 ชั้น และต้องทำทั้ง 2 ชั้น ในชั้นที่ 1 ผู้เข้าสอบต้องเขียนบรรยายข้อมูลจากการดูภาพ (Visual information) โดยใช้คำพูดของตนเอง (กราฟ ตาราง แผนภาพ หรือไดอะแกรม) ผู้เข้าสอบต้องเขียนอย่างน้อย 150 คำ และในชั้นที่ 2 ผู้เข้าสอบจะได้รับ ความคิดเห็นข้อโต้แย้ง หรือปัญหาที่ผู้เข้าสอบจำเป็นต้องอภิปราย ผู้เข้าสอบจำเป็นต้องเขียนให้ได้อย่างน้อย 250 คำ

การตรวจให้คะแนน (Scoring Technique) แบบทดสอบวัดทักษะการเขียนของแบบทดสอบ IELTS (International English Language Testing System) ประเมินภาระงานที่ 1 ทดสอบความสามารถในการเขียนเล่าเรื่องในภาพรวมได้อย่างมีระบบจากการดูรูปภาพ โดยให้คะแนนโดยใช้เกณฑ์ 5 ด้าน คือ การใช้ภาษาที่เหมาะสมกับหัวเรื่องและรูปแบบ ความสามารถในการจัดระบบ นำเสนอ และอาจจะเปรียบเทียบข้อมูล บรรยายขั้นตอนของกระบวนการหรือวิธีการ บรรยายวัตถุ เหตุการณ์หรือลำดับของเหตุการณ์ และอธิบายวิธีการทำงานของบางสิ่งอย่าง และภาระงานที่ 2 ทดสอบความสามารถเขียนเชิงโต้แย้งที่จัดระบบอย่างดี ชัดเจน มีประโยชน์ แสดงหลักฐานหรือตัวอย่างเพื่อสนับสนุนความคิดและใช้ภาษาได้อย่างถูกต้อง โดยให้คะแนนโดยใช้เกณฑ์ 4 ด้าน คือ การนำเสนอวิธีแก้ปัญหาต่อปัญหาแต่ละอย่าง การนำเสนอและปรับเปลี่ยนความคิดเห็น การเปรียบเทียบความต่าง ความเหมือนของหลักฐาน ความคิดเห็นและข้อเสนอแนะ และประเมินและชี้แจงน้ำหนักความคิด หลักฐานหรือข้อโต้แย้ง ผู้ตรวจแบบทดสอบ IELTS ที่ได้รับการรับรองทำหน้าที่

ประเมินความสามารถของผู้เข้าสอบในแต่ละภาระงาน มีเกณฑ์การประเมิน 4 ด้าน (สิ่งที่ผู้ประเมินต้องระลึกไว้เมื่อตัดสินใจที่จะให้คะแนนงานของผู้เข้าสอบ) คือ การบรรลุภาระงาน/การตอบโจทย์ตรงประเด็น (Task achievement/ response) ความสัมพันธ์เชื่อมโยงของเนื้อหา (Coherence and cohesion) การใช้คำศัพท์ที่เหมาะสม (Lexical resource) และไวยากรณ์ที่ใช้และความถูกต้อง (Grammatical range and accuracy)

แบบทดสอบ FCE (First Certificate in English)

แบบทดสอบ FCE (First Certificate in English) (UCLES, 2014) มีเนื้อหาสาระแบบทดสอบในการทดสอบทักษะการฟัง การพูด การอ่าน การเขียน และการใช้ภาษา ดังรายละเอียดต่อไปนี้

การทดสอบทักษะการฟัง

จากการศึกษาและวิเคราะห์แบบทดสอบ FCE (First Certificate in English) (UCLES, 2014) ที่เกี่ยวกับทักษะการฟังได้เนื้อหาสาระแบบทดสอบ คือ 1) จุดประสงค์ (Purpose) 2) สาระภาษา (Language Knowledge/Subskills) 3) 46 หัวเรื่อง (Topical Knowledge) 4) รูปแบบการทดสอบ (Test Type) และ 5) การตรวจให้คะแนน (Scoring Technique) โดยมีรายละเอียด ดังนี้

จุดประสงค์ (Purpose) ในการทดสอบทักษะการฟัง คือ เพื่อวัดความสามารถในการฟังเพื่อหาความรู้สึกรู้สึก ทักษะการฟังได้เนื้อหาสาระแบบทดสอบ คือ 1) จุดประสงค์ หน้าที่ ข้อตกลง ประเด็นสำคัญ รายละเอียด และข้อมูลจำเพาะจากสาระที่ฟัง

สาระภาษา (Language Knowledge/Subskills) ในการทดสอบทักษะการฟังครอบคลุมความสามารถในการบอก 1) ความรู้สึก ทักษะการฟังได้เนื้อหาสาระแบบทดสอบ คือ 1) จุดประสงค์และหน้าที่ของภาษา 3) ประเด็นสำคัญ และ 4) รายละเอียดหรือมูลเฉพาะ จากสาระที่ฟัง 46 หัวเรื่อง (Topical Knowledge) ของบทฟังเกี่ยวกับบทพูดสั้นๆ ที่ตัดตอนมาจากการพูดของคน ๑ เดียว หรือการสนทนาระหว่างผู้พูดที่กำลังคุยกัน เสียงที่บันทึกอาจจะได้มาจาก บทพูดคนเดียว (Monologues) เช่น ข้อความจากเครื่องรับโทรศัพท์ (Answer Phone Messages) รายการออกอากาศและสารคดีทางวิทยุ (Radio Broadcasts and Features) ข่าว (News) ประกาศสำหรับสาธารณะ (Public Announcements) เรื่องเล่าและเกร็ดความรู้ (Stories and Anecdotes) และการบรรยายและการพูดสั้น ๆ (Lectures and Talks) หรือการพูดที่มีปฏิสัมพันธ์ (Interacting Speakers) เช่น บทสนทนา (Conversations) การสัมภาษณ์ (Interviews) การอภิปราย (Discussions) และบทละครวิทยุ

รูปแบบการทดสอบ (Test Type) ในการวัดทักษะการฟัง ใช้ลักษณะข้อคำถาม 3 แบบ คือ แบบเลือกตอบ (Multiple-Choice) แบบเติมประโยคให้สมบูรณ์ (Sentence Completion) และแบบจับคู่หลายตัวเลือก (Multiple Matching)

การตรวจให้คะแนน (Scoring Technique) เป็นแบบตอบถูกได้ 1 หรือ ตอบผิดได้ 0

การทดสอบทักษะการพูด

จากการศึกษาและวิเคราะห์แบบทดสอบ FCE (First Certificate in English) (UCLES, 2014) ที่เกี่ยวกับทักษะการพูดได้เนื้อหาสาระแบบทดสอบ คือ 1) จุดประสงค์ (Purpose) 2) สาระภาษา (Language Knowledge/Subskills) 3) ชนิด 47 ภาระงาน (Task Type) และ 4) การตรวจให้คะแนน (Scoring Technique) โดยมีรายละเอียด ดังนี้

จุดประสงค์ (Purpose) ในการทดสอบทักษะการพูด คือ เพื่อวัดความสามารถในการพูดตอบคำถาม การแสดงความคิดเห็นกับคู่สนทนา และการอภิปรายกลุ่ม

สาระภาษา (Language Knowledge/Subskills) สำหรับการทดสอบทักษะการพูด ครอบคลุมความสามารถในการ 1) ให้ข้อมูลเกี่ยวกับตนเองและแสดงความคิดเห็นเกี่ยวกับหัวข้อต่าง ๆ 2) พูดเกี่ยวกับบางสิ่งบางอย่าง เปรียบเทียบ บรรยาย แสดงความคิดเห็น และคาดการณ์ และ 3) แลกเปลี่ยนความคิด แสดงและปรับเปลี่ยนความคิดเห็น แสดงการเห็นด้วยและ/หรือไม่เห็นด้วย ให้คำแนะนำ พยากรณ์ ประเมิน เพื่อนำไปสู่การตัดสินใจตลอดการเจรจา เป็นต้น

ชนิด 47 ภาระงาน (Task Type) ในการทดสอบทักษะการพูด ประกอบด้วย 4 ตอนย่อย และต้องสอบพร้อมผู้เข้าสอบคนอื่น ๆ มีกรรมการสอบ 2 คน มีภาระงาน 4 ชนิด คือ ภาระงานที่ 1 การสัมภาษณ์ (Interview) เป็นการสนทนากับกรรมการ ผู้เข้าสอบต้องสามารถให้ข้อมูลเกี่ยวกับตนเองและแสดงความคิดเห็นเกี่ยวกับหัวข้อต่าง ๆ ภาระงานที่ 2 การสลับกันพูดกับเพื่อน (Long turn) โดยการดูข้อมูลจากรูปภาพที่ได้รับ ผู้เข้าสอบพูดอธิบายและแสดงความคิดเห็นต่อรูปภาพ พร้อมทั้งสอบถามความคิดเห็นจากเพื่อนด้วย ภาระงานนี้วัดความสามารถในการพูดเกี่ยวกับบางสิ่งบางอย่าง การเปรียบเทียบ การบรรยาย การแสดงความคิดเห็น และการเดา ภาระงานที่ 3 ภาระงานที่ต้องร่วมมือกัน (Collaborative Task) ผู้เข้าสอบต้องสนทนากับผู้เข้าสอบคนอื่น กรรมการจะให้วัสดุและภาระงานที่ต้องทำ ผู้เข้าสอบพูดคุยกับสมาชิกคนอื่นและช่วยกันตัดสินใจ ภาระงานนี้วัดความสามารถในการแลกเปลี่ยนความคิด แสดงและปรับเปลี่ยนความคิดเห็น การเห็นด้วยและ/หรือไม่เห็นด้วย การให้คำแนะนำ การเดา การประเมิน การนำไปสู่การตัดสินใจตลอดการเจรจา เป็นต้น และ ภาระงานที่ 4 การอภิปราย (Discussion) เป็นการพูดคุยอย่างต่อเนื่องกับสมาชิกอื่น โดยได้รับคำถามนำจากกรรมการเกี่ยวกับหัวเรื่องหรือประเด็นที่พูดกันในภาระงานที่ 3 เพื่อวัดความสามารถในการแสดงและปรับเปลี่ยนความคิดเห็น การเห็นด้วยและ/หรือไม่เห็นด้วย

การตรวจให้คะแนน (Scoring Technique) การพูดของแบบทดสอบ FCE (First Certificate in English) การประเมินผลการพูด โดยให้ผู้ให้คะแนน 2 คน เป็นคู่สนทนา 1 คน ผู้ซึ่งพูดกับผู้เข้าสอบและดำเนินการทดสอบ คนที่เป็นคู่สนทนาจะคอยให้คะแนนผู้เข้าสอบในผลสัมฤทธิ์ในภาพรวม (คะแนนที่แสดงว่าผู้เข้าสอบทำได้ในระดับใดในการสอบทั้งหมด) อีกคนหนึ่งจะเป็นผู้ประเมินที่นั่งฟังอย่างเดียว กรรมการที่ทำหน้าที่ประเมินใช้เกณฑ์การให้คะแนน 4 ด้าน ประกอบด้วย ไวยากรณ์และคำศัพท์ (Grammar and Vocabulary) การจัดการเนื้อความ (Discourse Management) การออกเสียง (Pronunciation) และการสื่อสารเชิงปฏิสัมพันธ์ (Interactive Communication)

การทดสอบทักษะการอ่าน

จากการศึกษาและวิเคราะห์แบบทดสอบ FCE (First Certificate in English) (UCLES, 2014) ที่เกี่ยวกับทักษะการอ่านได้เนื้อหาสาระแบบทดสอบ คือ 1) จุดประสงค์ (Purpose) 2) สาระภาษา (Language Knowledge/Subskills) 3) 48 หัวเรื่อง (Topical Knowledge) 4) รูปแบบการทดสอบ (Test Type) และ 5) การตรวจให้คะแนน (Scoring Technique) โดยมีรายละเอียด ดังนี้

จุดประสงค์ (Purpose) ในการทดสอบทักษะการอ่าน คือ เพื่อวัดความสามารถของผู้เข้าสอบในการเข้าใจบทความชนิดต่าง ๆ

สาระภาษา (Language Knowledge/Subskills) ในการทดสอบทักษะการอ่านครอบคลุมความสามารถในการ 1) ทหารายละเอียด ข้อมูลจำเพาะ 2) ระบุความคิดหลัก 3) ระบุความคิดเห็น น้ำเสียง จุดประสงค์ ของผู้เขียน 4) อนุมานความ (implication) และทัศนคติ และ 5) เข้าใจโครงสร้างและรูปแบบการพัฒนาของตัวบท

48 หัวเรื่อง (Topical Knowledge) ของบทอ่านใช้ตัวเนื้อความต่าง ๆ ในแบบทดสอบอ่านที่นำมาจากหนังสือพิมพ์และนิตยสาร วารสาร หนังสือ ทั้งแบบนิยาย และไม่ใช่นิยาย และสื่อส่งเสริมการขายและให้ข้อมูลข่าวสาร

รูปแบบการทดสอบ (Test Type) ในการวัดทักษะการอ่าน มีรูปแบบข้อคำถามที่ใช้การทดสอบอ่าน 3 แบบ คือ คำถามแบบให้เลือกคำตอบ (Multiple Choice) การเติมคำในช่องว่างของเนื้อความ (Gapped Text) และการจับคู่โดยเลือกจากคำตอบที่มีหลายตัวเลือก (Multiple Matching)

การตรวจให้คะแนน (Scoring Technique) เป็นแบบตอบถูกได้ 1 หรือ ตอบผิดได้ 0

การทดสอบทักษะการเขียน

จากการศึกษาและวิเคราะห์แบบทดสอบ FCE (First Certificate in English) (UCLES, 2014) ที่เกี่ยวกับทักษะการเขียนได้เนื้อหาสาระแบบทดสอบ คือ 1) จุดประสงค์ (Purpose) 2) สาระภาษา (Language Knowledge/Subskills) 3) ชนิดภาระงาน (Task Type) และ 4) การตรวจให้คะแนน (Scoring Technique) โดยมีรายละเอียด ดังนี้

จุดประสงค์ (Purpose) ในการทดสอบทักษะการเขียน คือ เพื่อวัดความสามารถของผู้เข้าสอบในการเขียนบทความภาษาอังกฤษชนิดต่าง ๆ

สาระภาษา (Language Knowledge/Subskills) สำหรับการทดสอบทักษะการเขียนครอบคลุมความสามารถในการ 1) รู้หน้าที่ภาษา (Language Functions) เช่น การประเมิน (Evaluating) การแสดงความคิดเห็น (Expressing Opinions) การตั้งสมมติฐาน (Hypothesizing) การปรับเปลี่ยน (Justifying) และการชักชวน (Persuading) และ 2) เขียนตัวบทชนิดต่าง ๆ ที่กำหนดไว้ในแบบทดสอบ

ชนิด 48 ภาระงาน (Task Type) ในการทดสอบทักษะการเขียนวัดความสามารถในการเขียนของผู้เข้าสอบภาระงาน 2 ชนิด คือ ภาระงานอาจจะเป็น Articles, Email, Essay, Letter, Report,

Review ภาระงานที่ 1 เป็นคำถามบังคับ โดยผู้เข้าสอบจะได้รับหัวข้อเรียงความ และความคิด 2 อย่าง ที่สัมพันธ์กับหัวข้ออย่างชัดเจน ผู้เข้าสอบเขียนเรียงความโดยแสดงความคิดเห็นเกี่ยวกับหัวข้อเรื่อง ใช้ความคิดที่กำหนดให้ และต้องเพิ่มเติมความคิดที่ 3 ที่เป็นความคิดเห็นของตนเองที่สัมพันธ์กับหัวข้อเรื่องด้วย หัวข้อที่กำหนดให้อยู่ในความสนใจทั่วไป ไม่จำเป็นต้องมีความรู้พิเศษหรือเฉพาะด้าน ความยาวของเรียงความควรอยู่ระหว่าง 140 – 190 คำ ความสามารถที่ผู้เข้าสอบควรแสดงให้เห็นคือ การหน้าที่ภาษา (Language Functions) เช่น การประเมิน (Evaluating) การแสดงความคิดเห็น (Expressing Opinions) การตั้งสมมติฐาน (Hypothesizing) การปรับเปลี่ยน (Justifying) และการชักชวน (Persuading) และภาระงานที่ 2 การเขียนตามสถานการณ์ เป็นคำถามให้เลือก 1 จาก 3 ตัวเลือก ผู้เข้าสอบเขียนตัวบทโดยการเลือก 1 อย่างจากชนิดของงานเขียน เช่น Article, Email/letter, Report or Review ผู้เข้าสอบจะได้รับบริบท จุดประสงค์ของหัวข้อเรื่อง และผู้อ่านเป้าหมาย ความยาวของเรียงความควรอยู่ระหว่าง 140 – 190 คำ ความสามารถที่ผู้เข้าสอบควรแสดงให้เห็น คือ การเขียนตัวบทชนิดต่าง ๆ ที่กำหนดไว้ในแบบทดสอบ

การตรวจให้คะแนน (Scoring Technique) ของแบบทดสอบ FCE (First Certificate in English) การประเมินงานเขียนใช้เกณฑ์ 4 ด้าน คือ Content, Communicative Achievement, Organisation and Language ในด้าน Content เน้นว่าผู้เข้าสอบได้ผลิตภาระงานได้สมบูรณ์เพียงใด หรือว่าผู้เข้าสอบได้ทำในสิ่งที่กำหนดให้ทำหรือไม่ ด้าน Communicative Achievement จะตรวจสอบว่างานเขียนเหมาะสมกับภาระงานหรือไม่ (ตัวอย่างเช่น รูปแบบการเขียนตรงกับการเขียนบทความในนิตยสารหรือไม่) และเลือกใช้ภาษาได้ถูกต้องหรือไม่ เช่น เป็นทางการหรือไม่เป็นทางการ ด้านการจัดระบบเนื้อความ เน้นในเรื่องการจัดเรียงประโยค หรือว่าสมเหตุสมผลและตามลำดับหรือไม่ และด้านภาษา เน้นที่คำศัพท์และไวยากรณ์ รวมถึงความหลากหลายของภาษา เช่นเดียวกับความถูกต้องด้วย (UCLES, 2014)

การทดสอบทักษะการใช้ภาษา

จากการศึกษาและวิเคราะห์แบบทดสอบ FCE (First Certificate in English) (UCLES, 2014) ที่เกี่ยวกับทักษะการใช้ภาษาได้เนื้อหาสาระแบบทดสอบ คือ 1) จุดประสงค์ (Purpose) 2) สาระภาษา (Language Knowledge/Subskills) 3) หัวเรื่อง (Topical Knowledge) 4) รูปแบบการทดสอบ (Test Type) และ 5) การตรวจให้คะแนน (Scoring Technique) โดยมีรายละเอียด ดังนี้

จุดประสงค์ (Purpose) ในการทดสอบทักษะการใช้ภาษา คือ มีจุดประสงค์เพื่อกำหนดให้ผู้เข้าสอบ ได้อ่านข้อความหลากหลายชนิดและทำภาระงานที่เกี่ยวกับไวยากรณ์และคำศัพท์

สาระภาษา (Language Knowledge/Subskills) ในการทดสอบทักษะการใช้ภาษา ครอบคลุมความสามารถในการอ่านข้อความหลากหลายชนิดและทำภาระงานที่เกี่ยวกับไวยากรณ์และคำศัพท์ เช่น สำนวน (Idioms) คำที่ใช้ร่วมกัน (Collocations) ความหมายในแง่มุมต่าง ๆ (Shades of Meaning) กริยาวลี (Phrasal Verbs) วลีตายตัว (Fixed Phrases) เป็นต้น

50 หัวเรื่อง (Topical Knowledge) ในการทดสอบการใช้ภาษามีสาระที่เกี่ยวข้องกับการอ่านข้อความหลายชนิดและทำภาระงานที่เกี่ยวกับไวยากรณ์และคำศัพท์

รูปแบบการทดสอบ (Test Type) ในการวัดทักษะการใช้ภาษามีลักษณะข้อคำถามหลายชนิด เช่น เติมคำในเนื้อเรื่องโดยเลือกจากตัวเลือก (Multiple-Choice Cloze) เติมคำในช่องว่างของเนื้อเรื่องโดยคิดคำเอง (Open Cloze) การสร้างคำ (Word Formation) และการเปลี่ยนรูปของคำสำคัญ (Key Word Transformations)

การตรวจให้คะแนน (Scoring Technique) เป็นแบบตอบถูกได้ 1 หรือ ตอบผิดได้ 0

แบบทดสอบ telc B1-B2 English

แบบทดสอบ telc B1-B2 English (telc GmbH, 2012) มีเนื้อหาสาระแบบทดสอบในการทดสอบทักษะการฟัง การพูด การอ่าน การเขียน และการใช้ภาษา ดังรายละเอียดต่อไปนี้

การทดสอบทักษะการฟัง

จากการศึกษาและวิเคราะห์แบบทดสอบ telc B1-B2 English (telc GmbH, 2012) ที่เกี่ยวกับทักษะการฟัง ได้เนื้อหาสาระแบบทดสอบ คือ 1) จุดประสงค์ (Purpose) 2) สาระภาษา (Language Knowledge/Subskills) 3) 50 หัวเรื่อง (Topical Knowledge) 4) รูปแบบการทดสอบ (Test Type) และ 5) การตรวจให้คะแนน (Scoring Technique) โดยมีรายละเอียด ดังนี้

จุดประสงค์ (Purpose) ในการทดสอบทักษะการฟัง คือ เพื่อวัดความสามารถในการฟังเพื่อจับรายละเอียด และการฟังเพื่อความเข้าใจโดยรวม (telc GmbH, 2012)

สาระภาษา (Language Knowledge/Subskills) ในการทดสอบทักษะการฟังครอบคลุมความสามารถในการ 1) จับรายละเอียด และ 2) จับใจความสำคัญของสาระที่ฟัง

50 หัวเรื่อง (Topical Knowledge) ของบทฟังได้มาจากข้อความเสียง (Voice Mail Messages) การสนทนาประจำวัน (Everyday Conversation) การสัมภาษณ์ (An Interview) และความคิดเห็นเกี่ยวกับหัวเรื่อง (Opinions about a Topic) ลักษณะของเสียงบันทึก (Text type) เป็นแบบ ผู้พูดคนเดียว (Monologues) ได้แก่ ข้อความเสียง (Voice Mail Messages) ทั้งในบริบทที่เกี่ยวกับเรื่องส่วนตัวและการทำงาน และรายการที่มีคนแสดงความคิดเห็นในประเด็นเฉพาะเรื่อง (Programme with people stating their opinions on a particular topic) และประเภทบทสนทนา (Dialogues) ได้แก่ บทสนทนาในสถานการณ์ประจำวัน (Conversations in everyday situations) ทั้งในบริบทที่เกี่ยวกับเรื่องส่วนตัวและการทำงาน และรายการวิทยุหรือบทสัมภาษณ์จาก podcast ส่วนใหญ่นำมาจากการสื่อสารในชีวิตประจำวันทั่วไป และสถานที่ทำงาน

รูปแบบการทดสอบ (Test Type) ในการวัดทักษะการฟังใช้รูปแบบข้อคำถาม 3 แบบ คือ แบบเลือกตอบ (Multiple-Choice Items) แบบถูกหรือผิด (True/False Items) และแบบจับคู่ (Matching Items)

การตรวจให้คะแนน (Scoring Technique) เป็นแบบตอบถูกได้ 1 หรือ ตอบผิดได้ 0

การทดสอบทักษะการพูด

จากการศึกษาและวิเคราะห์แบบทดสอบ telc B1-B2 English (telc GmbH, 2012) ที่เกี่ยวกับทักษะการพูดได้เนื้อหาสาระแบบทดสอบ คือ 1) จุดประสงค์ (Purpose) 2) สาระภาษา (Language Knowledge/Subskills) 3) ชนิด 51 ภาระงาน (Task Type) และ 4) การตรวจให้คะแนน (Scoring Technique) โดยมีรายละเอียด ดังนี้

จุดประสงค์ (Purpose) ในการทดสอบทักษะการพูด คือ เพื่อวัดความสามารถผู้เข้าสอบในการพูดเกี่ยวกับประสบการณ์และความคิดเห็น (Talking about experiences and opinions) การนำเสนอ (Presentation) การตอบคำถามจากกรรมการ (Answering follow-up questions) การพูดคุยกัน

สาระภาษา (Language Knowledge/Subskills) สำหรับการทดสอบทักษะการพูดครอบคลุมความสามารถในการ 1) พูดเกี่ยวกับประสบการณ์และความคิดเห็น (Talking about experiences and opinions) การนำเสนอ (Presentation) การตอบคำถามจากกรรมการ (Answering follow-up questions) การพูดคุยกัน (Discussion) 2) บรรยายสถานการณ์ในรูปภาพได้ละเอียด ชัดเจน 3) ตอบคำถามของผู้ประเมินได้อย่างละเอียด 4) ทำการนำเสนอที่มีโครงสร้างชัดเจนเพื่อให้เหตุผลสนับสนุนความคิดเห็นของตนเอง 5) ตอบคำถามของผู้ประเมินได้อย่างละเอียด และ 6) แลกเปลี่ยนความคิดเห็นกับคู่สนทนาและทำให้การสนทนาดำเนินต่อไปได้

ชนิด 51 ภาระงาน (Task Type) ในการทดสอบทักษะการพูด ผู้เข้าสอบจะถูกจัดให้สอบเป็นคู่ พร้อมกับมีกรรมการดำเนินการสอบ 2 คน ทำหน้าที่ประเมินความสามารถของแต่ละคู่ ถ้ามีผู้เข้าสอบที่ขาดคู่ ผู้เข้าสอบจะได้สอบกับกรรมการเพียงคนเดียว ส่วนกรรมการอีกคนจะทำหน้าที่เป็นคู่สอบให้ ก่อนการสอบ ผู้เข้าสอบจะมีเวลาเตรียมตัว 20 นาที เมื่อผู้เข้าสอบเข้าไปในห้องเตรียมตัว จะได้รับบัตรภาระงาน สำหรับตอนที่ 2 และ 3 ในระหว่าง 20 นาที ผู้เข้าสอบสามารถคิดถึงสิ่งที่จะพูดหรืออาจจะจดบันทึกในกระดาษอื่นได้ แต่จะไม่อนุญาตให้ใช้พจนานุกรมหรือสื่อสารกับคู่ของตน โดยมีรายละเอียดในแต่ละภาระงาน คือ ใน Part 1A: กำหนดให้ผู้เข้าสอบบรรยายสถานการณ์ในรูปภาพได้อย่างละเอียดและชัดเจน ใน Part 1B: กำหนดให้ผู้เข้าสอบตอบคำถามของผู้ประเมินได้อย่างถูกต้อง ใน Part 2A: กำหนดให้ผู้เข้าสอบทำการนำเสนอที่มีโครงสร้างชัดเจนเพื่อให้เหตุผลสนับสนุนความคิดเห็นของตนเอง ใน Part 2B: กำหนดให้ผู้เข้าสอบตอบคำถามของผู้ประเมินได้อย่างถูกต้อง และใน Part 3: กำหนดให้ผู้เข้าสอบแลกเปลี่ยนความคิดเห็นกับคู่สนทนาและทำให้การสนทนาดำเนินต่อไปได้ (telc GmbH, 2012)

การตรวจให้คะแนน (Scoring Technique) ของแบบทดสอบ telc B1-B2 English ใช้เกณฑ์การประเมินการทดสอบพูด 5 ด้าน คือ การจัดการภาระงาน (Task Management) การออกเสียงและการใช้เสียงสูงต่ำ (Pronunciation and Intonation) ความคล่องแคล่ว (Fluency) ความถูกต้อง (Accuracy) และคำศัพท์ (Vocabulary) ในด้านคือ การจัดการภาระงาน (Task Management) จะดูว่าผู้เข้าสอบแต่ละคนทำภาระงานได้ครบถ้วนตามที่กำหนดในแต่ละภาระงานย่อยหรือไม่ เช่น ใน Part 1A: ผู้เข้าสอบสามารถทำการบรรยายสถานการณ์ในรูปภาพได้ละเอียดชัดเจนหรือไม่ Part 1B: ผู้เข้าสอบสามารถตอบคำถามของผู้ประเมินได้อย่างละเอียด ใน Part 2A: ผู้

เข้าสอบทำการนำเสนอที่มีโครงสร้างชัดเจนเพื่อให้เหตุผลสนับสนุนความคิดเห็นของตนเอง ใน Part 2B: ผู้เข้าสอบสามารถตอบคำถามของผู้ประเมินได้อย่างละเอียด และใน Part 3: จะดูว่าผู้เข้าสอบสามารถแลกเปลี่ยนความคิดเห็นกับคู่สนทนาและทำให้การสนทนาดำเนินต่อไปได้หรือไม่ สำหรับในอีก 4 ด้าน จะมีรายละเอียดว่า ในด้านการออกเสียงและการใช้เสียงสูงต่ำ (Pronunciation and Intonation) จะดูว่าผู้เข้าสอบพูดได้ชัดเจนพอที่จะให้เข้าใจความหรือไม่ ในด้านความคล่องแคล่ว (Fluency) จะตรวจสอบว่าผู้เข้าสอบพูดได้โดยฉับพลัน มีการหยุดที่ผิดธรรมชาติเพียงเล็กน้อย ในด้านความถูกต้อง (Accuracy) ดูว่าผู้เข้าสอบใช้โครงสร้างไวยากรณ์ได้อย่างมีประสิทธิภาพ และในด้านคำศัพท์ (Vocabulary) จะดูว่าผู้เข้าสอบใช้คำศัพท์ที่หลากหลายเพื่อนำเสนอสิ่งที่ต้องการจะพูด

การทดสอบทักษะการอ่าน

จากการศึกษาและวิเคราะห์แบบทดสอบ telc B1-B2 English (telc GmbH, 2012) ที่เกี่ยวกับทักษะการอ่านได้เนื้อหาสาระแบบทดสอบ คือ 1) จุดประสงค์ (Purpose) 2) สาระภาษา (Language Knowledge/Subskills) 3) 52 หัวเรื่อง (Topical Knowledge) 4) รูปแบบการทดสอบ (Test Type) และ 5) การตรวจให้คะแนน (Scoring Technique) โดยมีรายละเอียด ดังนี้

จุดประสงค์ (Purpose) ในการทดสอบทักษะการอ่าน คือ เพื่อวัดความสามารถของผู้เข้าสอบโดยเน้นการอ่านเพื่อจับรายละเอียด (Reading for Detail) การอ่านเพื่อความเข้าใจโดยรวม (Reading for Gist) และการเลือกอ่านข้อมูลในชีวิตประจำวัน (Selective Reading)

สาระภาษา (Language Knowledge/Subskills) ในการทดสอบทักษะการอ่านครอบคลุมความสามารถในการอ่านเพื่อเข้าใจรายละเอียด และใจความหลักหรือประเด็นสำคัญของสาระที่อ่าน

52 หัวเรื่อง (Topical Knowledge) ของบทอ่านเป็นเนื้อความต่าง ๆ ในแบบทดสอบอ่านที่นำมาจาก อีเมลล์และบรรทัดบอกหัวเรื่อง (Emails and Subject Lines) คำถามและคำตอบที่นำมาจากส่วนที่เปิดให้แสดงความคิดเห็นทางอินเทอร์เน็ต (Questions and Answers from an Internet Forum) บทอ่านสาธารณะชนิดต่าง ๆ (Different Types of Published Texts) และบทอ่านประเภทให้ข้อมูลที่เป็นทางการ (Formal Informative Texts) ส่วนลักษณะของเนื้อหาบทอ่าน (Text type) จะครอบคลุมถึงอีเมลล์ส่วนตัวและกึ่งทางการ (Personal or Semi-Formal Emails) ข้อความที่โพสต์บนส่วนที่ให้แสดงความคิดเห็นบนอินเทอร์เน็ต (Messages Posted on an Internet Discussion Forum) บทอ่านประเภทให้ข้อมูลเกี่ยวกับหัวเรื่องพิเศษ (Informative Text about a Special Subject) บทความในหนังสือพิมพ์ (Newspaper article) โฆษณา (Advertisement) และบทอ่านประเภทให้ข้อมูลที่เป็นทางการ (Formal Informative Texts) (telc GmbH, 2012: 29-32) รูปแบบการทดสอบ (Test Type) ในการวัดทักษะการอ่านมีรูปแบบข้อคำถามที่ใช้การทดสอบอ่าน 3 แบบ คือ ข้อคำถามแบบจับคู่ (Matching Items) คำถามแบบให้เลือกคำตอบ (Multiple-Choice Items) และข้อคำถามแบบให้ระบุ ถูก/ผิด (True/False Items)

การตรวจให้คะแนน (Scoring Technique) เป็นแบบตอบถูกได้ 1 หรือ ตอบผิดได้ 0

การทดสอบทักษะการเขียน

จากการศึกษาและวิเคราะห์แบบทดสอบ telc B1-B2 English (telc GmbH, 2012) ที่เกี่ยวกับทักษะการเขียนได้นี้เนื้อหาสาระแบบทดสอบ คือ 1) จุดประสงค์ (Purpose) 2) สาระภาษา (Language Knowledge/Subskills) 3) ชนิดภาระงาน (Task Type) และ 4) การตรวจให้คะแนน (Scoring Technique) โดยมีรายละเอียด ดังนี้

จุดประสงค์ (Purpose) ในการทดสอบทักษะการเขียน คือ เพื่อวัดความสามารถในการสื่อสารด้วยการเขียนของผู้เข้าสอบ

สาระภาษา (Language Knowledge/Subskills) สำหรับการทดสอบทักษะการเขียน ครอบคลุมความสามารถในการสื่อสารด้วยการเขียน

ชนิด 53 ภาระงาน (Task Type) ในการทดสอบทักษะการเขียน ผู้เข้าสอบถูกกำหนดให้ทำภาระงาน 1 ชิ้น ซึ่งเป็นการเขียนอีเมลกึ่งทางการ (Writing Semi-Formal Emails) โดยเลือก 1 ใน 2 หัวข้อที่กำหนดให้ มีจุดประสงค์เพื่อวัดความสามารถในการสื่อสารด้วยการเขียนของผู้เข้าสอบ สารรับเข้า (Input Text) ที่ผู้เข้าสอบจะได้รับ คือ คำบรรยายสั้น ๆ เกี่ยวกับสถานการณ์และประเด็นแนะนำ 3 อย่าง (A brief description of the situation and three guiding points) และสารส่งออก (Output text) ที่ผู้เข้าสอบจะต้องผลิต คือ ข้อความอีเมลอย่างเป็นทางการสั้น ๆ ที่ใช้สื่อสารกันโดยทั่วไป ส่วนตัว หรือเกี่ยวข้องกับบริบทในที่ทำงาน

การตรวจให้คะแนน (Scoring Technique) ของแบบทดสอบ telc B1-B2 English ประเมินการเขียนโดยใช้เกณฑ์ 4 ด้าน คือ Task Management ซึ่งจะดูว่าผู้เข้าสอบได้กล่าวถึงประเด็นที่กำหนดให้ครบถ้วนหรือไม่ ด้าน Communicative Design จะดูว่าผู้เข้าสอบได้ใช้ภาษาเป็นทางการหรือไม่เป็นทางการเหมาะสมหรือไม่ ได้เชื่อมโยงความคิดเพื่อให้ได้เนื้อความที่มีโครงสร้างดี ชัดเจนหรือไม่ ด้าน Accuracy จะดูว่า ไวยากรณ์และการสะกดคำถูกต้องโดยส่วนใหญ่หรือไม่ และด้าน Vocabulary จะดูว่าได้ใช้คำศัพท์ที่หลากหลายเพื่อสื่อสารของคุณหรือไม่ (telc GmbH, 2012)

การทดสอบทักษะการใช้ภาษา

จากการศึกษาและวิเคราะห์แบบทดสอบ telc B1-B2 English (telc GmbH, 2012) ที่เกี่ยวกับทักษะการใช้ภาษาได้นี้เนื้อหาสาระแบบทดสอบ คือ 1) จุดประสงค์ (Purpose) 2) สาระภาษา (Language Knowledge/Subskills) 3) หัวเรื่อง (Topical Knowledge) 4) รูปแบบการทดสอบ (Test Type) และ 5) การตรวจให้คะแนน (Scoring Technique) โดยมีรายละเอียด ดังนี้

จุดประสงค์ (Purpose) ในการทดสอบทักษะการใช้ภาษา คือ เพื่อวัดความเข้าใจของผู้เข้าสอบในการใช้กลุ่มคำที่เหมาะสมในบทสนทนา (Appropriate Phrases in a Conversation) และในจดหมายหรืออีเมลกึ่งทางการ (Appropriate Phrases in a Semi-Formal Letter or Email)

สาระภาษา (Language Knowledge/Subskills) ในการทดสอบทักษะการใช้ภาษา ครอบคลุมความสามารถในการเข้าใจการใช้กลุ่มคำที่เหมาะสมในบทสนทนา (Appropriate Phrases

in a Conversation) และในจดหมายหรืออีเมลกึ่งทางการ (Appropriate Phrases in a Semi-Formal Letter or Email)

54 หัวเรื่อง (Topical Knowledge) ในการทดสอบทักษะการใช้ภาษามีสาระที่เกี่ยวข้องกับการใช้กลุ่มคำที่เหมาะสมในบทสนทนา (Appropriate Phrases in a Conversation) และในจดหมายหรืออีเมลกึ่งทางการ (Appropriate Phrases in a Semi-Formal Letter or email) รูปแบบการทดสอบ (Test Type) ในการวัดทักษะการใช้ภาษาใช้รูปแบบข้อคำถามแบบจับคู่ (Matching Items) และ แบบเลือกตอบ (Multiple-Choice Items) ส่วนลักษณะของเนื้อหาบทอ่าน (Text type) เป็น บทสนทนา (Conversation) และ จดหมายสั้น ๆ หรืออีเมลในบริบททั่ว ๆ หรือที่สัมพันธ์กับที่ทำงาน (Short letters or emails in a public or work-related context)

การตรวจให้คะแนน (Scoring Technique) เป็นแบบตอบถูกได้ 1 หรือ ตอบผิดได้ 0

2.1.2.4 แบบทดสอบวัดความสามารถภาษาอังกฤษระดับอุดมศึกษา (University English Proficiency Test)

จากการศึกษาและวิเคราะห์โครงสร้างแบบทดสอบวัดความสามารถภาษาอังกฤษระดับอุดมศึกษา 3 ชนิด ได้แก่ แบบทดสอบ CET (College English Test) (Zheng Ying & Cheng Liying, 2008) แบบทดสอบ MUET (Malaysian University English Test) (Malaysian Examinations Council, 2006) และ แบบทดสอบ CU-TEP (Chulalongkorn University Test of English Proficiency) (สถาบันภาษา จุฬาลงกรณ์มหาวิทยาลัย, 2552) ได้เนื้อหาสาระแบบทดสอบที่จะนำไปใช้ออกแบบทดสอบในแต่ละส่วน 7 ประการ คือ 1) จุดประสงค์ (Purpose) 2) สาระภาษา (Language Knowledge/Subskills) 3) 54 หัวเรื่อง (Topical Knowledge) 4) ชนิดของเนื้อความ (Text Type) 5) รูปแบบภาระงาน (Task Type) 6) รูปแบบการทดสอบ (Test Type) และ 7) การตรวจให้คะแนน (Scoring Technique) โดยมีรายละเอียดในแต่ละชนิดแบบทดสอบ ดังต่อไปนี้

แบบทดสอบ CET (College English Test)

แบบทดสอบ CET (College English Test) (Zheng Ying & Cheng Liying, 2008) มีเนื้อหาสาระแบบทดสอบในการทดสอบทักษะการฟัง การพูด การอ่าน การเขียน และการใช้ภาษา ดังรายละเอียดต่อไปนี้

การทดสอบทักษะการฟัง

จากการศึกษาและวิเคราะห์แบบทดสอบ CET (College English Test) (Zheng Ying & Cheng Liying, 2008) ที่เกี่ยวกับทักษะการฟัง ได้เนื้อหาสาระแบบทดสอบ คือ 1) จุดประสงค์ (Purpose) 2) สาระภาษา (Language Knowledge/Subskills) 3) 54 หัวเรื่อง (Topical

Knowledge) 4) รูปแบบการทดสอบ (Test Type) และ 5) การตรวจให้คะแนน (Scoring Technique) โดยมีรายละเอียด ดังนี้

จุดประสงค์ (Purpose) ในการทดสอบทักษะการฟัง คือ เพื่อวัดความสามารถของนักศึกษา ในการเข้าใจการสนทนาด้วยการพูด (Oral Conversation) หรือบทความ (Passages) ตามมาตรฐาน ของการใช้ภาษาอังกฤษแบบ American English หรือ British English (Zheng Ying & Cheng Liying, 2008) สารระภาษา (Language Knowledge/Subskills) ในการทดสอบทักษะการฟัง ครอบคลุมความสามารถในการ 1) เข้าใจภาษาพูดในบทสนทนา และ 2) เข้าใจบทฟังที่มีสำเนียงพูด มาตรฐาน

55 หัวเรื่อง (Topical Knowledge) ของบทฟัง แบบทดสอบ CET (College English Test) มีบทฟังที่ประกอบด้วยบทสนทนาสั้นและยาว บทความ และการฟังเรื่องแล้วเขียน (Compound Dictation) จำนวน 1 บท (Zheng & Cheng, 2008: 410)

รูปแบบการทดสอบ (Test Type) แบบทดสอบ CET (College English Test) ใช้เทคนิคการสอบใน การวัดทักษะการฟัง 2 แบบ คือ แบบเลือกตอบ (Multiple-Choice) และแบบฟังแล้วเขียน (Compound Dictation) (Zheng & Cheng, 2008: 410)

การตรวจให้คะแนน (Scoring Technique) เป็นแบบตอบถูกได้ 1 หรือ ตอบผิดได้ 0

การทดสอบทักษะการพูด

จากการศึกษาและวิเคราะห์แบบทดสอบ CET (College English Test) (Zheng & Cheng, 2008) ที่เกี่ยวกับทักษะการพูดได้เนื้อหาสาระแบบทดสอบ คือ 1) จุดประสงค์ (Purpose) 2) สารระ ภาษา (Language Knowledge/Subskills) 3) ชนิด 55 ภาระงาน (Task Type) และ 4) การตรวจ ให้คะแนน (Scoring Technique) โดยมีรายละเอียด ดังนี้

จุดประสงค์ (Purpose) ในการทดสอบทักษะการพูด คือ เพื่อวัดความสามารถในการสื่อสารด้วยการ พูด (Zheng & Cheng, 2008: 412)

สารระภาษา (Language Knowledge/Subskills) สำหรับการทดสอบทักษะการพูด ครอบคลุมความสามารถในการ 1) มีปฏิสัมพันธ์ในการอภิปรายกลุ่ม 2) นำเสนอด้วยการพูด 3) ตอบ คำถาม 4) ออกเสียง ลงเสียงหนักเบา/สูงต่ำ และใช้ไวยากรณ์และคำศัพท์ได้ถูกต้อง 5) ใช้คำศัพท์และ โครงสร้างไวยากรณ์หลากหลาย 6) มีส่วนร่วมในการอภิปรายกลุ่ม 7) สามารถผลิตข้อความที่ขยาย เพิ่มและต่อเนื่องกัน 8) จัดการกับสถานการณ์และหัวข้อต่าง ๆ ได้ และ 9) ใช้ทรัพยากรทางภาษาตาม บริบทได้เหมาะสม

แบบทดสอบ CET (College English Test) กำหนดภาระงานในการทดสอบทักษะการพูด ใน 3 ชนิด คือ 1) การปฏิสัมพันธ์ระหว่างผู้เข้าสอบ จำนวน 3-4 คน และผู้ดำเนินการสอบ 2 คน (1 คนเป็นคู่สนทนาหรือผู้ประเมิน และอีก 1 คนเป็นผู้ประเมิน) ใช้เวลาสนทนาประมาณ 5 นาที 2) การ นำเสนอรายบุคคล ใช้เวลา 1.5 นาที และการอภิปรายกลุ่ม ใช้เวลา 4.5 นาที (ใช้เวลาทั้งหมด ประมาณ 10 นาที) แล 3) การตอบคำถามเพิ่มเติมจากผู้ดำเนินการสอบ เพื่อทดสอบความสามารถใน การพูดภาษาอังกฤษของผู้เข้าสอบ (ประมาณ 5 นาที) (Zheng & Cheng, 2008: 412)

แบบทดสอบ CET (College English Test) กำหนดเกณฑ์การให้คะแนนการทดสอบทักษะการพูด 6 ด้าน คือ 1) ความถูกต้อง (Accuracy) ในการออกเสียง ลงเสียงหนักเบา/สูงต่ำ และการใช้ไวยากรณ์และคำศัพท์ 2) ปริมาณ (Range) ของคำศัพท์และโครงสร้างไวยากรณ์ 3) ขนาด (Size) (เปอร์เซ็นต์) ของการมีส่วนร่วมในการอภิปรายกลุ่มของผู้เข้าสอบแต่ละคน 4) การจัดการระบบเนื้อความ (Discourse Management) ที่ได้จากความสามารภในการผลิตข้อความที่ขยายเพิ่มและต่อเนื่องกัน 5) ความยืดหยุ่น (Flexibility) ในการจัดการกับสถานการณ์และหัวข้อต่าง ๆ และ 6) ความเหมาะสม (Appropriateness) ในการใช้ทรัพยากรทางภาษาตามบริบท (Zheng & Cheng, 2008: 412)

การทดสอบทักษะการอ่าน

จากการศึกษาและวิเคราะห์แบบทดสอบ CET (College English Test) (Zheng & Cheng, 2008) ที่เกี่ยวกับทักษะการอ่านได้นี้เนื้อหาสาระแบบทดสอบ คือ 1) จุดประสงค์ (Purpose) 2) สาระภาษา (Language Knowledge/Subskills) 3) หัวเรื่อง (Topical Knowledge) 4) รูปแบบการทดสอบ (Test Type) และ 5) การตรวจให้คะแนน (Scoring Technique) โดยมีรายละเอียด ดังนี้ จุดประสงค์ (Purpose) ในการทดสอบทักษะการอ่าน คือ เพื่อวัดความสามารถของผู้เข้าสอบในการเข้าใจประเด็นสำคัญของเรื่องและหาข้อมูลเฉพาะ (Zheng & Cheng, 2008: 410-411)

สาระภาษา (Language Knowledge/Subskills) ในการทดสอบทักษะการอ่านครอบคลุมความสามารถในการ 1) เข้าใจประเด็นสำคัญของเรื่อง และ 2) ทหารายละเอียดเฉพาะในบทอ่าน

แบบทดสอบ CET (College English Test) ในส่วนการอ่านประกอบด้วยแบบทดสอบ 2 ชนิด คือ การอ่านเพื่อลงลึกในรายละเอียด (Reading in Depth) และการอ่านแบบ Skimming and Scanning (Zheng & Cheng, 2008: 410-411)

แบบทดสอบ CET (College English Test) ใช้เทคนิคการสอบสำหรับแบบทดสอบการอ่านลงรายละเอียด (Reading in Depth) เป็นคำถามแบบให้เลือกตอบ (Multiple-Choice Questions) และเติมคำในช่องว่างแบบมีตัวเลือกหรือคำถามให้ตอบสั้น ๆ (Banked Cloze/Short Answer Questions) ในส่วนของการอ่านคร่าว ๆ และการอ่านเพื่อหารายละเอียด (Skimming and Scanning) เป็นแบบทดสอบประเภท คำถามแบบให้เลือกตอบ (Multiple-Choice Questions) ข้อความที่ให้ระบุว่าถูกหรือผิด (True/False Statements) และการเติมคำในประโยคให้สมบูรณ์ (Sentence Completion) (Zheng & Cheng, 2008: 410-411)

การทดสอบทักษะการเขียน

จากการศึกษาและวิเคราะห์แบบทดสอบ CET (College English Test) (Zheng & Cheng, 2008) ที่เกี่ยวกับทักษะการเขียนได้นี้เนื้อหาสาระแบบทดสอบ คือ 1) จุดประสงค์ (Purpose) 2) สาระภาษา (Language Knowledge/Subskills) 3) ชนิด 56 ภาระงาน (Task Type) และ 4) การตรวจให้คะแนน (Scoring Technique) โดยมีรายละเอียด ดังนี้

จุดประสงค์ (Purpose) ในการทดสอบทักษะการเขียน คือ เพื่อประเมินความสามารถของนักศึกษาในการสื่อสารความคิดโดยใช้ภาษาเขียน (Zheng & Cheng, 2008: 411)

สาระภาษา (Language Knowledge/Subskills) สำหรับการทดสอบทักษะการเขียน ครอบคลุมความสามารถในการ 1) เขียนเรียงความโดยใช้ข้อมูลที่กำหนดให้ และ 2) แปลประโยคที่กำหนดให้จากภาษาจีนเป็นภาษาอังกฤษ

แบบทดสอบ CET (College English Test) กำหนดภาระงานในการทดสอบการเขียน 2 ชนิดคือ การเขียน และการแปล สำหรับการเขียน นักศึกษาถูกกำหนดให้เขียนเรียงความไม่น้อยกว่า 120 คำ สำหรับ CET-4 และ 150 คำ สำหรับ CET-6 โดยอาศัยข้อมูลที่กำหนดให้ เช่น ชื่อหัวเรื่อง โครงร่าง สถานการณ์ รูปภาพ หรือ กราฟ เป็นต้น และสำหรับการแปล นักศึกษาถูกกำหนดให้แปลประโยค จากภาษาจีนเป็นภาษาอังกฤษให้สมบูรณ์ 5 ประโยค (Zheng & Cheng, 2008: 411)

การทดสอบทักษะการใช้ภาษา

จากการศึกษาและวิเคราะห์แบบทดสอบ CET (College English Test) (Zheng & Cheng, 2008) ที่เกี่ยวกับทักษะการใช้ภาษาได้เนื้อหาสาระแบบทดสอบ คือ 1) จุดประสงค์ (Purpose) 2) สาระภาษา (Language Knowledge/Subskills) 3) 57 หัวเรื่อง (Topical Knowledge) 4) รูปแบบการทดสอบ (Test Type) และ 5) การตรวจให้คะแนน (Scoring Technique) โดยมีรายละเอียดดังนี้

จุดประสงค์ (Purpose) ในการทดสอบทักษะการใช้ภาษา คือ เพื่อประเมินความสามารถของนักศึกษาในการเข้าใจและตีความความหมายจากระดับคำศัพท์ ประโยค และย่อหน้าโดยใช้แบบทดสอบแบบ Cloze (Zheng & Cheng, 2008: 411)

สาระภาษา (Language Knowledge/Subskills) ในการทดสอบทักษะการใช้ภาษา ครอบคลุมเข้าใจและตีความความหมายจากระดับคำศัพท์ ประโยค และย่อหน้า

แบบทดสอบ CET (College English Test) มีสาระที่เกี่ยวข้องกับ คำศัพท์ ประโยค และข้อความในระดับย่อหน้า (Zheng & Cheng, 2008: 411)

แบบทดสอบ CET (College English Test) ใช้การทดสอบแบบเติมคำเรื่อง (Cloze) และการรับรู้ถึงข้อผิดพลาดในการใช้ภาษา (Error Recognition) ที่ให้ผู้เข้าสอบเลือกจากตัวเลือก (Multiple-Choice) (Zheng & Cheng, 2008: 411)

การตรวจให้คะแนน (Scoring Technique) เป็นแบบตอบถูกได้ 1 หรือ ตอบผิดได้ 0

แบบทดสอบ MUET (Malaysian University English Test)

แบบทดสอบ MUET (Malaysian University English Test) (Malaysia Examination Council, 2006) มีเนื้อหาสาระแบบทดสอบในการทดสอบทักษะการฟัง การพูด การอ่าน การเขียน และการใช้ภาษา ดังรายละเอียดต่อไปนี้

การทดสอบทักษะการฟัง

จากการศึกษาและวิเคราะห์แบบทดสอบ MUET (Malaysian University English Test) (Malaysia Examination Council, 2006) ที่เกี่ยวกับทักษะการฟัง ได้เนื้อหาสาระแบบทดสอบ คือ 1) จุดประสงค์ (Purpose) 2) สาระภาษา (Language Knowledge/Subskills) 3) 58 หัวเรื่อง (Topical Knowledge) 4) รูปแบบการทดสอบ (Test Type) และ 5) การตรวจให้คะแนน (Scoring Technique) โดยมีรายละเอียด ดังนี้

จุดประสงค์ (Purpose) ในการทดสอบทักษะการฟัง คือ เพื่อวัดความสามารถของผู้เข้าสอบ ในการเข้าใจข้อความจากเสียงพูด (Oral Text) ชนิดต่าง ๆ ที่มีความยาวและระดับความยากแตกต่างกันทั้งสาระและภาษา (Malaysian Examinations Council, 2006: 12)

สาระภาษา (Language Knowledge/Subskills) ในการทดสอบทักษะการฟังครอบคลุมความสามารถในการ 1) จดจำข้อมูลได้ 2) รู้จักความคิดสำคัญ 3) รู้จักรายละเอียดสนับสนุน 4) เข้าใจความหมายของคำ วลี และประโยคจากบริบท 5) ถอดความหรือเรียบเรียงภาษาใหม่ 6) ทำนายผลลัพธ์ที่อาจจะเกิดขึ้น 7) ปรับใช้แนวคิดในสถานการณ์ใหม่ 8) เข้าใจหน้าที่ภาษา 9) แยกแยะข้อมูลสำคัญออกจากข้อมูลที่ไม่สำคัญ 10) แยกแยะข้อเท็จจริงออกจากความคิดเห็น 11) ลงความเห็นจากหลักฐานที่ปรากฏ 12) ระบุบทบาทและความสัมพันธ์ของเนื้อความแต่ละตอน 13) ติดตามการดำเนินเรื่องของประเด็นหรือข้อโต้แย้ง 14) สรุปข้อมูล 15) ประเมินข้อมูล 16) ทำการตัดสินใจ 17) สรุปประเด็นสำคัญจากหลักฐานแวดล้อม และ 18) รู้จักและตีความความเห็น ทศนคติ หรือเจตนาของผู้พูด

58 หัวเรื่อง (Topical Knowledge) ของบทฟังได้มาจากการสรุปการบรรยาย (Lecture Briefing) การพูด (talk) การอภิปราย (Discussion) การสัมภาษณ์ (Interview) บทสนทนาทางโทรศัพท์ (Telephone Conversation) การประกาศ (Announcement) คำแนะนำ (Instructions) โฆษณา (Advertisement) ข่าว (News) การประชุม (Meeting) และสารคดี (Documentary) (Malaysian Examinations Council, 2006: 17)

แบบทดสอบ MUET (Malaysian University English Test) อาจจะทดสอบทักษะการฟังด้วยข้อคำถาม 3 แบบ คือ การถ่ายโอนข้อมูล (Information Transfer) การตอบคำถามสั้น (Short-Answer Questions) และแบบเลือกตอบ 3 ตัวเลือกและ 4 ตัวเลือก (3-Option multiple-Choice Questions and 4- Option Multiple-Choice Questions) (Malaysian Examinations Council, 2006: 17)

การตรวจให้คะแนน (Scoring Technique) เป็นแบบตอบถูกได้ 1 หรือ ตอบผิดได้ 0

การทดสอบทักษะการพูด

จากการศึกษาและวิเคราะห์แบบทดสอบ MUET (Malaysian University English Test) (Malaysian Examinations Council, 2006) ที่เกี่ยวกับทักษะการพูดได้เนื้อหาสาระแบบทดสอบ คือ 1) จุดประสงค์ (Purpose) 2) สาระภาษา (Language Knowledge/Subskills) 3) ชนิด 58 ภาระงาน (Task Type) และ 4) การตรวจให้คะแนน (Scoring Technique) โดยมีรายละเอียด ดังนี้

จุดประสงค์ (Purpose) ในการทดสอบทักษะการพูด คือ เพื่อประเมินความสามารถของผู้เข้าสอบในการนำเสนอรายบุคคลและมีส่วนร่วมในการอภิปรายกลุ่มในประเด็นร่วมสมัยที่หลากหลาย (Malaysian Examinations Council, 2006: 17)

สาระภาษา (Language Knowledge/Subskills) สำหรับการทดสอบทักษะการพูด ครอบคลุมความสามารถในการ 1) ใช้ภาษาได้ถูกต้องตามหลักไวยากรณ์ 2) ออกเสียง ลงเสียงหนักเบา และเสียงสูงต่ำที่ถูกต้อง 3) พูดด้วยความมั่นใจและคล่องแคล่ว 4) ใช้ภาษาเหมาะสมกับจุดประสงค์ที่ต้องการและผู้ฟัง 5) ใช้คำศัพท์และสำนวนภาษาที่หลากหลาย 6) ใช้โครงสร้างประโยคที่หลากหลาย 7) ใช้รูปภาษาที่เหมาะสมสำหรับสถานการณ์เฉพาะ 8) พัฒนาและจัดระบบความคิดได้เป็นอย่างดี 9) ใช้คำบอกลำดับและคำเชื่อมที่เหมาะสม 10) ใช้คำแทนคำซ้ำที่เหมาะสมพร้อมกับคำเชื่อมตัวอื่น ๆ 11) นิยาม บรรยาย อธิบาย 12) เปรียบเทียบ และระบุความต่าง 13) แบ่งประเภท 14) ให้เหตุผล 15) แสดงความคิดเห็น 16) แสดงความสัมพันธ์ของเหตุการณ์ 17) ให้ข้อเสนอแนะ และคำแนะนำ 18) แสดงการเห็นด้วยและไม่เห็นด้วย 19) ค้นหาความชัดเจน 20) ถามและให้ข้อมูล 21) เชิญชวน 22) ตั้งข้อสรุป 23) กล่าวและปรับเปลี่ยนความคิดเห็น 24) นำเสนอข้อโต้แย้ง 25) จัดการอภิปราย (การเริ่มต้น การสลับบทบาท การขัดจังหวะ การให้ตัวกระตุ้น การเจรจาต่อรอง การปิด) และ 26) ทำภาระงานได้ครบถ้วน สมบูรณ์ (นำเสนอข้อมูลที่เกี่ยวข้อง ให้เนื้อหาที่เพียงพอ แสดงการแก้ปัญหาได้ตามที่หัวข้อต้องการ)

แบบทดสอบ MUET (Malaysian University English Test) ทดสอบความสามารถในการพูดของผู้เข้าสอบที่เกี่ยวกับประเด็นร่วมสมัย (Contemporary Issues) ที่อาจจะมีหัวข้อเกี่ยวกับสังคมวัฒนธรรม เศรษฐกิจ วิทยาศาสตร์และเทคโนโลยี กีฬา สิ่งแวดล้อม การศึกษา สุขภาพ โดยใช้ภาระงาน 2 ชนิด คือ Task A การนำเสนอรายบุคคล (Individual Presentation) ซึ่งกำหนดเวลาในการเตรียมตัว 2 นาที และนำเสนอ 2 นาที และ Task B การอภิปรายกลุ่ม (Group Interaction) มีผู้เข้าสอบกลุ่มละ 4 คน ซึ่งกำหนดเวลาในการเตรียมตัว 2 นาที และอภิปราย 10 นาที (Malaysian Examinations Council, 2006: 17)

แบบทดสอบ MUET (Malaysian University English Test) กำหนดเกณฑ์การให้คะแนนการทดสอบพูดโดย อาศัยเกณฑ์ 7 ด้าน คือ 1) ความถูกต้อง (Accuracy) การใช้ภาษาถูกต้องตามหลักไวยากรณ์ การออกเสียงที่ถูกต้อง รวมถึงการลงเสียงหนักเบาและสูงต่ำ 2) ความคล่อง (Fluency) การพูดด้วยความมั่นใจและคล่องแคล่ว 3) ความเหมาะสม (Appropriacy) การใช้ภาษาที่เหมาะสมกับจุดประสงค์และผู้ฟัง การใช้คำศัพท์และสำนวนภาษาที่หลากหลาย การใช้โครงสร้างประโยคที่หลากหลาย และการสังเกตเห็นโครงสร้างภาษาที่เหมาะสมกับสถานการณ์จำเพาะ 4) ความสัมพันธ์และความต่อเนื่อง (Coherence and Cohesion) การพัฒนาและจัดระบบความคิด การใช้ตัวเว้นจังหวะ (Markers) และคำเชื่อม (Linking Devices) ที่เหมาะสม และการใช้ anaphora appropriately together with other cohesive devices 5) การใช้หน้าที่ภาษา (Use of Language Functions) การนิยาม การบรรยาย การอธิบาย, การเปรียบเทียบความเหมือน ความต่าง, การแบ่งประเภท, การให้เหตุผล, การแสดงความคิดเห็น, การแสดงความสัมพันธ์, การให้ข้อเสนอแนะและคำแนะนำ, การแสดงความคิดเห็นด้วยและไม่เห็นด้วย, การค้นหาความชัดเจน, การขอและให้ข้อมูล, การเชิญชวน, การสรุป, การกล่าวและปรับเปลี่ยนความคิด, การนำเสนอข้อโต้แย้ง 6)

การดำเนินการอภิปราย (Managing a Discussion) การเริ่มต้น (Initiating) การสลับบทบาท (Turn-taking) การขัดจังหวะ (Interrupting) Prompting การเจรจา (Negotiating) การจบการอภิปราย (Closing) และ 7) ความสมบูรณ์ของภาระงาน (Task Fulfillment) การนำเสนอข้อมูลที่เป็นประโยชน์ (Presenting relevant ideas) การนำเสนอสาระที่เพียงพอ (Providing Adequate Content) การแสดงการแก้ปัญหาในหัวข้อที่สมบูรณ์ (Showing a Mature Treatment of Topic) (Malaysian Examinations Council, 2006: 13-14)

การทดสอบทักษะการอ่าน

จากการศึกษาและวิเคราะห์แบบทดสอบ MUET (Malaysian University English Test) (Malaysian Examinations Council, 2006) ที่เกี่ยวกับทักษะการอ่านได้เนื้อหาสาระแบบทดสอบ คือ 1) จุดประสงค์ (Purpose) 2) สาระภาษา (Language Knowledge/Subskills) 3) 60 หัวเรื่อง (Topical Knowledge) 4) รูปแบบการทดสอบ (Test Type) และ 5) การตรวจให้คะแนน (Scoring Technique) โดยมีรายละเอียด ดังนี้

จุดประสงค์ (Purpose) ในการทดสอบทักษะการอ่าน คือ เพื่อวัดความสามารถของผู้เข้าสอบในการเข้าใจบทความชนิดต่าง ๆ ที่มีความยาวและความซับซ้อนแตกต่างกัน (Malaysian Examinations Council, 2006: 18)

สาระภาษา (Language Knowledge/Subskills) ในการทดสอบทักษะการอ่านครอบคลุมความสามารถในการ 1) อ่านคร่าว ๆ และอ่านหารายละเอียดเฉพาะ 2) ดึงข้อมูลเฉพาะ 3) ระบุความคิดสำคัญ 4) ระบุรายละเอียดสนับสนุน 5) ถอดความหมายจากคำ วลี ประโยคในบริบท 6) เข้าใจข้อความประเภทความเรียงและประเภทรูปภาพ 7) เข้าใจความสัมพันธ์ภายในประโยค ระหว่างประโยค 8) รู้จักประโยคที่ถอดความ 9) ทำนายผลลัพธ์ 10) ปรับใช้แนวคิดในสถานการณ์ใหม่ 11) เข้าใจหน้าที่ภาษา 12) ตีความข้อความที่เรียงเรียงตามลำดับและไม่ตามลำดับ 13) แยกแยะข้อมูลสำคัญออกจากข้อมูลที่ไม่สำคัญ 14) แยกแยะข้อเท็จจริงออกจากความคิดเห็น 15) ลงความเห็น 16) เชื่อมความคิดและแนวคิดภายในย่อหน้าและระหว่างย่อหน้า 17) ติดตามการดำเนินเรื่องของประเด็นหรือข้อโต้แย้ง 18) สรุปข้อมูล 19) ประเมินข้อมูล 20) ทำการตัดสินใจ 21) ดึงข้อสรุป และ 22) รู้จักและตีความความเห็น ทศนคติ หรือเจตนาของผู้พูด

แบบทดสอบ MUET (Malaysian University English Test) ทดสอบทักษะการอ่านโดยใช้บทความที่มาจากบทความในวารสาร หนังสือพิมพ์และนิตยสาร บทความวิชาการ และสาระที่นำมาจากแหล่งข้อมูลอิเล็กทรอนิกส์ และรูปแบบการนำเสนอของบทความ (Rhetorical Style) อาจเป็นแบบ วิเคราะห์ (Analytical) บรรยาย (Descriptive) จูงใจ (Persuasive) โต้แย้ง (Argumentative) และพรรณนา (Narrative) (Malaysian Examinations Council, 2006: 18)

แบบทดสอบ MUET (Malaysian University English Test) มีลักษณะข้อคำถามในส่วนแบบทดสอบที่วัดทักษะการอ่านชนิดเดียว คือ คำถามแบบให้เลือกตอบที่มีตัวเลือก 3 และ 4 ตัวเลือก (Malaysian Examinations Council, 2006: 18)

การตรวจให้คะแนน (Scoring Technique) เป็นแบบตอบถูกได้ 1 หรือ ตอบผิดได้ 0

การทดสอบทักษะการเขียน

จากการศึกษาและวิเคราะห์แบบทดสอบ MUET (Malaysian University English Test) (Malaysian Examinations Council, 2006) ที่เกี่ยวกับทักษะการเขียนได้เนื้อหาสาระแบบทดสอบ คือ 1) จุดประสงค์ (Purpose) 2) สาระภาษา (Language Knowledge/Subskills) 3) ชนิด 60 ภาระงาน (Task Type) และ 4) การตรวจให้คะแนน (Scoring Technique) โดยมีรายละเอียด ดังนี้

จุดประสงค์ (Purpose) ในการทดสอบทักษะการเขียน คือ เพื่อประเมินความสามารถของผู้เข้าสอบในการเขียนข้อความชนิดต่าง ๆ ที่ครอบคลุมรูปแบบการเขียนที่หลากหลาย (Malaysian Examinations Council, 2006: 16)

สาระภาษา (Language Knowledge/Subskills) สำหรับการทดสอบทักษะการเขียน ครอบคลุมความสามารถในการ 1) ใช้การสะกดคำและเครื่องหมายที่ถูกต้อง 2) ใช้ไวยากรณ์ถูกต้อง 3) ใช้โครงสร้างประโยคถูกต้อง 4) ใช้คำศัพท์และสำนวนภาษาที่หลากหลาย 5) ใช้โครงสร้างประโยคที่หลากหลายชัดเจน 6) ใช้ภาษาเหมาะสมกับจุดประสงค์ที่ต้องการและผู้รับสาร 7) สังเกตรูปภาษาที่เหมาะสมสำหรับสถานการณ์เฉพาะหรือชนิดของตัวบท 8) การพัฒนาและจัดระบบความคิด 9) ใช้คำบอกเครื่องหมายและคำเชื่อมที่เหมาะสม 10) ใช้คำแทนคำซ้ำที่เหมาะสมพร้อมกับคำเชื่อมตัวอื่น ๆ 11) นิยาม บรรยาย อธิบาย 12) เปรียบเทียบ และระบุความต่าง 13) แบ่งประเภท 14) ให้เหตุผล 15) แสดงความคิดเห็น 16) แสดงความสัมพันธ์ 17) ให้ข้อเสนอแนะ และคำแนะนำ 18) แสดงการเห็นด้วยและไม่เห็นด้วย 19) ค้นหาความชัดเจน 20) ถามและให้ข้อมูล 21) เชิญชวน 22) ดึงข้อสรุป 23) กล่าวและปรับเปลี่ยนความคิดเห็น 24) นำเสนอข้อโต้แย้ง 25) นำเสนอข้อมูลที่เกี่ยวข้อง 26) ให้เนื้อหาที่เพียงพอ และ 27) แสดงการแก้ปัญหาได้ตามที่หัวข้อต้องการ

แบบทดสอบ MUET (Malaysian University English Test) แบบทดสอบการเขียน ประกอบด้วยแบบทดสอบ 2 ข้อ คือ คำถามที่ 1 ให้ผู้เข้าสอบตีความข้อมูลโดยอาศัยตัวกระตุ้นเฉพาะที่กำหนดให้ (ไม่น้อยกว่า 150 คำ) และคำถามที่ 2 ให้ผู้เข้าสอบเขียนขยายความจากหัวข้อที่กำหนดให้ (ไม่น้อยกว่า 350 คำ) ชนิดของข้อความที่เขียนอาจจะเป็น รายงาน (Report) บทความ (Article) จดหมาย (Letter) และเรียงความ (Essay) และอาจจะเป็นรูปแบบการเขียน 4 รูปแบบ คือ เชิงวิเคราะห์ (Analytical) เชิงบรรยาย (Descriptive) เชิงจูงใจ (Persuasive) และเชิงโต้แย้ง (Argumentative) (Malaysian Examinations Council, 2006: 18)

แบบทดสอบ MUET (Malaysian University English Test) มีเกณฑ์การให้คะแนนการทดสอบเขียน 5 ด้าน คือ 1) ความถูกต้อง (Accuracy) การใช้สะกดคำและเครื่องหมายวรรคตอนต่าง ๆ ที่ถูกต้อง การใช้ไวยากรณ์ที่ถูกต้อง และการใช้โครงสร้างประโยคที่ถูกต้อง 2) ความเหมาะสม (Appropriacy) การใช้คำศัพท์และสำนวนภาษาที่หลากหลาย การใช้ประโยคหลากหลายที่ชัดเจน การใช้ภาษาเหมาะสมกับจุดประสงค์และผู้อ่านเป้าหมายที่กำหนด และการสังเกตโครงสร้างภาษาที่เหมาะสมกับสถานการณ์จำเพาะหรือชนิดของข้อความ 3) ความสัมพันธ์และความต่อเนื่อง (Coherence and Cohesion) การพัฒนาและจัดระบบความคิด การใช้ตัวเว้นจังหวะ (Markers) และคำเชื่อม (Linking Devices) ที่เหมาะสม 4) การใช้หน้าที่ภาษา (Use of Language Functions) การนิยาม การบรรยาย การอธิบาย, การเปรียบเทียบความเหมือน ความต่าง, การแบ่งประเภท, การ

ให้เหตุผล, การแสดงความคิดเห็น, การแสดงความสัมพันธ์, การให้ข้อเสนอแนะและคำแนะนำ, การแสดงความคิดเห็นด้วยและไม่เห็นด้วย, การค้นหาความชัดเจน, การขอและให้ข้อมูล, การเชิญชวน, การสรุป, การกล่าวและปรับเปลี่ยนความคิด, การนำเสนอข้อโต้แย้ง และ5) ความสมบูรณ์ของภาระงาน (Task Fulfillment) การนำเสนอข้อมูลที่เป็นประโยชน์ (Presenting Relevant Ideas) การนำเสนอสาระที่เพียงพอ (Providing Adequate Content) การแสดงการแก้ปัญหาในหัวข้อที่สมบูรณ์ (Showing a mature treatment of topic) (Malaysian Examinations Council, 2006: 16)

แบบทดสอบ CU-TEP (Chulalongkorn University Test of English Proficiency)

แบบทดสอบ CU-TEP (Chulalongkorn University Test of English Proficiency) (สถาบันภาษา จุฬาลงกรณ์มหาวิทยาลัย, 2552) มีเนื้อหาสาระแบบทดสอบในการทดสอบทักษะการฟัง การพูด การอ่าน การเขียน และการใช้ภาษา ดังรายละเอียดต่อไปนี้

การทดสอบทักษะการฟัง

จากการศึกษาและวิเคราะห์แบบทดสอบ CU-TEP (Chulalongkorn University Test of English Proficiency) (สถาบันภาษา จุฬาลงกรณ์มหาวิทยาลัย, 2552) ที่เกี่ยวกับทักษะการฟัง ได้เนื้อหาสาระแบบทดสอบ คือ 1) จุดประสงค์ (Purpose) 2) สาระภาษา (Language Knowledge/Subskills) 3) 62 หัวเรื่อง (Topical Knowledge) 4) รูปแบบการทดสอบ (Test Type) และ 5) การตรวจให้คะแนน (Scoring Technique) โดยมีรายละเอียด ดังนี้

จุดประสงค์ (Purpose) ในการทดสอบทักษะการฟัง คือ เพื่อทดสอบทักษะการฟังโดยให้ผู้เข้าสอบฟังเทปที่ใช้เสียงเจ้าของภาษาและตอบคำถามวัดความเข้าใจ (สถาบันภาษา จุฬาลงกรณ์มหาวิทยาลัย, 2552)

สาระภาษา (Language Knowledge/Subskills) ในการทดสอบทักษะการฟังครอบคลุมความสามารถในการ 1) แสดงนัย Implications 2) แสดงความหมาย (Meanings) 3) เกี่ยวกับการเสนอแนะ (Suggestions) 4) เกี่ยวกับการลงความเห็น (Inferences) 5) เกี่ยวกับการคาดการณ์หรือทำนาย (Predictions) 6) เกี่ยวกับการสมมติฐานหรือสันนิษฐาน (Assumptions) 7) เกี่ยวกับรายละเอียด (Details) 8) เกี่ยวกับใจความหลัก (Main Topic) 9) เกี่ยวกับความสัมพันธ์เกี่ยวข้องของเนื้อความ 10) เกี่ยวกับลำดับที่ถูกต้อง และ 11) เกี่ยวกับลักษณะภาพ

62 แบบทดสอบ CU-TEP (Chulalongkorn University Test of English Proficiency) มีบทฟังที่แบ่งออกได้เป็น 3 ประเภท คือ บทสนทนาสั้น ๆ (Short Conversations) ระหว่างผู้พูด 2 คน บทสนทนาที่ยาวขึ้น (Long Conversations) ระหว่างผู้พูด 2 คน และการคุยหรือพูดปาฐกถา (Talks/Lectures) โดยผู้พูดเพียงคนเดียว (สถาบันภาษา จุฬาลงกรณ์มหาวิทยาลัย, 2552)

แบบทดสอบ CU-TEP (Chulalongkorn University Test of English Proficiency) ทดสอบทักษะการฟังโดยใช้รูปแบบคำถามชนิดเดียว คือ แบบเลือกตอบ (Multiple-Choice) (สถาบันภาษา จุฬาลงกรณ์มหาวิทยาลัย, 2552)

การตรวจให้คะแนน (Scoring Technique) เป็นแบบตอบถูกได้ 1 หรือ ตอบผิดได้ 0

การทดสอบทักษะการพูด

จากการศึกษาและวิเคราะห์แบบทดสอบ CU-TEP (Chulalongkorn University Test of English Proficiency) (สถาบันภาษา จุฬาลงกรณ์มหาวิทยาลัย, 2552) ที่เกี่ยวกับทักษะการพูดได้ เนื้อหาสาระแบบทดสอบ คือ 1) จุดประสงค์ (Purpose) 2) สาระภาษา (Language Knowledge/Subskills) 3) ชนิด 63 ภาระงาน (Task Type) และ 4) การตรวจให้คะแนน (Scoring Technique) โดยมีรายละเอียด ดังนี้

จุดประสงค์ (Purpose) ในการทดสอบทักษะการพูด คือ เพื่อทดสอบความสามารถในการใช้ ภาษาพูดในการสื่อสาร (สถาบันภาษา จุฬาลงกรณ์มหาวิทยาลัย, 2552)

สาระภาษา (Language Knowledge/Subskills) สำหรับการทดสอบทักษะการพูด ครอบคลุมความสามารถในการ 1) พูดบรรยายรูปภาพที่กำหนดให้ 2) แสดงความคิดเห็นในประเด็นที่ กำหนดให้ และ 3) สรุปประเด็นของสาระจากการดูวิดีโอ

แบบทดสอบทักษะการพูดของ CU-TEP (Chulalongkorn University-Test of English Proficiency) ประกอบด้วย 3 ภาระงาน คือ 1) อธิบายรูปภาพต่าง ๆ ที่ปรากฏในหน้าจอ โปรเจคเตอร์ 2) แสดงความคิดเห็นตามข้อความปรากฏ และ 3) ย่อความจากวิดีโอที่ได้ดู ดำเนินการ สอบโดยพูดผ่านเครื่องบันทึกเสียง และจะใช้เวลาในการเตรียมพร้อมก่อนที่จะพูดเป็นเวลา 45 วินาที (สถาบันภาษา จุฬาลงกรณ์มหาวิทยาลัย, 2552)

การทดสอบทักษะการอ่าน

จากการศึกษาและวิเคราะห์แบบทดสอบ CU-TEP (Chulalongkorn University Test of English Proficiency) (สถาบันภาษา จุฬาลงกรณ์มหาวิทยาลัย, 2552) ที่เกี่ยวกับทักษะการอ่านได้ เนื้อหาสาระแบบทดสอบ คือ 1) จุดประสงค์ (Purpose) 2) สาระภาษา (Language Knowledge/Subskills) 3) 63 หัวเรื่อง (Topical Knowledge) 4) รูปแบบการทดสอบ (Test Type) และ 5) การตรวจให้คะแนน (Scoring Technique) โดยมีรายละเอียด ดังนี้

จุดประสงค์ (Purpose) ในการทดสอบทักษะการอ่าน คือ เพื่อทดสอบทักษะการอ่านโดยให้ ผู้เข้าสอบอ่านเรื่องที่กำหนดให้และตอบคำถามเพื่อวัดความเข้าใจ (สถาบันภาษา จุฬาลงกรณ์ มหาวิทยาลัย, 2552)

สาระภาษา (Language Knowledge/Subskills) ในการทดสอบทักษะการอ่านครอบคลุม ความสามารถในการ 1) เข้าใจคำถามเกี่ยวกับคำศัพท์จากเรื่อง 2) เข้าใจคำถามข้อเท็จจริงจากเรื่อง 3) เข้าใจคำถามใจความหลัก ชื่อเรื่อง จุดประสงค์ของเรื่อง 4) เข้าใจคำถามประเภทการลงความเห็น 5) เข้าใจคำถามที่ต้องใช้วิธีการพิเคราะห์ (Scanning) หรือกวาดภาพที่จะให้หาพบว่าสิ่งที่กล่าวถึงนั้นอยู่ ตรงส่วนใดของเรื่อง และ 6) เข้าใจคำถามอ้างอิงที่ใช้ถามเกี่ยวกับค่านามหรือสรรพนาม หรือข้อความ อื่น ๆ หมายถึงใครหรืออะไร

แบบทดสอบ CU-TEP (Chulalongkorn University–Test of English Proficiency) ทดสอบความสามารถในการอ่านของผู้เข้าสอบด้วยการอ่านที่มีลักษณะแตกต่างกัน 3 แบบ คือ การอ่านเรื่องแล้วเติมคำ (Cloze Reading) การอ่านบทอ่านสั้น (Reading Short Passages) และการอ่านบทอ่านยาว (Reading Long Passages) สำหรับประเภทของเรื่องที่น่าสนใจในการสอบมีความหลากหลายจาก 5 กลุ่ม คือ วิทยาศาสตร์และเทคโนโลยี (Science and Technology) 40 % ประวัติศาสตร์และวัฒนธรรม (History & Culture) 20 % สังคมวิทยา (Social Sciences) 10% ศิลปะ (Arts) 15% และชีวประวัติ (Biography) 15% (สถาบันภาษา จุฬาลงกรณ์มหาวิทยาลัย, 2552)

แบบทดสอบ CU-TEP (Chulalongkorn University – Test of English Proficiency) มีลักษณะข้อคำถามในแบบทดสอบส่วนการอ่านเพียงชนิดเดียว คือ ข้อคำถามแบบเลือกตอบ 4 ตัวเลือก (สถาบันภาษา จุฬาลงกรณ์มหาวิทยาลัย, 2552)

การตรวจให้คะแนน (Scoring Technique) เป็นแบบตอบถูกได้ 1 หรือ ตอบผิดได้ 0

การทดสอบทักษะการเขียน

จากการศึกษาและวิเคราะห์แบบทดสอบ CU-TEP (Chulalongkorn University Test of English Proficiency) (สถาบันภาษา จุฬาลงกรณ์มหาวิทยาลัย, 2552) ที่เกี่ยวกับทักษะการเขียนได้เนื้อหาสาระแบบทดสอบ คือ 1) จุดประสงค์ (Purpose) 2) สาระภาษา (Language Knowledge/Subskills) 3) ชนิด 64 ภาระงาน (Task Type) และ 4) การตรวจให้คะแนน (Scoring Technique) โดยมีรายละเอียด ดังนี้

จุดประสงค์ (Purpose) ในการทดสอบทักษะการเขียน คือ เพื่อวัดความสามารถของผู้เข้าสอบทางด้านการใช้ภาษาอังกฤษทั้งในส่วนของไวยากรณ์และคำศัพท์ (สถาบันภาษา จุฬาลงกรณ์มหาวิทยาลัย, 2552)

สาระภาษา (Language Knowledge/Subskills) สำหรับการทดสอบทักษะการเขียนครอบคลุมความสามารถในการระบุจุดบกพร่องของประโยคที่กำหนด

แบบทดสอบ CU-TEP (Chulalongkorn University Test of English Proficiency) มีสาระที่เกี่ยวข้องคือ การใช้ภาษาอังกฤษทั้งในส่วนของไวยากรณ์และคำศัพท์ (สถาบันภาษา จุฬาลงกรณ์มหาวิทยาลัย, 2552)

แบบทดสอบทดสอบทักษะการเขียนของ CU-TEP (Chulalongkorn University-Test of English Proficiency) เป็นทักษะการเขียนที่เน้นไวยากรณ์ เน้นความรู้ทางด้านการใช้ภาษาอังกฤษทั้งในส่วนของไวยากรณ์และคำศัพท์ มี 2 รูปแบบ คือ แบบการระบุข้อผิดพลาดในประโยค (Error Identification) และแบบเติมประโยคให้สมบูรณ์ (Sentence Completion) รูปแบบของข้อคำถามเป็นแบบเลือกคำตอบ (Multiple-Choice) (สถาบันภาษา จุฬาลงกรณ์มหาวิทยาลัย, 2552)

เนื่องจากแบบทดสอบวัดทักษะการเขียนของ CU-TEP (Chulalongkorn University-Test of English Proficiency) เป็นรูปแบบของข้อคำถามแบบเลือกคำตอบ (Multiple-Choice) การตรวจให้คะแนนจึงเป็นแบบตอบถูกได้ 1 และตอบผิด ได้ 0

2.1.2.5 สรุปเนื้อหาสาระสำหรับการสร้างแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test)

สาระสำหรับการสร้างแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Exit Examination) ได้มาจากการสังเคราะห์ข้อมูล 4 แหล่ง คือ 1) หลักสูตรภาษาอังกฤษระดับอุดมศึกษา (มหาวิทยาลัยทักษิณ, 2544) 2) กรอบอ้างอิงความสามารถทางภาษาของสภายุโรป (CEFR) (Council of Europe, 2003) 3) แบบทดสอบวัดความสามารถภาษาอังกฤษมาตรฐานสากล ซึ่งประกอบด้วย แบบทดสอบ TOEFL iBT (Test of English as a Foreign Language internet-Based Test) (Educational Testing Service, 2010) แบบทดสอบ TOEFL ITP (Test of English as a Foreign Language – Institutional Testing Program) (Educational Testing Service, 2014) แบบทดสอบ IELTS (International English Language Testing System) (IELTS, 2015) แบบทดสอบ FCE (First Certificate of English) (UCLES, 2014) และแบบทดสอบ telc B1-B2 English (telc GmbH, 2012) และ 4) แบบทดสอบวัดความสามารถภาษาอังกฤษระดับอุดมศึกษา ซึ่งประกอบด้วย แบบทดสอบ CET (College English Test) (Zheng Ying & Cheng Liying, 2008) แบบทดสอบ MUET (Malaysian University English Test) (Malaysia Examination Council, 2006) และแบบทดสอบ CU-TEP (Chulalongkorn University-Test of English Proficiency) (สถาบันภาษา จุฬาลงกรณ์มหาวิทยาลัย, 2552) โดยผู้วิจัยได้สังเคราะห์ข้อมูลเพื่อให้ได้สาระเกี่ยวกับ 1) จุดประสงค์ (Purpose) 2) สาระภาษา (Language Knowledge/Subskills) 3) 65 หัวเรื่อง (Topical Knowledge) 4) ชนิด 65 ภาระงาน (Task Type) 5) รูปแบบการทดสอบ (Test Type) และ 6) การตรวจให้คะแนน (Scoring Technique) ของการทดสอบแต่ละทักษะ ดังแสดงในตารางสังเคราะห์ตามลำดับ คือ การทดสอบทักษะการฟัง การทดสอบทักษะการพูด การทดสอบทักษะการอ่าน การทดสอบทักษะการเขียน และการทดสอบทักษะการใช้ภาษา

สาระสำหรับการทดสอบทักษะการฟัง

สาระสำหรับการทดสอบทักษะการฟังได้มาจากการสังเคราะห์ข้อมูลดังกล่าวข้างต้น เพื่อให้ได้ 1) จุดประสงค์ (Purpose) 2) สาระภาษา (Language Knowledge/Subskills) 3) 65 หัวเรื่อง (Topical Knowledge) 4) รูปแบบการทดสอบ (Test Type) และ 5) การตรวจให้คะแนน (Scoring Technique) สำหรับการทดสอบทักษะการฟัง ดังแสดงในตารางสังเคราะห์ต่อไปนี้

ตารางที่ 1 การสังเคราะห์สาระแบบทดสอบวัดทักษะการฟัง

แหล่งข้อมูล (Sources)	หลักฐานการฟังของ ผู้เข้าสอบ	CEFR (Council of Europe, 2003)
<p>สาระแบบทดสอบ (Test Content)</p> <p>จุดประสงค์ (Purpose)</p> <p>สาระภาษา (Language Knowledge/Subskills)</p>	<p>วัดความสามารถในการฟังของ</p> <p>ผู้เข้าสอบ</p> <p>เข้าใจสาระของบทฟังและสามารถจับประเด็นสำคัญและรายละเอียดที่จำเป็นได้</p>	<p>วัดความสามารถของผู้เข้าสอบในการฟังเพื่อจับใจความสำคัญและเข้าใจรายละเอียดของเสียงพูดที่ได้บันทึกไว้</p> <p>1) ติดตามประเด็นสำคัญของบทพูดแบบต่อเนื่องรอบ ๆ ตัว โดยบทพูดใช้สำเนียงมาตรฐานและออกเสียงชัดเจน-B1;</p> <p>2) จับใจความสำคัญการสนทนาของภาพแอนิเมชันระหว่างผู้พูดเจ้าของภาษา;</p> <p>3) จับใจความสำคัญของบทพูดรอบ ๆ ตัวได้เป็นส่วนใหญ่;</p> <p>4) ติดตามการบรรยาย หรือการพูดในสาขาของตนเอง;</p> <p>5) ติดตามการพูดสั้น ๆ ที่มีโครงสร้างชัดเจนในหัวข้อที่คุ้นเคย และนำเสนอด้วยสำเนียงพูดมาตรฐานและออกเสียงชัดเจน ;</p> <p>6) ติดตามประเด็นสำคัญ ๆ ของการบรรยาย การพูด และการรายงาน และรูปแบบอื่น ๆ ของการนำเสนอทางวิชาการ และวิชาชีพ ซึ่งซับซ้อนทางความหมายและทางภาษา ;</p> <p>7) เข้าใจข้อมูลด้านเทคนิคง่าย ๆ เช่น คำแนะนำการใช้ชีวิตประจำวัน เป็นต้น ;</p> <p>8) ติดตามการบอกทิศทางที่มีรายละเอียดชัดเจน;</p> <p>9) เข้าใจการประกาศและข่าวสารในหัวข้อที่เกี่ยวข้องกับรูปธรรมและนามธรรมที่พูดด้วยสำเนียงมาตรฐานและความเร็วปกติ ;</p> <p>10) เข้าใจสาระสำคัญโดยทั่วไปของสื่อระบบเสียงที่บันทึกไว้และถ่ายทอดสด ;</p> <p>11) เข้าใจประเด็นสำคัญของบทสรุปข่าวทางวิทยุและสื่อบันทึก ง่าย ๆ ;</p> <p>12) เข้าใจสิ่งที่ได้บันทึกไว้ด้วยสำเนียงมาตรฐานมักจะพบเจอได้ในการดำเนินชีวิตในสังคม การงานอาชีพ หรือแวดวงวิชาการ และแสดงความคิดเห็น บอกทัศนคติ และสาระสำคัญของสารได้ ;</p> <p>13) เข้าใจสารคดีทางวิทยุและสื่อระบบเสียงที่บันทึกไว้และถ่ายทอดสดอื่น ๆ ที่ส่งผ่านโดยสำเนียงมาตรฐานได้เป็นส่วนใหญ่ และสามารถถอดถอดออกได้</p>

ตารางที่ 1 การสังเคราะห์สาระแบบทดสอบวัดทักษะการฟัง (ต่อ)

แหล่งข้อมูล (Sources)	<p>ทลีสูตร (มหาวิทยาลัยทักษิณ, 2544)</p>	<p>CEFR (Council of Europe, 2003)</p>
สาระแบบทดสอบ (Test Content)	<p>หัวข้อเรื่อง (Topical Knowledge) การใช้ภาษาอังกฤษเพื่อการสื่อสาร)</p>	<p>หัวข้อเรื่อง (Topical Knowledge) ความเข้าใจบทสนทนาระหว่างเจ้าของภาษา 1) การพูดคุยแบบต่อเนื่องรอบ ๆ ตัว ซึ่งเป็นบทพูดที่มีการออกเสียงชัดเจนและใช้สำเนียงมาตรฐาน; 2) การสนทนาของภาษาแอนิเมชันระหว่างผู้พูดเจ้าของภาษา การฟังในฐานะสมาชิกคนหนึ่งของกลุ่มผู้ฟัง 3) การบรรยายที่มีเนื้อหาสาระเป็นที่คุ้นเคยและนำเสนอโดยใช้ภาษาที่ไม่ซับซ้อนและจัดวางโครงสร้างอย่างชัดเจน; 4) การพูดสั้น ๆ ที่มีการวางโครงสร้างชัดเจนในหัวข้อที่คุ้นเคย; 5) การบรรยาย การพูด และการรายงาน และรูปแบบอื่น ๆ ของการนำเสนอทางวิชาการและวิชาชีพซึ่งมีความซับซ้อนทั้งด้านความหมายและภาษา การฟังประกาศและการให้คำแนะนำ 6) คำแนะนำการใช้อุปกรณ์ในชีวิตประจำวัน; 7) การบอกทิศทาง; 8) การประกาศและข่าวสารในหัวข้อที่เกี่ยวข้องกับบูรณธรรมและนามธรรมที่พูดด้วยสำเนียงมาตรฐานและความเร็วปกติ การฟังสื่อทางเสียงและสิ่งที่ได้บันทึกไว้ 9) หัวข้อที่เป็นความสนใจส่วนตัวและนำเสนอด้วยเสียงพูดมาตรฐาน ออกเสียงชัดเจน -B1; 10) บทสรุปข่าวทางวิทยุและสิ่งอื่นที่ง่าย ๆ ในหัวข้อที่คุ้นเคยและนำเสนออย่างค่อนข้างชัดและชัดเจน -B1; 11) สิ่งที่ได้บันทึกไว้ด้วยสำเนียงมาตรฐานและมีลักษณะเฉพาะที่ได้ในการดำเนินชีวิตในสังคม การงานอาชีพ หรือแวดวงวิชาการ ; 12) สารคดีทางวิทยุ; 13) สื่อระบบเสียงที่บันทึกไว้และถ่ายทอดสดอื่น ๆ ที่นำเสนอด้วยสำเนียงมาตรฐาน</p>

ตารางที่ 1 การสังเคราะห์สาระแบบทดสอบวัดทักษะการฟัง (ต่อ)

แหล่งข้อมูล (Sources) เนื้อหาแบบทดสอบ (Test Content) จุดประสงค์ (Purpose)	CET (Zheng & Cheng, 2008)	MUET (Malaysian Examinations Council, 2006)	CU-TEP (สถาบันภาษา จุฬาลงกรณ์มหาวิทยาลัย, 2552)	TOEFL iBT (Educational Testing Service, 2010)	TOEFL ITP (Educational Testing Service, 2014)
	วัดความสามารถของนักศึกษาในการเข้าใจการสนทนาด้วยการพูด (Oral Conversation) หรือบทความ (Passages) ตามมาตรฐานของการใช้ภาษาอังกฤษแบบ American English หรือ British English	วัดความสามารถของผู้เข้าสอบในการเข้าใจข้อความจากเสียงพูด (Oral Text) ชนิดต่าง ๆ ที่มีควมยาวและระดับความยากแตกต่างกันทั้งสาระและภาษา	ทดสอบทักษะการฟังโดยใช้ผู้เข้าสอบฟังที่ใช้เสียงเจ้าของภาษาและตอบคำถามวัดความเข้าใจ	วัดความสามารถของผู้เข้าสอบในการเข้าใจภาษาพูดในบริบทวิชาการที่ผู้ใช้ภาษาต้องสามารถฟังการบรรยายและบทสนทนาได้	วัดความสามารถของผู้เข้าสอบในการฟังสาระที่เกี่ยวข้องกับชีวิตประจำวันและเป็นที่คุ้นเคย
สาระภาษา (Language Knowledge/Subskills)	1) เข้าใจภาษาพูดในบทสนทนา 2) เข้าใจบทฟังที่มีสำเนียงพูดมาตรฐาน	1) จัดจำข้อมูได้ 2) รู้จักความคิดสำคัญ 3) รู้จักรายละเอียดสนับสนุน 4) เข้าใจความหมายของคำวลี และประโยคจากบริบทภาษาใหม่ 5) ถอดความหรือเรียบเรียงภาษาใหม่ 6) ทำนายผลลัพธ์ที่อาจจะเกิดขึ้น 7) ปรับใช้แนวคิดในสถานการณ์ใหม่	ผู้เข้าสอบควรมีความสามารถในการเข้าใจคำถาม 1) แสดงนัย Implications 2) แสดงความหมาย (Meanings) 3) เกี่ยวกับการเสนอแนะ (Suggestions) 4) เกี่ยวกับการลงความเห็น (Inferences) 5) เกี่ยวกับการคาดการณ์หรือทำนาย (Predictions)	1) เข้าใจความคิดหลัก ประเด็นหลัก และรายละเอียดที่สำคัญที่สัมพันธ์กับความคิดหลัก 2) ตระหนักถึงทัศนคติของผู้พูดและระดับความประสงคของผู้พูด 3) ตระหนักถึงหน้าที่หรือความประสงค์ของผู้พูด 4) ตระหนักถึงการจัดลำดับของเนื้อความที่ 5) เข้าใจความสัมพันธ์ระหว่างความคิดที่นำเสนอ (ตัวอย่างเช่น	Identify 1) the meaning of what the second speaker said, 2) the synonyms of the difficult words, 3) words having similar sounds, 4) the passive forms, 5) Who and What with multiple nouns, 6) negative expressions,

ตารางที่ 1 การสังเคราะห์สาระแบบทดสอบวัดทักษะการฟัง (ต่อ)

แหล่งข้อมูล (Sources)	CET (Zheng & Cheng, 2008)	MUET (Malaysian Examinations Council, 2006)	CU-TEP (สถาบันภาษา จุฬาลงกรณ์มหาวิทยาลัย, 2552)	TOEFL iBT (Educational Testing Service, 2010)	TOEFL ITP (Educational Testing Service, 2014)
เนื้อหาแบบทดสอบ (Test Content)		<p>8) เข้าใจหน้าที่ที่ภาษา</p> <p>9) แยกแยะข้อมูลสำคัญออกจากข้อมูลที่ไม่สำคัญ</p> <p>10) แยกแยะข้อเท็จจริงออกจากความคิดเห็น</p> <p>11) ลงความเห็นจากหลักฐานที่ปรากฏ</p> <p>12) ระบุบทบาทและความสัมพันธ์ของข้อความแต่ละตอน</p> <p>13) ติดตามการดำเนินเรื่องของประเด็นหรือข้อโต้แย้ง</p> <p>14) สรุปข้อมูล</p> <p>15) ประเมินข้อมูล</p> <p>16) ทำการตัดสิน</p> <p>17) สรุปประเด็นสำคัญจากหลักฐานแวดล้อม</p> <p>18) รู้จำและตีความความเห็นที่คนคิด หรือเจตนาของผู้พูด</p>	<p>6) เกี่ยวกับสมมุติฐานหรือสันนิษฐาน (Assumptions)</p> <p>7) เกี่ยวกับรายละเอียด (Details)</p> <p>8) เกี่ยวกับใจความหลัก (Main Topic)</p> <p>9) เกี่ยวกับความสัมพันธ์ที่เกี่ยวข้องของเนื้อหา</p> <p>10) เกี่ยวกับลำดับที่ถูกต้อง</p> <p>11) เกี่ยวกับลักษณะภาพ</p>	<p>ความเหมือน/ความต่าง เหตุ/ผล หรือ ขั้นตอนในกระบวนการ</p> <p>6) ทำนายหรือข้อสรุปโดยอาศัยสิ่งที่ถูกกล่าวเป็นนัย ๆ ในเรื่องทั้ง</p> <p>7) ทำการเชื่อมโยงระหว่างส่วนต่างๆ ของข้อมูลในบทความหรือการบรรยาย</p> <p>8) ตระหนักถึงการเปลี่ยนหัวข้อเรื่องในการบรรยายและบทสนทนา และตระหนักถึงเป้าหมายและบทสรุปของการบรรยาย</p>	<p>7) double negative expressions,</p> <p>8) “almost negative” expressions,</p> <p>9) negative with comparatives, expressions of agreement,</p> <p>10) expressions of uncertainty and suggestion,</p> <p>11) emphatic expressions of surprise,</p> <p>12) untrue conditions,</p> <p>13) two-and three- part words,</p> <p>14) idioms, and</p> <p>15) draw conclusions about who, what, where,</p>

ตารางที่ 1 การสังเคราะห์สาระแบบทดสอบวัดทักษะการฟัง (ต่อ)

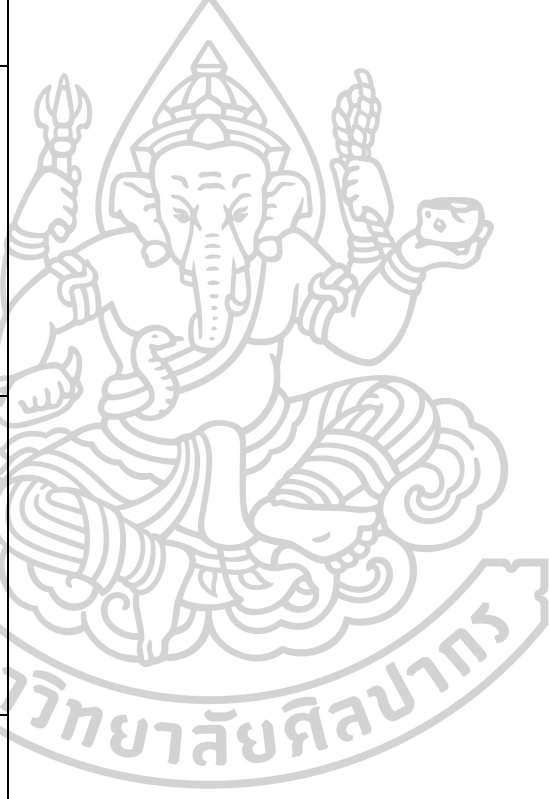
แหล่งข้อมูล (Sources)	CET (Zheng & Cheng, 2008)	MUET (Malaysian Examinations Council, 2006)	CU-TEP (สถาบันภาษา จุฬาลงกรณ์มหาวิทยาลัย, 2552)	TOEFL iBT (Educational Testing Service, 2010)	TOEFL ITP (Educational Testing Service, 2014)
เนื้อหาแบบทดสอบ (Test Content)					16) to anticipate the topics, 17) to anticipate the questions, 18) to determine the topic, 19) to draw conclusions about who, what, when, where.
บทฟัง (Listening Texts)	หัวข้อเรื่อง (Topical Knowledge) 1) บทสนทนาสั้น 2) บทสนทนายาว 3) บทความ	หัวข้อเรื่อง (Topical Knowledge) 1) การสรุปการบรรยาย (Lecture-Briefing) 2) การพูด (talk) 3) การอภิปราย (Discussion) 4) การสัมภาษณ์ (Interview)	หัวข้อเรื่อง (Topical Knowledge) 1) บทสนทนาสั้น ๆ (Short Conversations) ระหว่างผู้พูด 2 คน 2) บทสนทนาที่ยาวขึ้น (Long Conversations) ระหว่างผู้พูด 2 คน และ	หัวข้อเรื่อง (Topical Knowledge) 1) การบรรยายของอาจารย์ในห้องเรียน 2) บทสนทนาในห้องเรียน ระหว่างอาจารย์กับนักศึกษา 3) การประชุมในสำนักงานกับอาจารย์ หรือผู้ช่วยสอน อาจเป็นเรื่องด้าน	เรื่องราวในชีวิตประจำวัน และเป็นที่คุ้นเคย วิทยาศาสตร์ และ ประวัติศาสตร์ แหล่งที่มา ของข้อมูล เป็น บทสนทนา สั้น ๆ บทสนทนา และการบรรยายหรืออธิบาย

ตารางที่ 1 การสังเคราะห์สาระแบบทดสอบวัดทักษะการฟัง (ต่อ)

แหล่งข้อมูล (Sources)	CET (Zheng & Cheng, 2008)	MUET (Malaysian Examinations Council, 2006)	CU-TEP (สถาบันภาษา จุฬาลงกรณ์มหาวิทยาลัย, 2552)	TOEFL IBT (Educational Testing Service, 2010)	TOEFL ITP (Educational Testing Service, 2014)
เนื้อหาแบบทดสอบ (Test Content)		5) บทสนทนาทางโทรศัพท์ (Telephone Conversation) 6) การประกาศ (Announcement) 7) คำแนะนำ (Instructions) 8) โฆษณา (Advertisement) 9) ข่าว (News) 10) การประชุม (Meeting) และ 11) สารคดี (Documentary)	3) การคุยหรือพูดปาฐกถา (Talks/Lectures) โดยผู้พูดเพียงคนเดียว	วิชาการหรือสัมมนาที่กำหนดในรายวิชา 4) การติดต่อที่เคาน์เตอร์บริการกับเจ้าหน้าที่ของมหาวิทยาลัย เช่น การจ่ายค่าที่พัก การลงทะเบียนเรียน การสอบถามข้อมูลเบื้องต้น	
รูปแบบการทดสอบ (Test Type)	1) แบบเลือกตอบ (Multiple-Choice) และ 2) แบบฟังแล้วเขียน (Compound Dictation)	1) การถ่ายโอนข้อมูล (Information Transfer) 2) การตอบคำถามสั้น (Short-Answer Questions) และ 3) แบบเลือกตอบ 3 ตัวเลือกและ 4) ตัวเลือก (3-Option multiple Choice Questions and 4-Option Multiple-Choice Questions)	1) แบบเลือกตอบ (Multiple-Choice)	1) แบบเลือกตอบแบบ 4 ตัวเลือก (คำตอบถูก 1 และ 2 ตัวเลือก) 2) แบบเรียงลำดับขั้นตอนหรือเหตุการณ์ 3) แบบจับคู่	1) แบบเลือกตอบ (Multiple-Choice)

ตารางที่ 1 การสังเคราะห์สาระแบบทดสอบวัดทักษะการฟัง (ต่อ)

แหล่งข้อมูล (Sources)	CET (Zheng & Cheng, 2008)	MUET (Malaysian Examinations Council, 2006)	CU-TEP (สถาบันภาษา จุฬาลงกรณ์มหาวิทยาลัย, 2552)	TOEFL iBT (Educational Testing Service, 2010)	TOEFL ITP (Educational Testing Service, 2014)
เกณฑ์การตรวจให้คะแนน (Scoring Criteria)	ตอบถูกได้ 1/ตอบผิดได้ 0	ตอบถูกได้ 1/ตอบผิดได้ 0	ตอบถูกได้ 1/ตอบผิดได้ 0	1) ตอบถูกได้ 1/ตอบผิดได้ 0 2) ให้คะแนนบางส่วน (Partial Credit)	ตอบถูกได้ 1/ตอบผิดได้ 0



ตารางที่ 1 การสังเคราะห์สาระแบบทดสอบวัดทักษะการฟัง (ต่อ)

แหล่งข้อมูล (Sources) เนื้อหาแบบทดสอบ (Test Content)	IELTS (IELTS, 2015)	FCE (UCLES, 2014)	telc B1-B2 English (telc GmbH, 2012)	สรุปสาระที่จะนำไปสร้างแบบทดสอบ
จุดประสงค์ (Purpose)	วัดความเข้าใจอย่างละเอียดในประเด็นเฉพาะต่าง ๆ หรือ ความเข้าใจโดยทั่วไป ในประเด็นหลัก ๆ ของเสียงบันทึก	วัดความสามารถในการฟังเพื่อหาความรู้สึกรู้สึกที่คนคิดเห็น จุดประสงค์ หน้าที่ ข้อตกลง ประเด็นสำคัญ รายละเอียด และข้อมูลจำเพาะจากสาระที่ฟัง	วัดความสามารถในการฟังเพื่อจับรายละเอียด และการฟังเพื่อความเข้าใจโดยรวม	วัดความสามารถของผู้เข้าสอบในการฟังเพื่อจับใจความสำคัญและเข้าใจรายละเอียดของเสียงพูดที่จับใจได้
สาระภาษา (Language Knowledge/Subskills)	<ol style="list-style-type: none"> 1) เข้าใจประเด็นเฉพาะและข้อมูลโดยทั่วไปของเสียงบันทึกอย่างถ่องแท้ 2) ฟังเพื่อหาข้อมูลรายละเอียด 3) ติดตามการสนทนาระหว่างคน 2 คน 4) ตระหนักรู้ในความสัมพันธ์ของข้อเท็จจริงในเสียงบันทึก 5) เข้าใจคำบรรยาย เช่น การบรรยายสถานที่ ควรบอกได้ว่าคำบรรยายนี้สัมพันธ์กับรูปภาพอย่างไร 6) เข้าใจคำอธิบายว่าสิ่งต่าง ๆ อยู่ที่ไหน และสามารถเดินตามเส้นทางเพื่อไปยังจุดหมาย (เช่น 	<ol style="list-style-type: none"> 1) บอกความรู้สึกรู้สึกที่คนคิดเห็นของผู้พูด 2) บอกจุดประสงค์และหน้าที่ของภาษา 3) จับประเด็นสำคัญ 4) จับรายละเอียดหรือจุดเฉพาะจากสาระที่ฟัง 	<ol style="list-style-type: none"> 1) จับรายละเอียดและ 2) จับใจความสำคัญของสาระที่ฟัง 	<ol style="list-style-type: none"> 1) เข้าใจความคิดและประเด็นหลัก รายละเอียดและข้อเท็จจริงที่สำคัญ 2) ตระหนักรู้ในความคิดเห็น ความรู้สึกทัศนคติ หรือเจตนาของผู้พูด 3) ตระหนักรู้ในการจัดลำดับเนื้อหาของเสียงบันทึก 4) เข้าใจความสัมพันธ์ของเนื้อหาเสนอ ในด้านความเหมือน-ความต่าง เหตุ-ผล หรือเป็นขั้นตอนในกระบวนการ 5) ทำนายหรือหาข้อสรุปจากข้อมูลหรือหลักฐานที่ถูกต้องกล่าวเป็นนัยจากสาระที่ฟัง 6) เชื่อมโยงความคิด เหตุการณ์ ข้อมูลของสาระที่นำเสนอ 7) ตระหนักรู้ในการเปลี่ยนหัวเรื่อง บทนำ และบทสรุปของเนื้อหาเสนอ

ตารางที่ 1 การสังเคราะห์สาระแบบทดสอบวัดทักษะการฟัง (ต่อ)

แหล่งข้อมูล (Sources) เนื้อหาแบบทดสอบ (Test Content)	IELTS (IELTS, 2015)	FCE (UCLES, 2014)	telc B1-B2 English (telc GmbH, 2012)	สรุปสาระ ที่จะนำไปสร้างแบบทดสอบ
	<p>เดินตรงไป หรือ ผ่านประตูที่อยู่ไกลออกไป)</p> <p>7) ระบุข้อมูลที่สำคัญในเสียงบันทึก อาจจะต้องเข้าใจความสัมพันธ์ระหว่างความคิด ข้อเท็จจริง เหตุการณ์ เช่น เหตุและผล เป็นต้น และ</p> <p>8) ฟังเพื่อหาข้อเท็จจริง เช่น สถานที่ ราคา เวลา ที่ได้ยินในเสียงบันทึก</p>			
<p>บทฟัง (Listening Texts)</p>	<p>หัวข้อเรื่อง (Topical Knowledge)</p> <p>1) สถานการณ์ในชีวิตประจำวันและในสังคม</p> <ul style="list-style-type: none"> - บทสนทนาระหว่างผู้พูด 2 คน เช่น การสนทนาเกี่ยวกับการเตรียมการในการเดินทาง - บทพูดที่มีผู้พูดคนเดียว เช่น การกล่าวสุนทรพจน์เกี่ยวกับสิ่งอำนวยความสะดวกในท้องถิ่น 	<p>หัวข้อเรื่อง (Topical Knowledge)</p> <p>1) บทพูดคนเดียว (Monologues) เช่น</p> <ul style="list-style-type: none"> - ข้อความจากเครื่องรับโทรศัพท์ (Answer Phone Messages) - รายการออกอากาศและสารคดีทางวิทยุ (Radio Broadcasts and Features) - ข่าว (News) 	<p>หัวข้อเรื่อง (Topical Knowledge)</p> <p>1) ผู้พูดคนเดียว (Monologues) ได้แก่</p> <ul style="list-style-type: none"> - ข้อความเสียง (Voice Mail Messages) ที่รับเกี่ยวกับเรื่องส่วนตัวและการทำงาน และ - รายการที่มีคนแสดงความคิดเห็นในประเด็นเฉพาะเรื่อง (Programme with people stating their opinions on a particular topic) 	<p>หัวข้อเรื่อง (Topical Knowledge)</p> <p>1) บทฟังที่มีผู้พูดคนเดียว ได้แก่</p> <ul style="list-style-type: none"> - ข้อความเสียงจากเครื่องรับโทรศัพท์ที่เป็นเรื่องส่วนตัวและเกี่ยวข้องกับ - เสียงประกาศในที่สาธารณะ - บทสรุปข่าวทางวิทยุในหัวข้อที่คุ้นเคย - สารคดีทางวิทยุ - การบรรยายที่มีเนื้อหาเป็นต้นเคย

ตารางที่ 1 การสังเคราะห์สาระแบบทดสอบวัดทักษะการฟัง (ต่อ)

แหล่งข้อมูล (Sources) เนื้อหาแบบทดสอบ (Test Content)	IELTS (IELTS, 2015)	FCE (UCLES, 2014)	telc B1-B2 English (telc GmbH, 2012)	สรุปสาระ ที่จะนำไปสร้างแบบทดสอบ
	<p>2) สถานการณ์ทางการศึกษาและการฝึกอบรม</p> <ul style="list-style-type: none"> - บทสนทนาระหว่างผู้พูดหลัก 2 คน เช่น การพูดคุยของนักศึกษาในมหาวิทยาลัย 2 คน หรือนักศึกษามหาวิทยาลัย 2 คน อาจจะเป็นการแนะนำโดยตัวต่อตัว - บทพูดที่มีผู้พูดคนเดียว เช่น การพูดเกี่ยวกับเนื้อหาในรายวิชาทางวิชาการ 	<p>- ประกาศคำรับสารนิเทศ (Public Announcements)</p> <ul style="list-style-type: none"> - เรื่องเล่าและเกร็ดความรู้ (Stories and Anecdotes) และ - การบรรยายและการพูดสั้น ๆ (Lectures and Talks) 2) การพูดที่มีผู้สัมพัทธ์ (Interacting Speakers) เช่น - บทสนทนา (Conversations) - การสัมภาษณ์ (Interviews) - การอภิปราย (Discussions) - บทความวิทยุ (Radio Plays) 	<p>2) บทสนทนา (Dialogues) ได้แก่</p> <ul style="list-style-type: none"> - บทสนทนาในสถานการณ์ประจำวัน (Conversations in everyday situations) ทั้งในบริบทที่เกี่ยวข้องกับเรื่องส่วนตัวและการทำงาน และ - รายการวิทยุหรือบทสัมภาษณ์จาก podcast ส่วนใหญ่มีมาจากการสื่อสารในชีวิตประจำวันทั่วไป และสถานที่ทำงาน 	<ul style="list-style-type: none"> - คำแนะนำการใช้อุปกรณ์ในชีวิตประจำวัน - การพูดสั้น ๆ ในหัวข้อที่คุ้นเคย 2) บทสนทนา แยกได้เป็น 2 ประเภท คือ <ul style="list-style-type: none"> บทสนทนาสั้น ๆ ได้แก่ <ul style="list-style-type: none"> - การพูดคุยในสถานการณ์ประจำวัน ทั้งในบริบทเรื่องส่วนตัวและเกี่ยวข้องกับงาน บทสนทนาที่ยาวขึ้น ได้แก่ <ul style="list-style-type: none"> - บทสัมภาษณ์ที่เกี่ยวข้องกับการสื่อสารในชีวิตประจำวันและที่เกี่ยวข้องกับงาน - การอภิปรายที่เกี่ยวข้องกับการสื่อสารในชีวิตประจำวันและที่เกี่ยวข้องกับงาน - การประชุมที่เกี่ยวข้องกับการสื่อสารในชีวิตประจำวันและที่เกี่ยวข้องกับงาน

ตารางที่ 1 การสังเคราะห์สาระแบบทดสอบวัดทักษะการฟัง (ต่อ)

แหล่งข้อมูล (Sources)	IELTS (IELTS, 2015)	FCE (UCLES, 2014)	telc B1-B2 English (telc GmbH, 2012)	สรุปสาระที่จะนำไปสร้างแบบทดสอบ
เนื้อหาแบบทดสอบ (Test Content)	IELTS (IELTS, 2015) 1) แบบเลือกตอบ (Multiple Choice) 2) แบบจับคู่ (Matching) 3) การเติมความในแผนผัง/แผน ที่/ไดอะแกรม (Plan /Map/ Diagram Labeling) 4) การเติมบทสรุปจากแบบฟอร์ม/โน้ต/ตาราง/ชาร์ต (Form/ Note/Table/Flow chart/Summary Completion) 5) การเติมประโยคให้สมบูรณ์ (Sentence Completion) 6) คำถามให้ตอบสั้น ๆ (Short-Answer Questions)	FCE (UCLES, 2014) 1) แบบเลือกตอบ (Multiple-Choice) 2) แบบเติมประโยคให้สมบูรณ์ (Sentence Completion) 3) แบบจับคู่หลายตัวเลือก (Multiple-Matching)	telc B1-B2 English (telc GmbH, 2012) 1) แบบเลือกตอบ (Multiple-Choice Items) 2) แบบถูกหรือผิด (True/False Items) 3) แบบจับคู่ (Matching Items)	สรุปสาระที่จะนำไปสร้างแบบทดสอบ 1) แบบเลือกตอบ 2) แบบจับคู่ 3) แบบถ่ายโอนข้อมูล เช่น การเติมความในแผนผัง แผนที่ ไดอะแกรม 4) แบบเติมประโยคให้สมบูรณ์
เกณฑ์การตรวจให้คะแนน (Scoring Criteria)	ตอบถูกได้ 1/ตอบผิดได้ 0	ตอบถูกได้ 1/ตอบผิดได้ 0	ตอบถูกได้ 1/ตอบผิดได้ 0	ตอบถูกได้ 1/ตอบผิดได้ 0

ผลจากการสังเคราะห์สาระแบบทดสอบวัดทักษะการฟังในตารางที่ 1 สรุปได้ว่า แบบทดสอบด้านการฟังควรมีกรอบสาระด้าน 1) จุดประสงค์ (Purpose) 2) สาระภาษา (Language Knowledge/Subskills) 3) ลักษณะของบทฟัง (Listening Texts) 4) เทคนิคการสอบ (Test Types) และ 5) การตรวจให้คะแนน (Scoring Technique) ดังรายละเอียดต่อไปนี้

1) จุดประสงค์ของแบบทดสอบด้านการฟัง คือ เพื่อวัดความสามารถของผู้เข้าสอบในการฟังเพื่อจับใจความสำคัญและเข้าใจรายละเอียดของเสียงพูดที่ได้บันทึกไว้

2) สาระภาษา (Language Knowledge/Subskills) ที่เกี่ยวข้องกับทักษะการฟัง คือ 1) เข้าใจความคิดและประเด็นหลัก รายละเอียดและข้อเท็จจริงที่สำคัญ 2) ตระหนักรู้ในความคิดเห็น ความรู้สึก ทศนคติ หรือเจตนาของผู้พูด 3) ตระหนักรู้ในการจัดลำดับเนื้อความของเสียงบันทึก 4) เข้าใจความสัมพันธ์ของเนื้อความที่น่าเสนอ ในด้านความเหมือน-ความต่าง เหตุ-ผล หรือเป็นขั้นตอน ในกระบวนการ 5) ทำนายหรือหาข้อสรุปจากข้อมูลหรือหลักฐานที่ถูกกล่าวเป็นนัยจากสาระที่ฟัง 6) เชื่อมโยงความคิด เหตุการณ์ ข้อมูลของสาระที่น่าเสนอ และ 7) ตระหนักรู้ในการเปลี่ยนหัวเรื่อง บทนำ และบทสรุปของเนื้อความที่น่าเสนอ

3) ลักษณะของบทฟัง (Listening Texts) ที่สามารถนำมาออกแบบแบบทดสอบวัดความสามารถภาษาอังกฤษควรอยู่ในบริบทการใช้ภาษาอังกฤษเพื่อการสื่อสารในชีวิตประจำวันและในสังคมที่ครอบคลุมเรื่องส่วนตัวและการทำงานโดยสามารถแบ่งออกได้เป็น 2 ชนิด คือ บทพูดโดยคนคนเดียว และบทสนทนา โดยมีรายละเอียด คือ 1) บทพูดโดยคนคนเดียว อาจมาจาก 1) ข้อความเสียงจากเครื่องรับโทรศัพท์ที่เป็นเรื่องส่วนตัวและเกี่ยวกับงาน 2) เสียงประกาศในที่สาธารณะ 3) บทสรุปข่าวทางวิทยุในหัวข้อที่คุ้นเคย 4) สารคดีทางวิทยุ 5) การบรรยายที่มีเนื้อหาเป็นที่คุ้นเคย 6) คำแนะนำการใช้อุปกรณ์ในชีวิตประจำวัน และ 7) การพูดสั้น ๆ ในหัวข้อที่คุ้นเคย และ 2) บทสนทนาประกอบด้วยบทสนทนาสั้น ๆ และบทสนทนาที่ยาวขึ้น บทสนทนาสั้น ๆ อาจจะได้มาจากการพูดคุยในสถานการณ์ประจำวันทั้งในบริบทเรื่องส่วนตัวและเกี่ยวข้องกับงาน ส่วนบทสนทนายาว อาจเป็นบทฟังที่มาจาก 1) บทสัมภาษณ์ 2) การอภิปราย และ 3) การประชุม ที่เกี่ยวข้องกับการสื่อสารในชีวิตประจำวันและที่เกี่ยวข้องกับงาน

4) เทคนิคการสอบหรือรูปแบบข้อคำถาม (Test Types) ที่สามารถนำมาใช้ทดสอบทักษะการฟังของแบบทดสอบวัดความสามารถภาษาอังกฤษควรประกอบด้วย 1) แบบเลือกตอบ 2) แบบจับคู่ 3) แบบถ่ายโอนข้อมูล เช่น การเติมความในแผนผัง แผนที่ ไดอะแกรม) และ 4) แบบเติมประโยคให้สมบูรณ์

5) การตรวจให้คะแนน (Scoring Technique) แบบทดสอบการฟังจะเป็นรูปแบบการตอบถูกได้ 1 หรือ ตอบผิด ได้ 0

สรุปสาระสำหรับการทดสอบทักษะการพูด

สาระสำหรับการทดสอบทักษะการพูดได้มาจากการสังเคราะห์ข้อมูลดังกล่าวข้างต้น เพื่อให้ได้ 1) จุดประสงค์ (Purpose) 2) สาระภาษา (Language Knowledge/Subskills) 3) ชนิดของ 77 ภาระงาน (Task Type) และ 4) การตรวจให้คะแนน (Scoring Technique) สำหรับการทดสอบทักษะการพูด ดังแสดงในตารางสังเคราะห์ต่อไปนี้

ตารางที่ 2 การสังเคราะห์สาระแบบทดสอบวัดทักษะการพูด

แหล่งข้อมูล (Sources) สาระแบบทดสอบ (Test Content)	หลักสูตร (มหาวิทยาลัยทักษิณ, 2544)	CEFR (Council of Europe, 2003)
จุดประสงค์ (Purpose)	วัดความสามารถในการพูดของผู้เข้าสอบ	วัดความสามารถของผู้เข้าสอบในการใช้ภาษาพูดเพื่อการสื่อสาร
สาระภาษา (Language Knowledge/Subskills)	สื่อสารโดยใช้ภาษาพูดได้อย่างถูกต้องและเหมาะสมกับจุดประสงค์และผู้ฟัง	<ol style="list-style-type: none"> 1) เข้าสู่อารมณ์สนทนาในหัวข้อที่คุ้นเคยได้โดยไม่ต้องเตรียมตัว ; 2) ติดตามการพูดที่ออกเสียงซึ่งชัดเจนในการสนทนาประจำวัน แม้ว่าบางครั้งจะมีการขอให้กล่าวซ้ำสำหรับบางคำหรือวลี; 3) ดำเนินการสนทนาหรือพูดคุยให้ต่อเนื่องได้ แม้บางครั้งจะจับคู่สนทนาต้องการสื่อได้ยาก ; 4) แสดงออกและตอบสนององค์ความรู้ที่ถูกต้องของตนเอง เช่น ความประหลาดใจ ความสุข ความสนใจ และความไม่แตกต่างกัน ; 5) ร่วมการสนทนาต่อเนื่องในหัวข้อต่างๆ ได้เป็นส่วนใหญ่ ในลักษณะการมีส่วนร่วมได้อย่างลงตัว แม้จะอยู่ภาวะแวดล้อมที่เสียงดัง ; 6) รักษาความสัมพันธ์กับผู้พูดเจ้าของภาษาโดยไม่ได้ชักขึ้นหรือล่อเลียนพวกเขาอย่างไม่ตั้งใจ หรือขอให้พวกเขาแสดงพฤติกรรมที่นอกเหนือจากที่เขาจะทำได้กับเจ้าของภาษาด้วยกัน ; 7) สื่อระดับอารมณ์และแสดงจุดเด่นของความสำคัญของการสนทนาและประสบการณ์ส่วนตัว ; 8) ติดตามสิ่งที่พูดคุยกันในชีวิตประจำวัน ไปได้เป็นส่วนใหญ่ โดยที่คู่สนทนาใช้ภาษาที่เข้าใจง่ายและออกเสียงโดยชัดเจน ; 9) แสดงความคิดเห็นของตนเกี่ยวกับหัวข้อที่เป็นนามธรรมหรือเกี่ยวข้องกับวัฒนธรรม เช่น ดนตรี ภาพยนตร์ และสามารถอธิบายว่าเหตุใดบางอย่างถึงเป็นปัญหา ; 10) ให้ข้อเสนอแนะสั้นๆ ต่อความคิดเห็นของบุคคลอื่น ; 11) เปรียบเทียบและชี้ให้เห็นความแตกต่างของทางเลือก อภิปรายสิ่งที่จะต้องทำ สถานที่ที่จะไป ใครหรือสิ่งไหนที่ต้องเลือก เป็นต้น; 12) ติดตามประเด็นสำคัญในการพูดคุยแบบไม่เป็นทางการกับเพื่อนได้เป็นส่วนใหญ่ โดยคู่สนทนาออกเสียงที่ชัดเจนด้วยสำเนียงมาตรฐาน ;

ตารางที่ 2 การสังเคราะห์สาระแบบทดสอบวัดทักษะการพูด (ต่อ)

แหล่งข้อมูล (Sources)	หลักสูตร (มหาวิทยาลัยทักษิณ, 2544)	CEFR (Council of Europe, 2003)
สาระแบบทดสอบ (Test Content)		<p>13) ให้หรือถามหาแนวคิดหรือความคิดเห็นส่วนตัวในการอภิปรายหัวข้อที่สนใจ ;</p> <p>14) ทำให้ความคิดเห็นของตนเองเป็นที่เข้าใจเพื่อเป็นข้อเสนอแนะในการแก้ปัญหา ;</p> <p>15) แสดงความเชื่อ ความคิดเห็น การเห็นด้วยหรือไม่เห็นด้วย อย่างสุภาพ ;</p> <p>16) พูดได้ในอัตราเดียวกันกับการพูดอย่างมีระดับระหว่างผู้พูดเจ้าของภาษา ;</p> <p>17) แสดงความคิดเห็นของตนเองที่มีความชัดเจน นำเสนอและตอบสนองต่อข้อโต้แย้งที่ซับซ้อนได้อย่างน่าเชื่อถือ ;</p> <p>18) มีส่วนร่วมอย่างกระฉับกระเฉงในการพูดคุยแบบไม่เป็นทางการในบริบทที่คุ้นเคย ;</p> <p>19) แสดงความคิดเห็น ให้ประเด็นความเห็นได้อย่างชัดเจน ประเมินข้อเสนอกที่เป็นทางเลือก พร้อมทั้งสร้างและตอบสมมติฐาน ;</p> <p>20) จับประเด็นที่พูดด้วยกันได้เป็นส่วนมากโดยต้องใช้ความพยายามอยู่บ้าง แต่อาจจะพบว่ามีส่วนร่วมอย่างมีประสิทธิภาพในการพูดคุยกับผู้พูดเจ้าของภาษาจำนวนมาก ซึ่งไม่รับแต่งภาษาของพวกเขาเลย ;</p> <p>21) ออกและยืนยันความเห็นของตนเองในการพูดคุย โดยการให้คำอธิบายที่เป็นประโยชน์ ข้อโต้แย้ง และข้อเสนอแนะ ; ในด้านการแลกเปลี่ยนข้อมูล ผู้เข้าสอบควรมีความสามารถในการ</p> <p>22) แลกเปลี่ยน ตรวจสอบ ตรวจสอบ และยืนยันข้อมูลที่เป็นความจริงที่เห็นหรือเรื่องประจำวันที่คุ้นเคย และเกี่ยวกับสาขาของตนด้วยความมั่นใจ ;</p> <p>23) บรรยายวิธีการทำอะไรบางอย่าง พร้อมให้คำแนะนำอย่างละเอียด ;</p> <p>24) ค้นหาและถ่ายทอดข้อมูลที่เป็นข้อเท็จจริงอย่างตรงไปตรงมา ;</p> <p>25) สอบถามข้อมูลและไปตามการบอกทิศทางที่ละเอียดได้ ;</p> <p>26) รับข้อมูลที่มีความละเอียดได้มากขึ้น ;</p> <p>27) เข้าใจและแลกเปลี่ยนข้อมูลที่ซับซ้อนและคำแนะนำในเรื่องต่าง ๆ ที่สัมพันธ์กับบทบาททางอาชีพของตน ;</p>

ตารางที่ 2 การสังเคราะห์สาระแบบทดสอบวัดทักษะการพูด (ต่อ)

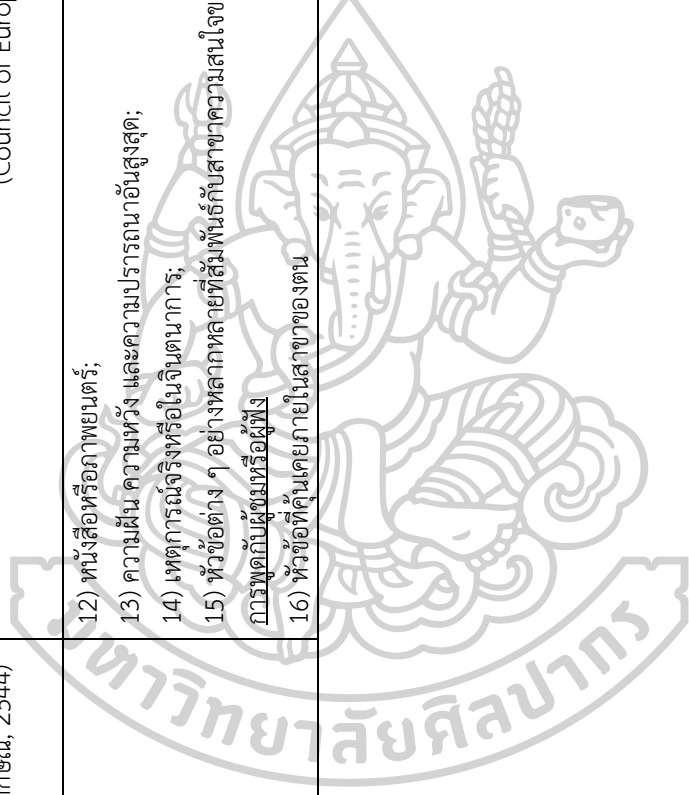
แหล่งข้อมูล (Sources) สารแบบทดสอบ (Test Content)	หลักสูตร (มหาวิทยาลัยทักษิณ, 2544)	CEFR (Council of Europe, 2003)
		<p>28) ส่งผ่านข้อมูลที่มีรายละเอียดได้อย่างน่าเชื่อถือ ;</p> <p>29) ให้คำอธิบายที่ละเอียดชัดเจนของการดำเนินการตามขั้นตอน ;</p> <p>30) สังเคราะห์และรายงานข้อมูลและข้อโต้แย้งจากแหล่งข้อมูลหลายแหล่ง ;</p> <p>31) ให้คำอธิบายหัวข้อต่าง ๆ ที่คุ้นเคยภายในสาขาความสนใจของตนได้อย่างตรงไปตรงมา ;</p> <p>32) เชื่อมโยงความสัมพันธ์ของการเล่าเรื่องอย่างตรงไปตรงมาหรือการอธิบายตามลำดับประเด็นได้อย่างคล่องแคล่วและสมเหตุสมผล สามารถเล่าประสบการณ์ได้อย่างละเอียด พร้อมทั้งอธิบายความรู้สึกและปฏิกิริยาตอบสนอง ;</p> <p>33) เชื่อมโยงรายละเอียดของเหตุการณ์ที่ไม่สามารถทำนายได้ เช่น อุบัติเหตุ เป็นต้น ;</p> <p>34) เชื่อมโยงโครงเรื่องหรือหนังสือหรือภาพยนตร์และบรรยายปฏิกิริยาตอบสนองของตนเอง ;</p> <p>35) บรรยายความฝัน ความหวัง และความปรารถนาอันสูงสุด ;</p> <p>36) สามารถบรรยายเหตุการณ์จริงหรือจินตนาการ ;</p> <p>37) สามารถเล่าเรื่องได้ ;</p> <p>38) ให้คำอธิบายที่ละเอียด ชัดเจน ในหัวข้อต่าง ๆ อย่างหลากหลายที่สัมพันธ์กับสาขาความสนใจของตน ;</p> <p>39) ทำการนำเสนออย่างตรงไปตรงมาโดยที่ผู้ฟังต้องเตรียมมาก่อนในหัวข้อที่คุ้นเคยภายในสาขาของตน มีความชัดเจนพอที่จะติดตามได้โดยไม่ยกลำบากเกินไปตลอดเวลา และประเด็นสำคัญได้รับการอธิบายด้วยการตัดสินใจที่สมเหตุสมผล ;</p> <p>40) จัดการกับคำถามที่สอบถามข้อมูลเพิ่มเติม แต่อาจจะขอให้มีการกล่าวซ้ำถ้าการพูดนั้นเร็วเกินไป ;</p> <p>41) ทำการนำเสนอที่ได้รับการพัฒนาอย่างเป็นระบบ ชัดเจน โดยมีการนำเสนอประเด็นสำคัญและรายละเอียดสนับสนุนที่สำคัญ ;</p> <p>42) ละสายตาจากเอกสารที่เตรียมมาได้โดยทันทีและติดตามประเด็นที่น่าสนใจที่หยิบยกโดยสมาชิกที่เป็นผู้ฟัง มักจะแสดงให้เห็นถึงความคล่องแคล่วและสะดวกที่จะแสดงความเห็นอย่างสังเกตเห็นได้ ;</p>

ตารางที่ 2 การสังเคราะห์สาระแบบทดสอบวัดทักษะการพูด (ต่อ)

แหล่งข้อมูล (Sources)	หลักสูตร มหาวิทยาลัยทักษิณ, 2544)	CEFR (Council of Europe, 2003)
สาระแบบทดสอบ (Test Content)		<p>43) นำเสนอที่ได้รับบริการเตรียมตัวชัดเจน ให้เหตุผลในการสนับสนุนหรือขัดแย้งบางประเด็นความคิดเห็น และแสดงให้เห็นถึงประโยชน์และข้อดีของทางเลือกหลายอย่าง ;</p> <p>44) จัดการกับชุดคำถามที่เกิดขึ้นด้วยระดับความคล่องแคล่วและความทันที่ทันใดซึ่งไม่ได้สร้างความคิดเตรียมใจให้กับตนเองหรือผู้ฟัง</p>
ชนิดของภาระงาน (Task Type)	หัวข้อเรื่อง (Topical Knowledge) การสนทนา การใช้ภาษาอังกฤษเพื่อการสื่อสาร	<p>หัวข้อเรื่อง (Topical Knowledge) การสนทนา</p> <ol style="list-style-type: none"> 1) หัวข้อที่คุ้นเคยและเรื่องราวประจำวัน; 2) การสนทนาต่อเนื่องในหัวข้อทั่วไป 3) หัวข้อที่เป็นนามธรรมหรือเกี่ยวข้องกับวัฒนธรรม เช่น ดนตรี ภาพยนตร์ เป็นต้น; 4) การพูดคุยแบบไม่เป็นทางการ (กับเพื่อน) 5) การอภิปรายในหัวข้อที่สนใจ 6) การแลกเปลี่ยนข้อมูล 7) วิธีการทำอะไรบางอย่าง; 8) การสืบค้นข้อมูลและส่งต่อข้อมูล; 9) การสอบถามทิศทาง; 10) ข้อมูลที่ซับซ้อนและคำแนะนำในเรื่องราวต่าง ๆ ที่สัมพันธ์กับบทบาททางอาชีพของตนเอง <p>การพูดคนเดียว (การบรรยายประสบการณ์)</p> <p>11) หัวข้อที่หลากหลายของเรื่องที่เคยภายในสาขาความสนใจของตน;</p>

ตารางที่ 2 การสังเคราะห์สาระแบบทดสอบวัดทักษะการพูด (ต่อ)

แหล่งข้อมูล (Sources)	หลักสูตร (มหาวิทยาลัยทักษิณ, 2544)	CEFR B1-B2 (Council of Europe, 2003)
สาระแบบทดสอบ (Test Content)		12) หนังสือหรือภาพยนตร์; 13) ความฝัน ความหวัง และความปรารถนาอันสูงสุด; 14) เหตุการณ์จริงหรือในจินตนาการ; 15) หัวข้อต่าง ๆ อย่างหลากหลายที่สัมพันธ์กับสาขาความสนใจของตน; 16) หัวข้อที่คุ้นเคยภายในสาขาของตน



ตารางที่ 2 การสังเคราะห์สาระแบบทดสอบวัดทักษะการพูด (ต่อ)

แหล่งข้อมูล (Sources)	CET (Zheng & Cheng, 2008)	MUET (Malaysian Examinations Council, 2006)	CU-TEP (สถาบันภาษา จุฬาลงกรณ์มหาวิทยาลัย, 2552)	TOEFL iBT (Educational Testing Service, 2010)
เนื้อหาแบบทดสอบ (Test Content)	วัดความสามารถในการสื่อสารด้วยการพูด	ประเมินความสามารถของผู้เข้าสอบในการนำเสนอรายบุคคลและมีส่วนร่วมในการอภิปรายกลุ่มในประเด็นร่วมสมัยที่หลากหลาย	ทดสอบความสามารถในการใช้ภาษาพูดในการสื่อสาร	วัดความสามารถของผู้เข้าสอบในการพูดภาษาอังกฤษอย่างมีประสิทธิภาพในบริบททางวิชาการ ทั้งในระหว่างการเรียนรู้และนอกห้องเรียน
สาระภาษา (Language Knowledge/Subskills)	<ol style="list-style-type: none"> 1) มีปฏิสัมพันธ์ในการอภิปรายกลุ่ม 2) นำเสนอด้วยการพูด 3) ตอบคำถาม 4) ออกเสียง ลงเสียงหนักเบา/สูงต่ำ และใช้ไวยากรณ์และคำศัพท์ได้อย่างถูกต้อง 5) ใช้คำศัพท์และโครงสร้างไวยากรณ์หลากหลาย 6) มีส่วนร่วมในการอภิปรายกลุ่ม 7) สามารถผลิตข้อความที่ขยายเพิ่มและต่อเนื่องกัน 8) จัดการกับสถานการณ์และหัวข้อต่าง ๆ ได้ 9) ใช้ทรัพยากรทางภาษาได้เหมาะสมกับบริบท 	<ol style="list-style-type: none"> 1) ใช้ภาษาได้ถูกต้องตามหลักไวยากรณ์ 2) ออกเสียง ลงเสียงหนักเบา และเสียงสูงต่ำที่ถูกต้อง 3) พูดด้วยความมั่นใจและคล่องแคล่ว 4) ใช้ภาษาเหมาะสมกับจุดประสงค์ที่ต้องการและผู้ฟัง 5) ใช้คำศัพท์และสำนวนภาษาที่หลากหลาย 6) ใช้โครงสร้างประโยคที่หลากหลาย 7) ใช้รูปภาษาที่เหมาะสมสำหรับสถานการณ์เฉพาะ 8) พัฒนาและจัดระบบความคิดได้เป็นอย่างดี 9) ใช้คำบอกเล่าและคำเชื่อมที่เหมาะสม 10) ใช้คำแทนคำซ้ำที่เหมาะสมพร้อมทั้งคำเชื่อมอื่น ๆ 	<ol style="list-style-type: none"> 1) พูดบรรยายรูปภาพที่กำหนดให้ 2) แสดงความคิดเห็นในประเด็นที่กำหนดให้ 3) สรุปประเด็นของสารจากการดูวิดีโอ 	<ol style="list-style-type: none"> 1) ตอบคำถามและมีส่วนร่วมในการอภิปรายทางวิชาการ 2) สรุปสิ่งที่ได้อ่านหรือได้ฟังและแสดงความคิดเห็นในหัวข้อที่กำลังพูดคุยกัน 3) มีส่วนร่วมในการสนทนาตามปกติ 4) แสดงความคิดเห็นและสื่อสารกับบุคคลในสถานที่ต่าง ๆ เช่น ร้านอาหาร ห้องสมุด โรงอาหาร และสำนักงานที่พักร



ตารางที่ 2 การสังเคราะห์สาระแบบทดสอบวัดทักษะการพูด (ต่อ)

แหล่งข้อมูล (Sources) เนื้อหาแบบทดสอบ (Test Content)	CET (Zheng & Cheng, 2008)	MUET (Malaysian Examinations Council, 2006)	CU-TEP (สถาบันภาษา จุฬาลงกรณ์มหาวิทยาลัย, 2552)	TOEFL iBT (Educational Testing Service, 2010)
		11) นิยาม บรรยาย อธิบาย 12) เปรียบเทียบ และระบุความต่าง 13) แบ่งประเภท 14) ให้เหตุผล 15) แสดงความคิดเห็น 16) แสดงความสัมพันธ์ของเหตุการณ์ 17) ให้ข้อเสนอแนะ และคำแนะนำ 18) แสดงการเห็นด้วยและไม่เห็นด้วย 19) ค้นหาความจริง 20) ถามและให้ข้อมูล 21) เชิญชวน 22) ตั้งข้อสรุป 23) กล่าวและปรับเปลี่ยนความคิดเห็น 24) นำเสนอข้อโต้แย้ง 25) จัดการอภิปราย (การเริ่มต้น การ สลับบทบาท การจัดจังหวะ การให้ ตัวกระตุก การเจรจาต่อรอง การปิด)		

ตารางที่ 2 การสังเคราะห์สาระแบบทดสอบวัดทักษะการพูด (ต่อ)

แหล่งข้อมูล (Sources) เนื้อหาแบบทดสอบ (Test Content)	CET (Zheng & Cheng, 2008)	MUET (Malaysian Examinations Council, 2006)	CU-TEP (สถาบันภาษา จุฬาลงกรณ์มหาวิทยาลัย, 2552)	TOEFL iBT (Educational Testing Service, 2010)
ชนิดของภาระงาน (Task Type)	<p>1) การปฏิสัมพันธ์ระหว่างผู้เข้าสอบจำนวน 3-4 คน และผู้ดำเนินการสอบ 2 คน (1 คนเป็นผู้สนทนาหรือผู้ประเมิน และอีก 1 คนเป็นผู้ประเมิน) ใช้เวลาสนทนาประมาณ 5 นาที</p> <p>2) การนำเสนอรายบุคคล ใช้เวลา 1.5 นาที</p> <p>3) การตอบคำถามเพิ่มเติมจากผู้ดำเนินการสอบ เพื่อทดสอบความสามารถในการพูดภาษาอังกฤษของผู้เข้าสอบ ใช้เวลาประมาณ 5 นาที</p>	<p>26) ทำภาระงานได้ครบถ้วน สมบูรณ์ (นำเสนอข้อมูลที่เกี่ยวข้อง ให้เนื้อหาที่เพียงพอ แสดงการแก้ปัญหาได้ตามที่หัวข้อต้องการ)</p> <p>Task A: การนำเสนอรายบุคคล (Individual Presentation) ซึ่งกำหนดเวลาในการเตรียมตัว 2 นาที และนำเสนอ 2 นาที</p> <p>Task B: การอภิปรายกลุ่ม (Group Interaction) มีผู้เข้าสอบกลุ่มละ 4 คน ซึ่งกำหนดเวลาในการเตรียมตัว 2 นาที และอภิปราย 10 นาที</p> <p>หัวข้อเรื่องที่จะนำมาทดสอบเป็นประเด็นร่วมสมัย (Contemporary Issues) ที่อาจจะมีหัวข้อเกี่ยวกับสังคมวัฒนธรรม เศรษฐกิจ วิทยาศาสตร์และเทคโนโลยี กีฬา สิ่งแวดล้อม การศึกษา สุขภาพ</p>	<p>1) อธิบายรูปภาพต่าง ๆ ที่ปรากฏในหน้าจอโปรแกรมเตอร์</p> <p>2) แสดงความคิดเห็นตามข้อความปรากฏ</p> <p>3) ย่อความจากวิดีโอที่ได้ดู ดำเนินการสอบโดยพูดผ่านเครื่องบันทึกเสียง และจะให้เวลาในการเตรียมพร้อมก่อนที่จะพูดเป็นเวลา 45 วินาที</p>	<p>A: ภาระงาน</p> <p>ผู้เข้าสอบจะได้รับภารกิจให้พูดในหัวข้อที่หลากหลายโดยอาศัยประสบการณ์ส่วนบุคคล สถานการณ์ในมหาวิทยาลัยและเนื้อหาเชิงวิชาการ มีคำถาม 6 คำถาม 2 ข้อแรก มีไว้สำหรับ 1) การพูดโดยอิสระ (Independent Speaking Tasks) และ คำถามอีก 4 ข้อ สำหรับ 2) การพูดเชิงบูรณาการ (Integrated Speaking Tasks) ผู้เข้าสอบต้องบูรณาการทักษะภาษาอังกฤษของตนเอง เช่น ฟังแล้วพูด หรือ ฟัง อ่านแล้วพูด เหมือนกับที่ผู้เข้าสอบได้เจอในระหว่างการเรียนหรือนอกห้องเรียน</p>

ตารางที่ 2 การสังเคราะห์สาระแบบทดสอบวัดทักษะการพูด (ต่อ)

แหล่งข้อมูล (Sources) เนื้อหาแบบทดสอบ (Test Content)	CET (Zheng & Cheng, 2008)	MUET (Malaysian Examinations Council, 2006)	CU-TEP (สถาบันภาษา จุฬาลงกรณ์มหาวิทยาลัย, 2552)	TOEFL iBT (Educational Testing Service, 2010)
				<p>B: ระยะเวลาสอบประมาณ 20 นาที</p> <p>C: ตัวกระตุ้น เป็นคำถามจากกรรมการ</p> <p>D: หัวเรื่อง เกี่ยวข้องกับสถานการณ์ในชีวิตจริงที่นักศึกษาอาจจะเจอในระหว่างการเรียนรู้ และนอกห้องเรียน</p> <p>การมีส่วนร่วมในการสนทนาตามปกติ</p> <p>แสดงความคิดเห็นและสื่อสารกับบุคคลในสถานที่ต่าง ๆ เช่น ร้านหนังสือ ห้องสมุด ร้านอาหาร และสำนักงานที่พัก</p> <p>E: ความสามารถที่ต้องแสดง คือ</p> <p>ความสามารถในการตอบคำถาม การมีส่วนร่วมในการอภิปรายทางวิชาการ</p> <p>สรุปสิ่งที่ได้อ่านหรือได้ฟัง และแสดงความความคิดเห็นในหัวข้อที่กำลังพูดคุยกัน</p>
เกณฑ์การตรวจให้คะแนน (Scoring Criteria)	<p>1) ความถูกต้อง (Accuracy) ในการออกเสียง ลงเสียงหนักเบา/สูงต่ำ และการใช้ไวยากรณ์และคำศัพท์</p>	<p>1) ความถูกต้อง (Accuracy) การใช้ภาษาถูกต้องตามหลักไวยากรณ์ การออกเสียงที่ถูกต้อง รวมถึงการลงเสียงหนักเบาและสูงต่ำ</p>		<p>1) ความพึงพรั่ง (Delivery): การพูดชัดเจนและสั่นไหว รวมทั้งออกเสียงได้ดี การเว้นจังหวะ และใช้เสียงสูงต่ำที่เป็นธรรมชาติ</p>

ตารางที่ 2 การสังเคราะห์สาระแบบทดสอบวัดทักษะการพูด (ต่อ)

แหล่งข้อมูล (Sources) เนื้อหาแบบทดสอบ (Test Content)	CET (Zheng & Cheng, 2008)	MUET (Malaysian Examinations Council, 2006)	CU-TEP (สถาบันภาษา จุฬาลงกรณ์มหาวิทยาลัย, 2552)	TOEFL iBT (Educational Testing Service, 2010)
	<p>2) ปริมาณ (Range) ของคำศัพท์และโครงสร้างไวยากรณ์</p> <p>3) ขนาด (Size) (เปอร์เซ็นต์) ของการมีส่วนร่วมในการอภิปรายกลุ่มของผู้เข้าสอบแต่ละคน</p> <p>4) การจัดระบบเนื้อความ (Discourse Management) ที่ดูได้จากความสามารถในการผลิตข้อความที่ขยายเพิ่มเติมต่อเนื่องกัน</p> <p>5) ความยืดหยุ่น (Flexibility) ในการจัดการกับสถานการณ์และหัวข้อต่างๆ</p> <p>6) ความเหมาะสม (Appropriateness) ในการใช้ทรัพยากรทางภาษาตามบริบท</p>	<p>2) ความคล่อง (Fluency) การพูดด้วยความมั่นใจและคล่องแคล่ว</p> <p>3) ความเหมาะสม (Appropriacy) การใช้ภาษาที่เหมาะสมกับจุดประสงค์และผู้ฟัง</p> <p>การใช้คำศัพท์และส่วนภาษาที่หลากหลาย การใช้โครงสร้างประโยคที่หลากหลาย และการสังเกตเห็นโครงสร้างภาษาที่เหมาะสมกับสถานการณ์เฉพาะ</p> <p>4) ความสัมพันธ์และความต่อเนื่อง (Coherence and Cohesion) การพัฒนาและจัดระบบความคิด การใช้ตัวเว้นจังหวะ (Markers) และคำเชื่อม (Linking Devices) ที่เหมาะสม</p> <p>5) การใช้หน้าที่ภาษา (Use of Language Functions) การนิยาม การบรรยาย การอธิบาย การเปรียบเทียบความเหมือน ความต่าง การแบ่งประเภท การให้เหตุผล การแสดงความคิดเห็น การแสดงความสัมพันธ์ การให้ข้อเสนอแนะและคำแนะนำ, การแสดงความเห็นด้วยและ</p>		<p>2) การใช้ภาษา (Language Use): ใช้หลักไวยากรณ์และคำศัพท์เพื่อนำเสนอความคิดได้อย่างมีประสิทธิภาพ</p> <p>3) การนำเสนอสอดคล้องกับหัวข้อเรื่อง (Topic Development): ตอบตรงคำถามและนำเสนอความคิดได้อย่างต่อเนื่อง สัมพันธ์กัน</p>

ตารางที่ 2 การสังเคราะห์สาระแบบทดสอบวัดทักษะการพูด (ต่อ)

แหล่งข้อมูล (Sources) เนื้อหาแบบทดสอบ (Test Content)	CET (Zheng & Cheng, 2008)	MUET (Malaysian Examinations Council, 2006)	CU-TEP (สถาบันภาษา จุฬาลงกรณ์มหาวิทยาลัย, 2552)	TOEFL iBT (Educational Testing Service, 2010)
		<p>ไม่เห็นด้วย, การค้นหาความชัดเจน, การขอและให้ข้อมูล, การเชิญชวน, การสรุป, การกล่าวและปรับเปลี่ยนความคิด, การนำเสนอข้อโต้แย้ง</p> <p>6) การดำเนินการอภิปราย (Managing a Discussion) การเริ่มต้น (Initiating) การสลับบทบาท (Turn-taking) การขัดจังหวะ (Interrupting) Prompting การเจรจา (Negotiating) การจบการอภิปราย (Closing) และ</p> <p>7) ความสมบูรณ์ของภาระงาน (Task Fulfillment) การนำเสนอข้อมูลที่จำเป็น (Presenting relevant ideas) การนำเสนอสาระที่เพียงพอ (Providing Adequate Content) การแสดงการแก้ปัญหาในหัวข้อที่สมบูรณ์ (Showing a Mature Treatment of Topic)</p>		

ตารางที่ 2 การสังเคราะห์สาระแบบทดสอบวัดทักษะการพูด (ต่อ)

แหล่งข้อมูล (Sources) เนื้อหาแบบทดสอบ (Test Content) จุดประสงค์ (Purpose)	IELTS (IELTS, 2015)	FCE (UCLES, 2014)	telc B1-B2 English (telc GmbH, 2012)	สรุปสาระ ที่จะนำไปสร้างแบบทดสอบ
	ทดสอบความสามารถในการพูด สัมภาษณ์แบบเผชิญหน้า	วัดความสามารถในการพูดตอบคำถาม การแสดงความคิดเห็นกับคู่สนทนา และการอภิปรายกลุ่ม	วัดความสามารถของผู้เข้าสอบในการพูด การพูดเกี่ยวกับประสบการณ์ และความคิดเห็น การนำเสนอ การตอบคำถามจากกรรมการ และการพูดคุยกัน	วัดความสามารถของผู้เข้าสอบในการพูดเพื่อการสื่อสาร
สำรภาษา (Language Knowledge/Subskills) 1) พูดเพื่อตอบคำถามสั้น ๆ 2) พูดได้อย่างต่อเนื่องในหัวข้อที่คุ้นเคย และมีปฏิสัมพันธ์กับผู้ทดสอบ 3) แสดงความคิดเห็นและให้ข้อมูลเกี่ยวกับเรื่องราวในชีวิตประจำวันและประสบการณ์หรือสถานการณ์ต่างๆ ไป 4) พูดได้อย่างต่อเนื่อง ๆ ในหัวข้อที่กำหนดให้ โดยใช้ภาษาและจัดระบบความคิดได้อย่างเหมาะสม 5) อธิบายความคิดเห็น วิเคราะห์ อภิปราย และพยากรณ์เกี่ยวกับประเด็นที่พูดถึง	1) ให้ข้อมูลเกี่ยวกับตนเองและแสดงความคิดเห็นเกี่ยวกับหัวข้อต่าง ๆ 2) พูดเกี่ยวกับบางสิ่งบางอย่าง เปรียบเทียบ บรรยาย แสดงความคิดเห็น และคาดการณ์ 3) แลกเปลี่ยนความคิดเห็น แสดงและปรับเปลี่ยนความคิดเห็น ให้คำแนะนำ และ/หรือไม่เห็นด้วย ให้คำแนะนำ พยากรณ์ ประเมิน เพื่อนำไปสู่การตัดสินใจ ตลอดจนการเจรจา เป็นต้น	1) พูดเกี่ยวกับประสบการณ์และความคิดเห็น การนำเสนอ การตอบคำถามจากกรรมการและการพูดคุยกัน 2) บรรยายสถานการณ์ในรูปภาพได้ละเอียด ชัดเจน 3) ตอบคำถามของผู้ประเมินได้อย่างละเอียด 4) ทำการนำเสนอที่มีโครงสร้างชัดเจนเพื่อให้เหตุผลสนับสนุนความคิดเห็นของตนเอง 5) ตอบคำถามของผู้ประเมินได้อย่างละเอียด และ 6) แลกเปลี่ยนความคิดเห็นกับคู่สนทนาและทำให้การสนทนา	1) ใช้ภาษาถูกต้องตามหลักไวยากรณ์และเหมาะสมกับจุดประสงค์และผู้ฟัง 2) ใช้คำศัพท์และโครงสร้างไวยากรณ์ที่หลากหลาย 3) ออกเสียง ลงเสียงหนัก-เบา และสูง-ต่ำ ได้ถูกต้อง 4) มีส่วนร่วมในการสนทนาและอภิปรายอย่างกระตือรือร้น 5) จัดระบบความคิดในการนำเสนอข้อมูลได้เหมาะสม 6) ผลิตรอบความที่ขยายเพิ่มและต่อเนื่องกันในหัวข้อที่กำหนด 7) ตอบคำถาม ให้ข้อมูลเกี่ยวกับตนเอง และแสดงความคิดเห็นในเรื่องต่าง ๆ 8) นำเสนอหัวข้อที่กำหนดให้ได้อย่างมี	

ตารางที่ 2 การสังเคราะห์สาระแบบทดสอบวัดทักษะการพูด (ต่อ)

แหล่งข้อมูล (Sources) เนื้อหาแบบทดสอบ (Test Content)	IELTS (IELTS, 2015)	FCE (UCLES, 2014)	telc B1-B2 English (telc GmbH, 2012)	สรุปสาระ ที่จะนำไปสร้างแบบทดสอบ
			ดำเนินต่อไปได้	โครงสร้างชัดเจนและให้รายละเอียด สนับสนุนความคิดเห็นของตนเอง 9) แสดงและอธิบายความเห็น บรรยาย เปรียบเทียบ วิเคราะห์ อภิปราย และ พยากรณ์ประเด็นที่กำลังมาถึง 10) แลกเปลี่ยนความคิดเห็น แสดงและ ปรับเปลี่ยนความคิด เห็นด้วยและ/ หรือไม่เห็นด้วย ให้คำแนะนำ คัดการณ์ ประเมิน เพื่อนำไปสู่การตัดสินใจและ สรุปผล เป็นต้น
ชนิดของภาระงาน (Task Type)	1) การแนะนำตัวและการสัมภาษณ์ (Introduction and Interview) ภาระงาน: การแนะนำตัวเองและการ สัมภาษณ์ เวลา: 4-5 นาที ตัวกระตุ้น: คำถามจากผู้ประเมิน หัวข้อเรื่อง: เรื่องทั่วไป เช่น บ้าน ครอบครัว งาน การเรียนหรือความ สนใจ	1) การสัมภาษณ์ (Interview) ภาระงาน: การสนทนากับกรรมการ เวลา: 2 นาที ตัวกระตุ้น: คำถามจากกรรมการ หัวข้อเรื่อง: ข้อมูลเกี่ยวกับตนเองและ ประสบการณ์ของผู้เข้าสอบ ความสามารถที่ต้องแสดง: - การให้ข้อมูลเกี่ยวกับตนเอง - การแสดงความคิดเห็นเกี่ยวกับหัวข้อ ต่าง ๆ	Part 1A ภาระงาน: กำหนดให้ผู้เข้าสอบ บรรยายสถานการณ์ในรูปภาพได้ อย่างละเอียดและชัดเจน เวลา: 1.5 นาที ตัวกระตุ้น: รูปภาพ หัวข้อเรื่อง: พบได้ในชีวิตประจำวัน	1) การสัมภาษณ์ ภาระงาน: สอบสัมภาษณ์ แสดงความ คิดเห็น และตอบคำถามจากกรรมการ เวลา: 5 นาที ตัวกระตุ้น: คำถามจากกรรมการ หัวข้อเรื่อง: - ข้อมูลเกี่ยวกับตนเอง ประสบการณ์ และความคิดเห็นเกี่ยวกับหัวเรื่องและ สถานการณ์ที่เกี่ยวข้องกับตัวเอง

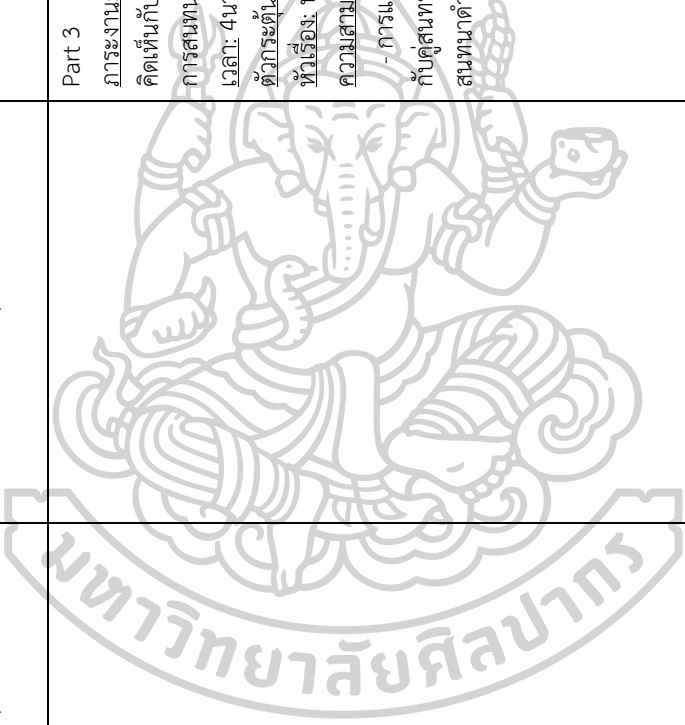
ตารางที่ 2 การสังเคราะห์สาระแบบทดสอบวัดทักษะการพูด (ต่อ)

แหล่งข้อมูล (Sources) เนื้อหาแบบทดสอบ (Test Content)	IELTS (IELTS, 2015)	FCE (UCLES, 2014)	telc B1-B2 English (telc GmbH, 2012)	สรุปสาระที่จะนำไปสร้างแบบทดสอบ
<p>ความสามารถที่ต้องแสดง:</p> <ul style="list-style-type: none"> - การแสดงความคิดเห็น - การให้ข้อมูลเกี่ยวกับเรื่องราวในชีวิตประจำวัน และประสบการณ์หรือสถานการณ์ทั่วไป โดยการตอบคำถามหลายข้อ <p>2) การพูดคนเดียวที่ยาวขึ้น</p> <p>ภาระงาน: การพูดคนเดียวที่ยาวขึ้น</p> <p>เวลา: 3 นาที</p> <p>ตัวกระตุ้น: บัตรภาระงาน</p> <p>หัวข้อ: ประสบการณ์ของตนเอง</p> <p>ความสามารถที่ต้องแสดง:</p> <ul style="list-style-type: none"> - การพูดยาว ๆ ในหัวข้อที่กำหนด - การใช้ภาษาที่เหมาะสม - การจัดระบบความคิดได้อย่างเหมาะสม <p>3) การอภิปรายร่วมกัน</p> <p>ภาระงาน: ผู้เข้าสอบและผู้ประเมินอภิปรายประเด็นที่สัมพันธ์กับหัวข้อในภาระงานที่ 2 โดยภาพกว้างๆ</p>	<p>2) การสนทนากับผู้ฟัง (Long turn)</p> <p>ภาระงาน: โดยการทำข้อมูลจากรูปภาพที่ได้รับ ผู้เข้าสอบพูดอธิบายและแสดงความคิดเห็นต่อรูปภาพ พร้อมทั้งแสดงความคิดเห็นจากเพื่อนด้วย</p> <p>เวลา: 4 นาที</p> <p>ตัวกระตุ้น: รูปภาพ</p> <p>หัวข้อ: สถานการณ์ในชีวิตประจำวัน</p> <p>ความสามารถที่ต้องแสดง:</p> <ul style="list-style-type: none"> - การพูดเกี่ยวกับบางสิ่งบางอย่าง - การเปรียบเทียบ การบรรยาย การแสดงความคิดเห็น และการเจรจา <p>3) ภาระงานที่ต้องร่วมมือกัน (Collaborative Task)</p> <p>ภาระงาน: ผู้เข้าสอบสนทนากับผู้เข้าสอบคนอื่น และช่วยกันตัดสินใจในประเด็นปัญหาที่ได้รับ</p> <p>เวลา: 4 นาที</p> <p>ตัวกระตุ้น: วัสดุและคำอธิบายภาระงาน</p>	<p>ความสามารถที่ต้องแสดง:</p> <ul style="list-style-type: none"> - บรรยายสถานการณ์ในรูปภาพได้อย่างละเอียดและชัดเจน <p>Part 1B</p> <p>ภาระงาน: กำหนดให้ผู้เข้าสอบตอบคำถามของผู้ประเมินได้อย่างถูกต้อง</p> <p>เวลา: 1.5 นาที</p> <p>ตัวกระตุ้น: คำถามของผู้ประเมิน</p> <p>หัวข้อ: พบได้ในชีวิตประจำวัน</p> <p>ความสามารถที่ต้องแสดง:</p> <ul style="list-style-type: none"> - ตอบคำถามของผู้ประเมินได้อย่างถูกต้อง <p>Part 2A</p> <p>ภาระงาน: กำหนดให้ผู้เข้าสอบทำการนำเสนอที่มีโครงสร้างชัดเจนเพื่อให้เหตุผลสนับสนุนความคิดเห็นของตนเอง</p>	<p>ความสามารถที่ต้องแสดง:</p> <ul style="list-style-type: none"> - ตอบคำถามผู้ประเมินอย่างละเอียด - ให้ข้อมูลเกี่ยวกับตนเอง <p>ประสบการณ์ที่เกี่ยวกับหัวข้อเรื่องและสถานการณ์ที่เกี่ยวข้องกับตัวเอง</p> <p>2) การบรรยายรูปภาพ</p> <p>ภาระงาน: พูดบรรยายสถานการณ์เหตุการณ์ในรูปภาพได้อย่างละเอียดครบถ้วน</p> <p>เวลา: 3 นาที</p> <p>ตัวกระตุ้น: รูปภาพ</p> <p>หัวข้อ: เรื่อง:</p> <ul style="list-style-type: none"> - สถานการณ์ เหตุการณ์ เรื่องราวที่เกิดขึ้นในชีวิตประจำวัน <p>ความสามารถที่ต้องแสดง:</p> <ul style="list-style-type: none"> - พูดได้ครอบคลุมประเด็นต่าง ๆ ในรูปภาพ - ใช้ภาษาได้ถูกต้องและเหมาะสม - จัดระบบความคิดในการนำเสนอได้ 	

ตารางที่ 2 การสังเคราะห์สาระแบบทดสอบวัดทักษะการพูด (ต่อ)

แหล่งข้อมูล (Sources) เนื้อหาแบบทดสอบ (Test Content)	IELTS (IELTS, 2015)	FCE (UCLES, 2014)	telc B1-B2 English (telc GmbH, 2012)	สรุปสาระที่จะนำไปสร้างแบบทดสอบ
<p>และในลักษณะที่เป็นนามธรรม และบางที่อาจคลุมเครือ</p> <p>เวลา: 4-5 นาที</p> <p>ตัวกระตุ้น: ประเด็นคำถาม</p> <p>หัวเรื่อง: ประเด็นที่สัมพันธ์กับหัวข้อในภาระงานที่ 2</p> <p>ความสามารถที่ต้องแสดง:</p> <ul style="list-style-type: none"> - การอธิบายความคิดเห็น - วิเคราะห์ อภิปราย และพยากรณ์เกี่ยวกับประเด็นที่พูดถึง 	<p>หัวเรื่อง: สถานการณ์ในชีวิตประจำวัน</p> <p>ความสามารถที่ต้องแสดง:</p> <ul style="list-style-type: none"> - การแลกเปลี่ยนความคิดเห็น และปรับเปลี่ยนความคิดเห็น - การเห็นด้วยและ/หรือไม่เห็นด้วย - การให้คำแนะนำ การเถียง การประเมิน <p>การนำไปสู่การตัดสินใจตลอดการเจรจา เป็นต้น</p> <p>4) การอภิปราย (Discussion)</p> <p>ภาระงาน: การพูดคุยอย่างต่อเนื่องกับสมาชิกอื่น โดยได้รับคำถามจากกรรมการเกี่ยวกับหัวเรื่องหรือประเด็นที่พูดกันในภาระงานที่ 3</p> <p>เวลา: 4 นาที</p> <p>ตัวกระตุ้น: คำถามนำจากกรรมการ</p> <p>หัวเรื่อง: หัวเรื่องหรือประเด็นที่พูดกันในภาระงานที่ 3</p> <p>ความสามารถที่ต้องแสดง:</p> <ul style="list-style-type: none"> - การแสดงและปรับเปลี่ยนความคิดเห็น การเห็นด้วยและ/หรือไม่เห็นด้วย 	<p>เวลา: 2 นาที</p> <p>ตัวกระตุ้น: ภาระงาน</p> <p>หัวเรื่อง: พบได้ในชีวิตประจำวัน</p> <p>ความสามารถที่ต้องแสดง:</p> <ul style="list-style-type: none"> - การนำเสนอที่มีโครงสร้างชัดเจนเพื่อให้เหตุผลสนับสนุนความคิดเห็นของตนเอง <p>Part 2B</p> <p>ภาระงาน: ตอบคำถามของผู้ประเมินได้อย่างถูกต้อง</p> <p>เวลา: 1 นาที</p> <p>ตัวกระตุ้น: คำถามของผู้ประเมิน</p> <p>หัวเรื่อง: พบได้ในชีวิตประจำวัน</p> <p>ความสามารถที่ต้องแสดง:</p> <ul style="list-style-type: none"> - การตอบคำถามของผู้ประเมินได้อย่างถูกต้อง 	<p>อย่างเป็นทางการและมีโครงสร้าง</p> <p>3) การนำเสนอรายบุคคล</p> <p>ภาระงาน: ผู้นำเสนอประเด็นตามสถานการณ์ที่กำหนดในภาระงานและตอบคำถามเพิ่มเติม</p> <p>เวลา: พูด 3 นาที (เตรียมตัว 2 นาที)</p> <p>ตัวกระตุ้น: รูปภาพและบัตรภาระงาน</p> <p>หัวเรื่อง:</p> <ul style="list-style-type: none"> - สถานการณ์ที่พบได้ในชีวิตประจำวัน <p>ความสามารถที่ต้องแสดง:</p> <ul style="list-style-type: none"> - พูดได้ครอบคลุมประเด็นต่าง ๆ ที่กำหนดในภาระงาน - ใช้ภาษาได้ถูกต้องและเหมาะสม - จัดระบบความคิดในการนำเสนอได้อย่างเป็นระเบียบและมีโครงสร้าง - ตอบคำถามกรรมการได้ถูกต้องและครบถ้วน <p>4) การอภิปรายกับคู่สนทนา</p> <p>ภาระงาน: ร่วมสนทนาและอภิปรายประเด็นที่ได้รับมอบหมายหรือปรากฏใน</p>	

ตารางที่ 2 การสังเคราะห์สาระแบบทดสอบวัดทักษะการพูด (ต่อ)

แหล่งข้อมูล (Sources) เนื้อหาแบบทดสอบ (Test Content)	IELTS (IELTS, 2015)	FCE (UCLES, 2014)	telc B1-B2 English (telc GmbH, 2012)	สรุปสาระ ที่จะนำไปสร้างแบบทดสอบ
		<p>Part 3</p> <p>ภาระงาน: แลกเปลี่ยนความคิดเห็นเกี่ยวกับสถานการณ์และทำให้การสนทนาดำเนินต่อไปได้</p> <p>เวลา: 4 นาที</p> <p>ตัวกระตุ้น: บัตรภาระงาน</p> <p>หัวข้อ: พบได้ในชีวิตประจำวัน</p> <p>ความสามารถที่ต้องแสดง:</p> <ul style="list-style-type: none"> - การแลกเปลี่ยนความคิดเห็นกับคู่สนทนาและทำให้การสนทนาดำเนินต่อไปได้ 	<p>ภาระงาน</p> <p>เวลา: 10 นาที</p> <p>ตัวกระตุ้น: บัตรภาระงาน</p> <p>หัวข้อ:</p> <ul style="list-style-type: none"> - ปัญหาหรือสถานการณ์ปัจจุบันในสังคม - ประเด็นที่อยู่ในความสนใจของตนเองและเพื่อนคู่สนทนาโดยการเลือกจากหัวข้อที่กำหนด <p>ความสามารถที่ต้องแสดง:</p> <p>ตอบคำถาม มีส่วนร่วมในการอภิปราย แลกเปลี่ยนความคิดเห็นกับคู่สนทนาและทำให้การสนทนาดำเนินต่อไปได้ โดย</p> <ul style="list-style-type: none"> - อธิบาย - ให้เหตุผล - วิเคราะห์ - พยากรณ์ - แสดงและปรับเปลี่ยนความคิดเห็น - สอบถามความคิดเห็นของผู้อื่น - เปรียบเทียบ 	

ตารางที่ 2 การสังเคราะห์สาระแบบทดสอบวัดทักษะการพูด (ต่อ)

แหล่งข้อมูล (Sources) เนื้อหาแบบทดสอบ (Test Content)	IELTS (IELTS, 2015)	FCE (UCLES, 2014)	telc B1-B2 English (telc GmbH, 2012)	สรุปสาระ ที่จะนำไปสร้างแบบทดสอบ
เกณฑ์การตรวจให้คะแนน (Scoring Criteria)	<ol style="list-style-type: none"> 1) ความคล่องแคล่วและความต่อเนื่อง (Fluency and Coherence) 2) การใช้คำศัพท์ที่เหมาะสม (Lexical Resource) 3) ไวยากรณ์ที่ใช้และความถูกต้อง (Grammatical Range and Accuracy) และ 4) การออกเสียง (Pronunciation) 	<ol style="list-style-type: none"> 1) ไวยากรณ์และคำศัพท์ (Grammar and Vocabulary) 2) การจัดการเนื้อความ (Discourse Management) 3) การออกเสียง (Pronunciation) และ 4) การสื่อสารเชิงปฏิสัมพันธ์ (Interactive Communication) 	<ol style="list-style-type: none"> 1) การจัดการภาระงาน (Task Management) 2) การออกเสียงและการใช้เสียงสูงต่ำ (Pronunciation and Intonation) 3) ความคล่องแคล่ว (Fluency) และ 4) ความถูกต้อง (Accuracy) และ 5) คำศัพท์ (Vocabulary) 	<ul style="list-style-type: none"> - บรรยาย - เห็นด้วยและ/หรือไม่เห็นด้วย - ให้คำแนะนำ - ประเมิน - สรุป และ - ตัดสินใจ <p>เกี่ยวกับปัญหาหรือประเด็นที่กำลังอภิปราย</p> <ol style="list-style-type: none"> 1) ความถูกต้อง (Accuracy) ในการออกเสียง ลงเสียงหนัก-เบา สูง-ต่ำ ใช้ภาษาถูกต้องตามหลักไวยากรณ์ 2) ความคล่องแคล่ว (Fluency) พูดด้วยความมั่นใจและคล่องแคล่ว 3) ไวยากรณ์และคำศัพท์ (Grammatical and Vocabulary Range) ใช้คำศัพท์และไวยากรณ์อย่างหลากหลาย 4) การดำเนินการอภิปราย (Managing a Discussion) การเริ่มต้น การสนทนา บทบาท การชี้แจงหวัะ การให้ตัวกระตุ่น

ตารางที่ 2 การสังเคราะห์สาระแบบทดสอบวัดทักษะการพูด (ต่อ)

แหล่งข้อมูล (Sources) เนื้อหาแบบทดสอบ (Test Content)	IELTS (IELTS, 2015)	FCE (UCLES, 2014)	telc B1-B2 English (telc GmbH, 2012)	สรุปสาระ ที่จะนำไปสร้างแบบทดสอบ
				การเจรจาต่อรอง การจบการอภิปราย และ 5) ความสมบูรณ์ของภาระงาน (Task Fulfillment) การนำเสนอสตอรี่กับหัวเรื่อง การจัดการงานได้ตามที่ระบุ การนำเสนอข้อมูลที่เป็นประโยชน์ การนำเสนอสาระที่เพียงพอ และการแก้ปัญหาได้อย่างสมบูรณ์

ผลจากการสังเคราะห์สาระแบบทดสอบวัดทักษะการพูดในตารางที่ 2 สรุปได้ว่าแบบทดสอบด้านการพูดควรมีกรอบสาระด้าน 1) จุดประสงค์ (Purpose) 2) สาระภาษา (Language Knowledge/Subskills) 3) 96 หัวเรื่อง (Topical Knowledge) 4) รูปแบบการทดสอบ (Test Type) และ 5) การตรวจให้คะแนน (Scoring Technique) ดังรายละเอียดต่อไปนี้

1) จุดประสงค์ของแบบทดสอบ คือ เพื่อวัดความสามารถของผู้เข้าสอบในการสื่อสารด้วยการพูด

2) สาระภาษา (Language Knowledge/Subskills) ที่เกี่ยวข้อง คือ 1) ใช้ภาษาถูกต้องตามหลักไวยากรณ์และเหมาะสมกับจุดประสงค์และผู้ฟัง 2) ใช้คำศัพท์และโครงสร้างไวยากรณ์ที่หลากหลาย 3) ออกเสียง ลงเสียงหนัก-เบา และสูง-ต่ำ ได้ถูกต้อง 4) มีส่วนร่วมในการสนทนาและอภิปรายอย่างกระตือรือร้น 5) จัดระบบความคิดในการนำเสนอข้อมูลได้เหมาะสม 6) ผลิตข้อความที่ขยายเพิ่มและต่อเนื่องกันในหัวข้อที่กำหนด 7) ตอบคำถาม ให้ข้อมูลเกี่ยวกับตนเอง และแสดงความคิดเห็นในหัวเรื่องต่าง ๆ ได้อย่างคล่องแคล่ว 8) นำเสนอหัวข้อที่กำหนดให้โดยมีโครงสร้างชัดเจนและให้รายละเอียดสนับสนุนความคิดเห็นของตนเอง ได้อย่างเหมาะสม 9) แสดงและอธิบายความเห็น บรรยาย เปรียบเทียบ วิเคราะห์ อภิปราย และพยากรณ์ประเด็นที่กล่าวถึง และ 10) แลกเปลี่ยนความคิดเห็น แสดงและปรับเปลี่ยนความคิดเห็น ด้วยและ/หรือไม่เห็นด้วย ให้คำแนะนำ คาดการณ์ ประเมิน เพื่อนำไปสู่การตัดสินใจและสรุปผล เป็นต้น

3) ลักษณะภาระงาน (Task Characteristics) อาจจะประกอบด้วย การสัมภาษณ์ การบรรยายรูปภาพ การนำเสนอรายบุคคล และการอภิปรายกับคู่สนทนา ดังรายละเอียดต่อไปนี้

การสัมภาษณ์ มีภาระงานคือ สอบสัมภาษณ์ แสดงความคิดเห็น และตอบคำถามจากกรรมการ ใช้เวลา 5 นาที มีตัวกระตุ้น คือ คำถามจากกรรมการ และมีหัวเรื่องเกี่ยวกับข้อมูลเกี่ยวกับตนเอง ประสบการณ์ และความคิดเห็นเกี่ยวกับหัวเรื่องและสถานการณ์ที่เกี่ยวข้องกับตัวเอง ความสามารถที่ต้องแสดง ได้แก่ ตอบคำถามผู้ประเมินอย่างละเอียด และให้ข้อมูลเกี่ยวกับตนเอง ประสบการณ์ และความคิดเห็นเกี่ยวกับหัวเรื่องและสถานการณ์ที่เกี่ยวข้องกับตัวเอง

การบรรยายรูปภาพ มีภาระงาน คือ พูดบรรยายสถานการณ์ เหตุการณ์ในรูปภาพได้อย่างละเอียดครบถ้วน ใช้เวลา 3 นาที มีตัวกระตุ้น คือ รูปภาพ หัวเรื่องที่เกี่ยวข้อง คือ สถานการณ์ เหตุการณ์ เรื่องราวที่เกิดขึ้นในชีวิตประจำวัน และความสามารถที่ต้องแสดง คือ พูดได้ครอบคลุมประเด็นต่าง ๆ ในรูปภาพ ใช้ภาษาได้ถูกต้องและเหมาะสม และจัดระบบความคิดในการนำเสนอได้อย่างเป็นระเบียบและมีโครงสร้าง

การนำเสนอรายบุคคล มีภาระงาน คือ พูดนำเสนอประเด็นตามสถานการณ์ที่กำหนดในภาระงานและตอบคำถามเพิ่มเติม ใช้เวลาในการพูด 3 นาที และเตรียมตัว 2 นาที มีตัวกระตุ้น คือ รูปภาพและบัตรภาระงาน หัวเรื่องที่เกี่ยวข้อง คือ สถานการณ์ที่พบได้ในชีวิตประจำวัน และความสามารถที่ต้องแสดง คือ พูดได้ครอบคลุมประเด็นต่าง ๆ ที่กำหนดในภาระงาน ใช้ภาษาได้ถูกต้องและเหมาะสม จัดระบบความคิดในการนำเสนอได้อย่างเป็นระเบียบและมีโครงสร้าง และตอบคำถามกรรมการได้ถูกต้องและครบถ้วน

การอภิปรายกับคู่สนทนา มีภาระงาน คือ ร่วมสนทนาและอภิปรายประเด็นที่ได้รับมอบหมายหรือปรากฏในภาระงาน ใช้เวลา 10 นาที มีตัวกระตุ้น คือ บัตรภาระงาน หัวเรื่อง

เกี่ยวข้องคือ ปัญหาหรือสถานการณ์ปัจจุบันในสังคม และ ประเด็นที่อยู่ในความสนใจของตนเองและเพื่อนคู่สนทนาโดยการเลือกจากหัวข้อที่กำหนด และความสามารถที่ต้องแสดง คือ ตอบคำถาม มีส่วนร่วมในการอภิปราย แลกเปลี่ยนความคิดเห็นกับคู่สนทนาและทำให้การสนทนาดำเนินต่อไปได้ โดย อธิบาย ให้เหตุผล วิเคราะห์ พยากรณ์ แสดงและปรับเปลี่ยนความคิดเห็น สอบถามความคิดเห็นของผู้อื่น เปรียบเทียบ บรรยาย เห็นด้วยและ/หรือไม่เห็นด้วย ให้คำแนะนำ ประเมิน สรุปและตัดสินใจเกี่ยวกับปัญหาหรือประเด็นที่กำลังอภิปราย

4) เกณฑ์การตรวจให้คะแนน ประกอบด้วย 1) ความถูกต้อง (Accuracy) ของการออกเสียง ลงเสียงหนัก-เบา สูง-ต่ำ และใช้ภาษาถูกต้องตามหลักไวยากรณ์ 2) ความคล่องแคล่ว (Fluency) ในการพูด 3) ใช้คำศัพท์และไวยากรณ์อย่างหลากหลาย (Grammatical and Vocabulary Range) 4) การดำเนินการอภิปราย (Managing a Discussion) การเริ่มต้น การสลับบทบาท การขัดจังหวะ การให้ตัวกระตุ้น การเจรจาต่อรอง การจบการอภิปราย และ 5) ความสมบูรณ์ของภาระงาน (Task Fulfillment) การนำเสนอสอดคล้องกับหัวเรื่อง การจัดการภาระงานได้ตามที่ระบุ การนำเสนอข้อมูลที่เป็นประโยชน์ การนำเสนอสาระที่เพียงพอ แสดงการแก้ปัญหาได้อย่างสมบูรณ์

สรุปสาระสำหรับการทดสอบทักษะการอ่าน

สาระสำหรับการทดสอบทักษะการอ่านได้มาจากการสังเคราะห์ข้อมูลดังกล่าวข้างต้น เพื่อให้ได้ 1) จุดประสงค์ (Purpose) 2) สาระภาษา (Language Knowledge/Subskills) 3) หัวเรื่อง (Topical Knowledge) 4) รูปแบบการทดสอบ (Test Type) และ 5) การตรวจให้คะแนน (Scoring Technique) สำหรับการทดสอบทักษะการอ่าน ดังแสดงในตารางสังเคราะห์ต่อไปนี้



ตารางที่ 3 การสังเคราะห์สาระแบบทดสอบวัดทักษะการอ่าน

แหล่งข้อมูล (Sources)	หลักสูตร (มหาวิทยาลัยทักษิณ, 2544)	CEFR (Council of Europe, 2003)
สาระแบบทดสอบ (Test Content)	วัดความสามารถในการอ่านของผู้เข้าสอบ	วัดความสามารถของผู้เข้าสอบในการอ่านเพื่อความเข้าใจบทความชนิดต่าง ๆ ที่ความเข้าใจประเด็นสำคัญและรายละเอียดเฉพาะของบทอ่าน
จุดประสงค์ (Purpose)	เข้าใจและจับประเด็นสำคัญ และรายละเอียดสนับสนุนของเนื้อหาสาระที่อ่านได้	<p>1) เข้าใจคำบรรยายเหตุการณ์ ความรู้สึก และความปรารถนาในจดหมายส่วนตัวได้เป็นอย่างดีเพื่อที่จะสื่อสารกับเพื่อนทางจดหมายได้อย่างเป็นปกติ ;2) อ่านจดหมายที่เกี่ยวข้องกับสาขาความสนใจของตนเองและจับความหมายที่สำคัญได้ทันที ;3) กวาดสายตาอ่านบทอ่านที่ยาวขึ้นเพื่อหาข้อมูลที่ต้องการ และเก็บข้อมูลจากส่วนต่าง ๆ ของบทอ่าน หรือจากบทอ่านต่าง ๆ เพื่อมาเติมเต็มภาระงานเฉพาะของของตน ;</p> <p>4) หาและเข้าใจข้อมูลสำคัญของวัสดุที่พบประจำวัน เช่น จดหมาย โบรชัวร์ และเอกสารทางราชการสั้น ๆ ;</p> <p>5) กวาดสายตาอย่างรวดเร็วในกรอ่านบทอ่านที่ยาวและซับซ้อน เพื่อหารายละเอียดที่สำคัญได้ ;</p> <p>6) ระบุอย่างรวดเร็วในสาระและความสำคัญของประเด็นข่าว บทความ และรายงาน ในหัวข้อที่เกี่ยวข้องกับสาขาอาชีพที่หลากหลาย ตัดสินใจได้ว่าการศึกษาอย่างละเอียดคุ้มค่าหรือไม่ ;</p> <p>7) ระบุข้อสรุปสำคัญของบทอ่านเชิงโต้แย้งที่มีการให้สถิติขมนี้ไว้อย่างชัดเจน ;</p> <p>8) ระบุบรรทัดของข้อโต้แย้งเพื่อแก้ปัญหาที่ถูกนำเสนอ โดยไม่จำเป็นต้องให้รายละเอียด ;</p> <p>9) บอกประเด็นสำคัญจากการอ่านบทความในหนังสือพิมพ์ที่นำเสนอหัวข้อที่คุ้นเคยและใช้ภาษาไม่ซับซ้อน ;</p> <p>10) รับข้อมูล ความคิด และความคิดเห็นจากแหล่งเฉพาะด้านมาก ๆ ในสาขาของตน ;</p> <p>11) เข้าใจบทความพิเศษนอกสาขาของตนเอง โดยที่สามารถใช้พจนานุกรมอยู่เป็นบางครั้งเพื่อยืนยันการตีความในคำศัพท์เฉพาะของตนเอง ;</p> <p>12) เข้าใจบทความและรายงานเกี่ยวกับปัญหาในปัจจุบันที่ผู้เขียนเอาตัวอย่างหรือมุมมองมานำเสนอ ;</p> <p>13) เข้าใจคำแนะนำการใช้สำหรับอุปกรณ์บางอย่างที่เข้าใจได้ไม่ยาก และเขียนไว้อย่างชัดเจน ;</p> <p>14) เข้าใจคำแนะนำการใช้ในสาขาของตนเองที่ซับซ้อนและมีความยาว รวมทั้งรายละเอียดในเงื่อนไขและคำเตือน โดยที่ต้องอ่านเข้าส่วนที่ยกอยู่บ้าง</p>

ตารางที่ 3 การสังเคราะห์สาระแบบทดสอบวัดทักษะการอ่าน (ต่อ)

แหล่งข้อมูล (Sources)	หลักสูตร มหาวิทยาลัยทักษิณ, 2544)	CEFR (Council of Europe, 2003)
สาระแบบทดสอบ (Test Content)	หัวข้อเรื่อง (Topical Knowledge) การใช้ภาษาอังกฤษเพื่อการสื่อสาร	หัวข้อเรื่อง (Topical Knowledge) การอ่านจดหมาย 1) จดหมายส่วนตัว; และ 2) จดหมายที่เกี่ยวข้องกับสาขาความสนใจของตนเอง การอ่านเพื่อจุดประสงค์เฉพาะ 3) บทอ่านต่าง ๆ ที่ยาวขึ้น ; 4) สื่อวัสดุที่พบประจำวัน เช่น จดหมาย โบรมัวร์ และเอกสารทางราชการอื่น ๆ ; 5) บทอ่านที่ยากและซับซ้อน ;และ 6) ข่าว บทความ และรายงาน ในหัวข้อที่เกี่ยวข้องกับสาขาอาชีพที่หลากหลาย การอ่านเพื่อหาข้อมูลและโต้แย้ง 7) บทอ่านเชิงโต้แย้งที่มีการให้สัญลักษณ์ไว้อย่างชัดเจน ; 8) บทความจากหนังสือพิมพ์ในหัวเรื่องที่คุ้นเคยที่นำเสนอด้วยภาษาที่ไม่ซับซ้อน; 9) ข้อมูลเฉพาะด้านอย่างแท้จริงในสาขาของตน ; 10) บทความพิเศษนอกสาขาของตนเอง ;และ 11) บทความและรายงานเกี่ยวกับปัญหาในปัจจุบัน การอ่านคำแนะนำการใช้ 12) คำแนะนำการใช้สำหรับอุปกรณ์บางอย่างที่เข้าใจได้ง่ายและเขียนไว้อย่างชัดเจน ;และ 13) คำแนะนำการใช้ในสาขาของตนเองที่ยากและซับซ้อน พร้อมทั้งมีรายละเอียดในเชิงและคำเตือน

ตารางที่ 3 การสังเคราะห์สาระแบบทดสอบวัดทักษะการอ่าน (ต่อ)

แหล่งข้อมูล (Sources)	CET (Zheng & Cheng, 2008)	MUET (Malaysian Examinations Council, 2006)	CU-TEP (สถาบันภาษา จุฬาลงกรณ์มหาวิทยาลัย, 2552)	TOEFL iBT (Educational Testing Service, 2010)	TOEFL ITP (Educational Testing Service, 2014)
เนื้อหาแบบทดสอบ (Test Content)	วัดความสามารถของผู้เข้าสอบในการเข้าใจประเด็นสำคัญของเรื่องและหาข้อมูลเฉพาะ	วัดความสามารถของผู้เข้าสอบในการเข้าใจบทความสั้นๆ ที่มีความยาวและความซับซ้อนแตกต่างกัน	ทดสอบทักษะการอ่านโดยให้ผู้เข้าสอบอ่านเรื่องที่กำหนดให้ และตอบคำถามเพื่อวัดความเข้าใจ	วัดความสามารถของผู้เข้าสอบในการเข้าใจบทความวิชาการและบทอ่าน อื่นๆ ในระดับมหาวิทยาลัย การอ่านเชิงวิชาการ	วัดความสามารถของผู้เข้าสอบในการอ่านสาระเชิงวิชาการ
สาระภาษา (Language Knowledge/Subskills)	1) เข้าใจประเด็นสำคัญของเรื่อง 2) หารายละเอียดเฉพาะในบทอ่าน	1) อ่านคร่าว ๆ และอ่านหารายละเอียดเฉพาะ 2) ดึงข้อมูลเฉพาะ 3) ระบุมุมคิดสำคัญ 4) ระบุรายละเอียดสนับสนุน 5) ถอดความหมายจากคำ วลี ประโยคในบริบท 6) เข้าใจข้อความประเภทความเรียงและประเภทรูปภาพ 7) เข้าใจความสัมพันธ์ภายในประโยค ระหว่างประโยค 8) รู้จักประโยคที่ถอดความ 9) ทำนายผลลัพธ์ 10) ปรับใช้แนวคิดในสถานการณ์ใหม่	1) เข้าใจคำถามเกี่ยวกับศัพท์จากรื่อง 2) เข้าใจคำถามข้อเท็จจริงจากรื่อง 3) เข้าใจคำถามใจความหลักของเรื่อง จุดประสงค์ของเรื่อง 4) เข้าใจคำถามประเภทการลงความเห็น 5) เข้าใจคำถามที่ต้องใช้วิธีการวิเคราะห์ (Scanning) หรือกวาดภาพที่จะให้หาพหุวาลีที่กล่าวถึงนั้นอยู่ตรงส่วนใดของเรื่อง และ 6) เข้าใจคำถามอ้างอิงที่ใช้ถามเกี่ยวกับคำนามหรือสรรพนาม	1) อ่านบทอ่านเพื่อค้นหาประเด็นหลักและข้อสรุปที่สำคัญได้อย่างมีประสิทธิภาพ 2) เพิ่มความคล่องแคล่วและความเร็วในการอ่าน 3) เข้าใจหัวข้อโดยทั่วไปหรือความคิดหลัก ประเด็นสำคัญ ข้อเท็จจริงที่สำคัญและรายละเอียด คำศัพท์และบริบทและการใช้คำสรรพนาม 4) เชื่อมโยงความเข้าใจเกี่ยวกับสิ่งที่ถูกกล่าวโดยนัยในบทอ่าน 5) จัดลำดับข้อความและจุดประสงค์ของบทอ่าน 6) เข้าใจความใจความสัมพันธ์	1) To answer main idea question correctly, 2) To use context to determine meanings of simple words, 3) To answer transition questions correctly, 4) To answer stated detail questions correctly, 5) To find pronoun referents, 6) To determine meanings from word parts.

ตารางที่ 3 การสังเคราะห์สาระแบบทดสอบวัดทักษะการอ่าน (ต่อ)

แหล่งข้อมูล (Sources) เนื้อหาแบบทดสอบ (Test Content)	CET (Zheng & Cheng, 2008)	MUET (Malaysian Examinations Council, 2006)	CU-TEP (สถาบันภาษา จุฬาลงกรณ์มหาวิทยาลัย, 2552)	TOEFL iBT (Educational Testing Service, 2010)	TOEFL ITP (Educational Testing Service, 2014)
		11) เข้าใจหน้าที่ภาษา 12) ดีความเข้าใจและไม่ตามลำดับตามลำดับและไม่ตามลำดับ 13) แยกแยะข้อมูลสำคัญออกจากข้อมูลที่ไม่สำคัญ 14) แยกแยะข้อเท็จจริงออกจากความคิดเห็น 15) ลงความเห็น 16) เชื่อมความคิดและแนวความคิดภายในย่อหน้าและระหว่างย่อหน้า 17) ติดตามการดำเนินเรื่องของประเด็นหรือข้อโต้แย้ง 18) สรุปข้อมูล 19) ประเมินข้อมูล 20) ทำการตัดสิน 21) ตั้งข้อสรุป 22) รู้จักและตีความความเห็นที่สนับสนุน หรือเจตนาของผู้พูด	หรือข้อความอื่น ๆ หมายถึงใครหรืออะไร	ระหว่างความคิด 7) จัดการข้อมูลเป็นชาร์ตเพื่อแบ่ง ประเภท หรือหาทศรูปเพื่อให้จำประเด็นหลัก ๆ และรายละเอียดที่สำคัญได้ 8) เข้าใจการเชื่อมโยงความคิดต่าง ๆ กันตลอดบทอ่าน	7) To answer implied detail question correctly และ 8) To determine where specific information is found

ตารางที่ 3 การสังเคราะห์สาระแบบทดสอบวัดทักษะการอ่าน (ต่อ)

แหล่งข้อมูล (Sources) เนื้อหาแบบทดสอบ (Test Content)	CET (Zheng & Cheng, 2008)	MUET (Malaysian Examinations Council, 2006)	CU-TEP (สถาบันภาษา จุฬาลงกรณ์มหาวิทยาลัย, 2552)	TOEFL iBT (Educational Testing Service, 2010)	TOEFL ITP (Educational Testing Service, 2014)
บทอ่าน (Reading Texts)	ชนิดของเนื้อความ (Text Type) 1) เจริญวิเคราะห์ (Analytical) 2) เจริญบรรยาย (Descriptive) 3) เจริญโน้มน้าวใจ (Persuasive) 4) เจริญโต้แย้ง (Argumentative) 5) เจริญพรรณนา (Narrative) แหล่งที่มาของบทอ่าน - บทความในวารสาร หนังสือพิมพ์และนิตยสาร - บทความวิชาการ - สารที่นำมาจากแหล่งข้อมูลอิเล็กทรอนิกส์	หัวข้อ - วิทยาศาสตร์และเทคโนโลยี (Science and Technology) 40 % - ประวัติศาสตร์และวัฒนธรรม (History & Culture) 20 % - สังคมวิทยา (Social Sciences) 10% - ศิลปะ (Arts) 15% - ชีวิตประวัติ (Biography) 15%	แหล่งที่มาของบทอ่าน - ส่วนตัดต่อจากตำราเรียนระดับมหาวิทยาลัยและสื่อวิชาการอื่น ๆ ที่เขียนเป็นภาษาอังกฤษ ซึ่งครอบคลุมภาษาอื่นในหลากหลายสาขาวิชา - นอกจากนี้บทอ่านมักจะนำเสนอข้อมูลเกี่ยวกับหัวเรื่องที่มีความสำคัญมากกว่าหนึ่งความคิดหรือแง่มุม	หัวข้อที่ใช้ในการทดสอบเกี่ยวข้องกับเทคโนโลยี ศิลปะ สัตว์ ธรรมชาติ และคน ข้อมูลจะเป็นบทอ่านสั้น ๆ จากตำรา หนังสือเชิงวิชาการ	รูปแบบคำถาม แบบเลือกตอบ 4 ตัวเลือก และมีคำตอบที่ถูกต้องเพียงตัวเดียว
รูปแบบการทดสอบ (Test Type)	ชนิดของภาระงาน (Task Type) 1) การอ่านเพื่อลึกลงในรายละเอียด (Reading in Depth) และ 2) การอ่านแบบ Skimming และ Scanning	รูปแบบคำถาม 1) คำถามแบบให้เลือกตอบที่มีตัวเลือก 3 และ 4 ตัวเลือก	ชนิดของภาระงาน (Task Type) 1) การอ่านเรื่องแล้วเติมคำ (Cloze Reading) 2) การอ่านบทอ่านสั้น (Reading Short Passages) และ	รูปแบบคำถาม 1) แบบเลือกตอบ 4 ตัวเลือก และมีคำตอบที่ถูกต้องเพียงตัวเดียว 2) แบบเลือกตอบมากกว่า 4 ตัวเลือกและคำตอบที่ถูกต้องมีได้มากกว่า 1 ตัวเลือก	รูปแบบคำถาม แบบเลือกตอบ 4 ตัวเลือก และมีคำตอบที่ถูกต้องเพียงตัวเดียว

ตารางที่ 3 การสังเคราะห์สาระแบบทดสอบวัดทักษะการอ่าน (ต่อ)

แหล่งข้อมูล (Sources)	CET (Zheng & Cheng, 2008)	MUET (Malaysian Examinations Council, 2006)	CU-TEP (สถาบันภาษา จุฬาลงกรณ์มหาวิทยาลัย, 2552)	TOEFL iBT (Educational Testing Service, 2010)	TOEFL ITP (Educational Testing Service, 2014)
เนื้อหาแบบทดสอบ (Test Content)	รูปแบบคำถาม 1) คำถามแบบให้เลือกตอบ (Multiple-Choice Questions) 2) เติมคำในช่องว่างแบบมีตัวเลือก (Banked Cloze Questions) 3) คำถามให้ตอบสั้น ๆ (Short Answer Questions) 4) ระบุถูกหรือผิด (True/False Statements) และ 5) การเติมคำในประโยคให้สมบูรณ์ (Sentence Completion)		3) การอ่านบทความยาว (Reading Long Passages) รูปแบบคำถาม 1) แบบเลือกตอบ 4 ตัวเลือก		
เกณฑ์การตรวจให้คะแนน (Scoring Criteria)	ตอบถูกได้ 1/ตอบผิดได้ 0	ตอบถูกได้ 1/ตอบผิดได้ 0	ตอบถูกได้ 1/ตอบผิดได้ 0	ตอบถูกได้ 1/ตอบผิดได้ 0 และให้คะแนนบางส่วน (Partial Credit)	ตอบถูกได้ 1/ตอบผิดได้ 0

ตารางที่ 3 การสังเคราะห์สาระแบบทดสอบวัดทักษะการอ่าน (ต่อ)

แหล่งข้อมูล (Sources) เนื้อหาแบบทดสอบ (Test Content)	IELTS (IELTS, 2015)	FCE (UCLES, 2014)	telc B1-B2 English (telc GmbH, 2012)	สรุปสาระ ที่จะนำไปสร้างแบบทดสอบ
จุดประสงค์ (Purpose)	ประเมินผู้เข้าสอบด้านความเข้าใจในรายละเอียดของประเด็นเฉพาะหรือความเข้าใจโดยทั่วไปเกี่ยวกับประเด็นหลักของเนื้อความ	วัดความสามารถของผู้เข้าสอบในการเข้าใจบทความชนิดต่าง ๆ	วัดความสามารถของผู้เข้าสอบโดยเน้นการอ่านเพื่อจับรายละเอียด การอ่านเพื่อความเข้าใจโดยรวม และการเลือกอ่านข้อมูลในชีวิตประจำวัน	วัดความสามารถของผู้เข้าสอบในการเข้าใจประเด็นสำคัญและรายละเอียดที่เกี่ยวข้องในการอ่านบทอ่านชนิดต่าง ๆ
สาระภาษา (Language Knowledge/Subskills)	<ol style="list-style-type: none"> เข้าใจในรายละเอียดของประเด็นเฉพาะหรือความเข้าใจโดยทั่วไปเกี่ยวกับประเด็นหลักของเนื้อความ จำแนกข้อมูลเฉพาะที่ไว้ในเนื้อความ จำแนกความคิดเห็นหรือประเด็นความแตกต่าง ๆ กวาดสายตาอ่านเนื้อความเพื่อที่จะหาข้อมูลเฉพาะ ระบุหัวข้อหลักของย่อหน้าหรือตอน จำแนกความแตกต่างระหว่างความคิดหลักกับความคิดสนับสนุน ระบุความสัมพันธ์และการเชื่อมโยงระหว่างข้อเท็จจริงในเนื้อความ 	<ol style="list-style-type: none"> ทราบรายละเอียด ข้อมูลเฉพาะ ระบุความคิดหลัก ระบุ ความคิดเห็น น้ำเสียง จุดประสงค์ ของผู้เขียน อนุมานความ (implication) และทัศนคติ เข้าใจโครงสร้างและรูปแบบการพัฒนาของตัวบท 	<ol style="list-style-type: none"> เข้าใจรายละเอียด และใจความหลักหรือประเด็นสำคัญของสาระที่อ่าน 	<ol style="list-style-type: none"> อ่านบทอ่านเพื่อหาประเด็นสำคัญและรายละเอียดที่เกี่ยวข้องได้อย่างมีประสิทธิภาพ เข้าใจความคิดหลัก ประเด็นสำคัญ และรายละเอียดเฉพาะในบทอ่าน เข้าใจศัพท์ วลี และประโยค ที่ใช้ในบริบทสาระที่อ่าน ระบุความสัมพันธ์และความเกี่ยวข้องระหว่างข้อเท็จจริงกับประเด็นหลักในบทอ่าน จัดลำดับเหตุการณ์หรือการดำเนินเรื่องหรือโครงสร้างในบทอ่าน เชื่อมโยงความเข้าใจ ลงความเห็นหรือทำนายเกี่ยวกับสิ่งที่ถูกกล่าวโดยนัยในบทอ่าน และ ระบุความเห็น ความรู้สึก

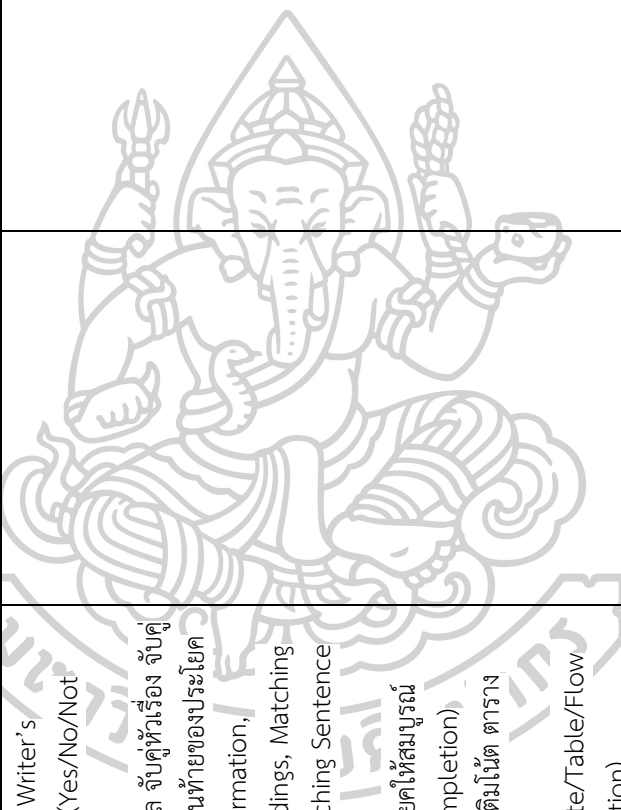
ตารางที่ 3 การสังเคราะห์สาระแบบทดสอบวัดทักษะการอ่าน (ต่อ)

แหล่งข้อมูล (Sources) เนื้อหาแบบทดสอบ (Test Content)	IELTS (IELTS, 2015)	FCE (UCLES, 2014)	telc B1-B2 English (telc GmbH, 2012)	สรุปสาระ ที่จะนำไปสร้างแบบทดสอบ
	<p>8) ระบุความคิดเห็น และทัศนคติ</p> <p>9) อ่านคร่าว ๆ และอ่านเจาะลึก</p> <p>เฉพาะที่เน้นเนื้อหาข้อมูล</p> <p>อย่างรวดเร็ว</p>			จุดประสงค์ และทัศนคติของผู้เขียน
องค์ประกอบของบทอ่าน (Components of Reading Texts)	<p>หัวข้อเรื่อง</p> <ul style="list-style-type: none"> - หัวข้อต่าง ๆ ที่อยู่ในความสนใจ <p>โดยทั่วไปของนักศึกษาระดับปริญญาตรีหรือระดับที่สูงกว่า</p> <p>ชนิดของเนื้อหา (Text Type)</p> <ol style="list-style-type: none"> 1) เชิงพรรณนา (Narrative) 2) เชิงบรรยาย (Descriptive) 3) เชิงโต้แย้ง (Argumentative) <p>แหล่งที่มาของบทอ่าน</p> <ul style="list-style-type: none"> - หนังสือ - วารสาร - นิตยสาร - หนังสือพิมพ์ <p>- แหล่งทรัพยากรออนไลน์ ที่เขียนสำหรับผู้อ่านที่ไม่ใช่ผู้เชี่ยวชาญเฉพาะด้าน</p>	<p>แหล่งที่มาของบทอ่าน</p> <ul style="list-style-type: none"> - หนังสือพิมพ์และนิตยสาร - วารสาร - หนังสือ ทั้งแบบนิยาย และไม่ใช่ <p>นิตยสาร</p> <ul style="list-style-type: none"> - สื่อส่งเสริมการขายและให้ข้อมูล <p>ข่าวสาร</p>	<p>หัวข้อเรื่อง</p> <ol style="list-style-type: none"> 1) อีเมลส่วนตัวและกึ่งทางการ (Personal or Semi-Formal Emails) 2) ข้อความที่โพสต์บนส่วนที่ให้ความคิดเห็นบนอินเทอร์เน็ต (Messages Posted on an Internet Discussion Forum) 3) บทความประเภทให้ข้อมูลเกี่ยวกับหัวข้อพิเศษ (Informative Text about a Special Subject) 4) บทความในหนังสือพิมพ์ (Newspaper article) 5) โฆษณา (Advertisement) และ 6) บทความประเภทให้ข้อมูลที่เป็นทางการ (Formal Informative Texts) 	<p>หัวข้อเรื่อง</p> <ol style="list-style-type: none"> 1) จดหมายหรืออีเมลส่วนตัวและที่เกี่ยวข้องกับงาน 2) ข้อความจากกากรโปสเตอร์แสดงความคิดเห็นบนอินเทอร์เน็ต 3) บทความประเภทให้ข้อมูลข่าวสารในหัวข้อเฉพาะต่าง ๆ 4) บทความในหนังสือพิมพ์ วารสาร และแหล่งทรัพยากรออนไลน์ 5) โฆษณา สื่อส่งเสริมการขายและให้ข้อมูลข่าวสาร <p>ประเภทของเนื้อหา</p> <ol style="list-style-type: none"> 1) โน้มน้าวใจ 2) เชิงบรรยาย 3) เชิงพรรณนา 4) เชิงโต้แย้ง

ตารางที่ 3 การสังเคราะห์สาระแบบทดสอบวัดทักษะการอ่าน (ต่อ)

แหล่งข้อมูล (Sources) เนื้อหาแบบทดสอบ (Test Content)	IELTS (IELTS, 2015)	FCE (UCLES, 2014)	telc B1-B2 English (telc GmbH, 2012)	สรุปสาระ ที่จะนำไปสร้างแบบทดสอบ
รูปแบบการทดสอบ (Test Type)	1) แบบให้เลือกคำตอบ (Multiple Choice) 2) การระบุข้อมูลถูกต้อง/ผิด/หรือไม่ ข้อมูลให้ (Identifying Information (True/False/Not Given) 3) การระบุความเห็นหรือการอ้างของ ผู้เขียนว่าถูก/ผิด/หรือไม่ใช่ข้อมูล	1) คำถามแบบให้เลือกคำตอบ (Multiple Choice) 2) การเติมคำในช่องว่างของเนื้อความ (Gapped Text) และ 3) การจับคู่โดยเลือกจากคำตอบที่มีหลายตัวเลือก (Multiple Matching)	1) แบบจับคู่ (Matching Items) 2) คำถามแบบให้เลือกคำตอบ (Multiple-Choice Items) และ 3) ข้อคำถามแบบให้ระบุ ถูก/ผิด (True/False Items)	1) คำถามแบบให้เลือกตอบ 2) เติมคำในประโยค 3) แบบจับคู่ และ 4) แบบระบุถูกหรือผิด
			แหล่งที่มาของบทอ่าน - อีเมลและบรรทัดบอกหัวเรื่อง (Emails and Subject Lines) - คำถามและคำตอบที่นำมาจากส่วนที่เปิดให้แสดงความคิดเห็นทางอินเทอร์เน็ต (Questions and Answers from an Internet Forum) - บทอ่านสารนิพนธ์ชนิดต่าง ๆ (Different Types of Published Texts) - บทอ่านประเภทให้ข้อมูลที่เป็นทางการ (Formal Informative Texts)	แหล่งที่มาของบทอ่าน - หนังสือและนิตยสาร - วารสาร - หนังสือพิมพ์ - สื่อส่งเสริมการขายและให้ข้อมูลข่าวสาร - แหล่งทรัพยากรออนไลน์

ตารางที่ 3 การสังเคราะห์สาระแบบทดสอบวัดทักษะการอ่าน (ต่อ)

แหล่งข้อมูล (Sources) เนื้อหาแบบทดสอบ (Test Content)	IELTS (IELTS, 2015)	FCE (UCLES, 2014)	telc B1-B2 English (telc GmbH, 2012)	สรุปสาระ ที่จะนำไปสร้างแบบทดสอบ
	<p>ให้ (Identifying Writer's Views/Claims (Yes/No/Not given)</p> <p>4) การจับคู่ข้อมูล จับคู่หัวเรื่อง จับคู่ลักษณะ จับคู่ส่วนท้ายของประโยค (Matching Information, Matching Headings, Matching Features, Matching Sentence Endings)</p> <p>5) การเติมประโยคให้สมบูรณ์ (Sentence Completion)</p> <p>6) บทสรุป การเติมโน้ต ตาราง ชาร์ตให้สมบูรณ์ (Summary/Note/Table/Flow Chart Completion)</p> <p>7) การเติมโตอะแกรมให้สมบูรณ์ (Diagram Label Completion)</p> <p>8) คำถามให้ตอบสั้น ๆ (Short-Answer Questions)</p>			

ตารางที่ 3 การสังเคราะห์สาระแบบทดสอบวัดทักษะการอ่าน (ต่อ)

แหล่งข้อมูล (Sources)	IELTS (IELTS, 2015)	FCE (UCLES, 2014)	telc B1-B2 English (telc GmbH, 2012)	สรุปสาระ ที่จะนำไปสร้างแบบทดสอบ
เนื้อหาแบบทดสอบ (Test Content) เกณฑ์การตรวจให้ คะแนน (Scoring Criteria)	ตอบถูกได้ 1/ตอบผิดได้ 0	ตอบถูกได้ 1/ตอบผิดได้ 0	ตอบถูกได้ 1/ตอบผิดได้ 0	ตอบถูกได้ 1/ตอบผิดได้ 0



ผลจากการสังเคราะห์สาระแบบทดสอบวัดทักษะการอ่านในตารางที่ 3 สรุปได้ว่าแบบทดสอบด้านการอ่านควรมีกรอบสาระด้าน 1) จุดประสงค์ (Purpose) 2) สาระภาษา (Language Knowledge/Subskills) 3) 109 หัวเรื่อง (Topical Knowledge) 4) รูปแบบการทดสอบ (Test Type) และ 5) การตรวจให้คะแนน (Scoring Technique) ดังรายละเอียดต่อไปนี้

1) จุดประสงค์ของแบบทดสอบ (Purpose) คือ เพื่อวัดความสามารถของผู้เข้าสอบในการเข้าใจประเด็นสำคัญและรายละเอียดที่เกี่ยวข้องในการอ่านบทอ่านชนิดต่าง ๆ

2) สาระภาษา (Language Knowledge/Subskills) ที่เกี่ยวข้อง คือ ความสามารถในการ 1) เข้าใจความคิดหลัก ประเด็นสำคัญ และรายละเอียดเฉพาะที่เกี่ยวข้องของบทอ่าน 2) เข้าใจคำศัพท์ วลี และประโยค ที่ใช้ในบริบทของสาระที่อ่าน 3) ระบุความสัมพันธ์และความเกี่ยวข้องระหว่างข้อเท็จจริงกับประเด็นหลักในบทอ่าน 4) จัดลำดับเหตุการณ์หรือการดำเนินเรื่องหรือโครงสร้างในบทอ่าน 5) เชื่อมโยงความเข้าใจ ลงความเห็น หรือทำนายเกี่ยวกับสิ่งที่ถูกกล่าวโดยนัยในบทอ่าน และ 6) ระบุความเห็น ความรู้สึก จุดประสงค์ และทัศนคติของผู้เขียน

3) ลักษณะของบทอ่าน (Listening Texts) ที่สามารถนำมาออกแบบแบบทดสอบวัดความสามารถภาษาอังกฤษควรอยู่ในบริบทการใช้ภาษาอังกฤษเพื่อการสื่อสารในชีวิตประจำวันและในสังคมที่ครอบคลุมเรื่องส่วนตัวและการทำงานโดยสามารถ อันได้แก่ 1) จดหมายหรืออีเมลส่วนตัว และที่เกี่ยวข้องกับงาน 2) ข้อความจากการโพสต์แสดงความคิดเห็นบนอินเทอร์เน็ต 3) บทอ่านประเภทให้ข้อมูลข่าวสารในหัวเรื่องเฉพาะต่าง ๆ 4) บทความในหนังสือพิมพ์ วารสาร และแหล่งทรัพยากรออนไลน์ และ 5) โฆษณา สื่อส่งเสริมการขายและให้ข้อมูลข่าวสาร สำหรับประเภทของเนื้อความอาจจะเป็นแบบ 1) โน้มน้าวใจ 2) เชิงบรรยาย 3) เชิงพรรณนา และ 4) เชิงโต้แย้ง และแหล่งที่มาของบทอ่านอาจมาจาก หนังสือและนิตยสาร วารสาร หนังสือพิมพ์ สื่อส่งเสริมการขาย และให้ข้อมูลข่าวสาร และแหล่งทรัพยากรออนไลน์

4) รูปแบบการทดสอบ (Test Type) ที่สามารถนำมาใช้ทดสอบทักษะการฟังของแบบทดสอบวัดความสามารถภาษาอังกฤษควรประกอบด้วย 1) แบบเลือกตอบ 2) แบบเติมคำในประโยค 3) แบบจับคู่ และ 4) แบบระบุถูกหรือผิด

5) การตรวจให้คะแนนจะเป็นรูปแบบการตอบถูกได้ 1 หรือ ตอบผิด ได้ 0

สรุปสาระสำหรับการทดสอบทักษะการเขียน

สาระสำหรับการทดสอบทักษะการเขียนได้มาจากการสังเคราะห์ข้อมูลดังกล่าวข้างต้นเพื่อให้ได้ 1) จุดประสงค์ (Purpose) 2) สาระภาษา (Language Knowledge/Subskills) 3) ชนิด 109 ภาระงาน (Task Type) และ 4) การตรวจให้คะแนน (Scoring Technique) สำหรับการทดสอบทักษะการเขียน ดังแสดงในตารางสังเคราะห์ต่อไปนี้

ตารางที่ 4 การสังเคราะห์สาระแบบทดสอบวัดทักษะการเขียน

แหล่งข้อมูล (Sources)	หลักสูตร (มหาวิทยาลัยทักษิณ, 2554)	CEFR (Council of Europe, 2003)
สาระแบบทดสอบ (Test Content)	วัดความสามารถในการเขียนของผู้เข้าสอบ	วัดความสามารถของผู้เข้าสอบในการสื่อสารโดยใช้ภาษาเขียน
จุดประสงค์ (Purpose)	สื่อสารโดยใช้ภาษาเขียนได้ อย่างถูกต้องและเหมาะสมกับ จุดประสงค์และผู้รับสาร	<ol style="list-style-type: none"> 1) เขียนจดหมายส่วนตัวเพื่อส่งข่าวและแสดงความคิดเห็นในหัวข้อที่เกี่ยวข้องกับนามธรรมหรือวัฒนธรรม ; 2) เขียนจดหมายส่วนตัวอธิบายประสบการณ์ ความรู้สึก และเหตุการณ์ โดยให้รายละเอียดพอสมควร ; 3) เขียนจดหมายเพื่ออธิบายของอารมณณ์และเสนอการให้ความสำคัญส่วนตัวต่อเหตุการณ์และประสบการณ์และแสดงความคิดเห็นต่อสิ่งใหม่และความคิดเห็นในตัวสาร ; 4) เขียนบรรยายที่มีรายละเอียดขั้นซ้อน ในหัวข้อต่าง ๆ ที่คุ้นเคย ภายใต้อาสาความสนใจของตน ; 5) เขียนบอกเล่าประสบการณ์ บรรยายความรู้สึก และปฏิกิริยาตอบสนอง ด้วยข้อความที่เกี่ยวข้องกันแบบง่าย ๆ ; 6) เขียนบรรยายเหตุการณ์ การเดินทางที่เพิ่งผ่านมา ตามความจริงหรือจินตนาการ ; 7) เขียนเล่าเรื่องได้ ; 8) เขียนบรรยายที่ละเอียด ชัดเจนของเหตุการณ์และประสบการณ์หรือจินตนาการ โดยแสดงความสัมพันธ์ระหว่างความคิดในข้อความที่ต่อเนื่องกันชัดเจน และปฏิบัติตามข้อตกลงที่กำหนดไว้ตามประเภทการเขียนที่เกี่ยวข้อง ; 9) เขียนบรรยายที่ละเอียด ชัดเจน ในหัวข้อที่หลากหลาย เกี่ยวกับสภาพความสนใจของตน ; และ 10) เขียนวิจารณ์ภาพยนตร์ หนังสือ หรือละคร

ตารางที่ 4 การสังเคราะห์สาระแบบทดสอบวัดทักษะการเขียน (ต่อ)

แหล่งข้อมูล (Sources)	<p>หลักสูตร มหาวิทยาลัยทักษิณ, 2554)</p>	<p>CEFR (Council of Europe, 2003)</p>
เนื้อหาแบบทดสอบ (Test Content)	<p>หัวข้อเรื่อง (Topical Knowledge) การใช้ภาษาอังกฤษเพื่อการสื่อสาร</p>	<p>หัวข้อเรื่อง (Topical Knowledge) การเขียนจดหมาย 1) การเขียนจดหมายส่วนตัวเพื่อส่งข่าวและแสดงความคิดเห็นในหัวข้อที่เกี่ยวข้องกับนามธรรมหรือวัฒนธรรม ; 2) การเขียนจดหมายส่วนตัวอธิบายประสบการณ์ ความรู้สึก และเหตุการณ์ โดยให้รายละเอียดพอสมควร ;และ 3) การเขียนจดหมายเพื่อสื่อระดับของอารมณ์และเสนอการให้ความช่วยเหลือต่อเหตุการณ์และประสบการณ์และแสดงความคิดเห็นต่อสิ่งใหม่และความคิดเห็นในตัวสาร การเขียนเชิงสร้างสรรค์ 4) การเขียนบรรยายเกี่ยวกับหัวข้อต่าง ๆ ที่คุ้นเคย ภายใต้อารมณ์ของตนเอง ; 5) การเขียนบอกเล่าประสบการณ์ บรรยายความรู้สึก และปฏิบัติการตอบสนอง ด้วยข้อความที่เกี่ยวข้องกันแบบง่าย ๆ ; 6) การเขียนบรรยายเหตุการณ์ เช่น การเดินทางที่เพิ่งผ่านมา โดยอาจเป็นเหตุการณ์จริงหรือตามจินตนาการ ; 7) การเขียนเล่าเรื่อง ; 8) การเขียนบรรยายเหตุการณ์และประสบการณ์จริงหรือตามจินตนาการที่ละเอียดและชัดเจน โดยแสดงความสัมพันธ์ระหว่างความคิดในข้อความที่ต่อเนื่องกันได้ชัดเจน และปฏิบัติตามข้อตกลงที่กำหนดไว้ตามแต่ละประเภทการเขียนที่เกี่ยวข้อง ; 9) การเขียนบรรยายในหัวเรื่องที่หลากหลาย เกี่ยวข้องกับสาขาความสนใจของตนเองละเอียดและชัดเจน-B2; 10) การเขียนวิจารณ์ภาพยนตร์ หนังสือ หรือละคร</p>

ตารางที่ 4 การสังเคราะห์สาระแบบทดสอบวัดทักษะการเขียน (ต่อ)

แหล่งข้อมูล (Sources)	CET (Zheng & Cheng, 2008)	MUET (Malaysian Examinations Council, 2006)	TOEFL iBT (Educational Testing Service, 2010)
เนื้อหาแบบทดสอบ (Test Content)	<p>ประเมินความสามารถของนักศึกษาในการสื่อสารความคิดโดยใช้ภาษาเขียน</p> <p>1) เขียนเรียงความโดยใช้ข้อมูลที่กำหนดให้ และ</p> <p>2) แปลประโยคที่กำหนดให้จากภาษาจีนเป็นภาษาอังกฤษ</p>	<p>ประเมินความสามารถของผู้เข้าสอบในการเขียนข้อความชนิดต่าง ๆ ที่ครอบคลุมรูปแบบการเขียนที่หลากหลาย</p> <p>1) ใช้การสะกดคำและเครื่องหมายที่ถูกต้อง</p> <p>2) ใช้ไวยากรณ์ถูกต้อง</p> <p>3) ใช้โครงสร้างประโยคถูกต้อง</p> <p>4) ใช้คำศัพท์และสำนวนภาษาที่หลากหลาย</p> <p>5) ใช้โครงสร้างประโยคที่หลากหลายชัดเจน</p> <p>6) ใช้ภาษาเหมาะสมกับจุดประสงค์ที่ต้องการและผู้รับสาร</p> <p>7) สังเกตดูภาษาที่เหมาะสมสำหรับสถานการณ์เฉพาะหรือชนิดของตัวบท</p> <p>8) การพัฒนาและจัดระบบความคิด</p> <p>9) ใช้คำบอกเครื่องหมายและคำเชื่อมที่เหมาะสม</p> <p>10) ใช้คำแทนคำซ้ำที่เหมาะสมพร้อมทั้งคำเชื่อมอื่น ๆ</p> <p>11) นิยาม บรรยาย อธิบาย</p> <p>12) เปรียบเทียบ และระบุความต่าง</p> <p>13) แบ่งประเภท</p> <p>14) ให้เหตุผล</p> <p>15) แสดงความคิดเห็น</p> <p>16) แสดงความสัมพันธ์</p> <p>17) ให้ข้อเสนอแนะ และคำแนะนำ</p>	<p>วัดความสามารถของผู้เข้าสอบในการเขียนภาษาอังกฤษในบริบทวิชาการ</p> <p>1) จัดบันทึกสิ่งที่ได้ยินและอ่าน และใช้บันทึกเพื่อจัดระเบียบข้อมูลก่อนการเขียน</p> <p>2) สรุป ดัดแปลงจากแหล่งข้อมูล</p> <p>3) เขียนในทำนองที่บอกกล่าวสิ่งที่ผู้เข้าสอบได้ยินสัมพันธ์กับสิ่งที่ผู้เข้าสอบได้อ่าน และเขียนเรียงความที่แสดงและสนับสนุนความคิดเห็น</p>
จุดประสงค์ (Purpose)			
สาระภาษา (Language Knowledge/Subskills)			

ตารางที่ 4 การสังเคราะห์สาระแบบทดสอบวัดทักษะการเขียน (ต่อ)

แหล่งข้อมูล (Sources) เนื้อหาแบบทดสอบ (Test Content)	CET (Zheng & Cheng, 2008)	MUET (Malaysian Examinations Council, 2006)	TOEFL iBT (Educational Testing Service, 2010)
<p>ภาระงานเขียน (Writing Tasks)</p> <p>ภาระงานที่ 1 เขียนเรียงความไม่น้อยกว่า 120 คำ สำหรับ CET-4 และ 150 คำ สำหรับ CET-6 กำหนดเวลา 30 นาที ตัวกระตุ้น ข้อมูลที่กำหนดให้ (เช่น ชื่อหัวเรื่อง โครงร่าง สถานการณ์ รูปภาพ หรือ กราฟ เป็นต้น) ผลิต (Output Text) 1) เรียงความไม่น้อยกว่า 120 คำ สำหรับ CET-4 และ 150 คำ สำหรับ CET-6</p>	<p>18) แสดงการเห็นด้วยและไม่เห็นด้วย 19) ค้นหาความชัดเจน 20) ถามและให้ข้อมูล 21) เชิญชวน 22) ดึงข้อสรุป 23) กลางและปรับเปลี่ยนความคิดเห็น 24) นำเสนอข้อโต้แย้ง 25) นำเสนอข้อมูลที่เกี่ยวข้อง 26) ให้อธิบายที่เพียงพอ และ 27) แสดงการแก้ปัญหาได้ตามที่หัวข้อต้องการ</p> <p>ข้อที่ 1 ภาระงาน ให้ผู้ใช้สอบตีความข้อมูล (ไม่น้อยกว่า 150 คำ) ตัวกระตุ้น ตัวกระตุ้นเฉพาะที่กำหนดให้ ข้อที่ 2 ภาระงาน ให้ผู้ใช้สอบเขียนข้อความจากหัวข้อที่กำหนดให้ (ไม่น้อยกว่า 350 คำ) ตัวกระตุ้น หัวข้อที่กำหนดให้ ผลิต (Output Text) รายงาน (Report) บทความ (Article) จดหมาย (Letter) และเรียงความ (Essay)</p>	<p>1) งานเขียนเชิงบูรณาการ (Integrated Writing Task) บูรณาการข้อมูลทั้งที่ศึกษาได้ฟังจากการบรรยายกับสิ่งที่พวกเขาได้อ่านในตำราหรือสื่ออื่น ๆ ตัวกระตุ้น - ข้อมูลจากการฟังบรรยาย - สิ่งที่พวกเขาได้อ่านในตำราหรือสื่ออื่น ๆ</p> <p>ผลิต (Output Text) - งานเขียนเชิงบูรณาการ</p>	

ตารางที่ 4 การสังเคราะห์สาระแบบทดสอบวัดทักษะการเขียน (ต่อ)

แหล่งข้อมูล (Sources) เนื้อหาแบบทดสอบ (Test Content)	CET (Zheng & Cheng, 2008)	MUET (Malaysian Examinations Council, 2006)	TOEFL iBT (Educational Testing Service, 2010)
	<p>ภาระงานที่ 2</p> <p>แปลประโยค 5 ประโยคจากภาษาจีนเป็นภาษาอังกฤษให้สมบูรณ์ กำหนดเวลา 5 นาที</p> <p>ตัวภาระต้น ประโยคภาษาจีน</p> <p>ผลผลิต (Output Text)</p> <p>ประโยคภาษาอังกฤษ</p>	<p>รูปแบบการเขียนเชิงวิเคราะห์ (Analytical) เชิงบรรยาย (Descriptive) เชิงใจใจ (Persuasive) และเชิงโต้แย้ง (Argumentative)</p>	<p><u>ความสามารถที่ต้องแสดง</u></p> <p>1) จัดบันทึกสิ่งที่ได้ยินและอ่าน</p> <p>2) ใช้บันทึกในจัดระเบียบข้อมูลก่อนการเขียน 3) สรุป ตัดแปลงสำนวนภาษา และอ้างอิงข้อมูลให้ถูกต้องจากแหล่งข้อมูล</p> <p>4) เขียนในทำนองที่บอกกล่าวให้ผู้เข้าสอบได้ยินสัมพันธ์กับสิ่งที่ผู้เข้าสอบได้อ่าน</p> <p>2) งานเขียนอิสระ (Independent Writing Task) เขียนเรียงความที่แสดงและสนับสนุนความคิดเห็น</p> <p><u>ตัวภาระต้น</u></p> <p>- ข้อกำหนดตามภาระงาน เช่น ประเด็นที่อยู่ในความขัดแย้งกัน (a controversial issue)</p> <p><u>ผลผลิต (Output Text)</u></p> <p>การเขียนแสดงความคิดเห็นและสนับสนุนความคิดเห็นด้วยความรู้และประสบการณ์ของตนเอง</p>

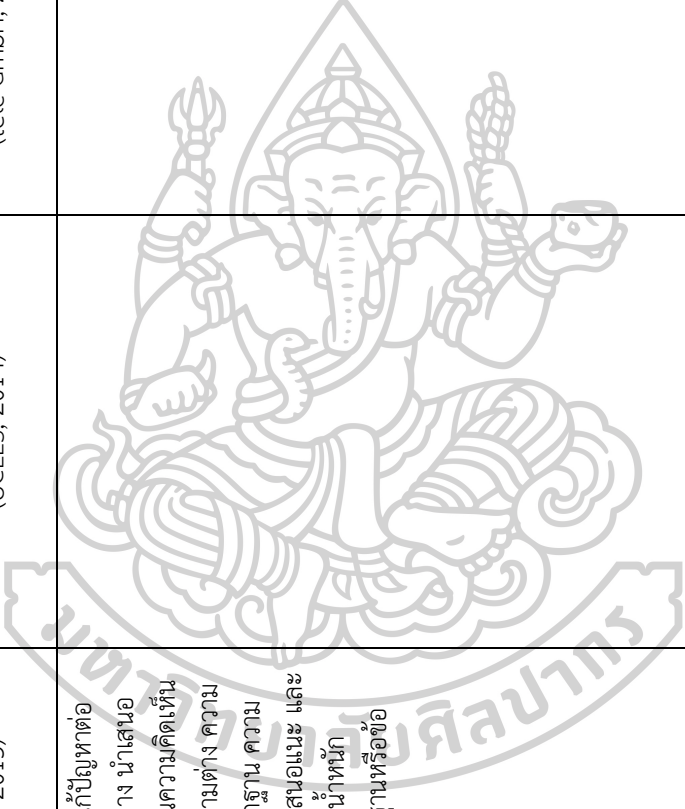
ตารางที่ 4 การสังเคราะห์สาระแบบทดสอบวัดทักษะการเขียน (ต่อ)

แหล่งข้อมูล (Sources) เนื้อหาแบบทดสอบ (Test Content) เกณฑ์การตรวจให้คะแนน (Scoring Criteria)	CET (Zheng & Cheng, 2008)	MUET (Malaysian Examinations Council, 2006)	TOEFL iBT (Educational Testing Service, 2010)
	<p>1) ความถูกต้อง (Accuracy) การใช้สะกดคำและเครื่องหมายวรรคตอนต่าง ๆ ที่ถูกต้อง การใช้ไวยากรณ์ที่ถูกต้อง และการใช้โครงสร้างประโยคที่ถูกต้อง</p> <p>2) ความเหมาะสม (Appropriacy) การใช้คำศัพท์และสำนวนภาษาที่หลากหลาย</p> <p>3) ความสัมพันธ์และความต่อเนื่อง (Coherence and Cohesion) การพัฒนาและจัดระบบความคิด การใช้ตัวเน้นจังหวะ (Markers) และคำเชื่อม (Linking Devices) ที่เหมาะสม</p> <p>4) การใช้หน้าที่ภาษา (Use of Language Functions) การนิยาม การบรรยาย การอธิบาย, การเปรียบเทียบ ความเหมือน ความต่าง, การแบ่งประเภท, การให้เหตุผล, การแสดงความคิดเห็น, การแสดงความสัมพันธ์ เป็นต้น</p> <p>5) ความสมบูรณ์ของภาระงาน (Task Fulfillment) การนำเสนอข้อมูลที่เป็นประโยชน์ (Presenting Relevant Ideas) การนำเสนอสาระที่เพียงพอ (Providing Adequate Content) การแสดงการแก้ปัญหาในหัวข้อที่สมบูรณ์ (Showing a mature treatment of topic)</p>	<p>1) ความถูกต้อง (Accuracy) การใช้สะกดคำและเครื่องหมายวรรคตอนต่าง ๆ ที่ถูกต้อง การใช้ไวยากรณ์ที่ถูกต้อง และการใช้โครงสร้างประโยคที่ถูกต้อง</p> <p>2) ความเหมาะสม (Appropriacy) การใช้คำศัพท์และสำนวนภาษาที่หลากหลาย</p> <p>3) ความสัมพันธ์และความต่อเนื่อง (Coherence and Cohesion) การพัฒนาและจัดระบบความคิด การใช้ตัวเน้นจังหวะ (Markers) และคำเชื่อม (Linking Devices) ที่เหมาะสม</p> <p>4) การใช้หน้าที่ภาษา (Use of Language Functions) การนิยาม การบรรยาย การอธิบาย, การเปรียบเทียบ ความเหมือน ความต่าง, การแบ่งประเภท, การให้เหตุผล, การแสดงความคิดเห็น, การแสดงความสัมพันธ์ เป็นต้น</p> <p>5) ความสมบูรณ์ของภาระงาน (Task Fulfillment) การนำเสนอข้อมูลที่เป็นประโยชน์ (Presenting Relevant Ideas) การนำเสนอสาระที่เพียงพอ (Providing Adequate Content) การแสดงการแก้ปัญหาในหัวข้อที่สมบูรณ์ (Showing a mature treatment of topic)</p>	<p>งานเขียนบูรณาการให้คะแนนตามคุณภาพของการเขียนตามเกณฑ์ดังนี้</p> <p>1) การจัดระบบความคิด Organization</p> <p>2) การใช้ไวยากรณ์และคำศัพท์ที่เหมาะสม และถูกต้อง Precise Use of Grammar, and Vocabulary)</p> <p>3) ความสมบูรณ์และถูกต้องของเนื้อหาสาระที่เขียน (The Completeness and Accuracy of the Content)</p> <p>งานเขียนเรียงความอิสระให้คะแนนตามคุณภาพงานเขียนโดยองค์รวมตามประเด็นต่อไปนี้</p> <p>1) การพัฒนาความคิด (Development)</p> <p>2) จัดระบบความคิด (Organization)</p> <p>3) การใช้ไวยากรณ์และคำศัพท์ที่เหมาะสม และถูกต้อง (Precise use of Grammar, and Vocabulary)</p>

ตารางที่ 4 การสังเคราะห์สาระแบบทดสอบวัดทักษะการเขียน (ต่อ)

แหล่งข้อมูล (Sources) เนื้อหาแบบทดสอบ (Test Content)	IELTS (IELTS, 2015)	FCE (UCLES, 2014)	telc B1-B2 English (telc GmbH, 2012)	สรุปสาระที่จะนำไปสร้างแบบทดสอบ
จุดประสงค์ (Purpose) สาระภาษา (Language Knowledge/Subskills)	ทดสอบความสามารถในการใช้ภาษาเขียนของผู้เข้าสอบ 1) เขียนเล่าเรื่องจากการดูรูปภาพได้อย่างมีระบบ 2) ใช้ภาษาที่เหมาะสมกับหัวข้อเรื่องและรูปแบบ จัดระบบนำเสนอ และอาจจะเปรียบเทียบข้อมูล บรรยายขั้นตอนของกระบวนการหรือวิธีการ บรรยายวัตถุประสงค์ การณ์หรือลำดับของเหตุการณ์ และอธิบายวิธีการทำงานของบางสิ่งอย่าง 3) เขียนเชิงโต้แย้งโดยจัดระบบข้อมูลได้อย่างดีชัดเจน มีประโยชน์ แสดงหลักฐานหรือตัวอย่างเพื่อสนับสนุนความคิดและใช้ภาษาได้อย่างถูกต้อง	วัดความสามารถของผู้เข้าสอบในการเขียนบทความภาษาอังกฤษชนิดต่าง ๆ 1) รู้หน้าที่ภาษา (Language functions) เช่น การประเมิน (evaluating) การแสดงความคิดเห็น (expressing opinions) การตั้งสมมติฐาน (hypothesizing) การปรับเปลี่ยน (justifying) และการชักชวน (persuading) 2) เขียนตัวบทชนิดต่าง ๆ ที่กำหนดไว้ในแบบทดสอบ	วัดความสามารถในการสื่อสารด้วยการเขียนของผู้เข้าสอบ ความสามารถในการสื่อสารด้วยการเขียน	วัดความสามารถของผู้เข้าสอบในการสื่อสารโดยใช้ภาษาเขียน 1) ทำภาระงานการเขียนในรูปแบบต่าง ๆ ตามข้อมูลที่กำหนดให้ เช่น แปลประโยค เขียนเรียงความ เขียนเล่าเรื่องเขียนเชิงโต้แย้ง 2) ใช้การสะกดคำและเครื่องหมายที่ถูกต้อง 3) ใช้คำบอกลำดับเหตุการณ์และคำเชื่อมความที่เหมาะสม 4) เลือกใช้คำแทนคำที่ต้องกรากกล่าวซ้ำได้เหมาะสม 5) ใช้ภาษาถูกต้องตามหลักไวยากรณ์ 6) ใช้ภาษาเหมาะสมกับจุดประสงค์ หัวเรื่อง สถานการณ์เฉพาะ ชนิดของเนื้อหา และผู้รับสาร 7) ใช้คำศัพท์และสำนวนภาษาที่หลากหลาย 8) ใช้โครงสร้างประโยคหลากหลายถูกต้อง และชัดเจน

ตารางที่ 4 การสังเคราะห์สาระแบบทดสอบวัดทักษะการเขียน (ต่อ)

แหล่งข้อมูล (Sources) เนื้อหาแบบทดสอบ (Test Content)	IELTS (IELTS, 2015)	FCE (UCLES, 2014)	telc B1-B2 English (telc GmbH, 2012)	สรุปสาระ ที่จะนำไปสร้างแบบทดสอบ
	<p>4) นำเสนอวิธีแก้ปัญหาต่อปัญหาแต่ละอย่าง นำเสนอและปรับเปลี่ยนความคิดเห็นเปรียบเทียบความต่าง ความเหมือนของหลักฐาน ความคิดเห็นและข้อเสนอแนะ และประเมินและชี้ให้เห็นที่ความคิด หลักฐานหรือข้อโต้แย้ง</p>			<p>9) พัฒนาและจัดระบบข้อมูล ความคิด การนำเสนอได้อย่างดี พร้อมทั้งแสดงข้อมูล หลักฐาน ตัวอย่างเพื่อสนับสนุนความคิดอย่างเพียงพอ และแสดงการแก้ปัญหาได้ตามที่ภาระงานกำหนด</p> <p>10) ใช้ภาษาในหน้าที่ต่าง ๆ ได้ถูกต้องเหมาะสม เช่น นิยาม อธิบาย บรรยาย แบ่งประเภท โน้มน้าว เชิญชวน นำเสนอ เปรียบเทียบ และระบุความต่าง แสดงความคิดเห็น และปรับเปลี่ยนความคิดเห็น แสดงความสัมพันธ์ ค้นหาคำชัดเจนน สอดถามและให้ข้อมูล ให้เหตุผล อ้างอิง ให้ข้อเสนอแนะและคำแนะนำ แสดงการเห็นด้วยและไม่เห็นด้วย พยายาม ตั้งสมมติฐาน ปรับเปลี่ยนสมมติฐาน ประเมิน และสรุปผล</p>

ตารางที่ 4 การสังเคราะห์สาระแบบทดสอบวัดทักษะการเขียน (ต่อ)

แหล่งข้อมูล (Sources)	IELTS (IELTS, 2015)	FCE (UCLES, 2014)	telc B1-B2 English (telc GmbH, 2012)	สรุปสาระ ที่จะนำไปสร้างแบบทดสอบ
เนื้อหาแบบทดสอบ (Test Content) ภาระงานเขียน (Writing Tasks)	<p>ภาระงานที่ 1 เขียนบรรยายข้อมูลจากการดูภาพ (Visual Information) โดยใช้คำพูดของตนเอง ความยาวอย่างน้อย 150 คำ ภายในเวลา 20 นาที</p> <p>ตัวกระตุ้น กราฟ ตาราง แผนภาพ หรือไดอะแกรม</p> <p>ผลิต (Output Text) การเขียนบรรยายข้อมูลจาก การดูภาพ (Visual Information) ความยาวอย่างน้อย 150 คำ</p> <p>ความสามารถที่ต้องแสดง ความสามารถในการเขียนเล่าเรื่องในภาพรวมได้อย่างมีระบบจากการดูรูปภาพ</p>	<p>ภาระงานที่ 1 เขียนเรียงความโดยแสดงความคิดเห็นเกี่ยวกับหัวข้อที่ได้รับโดยใช้ความคิดที่ กำหนดให้ และต้องเพิ่มเติมความคิดที่ 3 ที่เป็นความคิดเห็นของตนเองที่สัมพันธ์กับหัวเรื่องด้วย ความยาวของเรียงความควรอยู่ระหว่าง 140 – 190 คำ</p> <p>หัวเรื่อง หัวข้อที่กำหนดให้อยู่ในใจทั่วไป ผู้เข้าสอบไม่จำเป็นต้องมีความรู้พิเศษหรือเฉพาะด้าน</p> <p>ตัวกระตุ้น หัวข้อเรียงความ และความคิด 2 อย่างที่สัมพันธ์กับหัวข้อเรื่องอย่างชัดเจน</p> <p>ความสามารถที่ต้องแสดง - การประเมิน (Evaluating) - การแสดงความคิดเห็น (Expressing Opinions) - การตั้งสมมติฐาน (Hypothesizing) - การปรับเบี่ยน (Justifying)</p>	<p>ภาระงาน เขียนอีเมลถึงทางการ (Writing Semi-Formal Emails) โดยเลือก 1 ใน 2 หัวข้อที่กำหนดให้ มีจุดประสงค์เพื่อวัดความสามารถในการสื่อสารด้วยการเขียนของผู้เข้าสอบ</p> <p>ตัวกระตุ้น (Input Text) คำอธิบายที่เป็นสถานการณ์สั้น ๆ และประเด็นแนะนำ 3 อย่าง</p> <p>ผลิต (Output Text) อีเมลอย่างเป็นทางการสั้น ๆ ในบริบททั่วไป ส่วนตัว หรือที่เกี่ยวข้องกับงาน</p>	<p>1) การเขียนจดหมายส่วนตัว ภาระงาน: เขียนจดหมายส่วนตัวตามสถานการณ์ที่กำหนดให้ เวลา: 30 นาที</p> <p>ตัวกระตุ้น: สถานการณ์ที่กำหนดให้</p> <p>หัวเรื่อง: ข้อมูลทั่วไป ส่วนตัว และหัวข้อที่คุ้นเคย</p> <p>ผลิต (Output Text): จดหมายส่วนตัว</p> <p>2) การเขียนเรียงความแสดงความคิดเห็น ภาระงาน: เขียนเรียงความแสดงความคิดเห็นในหัวข้อที่กำหนด ความยาวอย่างน้อย 150 คำ</p> <p>เวลา: 30 นาที</p> <p>ตัวกระตุ้น: สถานการณ์หรือบริบท จุดประสงค์การเขียน และผู้อ่านเป้าหมาย</p> <p>หัวเรื่อง: หัวเรื่องที่คุ้นเคย พบเจอได้ใน</p>

ตารางที่ 4 การสังเคราะห์สาระแบบทดสอบวัดทักษะการเขียน (ต่อ)

แหล่งข้อมูล (Sources) เนื้อหาแบบทดสอบ (Test Content)	IELTS (IELTS, 2015)	FCE (UCLES, 2014)	telc B1-B2 English (telc GmbH, 2012)	สรุปสาระ ที่จะนำไปสร้างแบบทดสอบ
<p>ภาชนะที่ 2</p> <p>ผู้เข้าสอบเขียนอภิปราย พร้อมแสดงหลักการและเหตุผลสนับสนุนความคิดเห็นของตนเอง ความยาวอย่างน้อย 250 คำ เวลาประมาณ 40 นาที</p> <p>ตัวกระตุ้น</p> <p>ความคิดเห็น ข้อโต้แย้ง หรือปัญหาที่ผู้เข้าสอบจะใช้ในการอภิปราย</p> <p>ผลผลิต (Output Text)</p> <p>การเขียนเชิงอภิปราย</p> <p>ความสามารถที่ต้องแสดง</p> <p>ความสามารถเขียนเชิงโต้แย้งที่</p> <p>จัดระบบอย่างดี ชัดเจน มีประโยชน์ แสดงหลักฐานหรือตัวอย่างเพื่อสนับสนุนความคิด และใช้ภาษาได้อย่างถูกต้อง</p>	<p>- การชักชวน (Persuading)</p> <p>ผลผลิต (Output Text)</p> <p>Articles, Email, Essay, Letter, Report, Review</p> <p>ภาชนะที่ 2</p> <p>ผู้เข้าสอบเขียนเรียงความโดยการเลือกสถานการณ์ 1 ใน 3 ตัวเลือก ความยาวของเรียงความควรอยู่ระหว่าง 140 – 190 คำ</p> <p>ตัวกระตุ้น</p> <p>ผู้เข้าสอบจะได้รับบริบท จุดประสงค์ของทั้งเรื่อง และผู้อ่านเป้าหมาย</p> <p>ความสามารถที่ต้องแสดง</p> <p>การเขียนด้วยบทชนิดต่าง ๆ ที่กำหนดไว้ในแบบทดสอบ</p> <p>ผลผลิต (Output Text)</p> <p>Articles, Email, Essay, Letter, Report, Review</p>		<p>ชีวิตประจำวัน</p> <p>ผลผลิต (Output Text):</p> <p>เรียงความแสดงความคิดเห็น</p> <p>3) การเขียนที่เกี่ยวข้องกับงาน</p> <p>ภาชนะงาน: เขียนเอกสารหรือจดหมายที่เกี่ยวข้องกับงาน เช่น บันทึกข้อความ จดหมายสมัครงาน จดหมายส่งข้อ</p> <p>จดหมายร้องเรียน เป็นต้น เพื่อสื่อสารในหน่วยงาน ตามสถานการณ์ที่กำหนด</p> <p>เวลา: 40 นาที</p> <p>ตัวกระตุ้น:</p> <p>สถานการณ์หรือบริบท จุดประสงค์การเขียน และผู้อ่านเป้าหมาย</p> <p>หัวเรื่อง:</p> <p>หัวเรื่องที่เกี่ยวข้องกับงาน</p> <p>ผลผลิต (Output Text):</p> <p>บันทึกข้อความ จดหมายสมัครงาน จดหมายส่งข้อ จดหมายร้องเรียน เป็นต้น</p>	

ตารางที่ 4 การสังเคราะห์สาระแบบทดสอบวัดทักษะการเขียน (ต่อ)

แหล่งข้อมูล (Sources) เนื้อหาแบบทดสอบ (Test Content)	IELTS (IELTS, 2015)	FCE (UCLES, 2014)	telc B1-B2 English (telc GmbH, 2012)	สรุปสาระ ที่จะนำไปสร้างแบบทดสอบ
เกณฑ์การตรวจให้คะแนน (Scoring Criteria)	<p>1) การบรรลุภาระงาน/การตอบโต้ตรงประเด็น (Task Achievement/ Response)</p> <p>2) ความสัมพันธ์เชื่อมโยงของเนื้อหา (Coherence and Cohesion)</p> <p>3) การใช้คำศัพท์ที่เหมาะสม (Lexical Resource)</p> <p>4) ไวยากรณ์ที่ใช้และความถูกต้อง (Grammatical Range and Accuracy)</p>	<p>1) Content เน้นว่าผู้เข้าสอบได้ผลิตภาระงานได้สมบูรณ์เพียงใด หรือดูว่าผู้เข้าสอบได้ทำในสิ่งที่กำหนดให้ทำหรือไม่</p> <p>2) Communicative Achievement จะตรวจสอบว่างานเขียนเหมาะสมกับภาระงานหรือไม่ (ตัวอย่างเช่น รูปแบบการเขียนตรงกับการเขียนบทความในนิตยสารหรือไม่) และเลือกใช้ภาษาได้ถูกต้องหรือไม่ เช่น เป็นทางการหรือไม่เป็นทางการ</p> <p>3) การจัดระบบเนื้อความ เน้นในเรื่องการจัดเรียงประโยค หรือดูว่าสมเหตุสมผลและตามลำดับหรือไม่</p> <p>4) ภาษา เน้นที่คำศัพท์และไวยากรณ์ รวมถึงความหลากหลายของภาษา เช่นเดียวกับความถูกต้องด้วย</p>	<p>1) Task Management ซึ่งจะดูว่าผู้เข้าสอบได้กล่าวถึงประเด็นที่กำหนดให้ครบถ้วนหรือไม่</p> <p>2) Communicative Design จะดูว่าผู้เข้าสอบได้ใช้ภาษาเป็นทางการหรือไม่เป็นทางการ เหมาะสมหรือไม่ ได้เชื่อมโยงความคิดเพื่อให้ได้น้อยความที่มีโครงสร้างดี ชัดเจน หรือไม่</p> <p>3) ด้าน Accuracy จะดูว่าไวยากรณ์และการสะกดคำ ถูกต้องโดยส่วนใหญ่หรือไม่</p> <p>4) ด้าน Vocabulary จะดูว่าใช้คำศัพท์ที่หลากหลายเพื่อสื่อสารของคุณหรือไม่</p>	<p>1) ความสมบูรณ์ของภาระงาน (Task Achievement) ภาระงานสำเร็จตามข้อกำหนด ตอบโจทย์ทุกประเด็นครบถ้วน</p> <p>2) การจัดระบบความคิดและการนำเสนอความคิด (Organization) ได้แก่ การพัฒนาและจัดระบบความคิด ใช้คำบอกเครื่องหมายและคำเชื่อมที่เหมาะสม และใช้คำแทนคำซ้ำที่เหมาะสมพร้อมทั้งคำเชื่อมตัวอื่น ๆ</p> <p>3) ความถูกต้องและเหมาะสมของการใช้ภาษา (Language Use) ใช้คำศัพท์ไวยากรณ์หลากหลาย ถูกต้อง และเหมาะสมกับจุดประสงค์การเขียนและผู้อ่านเป้าหมาย</p>

ผลจากการสังเคราะห์สาระแบบทดสอบวัดทักษะการเขียนในตารางที่ 4 สรุปได้ว่าแบบทดสอบด้านการเขียนควรมีกรอบสาระด้าน 1) จุดประสงค์ (Purpose) 2) สาระภาษา (Language Knowledge/Subskills) 3) ชนิด 121 ภาระงาน (Task Type) และ 4) การตรวจให้คะแนน (Scoring Technique) สำหรับการทดสอบทักษะการเขียน ดังรายละเอียดต่อไปนี้

1) จุดประสงค์ของแบบทดสอบ (Purpose) คือ วัดความสามารถของผู้เข้าสอบในการสื่อสารด้วยการเขียน

2) สาระภาษา (Language Knowledge/Subskills) ที่เกี่ยวข้อง คือ 1) ทำภาระงานการเขียนในรูปแบบต่าง ๆ ตามข้อมูลที่กำหนดให้ เช่น แพลประโยค เขียนเรียงความ เขียนเล่าเรื่อง เขียนเชิงโต้แย้ง 2) ใช้การสะกดคำและเครื่องหมายที่ถูกต้อง 3) ใช้คำบอกลำดับเหตุการณ์และคำเชื่อมความได้ถูกต้อง 4) เลือกใช้คำแทนคำที่ต้องต้องการกล่าวซ้ำได้ถูกต้อง 5) ใช้ภาษาถูกต้องตามหลักไวยากรณ์ 6) ใช้ภาษาเหมาะสมกับจุดประสงค์ หัวเรื่อง สถานการณ์เฉพาะ ชนิดของเนื้อความ และผู้รับสาร 7) ใช้คำศัพท์และสำนวนภาษาที่หลากหลาย 8) ใช้โครงสร้างประโยคหลากหลาย ถูกต้อง และชัดเจน 9) พัฒนาและจัดระบบข้อมูล ความคิด การนำเสนอได้อย่างดี พร้อมทั้งแสดงข้อมูลหลักฐาน ตัวอย่างเพื่อสนับสนุนความคิดอย่างเพียงพอ และแสดงการแก้ปัญหาได้ตามที่ภาระงานกำหนด และ 10) ใช้ภาษาในหน้าที่ต่าง ๆ ได้ถูกต้อง เหมาะสม เช่น นิยาม อธิบาย บรรยาย แบ่งประเภท โน้มน้าว เชิญชวน นำเสนอ เปรียบเทียบ และระบุความต่าง แสดงความคิดเห็น และปรับเปลี่ยนความคิดเห็น แสดงความสัมพันธ์ ค้นหาความชัดเจน สอบถามและให้ข้อมูล ให้เหตุผล อ้างอิง ให้ข้อเสนอแนะและคำแนะนำ แสดงการเห็นด้วยและไม่เห็นด้วย พยากรณ์ ตั้งสมมติฐานปรับเปลี่ยนสมมติฐาน ประเมิน และสรุปผล

3) ลักษณะภาระงาน (Task Characteristics) ประกอบด้วย การเขียนจดหมายส่วนตัว การเขียนเรียงความแสดงความคิดเห็น การเขียนที่เกี่ยวข้องกับงาน ดังรายละเอียดต่อไปนี้

การเขียนจดหมายส่วนตัว มีภาระงาน คือ เขียนจดหมายส่วนตัวตามสถานการณ์ที่กำหนดให้ กำหนดเวลา 30 นาที มีตัวกระตุ้น คือ สถานการณ์ที่กำหนดให้ หัวเรื่องที่เกี่ยวข้อง คือ ข้อมูลทั่วไปส่วนตัว และหัวข้อที่คุ้นเคย และผลผลิต (Output Text) คือ จดหมายส่วนตัว

การเขียนเรียงความแสดงความคิดเห็น มีภาระงาน คือ เขียนเรียงความแสดงความคิดเห็นในหัวข้อที่กำหนด ความยาวอย่างน้อย 150 คำ กำหนดเวลา 30 นาที มีตัวกระตุ้น คือ สถานการณ์หรือบริบท จุดประสงค์การเขียน และผู้อ่านเป้าหมาย หัวเรื่องที่เกี่ยวข้อง คือ เรื่องที่คุ้นเคย พบเจอได้ในชีวิตประจำวัน และ ผลผลิต (Output Text) คือ เรียงความแสดงความคิดเห็น

การเขียนที่เกี่ยวข้องกับงาน มีภาระงาน: เขียนเอกสารหรือจดหมายที่เกี่ยวข้องกับงาน เช่น บันทึกข้อความ จดหมายสมัครงาน จดหมายสั่งซื้อ จดหมายร้องเรียน เป็นต้น เพื่อสื่อสารในหน่วยงาน ตามสถานการณ์ที่กำหนด กำหนดเวลา 40 นาที มีตัวกระตุ้น คือ สถานการณ์หรือบริบท จุดประสงค์การเขียน และผู้อ่านเป้าหมาย หัวเรื่องที่เกี่ยวข้อง คือ เรื่องที่เกี่ยวข้องกับงาน และ ผลผลิต (Output Text) คือ บันทึกข้อความ จดหมายสมัครงาน จดหมายสั่งซื้อ จดหมายร้องเรียน เป็นต้น

4) เกณฑ์การให้คะแนนภาระงานเขียน (Scoring Technique) ควรมีองค์ประกอบด้วย 3 ด้าน คือ 1) ความสมบูรณ์ของภาระงาน (Task Achievement) ภาระงานสำเร็จตามข้อกำหนด ตอบโจทย์ทุกประเด็น ครบถ้วน 2) การจัดระบบความคิดและการนำเสนอความคิด (Organization) 3)

ความถูกต้องและเหมาะสมของการใช้ภาษา (Language Use) ใช้คำศัพท์ ไวยากรณ์หลากหลาย ถูกต้อง และเหมาะสมกับจุดประสงค์การเขียนและผู้อ่านเป้าหมาย

สรุปสาระสำหรับการทดสอบทักษะการใช้ภาษา

สาระสำหรับการทดสอบทักษะการใช้ภาษาได้มาจากการสังเคราะห์ข้อมูลดังกล่าวข้างต้น เพื่อให้ได้ 1) จุดประสงค์ (Purpose) 2) สาระภาษา (Language Knowledge/Subskills) 3) หัวเรื่อง (Topical Knowledge) 4) ตัวกระตุ้น (Input Text) 5) รูปแบบการทดสอบ (Test Type) และ 6) การตรวจให้คะแนน (Scoring Technique) สำหรับการทดสอบทักษะการใช้ภาษา ดังแสดงในตารางสังเคราะห์ต่อไปนี้



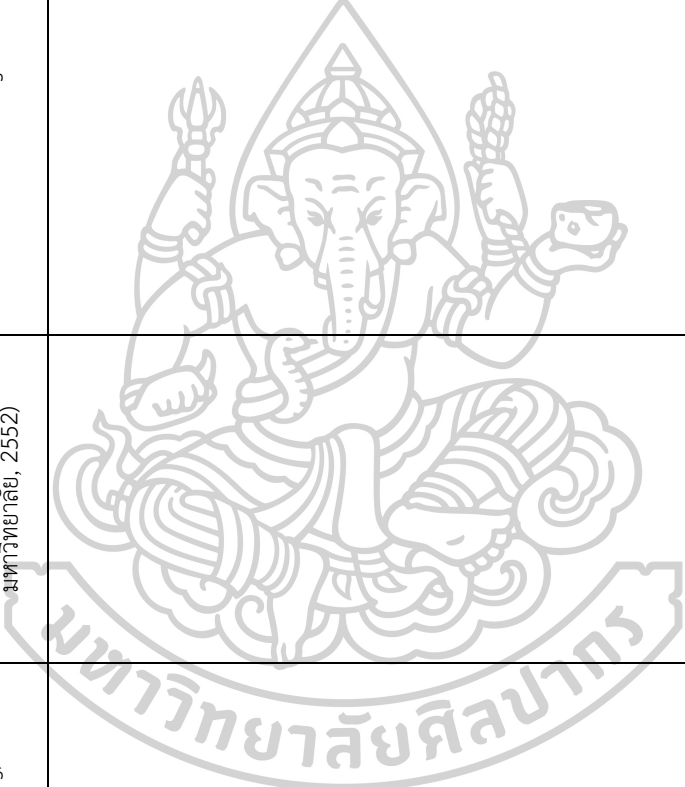
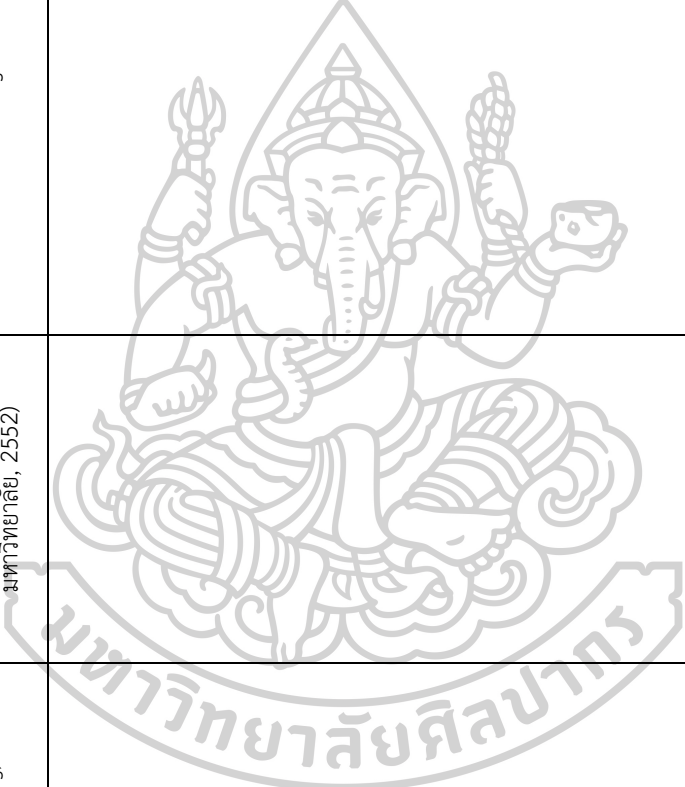
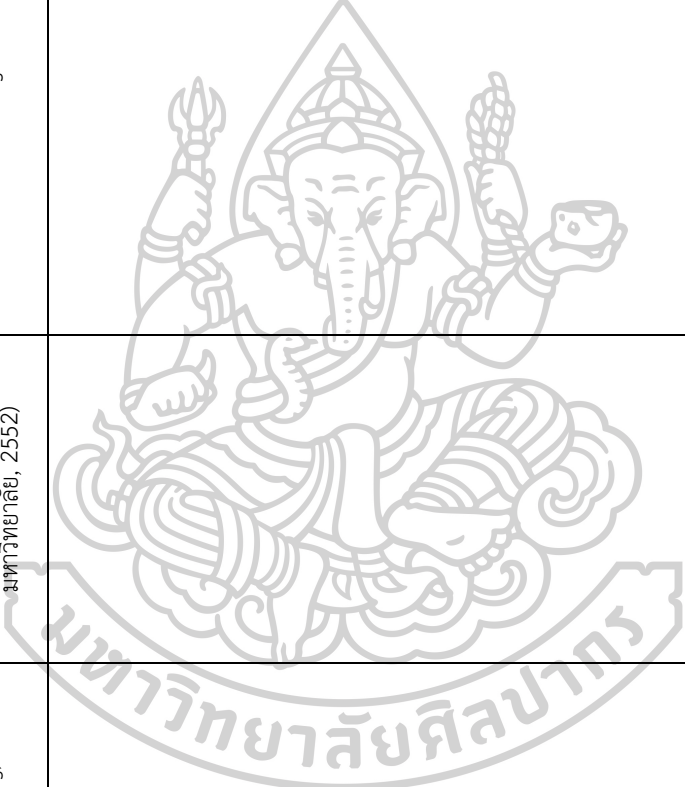
ตารางที่ 5 การสังเคราะห์สาระแบบทดสอบวัดทักษะการใช้ภาษา

แหล่งข้อมูล (Sources)	หลักสูตร (มหาวิทยาลัยทักษิณ, 2554)	CEFR (Council of Europe, 2003)
สาระแบบทดสอบ (Test Content)	วัดความสามารถในการใช้คำศัพท์และไวยากรณ์ของผู้เข้าสอบ	วัดความสามารถของผู้เข้าสอบในการใช้คำศัพท์และโครงสร้างไวยากรณ์ในการสื่อสารได้อย่างถูกต้องและเหมาะสม
จุดประสงค์ (Purpose)	วัดความสามารถในการใช้คำศัพท์และไวยากรณ์ได้	1) มีคำศัพท์เพียงพอที่จะสื่อสารได้เข้าใจในหัวข้อต่าง ๆ ที่สัมพันธ์กับชีวิตประจำวัน เช่น ครอบครัว งาน อุดมคติ ความสนใจ งาน การเดินทาง และเหตุการณ์ปัจจุบัน; 2) เปลี่ยนรูปของคำศัพท์เพื่อหลีกเลี่ยงการใช้คำซ้ำ; 3) ความถูกต้องของคำศัพท์ส่วนใหญ่อยู่ในระดับสูง; 4) สื่อสารได้ถูกต้องอย่างสมเหตุสมผลในบริบทที่คุ้นเคยโดยทั่วไป มีการควบคุมการใช้ภาษาได้ดี มีอิทธิพลของการใช้ภาษาแม่และข้อผิดพลาดที่สังเกตเห็นได้บ้าง แต่สามารถเข้าใจได้ว่าผู้พูดกำลังสื่อสารอะไร ; 5) ใช้ไวยากรณ์ได้สมเหตุสมผล อย่างถูกต้อง ในรูปแบบหรือโครงสร้างที่มักใช้ในชีวิตประจำวัน ที่สัมพันธ์กับสถานการณ์ที่สามารถทำนายได้ ; 6) ควบคุมการใช้ไวยากรณ์ได้ดี มีความผิดพลาดอยู่บ้างหรือเป็นความผิดพลาดที่ไม่เป็นระบบ ;และ 7) แสดงให้เห็นถึงระดับการควบคุมไวยากรณ์ที่ค่อนข้างสูง ไม่สร้างความผิดพลาดที่นำไปสู่การเข้าใจผิด
หัวข้อ (Topical Knowledge)	การคำศัพท์และไวยากรณ์ในการสื่อสาร	1) คำศัพท์เพื่อการสื่อสารในหัวข้อที่สัมพันธ์กับชีวิตประจำวัน เช่น ครอบครัว งานอดิเรก และความสนใจ การเดินทาง และเหตุการณ์ปัจจุบัน ; 2) คำศัพท์เพื่ออภิปรายหรือบรรยายหรือเนื้อหาที่เกี่ยวข้องกับสาขาของตน; 3) คำศัพท์เพื่อการสื่อสารในหัวข้อโดยทั่วไป –B2 และในด้านความถูกต้องตามหลักไวยากรณ์ มีสาระเกี่ยวกับโครงสร้างที่มักใช้ในชีวิตประจำวัน ที่สัมพันธ์กับสถานการณ์ที่สามารถทำนายได้ดีกว่า

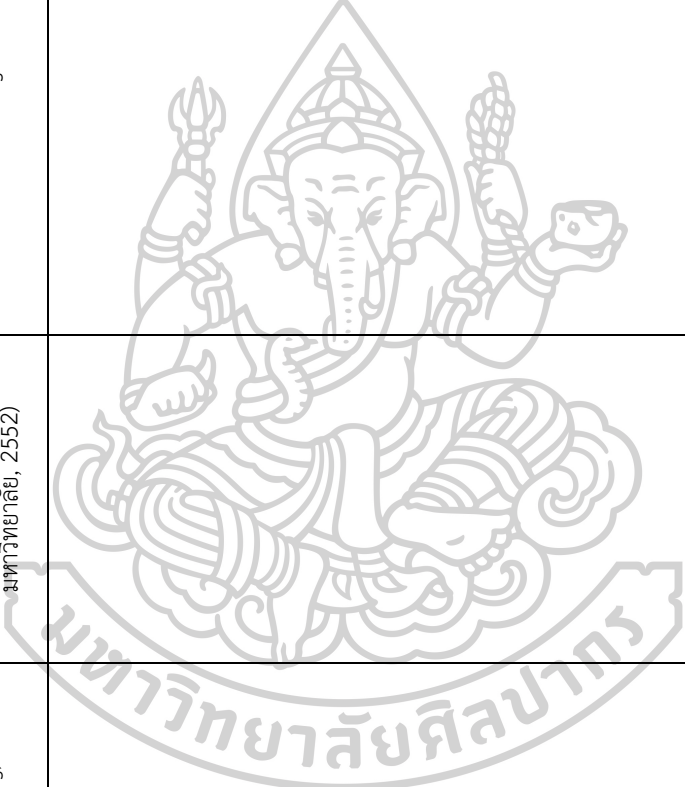
ตารางที่ 5 การสังเคราะห์สาระแบบทดสอบวัดทักษะการใช้ภาษา (ต่อ)

แหล่งข้อมูล (Sources) เนื้อหาแบบทดสอบ (Test Content)	CET (Zheng & Cheng, 2008)	CU-TEP (สถาบันภาษา จุฬาลงกรณ์มหาวิทยาลัย, 2552)	TOEFL iBT (Educational Testing Service, 2010)	TOEFL ITP (Educational Testing Service, 2014)
จุดประสงค์ (Purpose)	ประเมินความสามารถของนักศึกษาในการเข้าใจและตีความความหมายจากระดับคำศัพท์ ประโยค และย่อหน้า โดยใช้แบบทดสอบแบบ Cloze	วัดความสามารถของผู้เข้าสอบทางด้านการใช้ภาษาอังกฤษทั้งในส่วนของไวยากรณ์และคำศัพท์	วัดความสามารถของผู้เข้าสอบในการใช้คำศัพท์และโครงสร้างไวยากรณ์ในการสื่อสารได้อย่างถูกต้องและเหมาะสม	วัดความสามารถของผู้เข้าสอบในความรู้เชิงภาษาและความสามารถในการระบุข้อผิดพลาดในการใช้ภาษา
สำรภาษา (Language Knowledge/Subskills)	เข้าใจและตีความความหมายจากระดับคำศัพท์ ประโยค และย่อหน้า	ระบุจุดบกพร่องของประโยคที่กำหนด	ใช้คำศัพท์และโครงสร้างไวยากรณ์ในการสื่อสารได้อย่างถูกต้องและเหมาะสม	1) Be sure the sentence has a subject and a verb, 2) Be careful of objects of prepositions, 3) Be careful of present participles, 4) Use reduced adjective clauses correctly, 5) Use adverb Time and Cause connectors correctly, 6) Use noun clause connector/subjects correctly, 7) Use other adverb connectors correctly, 8) Invert the subject and verb with place expressions, 9) Be careful of past participles, 10) Use noun clause connectors correctly,

ตารางที่ 5 การสังเคราะห์สาระแบบทดสอบวัดทักษะการใช้ภาษา (ต่อ)

แหล่งข้อมูล (Sources) เนื้อหาแบบทดสอบ (Test Content)	CET (Zheng & Cheng, 2008)	CU-TEP (สถาบันภาษา จุฬาลงกรณ์มหาวิทยาลัย, 2552)	TOEFL iBT (Educational Testing Service, 2010)	TOEFL ITP (Educational Testing Service, 2014)
				11) Invert the subject and verb with conditionals, 12) Use reduced adverb clauses correctly, 13) Invert the subject and verb with negatives, 14) Use adjective clause connectors correctly, 15) Invert the subject and verb with comparisons, 16) After be, use the present participle or the past participle, 17) Form comparatives and superlatives correctly, 18) Make verbs agree after certain words, 19) After Will, Would, or other Modals, use the base form of the verb, 20) Use the correct tense with time expressions, 21) Distinguish subject and object pronouns, 22) Distinguish possessive adjectives and pronouns, 23) Use parallel structure with paired conjunctions, 24) Make verb agree

ตารางที่ 5 การสังเคราะห์สาระแบบทดสอบวัดทักษะการใช้ภาษา (ต่อ)

แหล่งข้อมูล (Sources) เนื้อหาแบบทดสอบ (Test Content)	CET (Zheng & Cheng, 2008)	CU-TEP (สถาบันภาษา จุฬาลงกรณ์มหาวิทยาลัย, 2552)	TOEFL iBT (Educational Testing Service, 2010)	TOEFL ITP (Educational Testing Service, 2014)
				<p>after prepositional phrases, 25) Use comparatives and superlatives correctly, 26) Distinguish A and An, 27) Use the correct tense with Will and Would, 28) Distinguish Like, Alike, Unlike, and Dislike, 29) Make inverted verbs agree, 30) Distinguish countable and uncountable nouns, 31) Distinguish specific and general ideas, 32) Position adjectives and adverbs correctly, 33) Recognize active and passive meanings, 34) Use articles with singular nouns, 35) Use predicate adjectives correctly, 36) Distinguish Make and Do, 37) Use the correct form of the passive, 38) Recognize when preposition have been omitted, 39) Distinguish the person from the thing, and 40) Use adjectives after linking verbs</p>

ตารางที่ 5 การสังเคราะห์สาระแบบทดสอบวัดทักษะการใช้ภาษา (ต่อ)

แหล่งข้อมูล (Sources)	CET (Zheng & Cheng, 2008)	CU-TEP (สถาบันภาษา จุฬาลงกรณ์มหาวิทยาลัย, 2552)	TOEFL iBT (Educational Testing Service, 2010)	TOEFL ITP (Educational Testing Service, 2014)
เนื้อหาแบบทดสอบ (Test Content) หัวเรื่อง (Topical Knowledge)	คำศัพท์ ประโยค และข้อความในระดับย่อหน้า	การใช้ภาษาอังกฤษทั้งในส่วนของไวยากรณ์และคำศัพท์	1) คำศัพท์เพื่อการสื่อสารในหัวข้อที่สัมพันธ์กับชีวิตประจำวัน เช่น ครอบครั้ว งานอดิเรก และความสนใจ งาน การเดินทาง และเหตุการณ์ปัจจุบัน ; 2) คำศัพท์เพื่อการสื่อสารในเรื่องราวหรือเนื้อหาที่เกี่ยวข้องกับสาขาของตน ; และ 3) คำศัพท์เพื่อการสื่อสารในหัวข้อโดยทั่วไป -B2 4) โครงสร้างไวยากรณ์ที่มีใช้เป็นประจำ ในชีวิตประจำวัน ที่สัมพันธ์กับสถานการณ์ที่สามารถทำนายได้มากกว่า เช่น จำนวนภาษา คำที่ใช้ร่วมกัน ความหมายในแง่มุมต่าง ๆ ของคำศัพท์ หรือกริยาลี	เนื้อหาเชิงวิชาการ
ตัวกระตุ้น (Input Text)			บทสนทนา จดหมายสั้น ๆ อีเมลส่วนตัว หรือที่เกี่ยวข้องกับงาน	ข้อมูลจะเป็นประโยคสั้น ๆ

ตารางที่ 5 การสังเคราะห์สาระแบบทดสอบวัดทักษะการใช้ภาษา (ต่อ)

แหล่งข้อมูล (Sources)	CET (Zheng & Cheng, 2008)	CU-TEP (สถาบันภาษา จุฬาลงกรณ์มหาวิทยาลัย, 2552)	TOEFL iBT (Educational Testing Service, 2010)	TOEFL ITP (Educational Testing Service, 2014)
เนื้อหาแบบทดสอบ (Test Content)	แบบเติมคำเรื่อง (Cloze) และการรับรู้ถึงข้อผิดพลาดในการใช้ภาษา (Error Recognition) ที่ให้ผู้เข้าสอบเลือกจากตัวเลือก (Multiple-Choice)	ข้อคำถาม 2 แบบ คือ การระบุข้อผิดพลาด (Error Identification) และการเติมประโยคให้สมบูรณ์ (Sentence Completion) โดยให้ผู้เข้าสอบเลือกคำตอบจากตัวเลือก (Multiple-Choice)	แบบเลือกตอบ การระบุข้อผิดพลาดในประโยคแบบจับคู่	แบบเลือกตอบแบบเติมเต็ม 4 ตัวเลือก และมีคำตอบถูกเพียง 1 ตัวเลือก
เกณฑ์การตรวจให้คะแนน (Scoring Criteria)	ตอบถูกได้ 1/ตอบผิดได้ 0	ตอบถูกได้ 1/ตอบผิดได้ 0	ตอบถูกได้ 1/ตอบผิดได้ 0	ตอบถูกได้ 1/ตอบผิดได้ 0

ตารางที่ 5 การสังเคราะห์สาระแบบทดสอบวัดทักษะการใช้ภาษา (ต่อ)

แหล่งข้อมูล (Sources) เนื้อหาแบบทดสอบ (Test Content)	FCE (UCLES, 2014: 2)	telc B1-B2 English (telc GmbH, 2012: 33-34)	สรุปสาระที่จะนำไปสร้างแบบทดสอบ
จุดประสงค์ (Purpose)	กำหนดให้ผู้เข้าสอบได้อ่านข้อความหลากหลายชนิดและทำภาระงานที่เกี่ยวข้องกับไวยากรณ์และคำศัพท์	วัดความเข้าใจของผู้เข้าสอบในการใช้กลุ่มคำที่เหมาะสมในบทสนทนา และในจดหมายหรืออีเมลถึงทางการ	วัดความสามารถของผู้เข้าสอบในการใช้คำศัพท์และโครงสร้างไวยากรณ์ในการสื่อสารได้อย่างถูกต้องและเหมาะสม
สาระภาษา (Language Knowledge/Subskills)	อ่านข้อความหลากหลายชนิดและทำภาระงานที่เกี่ยวข้องกับไวยากรณ์และคำศัพท์ เช่น จำนวน (Idioms) คำที่ใช้ร่วมกัน (Collocations) ความหมายในแง่มุมต่าง ๆ (Shades of Meaning) กริยาวิเศษณ์ (Phrasal Verbs) วลีตายตัว (Fixed Phrases) เป็นต้น	เข้าใจการใช้กลุ่มคำที่เหมาะสมในบทสนทนา และในจดหมายหรืออีเมลถึงทางการ	วัดความสามารถในการใช้คำศัพท์และไวยากรณ์ได้ถูกต้องและเหมาะสมกับบริบทและสถานการณ์
หัวข้อเรื่อง (Topical Knowledge)	การอ่านข้อความหลายชนิดและการทำภาระงานที่เกี่ยวข้องกับไวยากรณ์และคำศัพท์	การใช้กลุ่มคำที่เหมาะสมในบทสนทนา (Appropriate Phrases in a Conversation) และในจดหมายหรืออีเมลถึงทางการ (Appropriate Phrases in a Semi-Formal Letter or email)	1) สำนวนภาษาและคำที่ใช้ร่วมกันที่ใช้สื่อสารในชีวิตประจำวัน หรือเกี่ยวข้องกับงาน 2) ความหมายในแง่มุมต่าง ๆ ของคำศัพท์ หรือกริยาวิเศษณ์ ที่ใช้สื่อสารในชีวิตประจำวันหรือเกี่ยวข้องกับงาน 3) ตัวกระตุ้ม (Input Text): บทสนทนาในการสื่อสารในชีวิตประจำวันและที่ทำงาน จดหมายสั้น ๆ อีเมลส่วนตัวหรือที่เกี่ยวข้องกับงาน

ตารางที่ 5 การสังเคราะห์สาระแบบทดสอบวัดทักษะการใช้ภาษา (ต่อ)

แหล่งข้อมูล (Sources) เนื้อหาแบบทดสอบ (Test Content) รูปแบบการทดสอบ (Test Type)	FCE (UCLES, 2014: 2)	telc B1-B2 English (telc GmbH, 2012: 33-34)	สรุปสาระ ที่จะนำไปสร้างแบบทดสอบ
	เติมคำในเนื้อเรื่องโดยเลือกจากตัวเลือก (Multiple-Choice Cloze) เติมคำในช่องว่างของเนื้อเรื่องโดยคิดคำเอง (Open Cloze) การสร้างคำ (Word Formation) และการเปลี่ยนรูปของคำสำคัญ (Key Word Transformations)	แบบจับคู่ (Matching items) และ แบบเลือกตอบ (Multiple-Choice Items) ส่วนลักษณะของเนื้อหาบทอ่าน (Text type) เป็น บทสนทนา (Conversation) และ จดหมายสั้น ๆ หรืออีเมลในบริบททั่ว ๆ ไป หรือที่สัมพันธ์กับที่ทำงาน (Short letters or emails in a public or work-related context)	1) แบบเลือกตอบ 2) การระบุข้อผิดพลาดในประโยค 3) แบบจับคู่
เกณฑ์การตรวจให้คะแนน (Scoring Criteria)	ตอบถูกได้ 1/ตอบผิดได้ 0	ตอบถูกได้ 1/ตอบผิดได้ 0	ตอบถูกได้ 1/ตอบผิดได้ 0

ผลจากการสังเคราะห์สาระแบบทดสอบวัดทักษะการใช้ภาษาในตารางที่ 5 สรุปได้ว่าแบบทดสอบด้านการใช้ภาษาควรมีกรอบสาระด้าน 1) จุดประสงค์ (Purpose) 2) สาระภาษา (Language Knowledge/Subskills) 3) หัวเรื่อง (Topical Knowledge) 4) รูปแบบการทดสอบ (Test Type) และ 5) การตรวจให้คะแนน (Scoring Technique) ดังรายละเอียดต่อไปนี้

1) จุดประสงค์ของแบบทดสอบ (Purpose) คือ วัดความสามารถของผู้เข้าสอบในการใช้คำศัพท์และโครงสร้างไวยากรณ์เพื่อการสื่อสาร

2) สาระภาษา (Language Knowledge/Subskills) ที่เกี่ยวข้อง คือ ใช้คำศัพท์และไวยากรณ์ได้ถูกต้องและเหมาะสมกับบริบทและสถานการณ์

3) หัวเรื่อง (Topical Knowledge) ที่สามารถนำมาออกแบบแบบทดสอบวัดความสามารถภาษาอังกฤษควรอยู่ในบริบทการใช้ภาษาอังกฤษเพื่อการสื่อสารในชีวิตประจำวันและในสังคมที่ครอบคลุมเรื่องส่วนตัวและการทำงาน ได้แก่ 1) สำนวนภาษาและคำที่ใช้ร่วมกันที่ใช้สื่อสารในชีวิตประจำวันหรือเกี่ยวข้องกับงาน และ 2) ความหมายในแง่มุมต่าง ๆ ของคำศัพท์ หรือกริยาวลี ที่ใช้สื่อสารในชีวิตประจำวันหรือเกี่ยวข้องกับงาน โดยมีตัวกระตุ้น (Input Text) เป็น บทสนทนาในการสื่อสารในชีวิตประจำวันและที่ทำงาน จดหมายสั้น ๆ อีเมลส่วนตัวหรือที่เกี่ยวข้องกับงาน

4) รูปแบบการทดสอบ (Test Type) ที่สามารถนำมาใช้ทดสอบทักษะการใช้ภาษาของแบบทดสอบวัดความสามารถภาษาอังกฤษควรประกอบด้วย 1) แบบเลือกตอบ 2) การระบุข้อผิดพลาดในประโยคโดยมีตัวเลือก และ 3) แบบจับคู่

5) การตรวจให้คะแนน (Scoring Technique) มีรูปแบบการตอบถูกได้ 1 หรือ ตอบผิด ได้

0



2.2 การพัฒนาแบบทดสอบวัดความสามารถภาษาอังกฤษให้เป็นแบบทดสอบมาตรฐาน

การพัฒนาแบบทดสอบวัดความสามารถภาษาอังกฤษให้เป็นแบบทดสอบมาตรฐาน นักพัฒนาแบบทดสอบจำเป็นต้องมีความรู้ที่เกี่ยวข้องใน 2 หัวข้อ คือ 1) องค์ประกอบที่เกี่ยวข้องกับการพัฒนาแบบทดสอบ และ 2) การประเมินคุณภาพแบบทดสอบ ซึ่งรายละเอียดใน 2 ส่วนนี้ ถือว่ามีความสำคัญและเป็นพื้นฐานของการพัฒนาแบบทดสอบทุกชนิด นอกจากนี้ในแต่ละส่วนยังมีกิจกรรมย่อยที่นักพัฒนาแบบทดสอบจำเป็นต้องมีความรู้และสามารถนำไปปฏิบัติได้ ดังรายละเอียดต่อไปนี้

2.2.1 องค์ประกอบที่เกี่ยวข้องกับการพัฒนาแบบทดสอบ

ในการการพัฒนาแบบทดสอบวัดความสามารถภาษาอังกฤษให้เป็นแบบทดสอบมาตรฐาน นักพัฒนาแบบทดสอบจำเป็นต้องมีความรู้ในองค์ประกอบสำคัญที่จะนำมาใช้ในการพัฒนาแบบทดสอบ องค์ประกอบเหล่านั้น ได้แก่ 1) ขั้นตอนในการพัฒนาแบบทดสอบ 3) ตารางกำหนดลักษณะแบบทดสอบ และ 3) หลักการเขียนข้อคำถามในแบบทดสอบภาษาให้มีคุณภาพดี โดยแต่ละองค์ประกอบมีรายละเอียดดังนี้

2.2.1.1 ขั้นตอนในการพัฒนาแบบทดสอบ

ในการพัฒนาแบบทดสอบวัดความสามารถภาษาอังกฤษให้เป็นแบบทดสอบมาตรฐานและมีคุณภาพเป็นที่ยอมรับ นักพัฒนาแบบทดสอบควรปฏิบัติตามขั้นตอนในการพัฒนาแบบทดสอบมาตรฐานที่นักวิชาการด้านการทดสอบภาษาอังกฤษได้นำเสนอไว้ ซึ่งอาจจะมีจำนวนขั้นตอนที่แตกต่างกันบ้าง แต่มีจุดมุ่งหมายเดียวกันคือ การให้ได้มาซึ่งแบบทดสอบที่ได้มาตรฐานและมีคุณภาพ ตัวอย่างเช่น (Bachman & Lyle F. & Palmer, 1996) นำเสนอการพัฒนาแบบทดสอบภาษาอังกฤษด้วยขั้นตอนหลัก 3 ขั้นตอน (Genesee & Fred & Upshur, 1996) และ (Bachman & Lyle F, 2004) นำเสนอการพัฒนาแบบทดสอบมาตรฐานที่มี 5 และ 6 ขั้นตอน ตามลำดับ ขั้นตอนการพัฒนาแบบทดสอบโดยนักวิชาการทั้ง 3 กลุ่มนี้ มีรายละเอียดดังต่อไปนี้

ขั้นตอนการพัฒนาแบบทดสอบของ (Bachman & Lyle F. & Palmer, 1996)

Bachman and Lyle F. & Palmer (1996) ได้นำเสนอขั้นตอนในการพัฒนาแบบทดสอบ 3 ขั้นตอนประกอบด้วย 1) ขั้นตอนออกแบบ (Design Stage) 2) ขั้นตอนปฏิบัติการ (Operationalization Stage) และ 3) ขั้นตอนบริหารการสอบ (Test Administration) ซึ่งมีรายละเอียดในแต่ละขั้น ดังนี้

1) ขั้นตอนการออกแบบ (Design Stage): ในขั้นการออกแบบ นักพัฒนาแบบทดสอบควรอธิบายส่วนประกอบต่าง ๆ ในการออกแบบแบบทดสอบอย่างละเอียด เพื่อให้มั่นใจว่าผลของการได้ทำภาระงานต่าง ๆ ในแบบทดสอบมีสัมพันธ์อย่างใกล้ชิดเท่าที่จะเป็นไปได้กับการใช้ภาษาจริงของผู้เข้า

สอบ และยังหมายความรวมถึงว่า คะแนนในการสอบจะมีประโยชน์อย่างสูงสุดตามจุดประสงค์ที่ได้กำหนดไว้ โดยทั่วไปแล้วการออกแบบเป็นกระบวนการที่แสดงให้เห็นถึงลำดับของกิจกรรมต่าง ๆ ที่เกิดขึ้น แต่อาจจะมีกิจกรรมบางอย่างที่อาจเกิดขึ้นซ้ำ ๆ หลายครั้ง เช่น การตรวจสอบคุณภาพด้านประโยชน์ของแบบทดสอบและการจัดสรรทรัพยากรและการจัดการ ที่นักพัฒนาแบบทดสอบอาจจะต้องพิจารณาอย่างละเอียดในหลายขั้นตอนตลอดกระบวนการ เป็นต้น ผลลัพธ์ที่นักพัฒนาแบบทดสอบจะได้รับจากขั้นออกแบบ คือ A Design Statement ซึ่งเป็นเอกสารอย่างหนึ่งที่รวบรวม

- 1) คำบรรยายจุดประสงค์ของแบบทดสอบ
- 2) คำบรรยายปริเขตการใช้ภาษาเป้าหมายและชนิดของภาระงาน
- 3) คำบรรยายลักษณะของผู้ใช้แบบทดสอบหรือผู้ทำแบบทดสอบ
- 4) คำนิยามโครงสร้างภาษาที่จะถูกทดสอบ
- 5) แผนสำหรับการประเมินคุณภาพด้านผลประโยชน์ของแบบทดสอบ และ
- 6) รายการทรัพยากรที่ต้องการและที่มีอยู่แล้ว และแผนสำหรับการจัดสรรและจัดการ เพื่อให้ได้ผลลัพธ์ดังกล่าวข้างต้น นักพัฒนาแบบทดสอบจำเป็นต้องดำเนินกิจกรรมที่เกี่ยวข้อง 6 ชนิด โดยมีรายละเอียดในแต่ละชนิด ดังนี้

การบรรยายจุดประสงค์ของแบบทดสอบ: กิจกรรมนี้เป็นการกล่าวให้ชัดเจนว่าจะนำแบบทดสอบที่จะพัฒนาขึ้นไปใช้เพื่อจุดประสงค์ใด อ้างถึงความสามารถทางภาษาอย่างเจาะจง หรืออ้างถึงสมรรถนะในการใช้ภาษาที่นักพัฒนาแบบทดสอบต้องการ โดยอาศัยพื้นฐานจากคะแนนสอบและการตัดสินใจที่เฉพาะเจาะจงที่เกิดจากการอ้างอิงเหล่านี้ คำกล่าวเกี่ยวกับจุดประสงค์ที่ได้มานำไปสู่การคิดหาแนวทางเพื่อใช้พิจารณาผลกระทบที่มุ่งหวังของการใช้แบบทดสอบอีกด้วย

การระบุและบรรยายภาระงานในปริเขตของการใช้ภาษาเป้าหมาย: กิจกรรมนี้ช่วยทำภาระงานในปริเขตของการใช้ภาษาเป้าหมายให้กระจ่างขึ้น ซึ่งเกี่ยวกับการอ้างอิงความสามารถทางภาษาที่นักพัฒนาแบบทดสอบต้องการจะขยายผลและบรรยายชนิดของภาระงานในปริเขตของการใช้ภาษาเป้าหมายในแง่ของลักษณะที่แตกต่างกันออกไป กิจกรรมนี้ทำให้ได้ชุดของคำบรรยายเกี่ยวกับภาระงานในปริเขตการใช้ภาษาเป้าหมายแต่ละชนิด อันจะเป็นพื้นฐานสำหรับการพัฒนาภาระงานในแบบทดสอบ คำบรรยายเหล่านี้ยังช่วยให้ได้แนวทางสำหรับการพิจารณาความสมจริงที่คาดหวัง (The Potential Authenticity) และการมีปฏิสัมพันธ์ของภาระงานในแบบทดสอบ (Interactiveness of Test Tasks)

การบรรยายลักษณะของผู้ใช้แบบทดสอบหรือผู้ทำแบบทดสอบ: กิจกรรมนี้เป็นการบรรยายธรรมชาติของกลุ่มประชากรที่ใช้แบบทดสอบให้กระจ่างชัด ซึ่งเป็นกลุ่มเป้าหมายที่จะได้ทำแบบทดสอบชุดที่กำลังถูกออกแบบ คำบรรยายที่ได้นี้จะนำไปสู่แนวทางสำหรับการพิจารณาผลกระทบที่คาดหวังจากการใช้แบบทดสอบ

การนิยามโครงสร้างภาษาที่ต้องการจะวัด: กิจกรรมนี้ครอบคลุมการทำให้ธรรมชาติของความสามารถที่นักพัฒนาแบบทดสอบต้องการวัดกระจ่างชัดด้วยการนิยามไว้อย่างคร่าว ๆ ผลลัพธ์ของกิจกรรมนี้ คือ คำนิยามเชิงทฤษฎีของโครงสร้างที่จะถูกทดสอบ (Theoretical Definition of the Construct) อันจะนำไปสู่พื้นฐานสำหรับการพิจารณาและการตรวจสอบความตรงเชิงโครงสร้าง (Construct validity) ของการตีความที่นักพัฒนาแบบทดสอบได้พิจารณาไว้จากคะแนนสอบ คำนิยามเชิงทฤษฎีนี้ยังเป็นพื้นฐานสำหรับการพัฒนาในขั้นปฏิบัติการ (Operationalization Stage) ของภาระงานในแบบทดสอบอีกด้วย ในการทดสอบภาษานั้น คำนิยามเชิงโครงสร้างที่เกี่ยวกับทฤษฎี

(Theoretical Construct Definition) สามารถดัดแปลงมาจากทฤษฎีของความสามารถทางภาษา
โครงสร้างประมวลรายวิชา (A Syllabus Specification) หรือจากทั้งสองอย่าง

การพัฒนาแผนสำหรับการประเมินคุณภาพด้านประโยชน์ของแบบทดสอบ: แผนสำหรับการ
ประเมินคุณประโยชน์ของแบบทดสอบรวมอยู่ในกิจกรรมต่าง ๆ ที่อยู่ในทุกขั้นตอนของ
กระบวนการพัฒนาแบบทดสอบ แผนดังกล่าวนี้จะรวมถึงการพิจารณาความสมดุลที่เหมาะสม
(Appropriate Balance) ของคุณภาพด้านผลประโยชน์เบื้องต้น 6 ประการ และการกำหนดผลที่
ยอมรับได้ขั้นต่ำสำหรับแต่ละประเด็น และแบบตรวจรายการ (Checklist) สำหรับคำถามที่นักพัฒนา
แบบทดสอบจะถามเกี่ยวกับภาระงานในแบบทดสอบแต่ละอย่าง การประเมินคุณประโยชน์ในการ
ทดลองใช้แบบทดสอบ และการใช้แบบทดสอบจริงจะครอบคลุมถึงการเก็บข้อมูลย้อนกลับด้วย การ
เก็บข้อมูลย้อนกลับ (Feedback) เกี่ยวข้องกับข้อมูลหลายอย่าง ทั้งเชิงปริมาณ (Quantitative) เช่น
คะแนนสอบรวมและคะแนนในแต่ละภาระงานย่อย และเชิงคุณภาพ (Qualitative) เช่น การบรรยาย
ของผู้สังเกตการณ์ การรายงานตนเองด้วยคำพูด (Verbal Self-Report) ของนักเรียนเกี่ยวกับ
กระบวนการทำแบบทดสอบ ประการสุดท้าย แผนจะครอบคลุมถึงวิธีการสำหรับการวิเคราะห์ข้อมูลที่ได้
ได้เก็บรวบรวมไว้ ซึ่งจะเกี่ยวข้องกับกระบวนการต่างๆ เช่น การวิเคราะห์เชิงบรรยาย (Descriptive
Analysis) ของคะแนนสอบ การประมาณค่าความเชื่อมั่น (Estimates of Reliability) และการ
วิเคราะห์ข้อมูลเชิงคุณภาพด้วยวิธีการที่เหมาะสม

การระบุทรัพยากรและการพัฒนาแผนสำหรับการจัดสรรและจัดการ: กิจกรรมในขั้นนี้เป็น
การกำหนดทรัพยากร (บุคลากร วัสดุ เวลา) ที่เกี่ยวข้องให้ชัดเจน เตรียมให้มีพร้อมสำหรับการดำเนิน
กิจกรรมต่าง ๆ ในระหว่างการพัฒนาแบบทดสอบและจัดทำแผนสำหรับการจัดสรร และจัดการ
ทรัพยากรตลอดกระบวนการพัฒนาแบบทดสอบ กิจกรรมนี้ทำให้ได้ข้อมูลพื้นฐานในการพิจารณาการ
นำแบบทดสอบไปใช้ในอนาคต (The Potential Practicality of the Test) และสำหรับการ
ตรวจสอบติดตามสิ่งนี้ตลอดกระบวนการพัฒนาแบบทดสอบ

2) ขั้นปฏิบัติการ (Operationalization Stage) ขั้นนี้เกี่ยวข้องกับการพัฒนาโครงสร้าง
ลักษณะของภาระงานในแบบทดสอบ (Test Task Specifications) เพื่อกำหนดภาระงานต่าง ๆ ที่จะ
บรรจุไว้ในแบบทดสอบ และพิมพ์เขียว (Blueprint) ให้ชัดเจน พร้อมทั้งมีการบรรยายถึงหลักการที่ใช้
ในการจัดเรียงภาระงานแต่ละชนิดในแบบทดสอบให้เป็นแบบทดสอบจริง นอกจากนี้ภารกิจในขั้น
ปฏิบัติการยังเกี่ยวข้องกับการพัฒนาและการเขียนภาระงานสำหรับแบบทดสอบฉบับจริง การเขียน
คำสั่งและการระบุวิธีตรวจให้คะแนนแบบทดสอบ ดังรายละเอียดต่อไปนี้

การพัฒนาภาระงานในแบบทดสอบและพิมพ์เขียว: ในการพัฒนาภาระงานในแบบทดสอบ
นักพัฒนาแบบทดสอบจะเริ่มด้วยการบรรยายภาระงานชนิดต่าง ๆ ที่เกี่ยวข้องกับการใช้
ภาษาเป้าหมายซึ่งได้กล่าวไว้แล้วในข้อความบรรยายการออกแบบ (The Design Statement) และ
ปรับปรุงสิ่งเหล่านี้ อีกทั้งพิจารณาคุณภาพด้านประโยชน์ เพื่อจัดทำโครงสร้างภาระงานในการสอบ
โครงสร้างดังกล่าวประกอบด้วย คำบรรยายโดยละเอียดของลักษณะภาระงานที่สำคัญ ซึ่งเป็นข้อมูล
พื้นฐานสำหรับการเขียนภาระงานในแบบทดสอบจริงต่อไป ส่วนพิมพ์เขียวของแบบทดสอบ
ประกอบด้วยลักษณะที่นำไปสู่โครงสร้างหรือการเรียงลำดับโดยรวมของแบบทดสอบ พร้อมกับการ
กำหนดโครงสร้างของภาระงานแต่ละชนิดที่จะบรรจุไว้ในแบบทดสอบ ข้อความบรรยายการออกแบบ

(A Design Statement) บรรยายปัจจัยที่เกี่ยวข้องโดยทั่วไปสำหรับการออกแบบแบบทดสอบ รวมถึงจุดประสงค์ของแบบทดสอบ ปรัชญาของการใช้ภาษาเป้าหมายที่แบบทดสอบถูกออกแบบให้วัด บุคคลผู้ที่จะต้องทำแบบทดสอบ สิ่งที่แบบทดสอบมุ่งหวังจะทดสอบ และอื่น ๆ ส่วนพิมพ์เขียวของแบบทดสอบ (Blueprint) บรรยายว่าภาระงานต่าง ๆ ในแบบทดสอบจะถูกสร้างและจัดเรียงเป็นแบบทดสอบอย่างไร

การเขียนคำสั่ง: การเขียนคำสั่งเกี่ยวข้องกับการบรรยายโครงสร้างของแบบทดสอบอย่างเต็มรูปแบบและชัดเจน ธรรมชาติของภาระงานที่ผู้เข้าสอบจะได้ทำ และผู้เข้าสอบจะต้องตอบสนองต่อภาระงานอย่างไร คำสั่งบางอย่างมีลักษณะโดยทั่วไป สามารถใช้ได้กับแบบทดสอบทั้งฉบับ แต่บางคำสั่งอาจจะเกี่ยวข้องกับภาระงานเฉพาะอย่าง

การระบุวิธีตรวจให้คะแนน: การระบุวิธีตรวจให้คะแนน ประกอบด้วย 2 ขั้นตอน คือ การนิยามเกณฑ์ที่จะตัดสินคุณภาพของคำตอบ และการตัดสินเกี่ยวกับกระบวนการอันเป็นที่มาของคะแนน

3) ขั้นตอนการบริหารการสอบ (Test Administration) การบริหารการสอบเกี่ยวข้องกับการจัดสอบกับกลุ่มเป้าหมาย การเก็บรวบรวมข้อมูล และการวิเคราะห์ข้อมูล โดยมีจุดประสงค์หลัก 2 ประการ คือ 1) การประเมินคุณประโยชน์ของแบบทดสอบ และ 2) การอ้างอิงหรือตัดสินใจตามที่ได้รับผลการระบุไว้ในขั้นตอนการออกแบบ สำหรับการบริหารการสอบมักจะแบ่งออกเป็น 2 ระยะ คือ การทดลองใช้แบบทดสอบ (Try-Out Testing) และการทดสอบจริง (Operational Testing) โดยระยะทดลองใช้ (Try Out) เกี่ยวข้องกับการบริหารการสอบ มีจุดประสงค์เพื่อเก็บข้อมูลเกี่ยวกับความมีประโยชน์ของแบบทดสอบ และเพื่อปรับปรุงแบบทดสอบและกระบวนการในการสอบ การปรับปรุงที่เป็นผลมาจากการได้รับข้อมูลย้อนกลับในระยะทดลอง อาจจะเป็นเรื่องธรรมดาที่ไม่ซับซ้อน และอาจจะประกอบด้วยการแก้ไขแบบทดสอบเพียงเล็กน้อย หรือการวิเคราะห์ผลของการทดลองใช้อาจจะชี้ให้เห็นว่า การแก้ไขในภาพรวมเป็นสิ่งจำเป็น อาจเกี่ยวข้องกับการย้อนกลับไปยังขั้นออกแบบและคิดทบทวนอีกครั้งในบางส่วนของ A Design Statement ในความพยายามของการทดสอบจริงแบบทดสอบหรือภาระงานสอบส่วนมากมักจะถูกทดลองใช้ก่อนนำไปใช้จริง สำหรับแบบทดสอบที่ใช้กันอยู่ในห้องเรียนโดยทั่วไป มักจะไม่ได้มีการทดลองก่อนนำไปใช้จริง แม้ว่านักพัฒนาแบบทดสอบมักจะแนะนำอย่างเอาใจจริงเอาใจว่า ควรให้นักเรียนที่คัดเลือกไว้ได้ทดลองทำแบบทดสอบหรือให้เพื่อนครูได้ตรวจทานแบบทดสอบล่วงหน้า เนื่องจากการกระทำดังกล่าวสามารถทำให้นักพัฒนาแบบทดสอบมีข้อมูลที่สามารถเป็นประโยชน์ในการปรับปรุงแบบทดสอบและภาระงานก่อนการนำไปใช้จริงได้ การใช้แบบทดสอบจริง (Operational Test Use) เกี่ยวข้องกับการบริหารการสอบ โดยมีเป้าหมายพื้นฐานเพื่อใช้แบบทดสอบตามวัตถุประสงค์ที่ได้กำหนดไว้ให้สมบูรณ์ และเพื่อเก็บข้อมูลเกี่ยวกับคุณประโยชน์ของแบบทดสอบ ในการพัฒนาแบบทดสอบทุกชนิด มักจะประกอบด้วย การจัดสอบและตรวจให้คะแนนแบบทดสอบ แล้วตามด้วยการวิเคราะห์ผลตามความเหมาะสม ซึ่งขึ้นอยู่กับความต้องการในแต่ละสถานการณ์ด้วย

นอกจากกิจกรรมหลัก 2 อย่างแล้ว ในขั้นตอนการบริหารการสอบนี้ ยังมีกิจกรรมย่อยอีกหลายอย่างที่สอดแทรกอยู่ในกิจกรรมหลัก เช่น 1) กระบวนการสำหรับการบริหารการสอบและการเก็บข้อมูลย้อนกลับ การบริหารการสอบเกี่ยวข้องกับการจัดเตรียมสภาพแวดล้อม เก็บรวบรวมวัสดุ

แบบทดสอบ การอบรมผู้ดูแลการสอบ และการจัดสอบจริง และกระบวนการในการบริหาร จำเป็นต้องได้รับการพัฒนาเพื่อให้ใช้ได้ทั้งในระยะทดลองและการสอบจริง ส่วนการเก็บข้อมูลย้อนกลับเกี่ยวข้องกับการได้รับข้อมูลเชิงคุณภาพ และปริมาณที่เกี่ยวกับคุณประโยชน์ของแบบทดสอบจากผู้ทำและผู้ใช้แบบทดสอบ โดยข้อมูลย้อนกลับจะถูกเก็บครั้งแรกในระหว่างการทดลอง และถูกเก็บอีกครั้งในการใช้แบบทดสอบจริง 2) กระบวนการสำหรับการวิเคราะห์คะแนนสอบซึ่งจะประกอบด้วย การบรรยายคะแนนสอบ (Describing Test Scores) โดยใช้สถิติเชิงบรรยาย (Descriptive Statistics) เพื่ออธิบายคุณลักษณะเชิงปริมาณ (Quantitative Characteristics) ของแบบทดสอบ การรายงานคะแนนสอบ (Reporting Test Scores) โดยการวิเคราะห์ทางสถิติสำหรับตัดสินว่าจะรายงานคะแนนสอบให้มีประสิทธิภาพมากที่สุดได้อย่างไร การวิเคราะห์แบบทดสอบรายข้อ (Item Analysis) โดยการใช่วิธีการทางสถิติที่หลากหลายสำหรับการวิเคราะห์และปรับปรุงคุณภาพของภาระงานในแบบทดสอบหรือแบบทดสอบแต่ละข้อ การหาค่าความเชื่อมั่นของคะแนนสอบ (Estimating Reliability of Test Scores) โดยการใช่วิธีการทางสถิติหลายอย่างสำหรับการคำนวณความคงที่ของคะแนนสอบจากเงื่อนไขการใช้แบบทดสอบเฉพาะอย่างที่แตกต่างกัน การหาค่าความตรงของการใช้แบบทดสอบ (Investigating the Validity of Test Use) เกี่ยวข้องกับการพิจารณาอย่างมีเหตุผล (Logical considerations) และวิธีการเชิงประจักษ์ (Empirical Procedures) หลายอย่าง ทั้งเชิงปริมาณและคุณภาพ สำหรับการค้นหาความตรงของการกล่าวอ้างที่มาจากคะแนนสอบ ภายใต้เงื่อนไขเฉพาะอย่างของการใช้แบบทดสอบ และ 3) การเก็บรักษาแบบทดสอบ (Archiving): การเก็บรักษาแบบทดสอบเกี่ยวข้องกับการสร้างคลังหรือธนาคารขนาดใหญ่สำหรับเก็บรวบรวมภาระงานการสอบ เพื่ออำนวยความสะดวกในการพัฒนาแบบทดสอบในครั้งต่อ ๆ ไป การเก็บรักษาแบบทดสอบทำให้เกิดความเป็นไปได้ให้แบบทดสอบได้รับการดัดแปลงหรือให้เหมาะกับผู้ใช้สอบเฉพาะกลุ่มได้ง่ายขึ้น โดยทั่วไปแล้ว กระบวนการในการเก็บรักษาได้รับการออกแบบที่จะอนุญาตให้เกิดการนำภาระงานหรือข้อมูลสำคัญเกี่ยวกับภาระงานมาใช้ได้ง่าย การเก็บรักษาช่วยทำให้เกิดความปลอดภัยของแบบทดสอบได้อีกด้วย ประการสุดท้าย กระบวนการเก็บรักษาแบบทดสอบอาจจะใช้เพื่ออำนวยความสะดวกในการเลือกภาระงานที่มีลักษณะเฉพาะออกมาใช้

ขั้นตอนการพัฒนาแบบทดสอบของ (Genesee & Fred & Upshur, 1996)

Genesee and Fred & Upshur (1996) นำเสนอการพัฒนาแบบทดสอบมาตรฐาน (Standardized Test Development) ด้วยขั้นตอนหลัก 5 ขั้นตอน คือ 1) การกำหนดจุดประสงค์การสอบ 2) การจัดทำโครงสร้างแบบทดสอบและการสร้างชุดข้อคำถาม 3) การทดลองใช้แบบทดสอบกับกลุ่มทดลอง 4) การวิเคราะห์แบบทดสอบเบื้องต้นและการคัดเลือกแบบทดสอบเพื่อสร้างแบบทดสอบฉบับจริง และ 5) การกำหนดเกณฑ์การประเมินผลการสอบ โดยมีรายละเอียดในแต่ละขั้นตอนดังนี้

1) การกำหนดจุดประสงค์การสอบ การพัฒนาแบบทดสอบมาตรฐานเหมือนกับการพัฒนาแบบทดสอบชนิดอื่น ๆ ที่มักจะเริ่มต้นด้วยการกำหนดคำนิยามจุดประสงค์แบบทดสอบ สำหรับแบบทดสอบที่ครูสร้างขึ้นถูกจัดทำเพื่อประเมินผลสัมฤทธิ์ที่เป็นผลมาจากการได้รับกิจกรรมการเรียนรู้

การสอนเฉพาะอย่าง ดังนั้น สารของแบบทดสอบส่วนใหญ่จึงถูกกำหนดด้วยวัตถุประสงค์ แผน และการฝึกปฏิบัติที่มาจาก การเรียนการสอนเฉพาะอย่าง ในทางกลับกัน แบบทดสอบมาตรฐานไม่ได้ถูกสร้างเพื่อวัดผลสัมฤทธิ์ในหลักสูตรหรือยึดพื้นฐานจากวิชาใด แต่ถูกสร้างขึ้นเพื่อวัดทักษะหรือความรู้ด้านภาษาโดยทั่วไปมากกว่า ดังนั้นข้อความบอกจุดประสงค์มักจะถูกกล่าวไว้อย่างค่อนข้างทั่วไปมากกว่าแบบทดสอบที่ใช้วัดความรู้ในห้องเรียน ดังตัวอย่างจุดประสงค์ของแบบทดสอบมาตรฐานที่กล่าวว่า เพื่อประเมินความสามารถของผู้เข้าสอบในการผลิตลักษณะภาษาที่ถูกต้องตามหลักไวยากรณ์และการออกเสียงในการพูด (Michigan Oral Productive Test, 1970 อ้างถึงใน Genesee & Upshur, 1996: 236) หรือ เพื่อประเมินความสามารถในการพูดภาษาอังกฤษของบุคคลที่ไม่ได้พูดภาษาอังกฤษเป็นภาษาแม่ (Test of Spoken English, 1982 อ้างถึงใน Genesee & Upshur, 1996: 236)

2) การจัดทำโครงสร้างแบบทดสอบและการสร้างชุดข้อคำถาม เนื่องจากแบบทดสอบมาตรฐานไม่ได้อิงวัตถุประสงค์ และไม่มีสาระภาษาหรือชุดของกิจกรรมการเรียนการสอนที่คิดไว้ก่อนล่วงหน้าให้ได้เลือก เพื่อตัดสินใจในการบรรจุทักษะภาษา และรูปแบบภาระงานสอบชนิดใดไว้ในแบบทดสอบ การที่จะจัดทำโครงสร้างแบบทดสอบมาตรฐานให้สำเร็จ นักพัฒนาแบบทดสอบควรขอความร่วมมือจากผู้เชี่ยวชาญในการพัฒนาโครงสร้างแบบทดสอบสำหรับสร้างข้อคำถามให้สอดคล้องกับจุดประสงค์ทั่วไปของแบบทดสอบที่จะพัฒนา โครงสร้างแบบทดสอบที่พัฒนาขึ้นทำหน้าที่หลัก 2 ประการ คือ 1) สื่อสารกับผู้ที่ใช้แบบทดสอบว่าแบบทดสอบชุดนี้วัดทักษะอะไร และ 2) แสดงรายละเอียดของปริศนาภาษาที่กำลังถูกประเมินเพื่อให้นักเขียนข้อคำถามสามารถคิดชุดของข้อคำถามที่มีลักษณะคล้าย ๆ กันได้ ลำดับต่อมากลุ่มของผู้เขียนแบบทดสอบจะจัดทำชุดข้อคำถามตามข้อกำหนดที่ระบุไว้ในโครงสร้างแบบทดสอบ ชุดข้อคำถามที่จัดเตรียมจะมีจำนวนข้อเกินกว่าที่จะใช้ในแบบทดสอบฉบับจริง เนื่องจากสันนิษฐานว่าข้อคำถามไม่ได้ใช้การได้ดีทุกข้อ จึงสร้างจำนวนข้อให้มากกว่าที่ต้องการใช้

3) การทดลองใช้แบบทดสอบกับกลุ่มทดลอง หลังจากสร้างข้อคำถามเสร็จแล้ว นักพัฒนาแบบทดสอบนำชุดข้อคำถามที่ผลิตได้ไปทดลองสอบกับกลุ่มทดลอง ภายใต้สภาพแวดล้อมที่ใกล้เคียงที่สุดกับสถานการณ์ที่แบบทดสอบจะถูกใช้จริง กลุ่มทดลองที่ถูกเลือกควรเป็นตัวแทนหรือเหมาะสมเท่าที่จะเป็นไปได้กับกลุ่มตัวอย่างจริง หลังจากนั้นคำตอบของกลุ่มทดลองในแบบทดสอบแต่ละข้อจะถูกนำไปวิเคราะห์ ตามลักษณะภาคบังคับพื้นฐาน 2 ประการ คือ 1) แต่ละข้อมีความยากระดับใด (ความยากรายข้อ) และ 2) แต่ละข้อจำแนกผู้เรียนตามระดับกลุ่มความสามารถได้ดีเพียงใด (ค่าอำนาจจำแนกรายข้อ) การวัดความยากของแบบทดสอบที่ง่ายที่สุดวิธีหนึ่ง คือ สัดส่วนของกลุ่มทดลองที่ตกหรือตอบผิดในแต่ละข้อต่อจำนวนคนสอบทั้งหมดในข้อนั้น ส่วนค่าอำนาจจำแนกถูกตัดสินด้วยวิธีการที่แตกต่างออกไป โดยการเปรียบเทียบคะแนนของแบบทดสอบแต่ละข้อที่ผู้เข้าสอบทำได้ และทำไม่ได้ ตัวอย่างเช่น ถ้าแบบทดสอบมีความยากอยู่ในระดับปานกลาง นักเรียนที่มีความสามารถน้อยส่วนใหญ่ควรตอบข้อนั้นไม่ได้ ประมาณกึ่งหนึ่งของนักเรียนที่มีความสามารถปานกลางควรผ่านข้อนั้น หรือมองอีกมุมหนึ่ง ความเป็นไปได้ที่ผู้เข้าสอบจะผ่านแบบทดสอบข้อใดควรขึ้นอยู่กับความสามารถของเขาเมื่อเทียบกับความยากของแบบทดสอบ เมื่อสิ่งนี้เป็นจริง แบบทดสอบจะถูกเรียกว่ามีค่าอำนาจจำแนกสูง แต่ใช้ไม่ได้กับทุกข้ออย่างเสมอไป แบบทดสอบบางข้ออาจมี

ลักษณะเฉพาะที่ผู้ที่มีความสามารถสูงเท่านั้นที่ทำได้ หรือบางข้ออาจจะต้องการข้อมูลคำตอบที่เป็นข้อเท็จจริงที่ไม่ซับซ้อนโดยผู้เข้าสอบที่มีความสามารถต่ำกว่าคนน่าจะรู้คำตอบ

4) การวิเคราะห์แบบทดสอบเบื้องต้นและการคัดเลือกแบบทดสอบเพื่อสร้างแบบทดสอบฉบับจริง หลังจากการวิเคราะห์แบบทดสอบรายชื่อเสร็จสมบูรณ์ นักพัฒนาแบบทดสอบคัดเลือกแบบทดสอบที่จะบรรจุไว้ในแบบทดสอบฉบับจริง และข้อใดบ้างที่ควรตัดออกหรือปรับปรุง หลักการที่มักใช้ในการเลือกแบบทดสอบ คือ การเก็บข้อที่ไม่ยากและไม่ง่ายจนเกินไปจากการทดลองใช้มาบรรจุไว้ในแบบทดสอบฉบับจริง นอกจากนี้ ข้อคำถามที่จะบรรจุไว้ในแบบทดสอบฉบับจริงต้องมีอำนาจจำแนกได้ดีด้วย โดยทั่วไปแล้วแบบทดสอบข้อใดที่มีอำนาจจำแนกดีมาก ถือเป็นข้อคำถามที่ดีมากเช่นกัน

5) การกำหนดเกณฑ์การประเมินผลการสอบ ภารกิจสุดท้ายในการพัฒนาแบบทดสอบมาตรฐานแบบอิงกลุ่ม คือ การกำหนดเกณฑ์ระดับคะแนน นักพัฒนาแบบทดสอบนำแบบทดสอบฉบับจริงไปใช้สอบกับผู้เข้าสอบอีกกลุ่มหนึ่งหรือหลายกลุ่ม กลุ่มเหล่านี้จะถูกเรียกว่ากลุ่มบรรทัดฐาน (Norming Groups) กลุ่มผู้เข้าสอบเหล่านี้ถูกคัดเลือกจากผู้ที่มีลักษณะไม่แตกต่างจากผู้เข้าสอบที่เป็นกลุ่มเป้าหมายของแบบทดสอบ คะแนนของกลุ่มบรรทัดฐานจะถูกวิเคราะห์เพื่ออธิบายค่าเฉลี่ยและช่วงห่างของคะแนนผลการสอบ หลังจากนั้นผลคะแนนเหล่านี้จะถูกใช้เพื่อตีความคะแนนสอบของกลุ่มเป้าหมายเมื่อแบบทดสอบถูกนำไปใช้จริง

ขั้นตอนการพัฒนาแบบทดสอบของ (H. D. Brown, 2004)

H. D. Brown (2004) นำเสนอการพัฒนาแบบทดสอบมาตรฐานด้วยขั้นตอนหลัก 6 ขั้นตอน คือ 1) กำหนดจุดประสงค์และวัตถุประสงค์การสอบ 2) การออกแบบโครงสร้างแบบทดสอบ 3) การออกแบบ เลือก และจัดเรียงภาระงานหรือข้อคำถามในแบบทดสอบ 4) การประเมินข้อคำถามชนิดต่าง ๆ ด้วยวิธีการที่เหมาะสม 5) การระบุกระบวนการให้คะแนนและรูปแบบการรายงานผล และ 6) การศึกษาเพื่อประเมินความตรงของแบบทดสอบ โดยมีรายละเอียดในแต่ละขั้นตอน ดังต่อไปนี้

1) การกำหนดจุดประสงค์และวัตถุประสงค์การสอบ แบบทดสอบมาตรฐานส่วนใหญ่ถูกคาดหวังให้สามารถนำไปใช้ได้จริงสูงทั้งในการจัดสอบและการตรวจให้คะแนนโดยไม่ต้องกังวลเรื่องการตรวจสอบความตรงของแบบทดสอบ การวางแผนเบื้องต้นด้านเวลาและงบประมาณสำหรับแบบทดสอบชนิดนี้เป็นเรื่องสำคัญ แต่แบบทดสอบสามารถใช้ซ้ำได้หลายครั้ง ดังนั้นจึงเป็นเรื่องสำคัญที่จะต้องเขียนจุดประสงค์และวัตถุประสงค์การสอบไว้อย่างเฉพาะเจาะจงและชัดเจน

2) การออกแบบโครงสร้างแบบทดสอบ เป็นขั้นตอนที่ยากที่สุดในการพัฒนาแบบทดสอบ ในขั้นนี้จำเป็นต้องมีการตัดสินใจว่าจะกำหนดโครงสร้างของแบบทดสอบอย่างไร ก่อนการกำหนดโครงสร้างแบบทดสอบ ควรมีการศึกษา ทำความเข้าใจเกี่ยวกับชุดของโครงสร้างความสามารถที่อยู่ภายใต้แบบทดสอบชนิดที่ต้องการจะพัฒนา ขั้นตอนของการวางรากฐานสำคัญอันนี้อาจใช้เวลาเป็นหลายสัปดาห์ หลายเดือนหรือหลายปีที่ต้องใช้ความพยายามอย่างหนัก และแบบทดสอบมาตรฐานที่ไม่ได้คุณภาพมักจะเป็นผลผลิตของการประเมินความตรงเชิงโครงสร้างแบบเร่งรีบ

3) การออกแบบ เลือกลง และจัดเรียงภาระงานหรือข้อคำถามในแบบทดสอบ เมื่อจัดทำโครงสร้างแบบทดสอบสำหรับแบบทดสอบมาตรฐานเรียบร้อยแล้ว บางครั้งกระบวนการออกแบบอาจมีภาระงานอื่นต่อเนื่อง เช่น การเลือกและการจัดเรียงข้อคำถาม โครงสร้างแบบทดสอบทำหน้าที่เหมือนเป็นพิมพ์เขียวในการกำหนดจำนวนและชนิดของข้อคำถามที่จะต้องสร้างขึ้นในแบบทดสอบ

4) การประเมินข้อคำถามชนิดต่าง ๆ ด้วยวิธีการที่เหมาะสม การประเมินข้อคำถามแบบเลือกตอบในแบบทดสอบมาตรฐานมีการประเมินพื้นฐานอยู่ 3 ชนิด คือ การประเมินความยากง่ายของแบบทดสอบรายข้อ (Item Facility) ค่าอำนาจจำแนกของแบบทดสอบรายข้อ (Item Discrimination) และการวิเคราะห์ตัวลวง (Distractor Analysis) หากเป็นการสร้างแบบทดสอบเพื่อใช้ในห้องเรียน การประเมินแบบทดสอบทั้ง 3 ชนิด อาจดูไม่คุ้มค่าเพราะใช้เวลามากและเป็นแบบทดสอบที่สร้างขึ้นเพื่อใช้เพียงครั้งเดียว แต่สำหรับแบบทดสอบมาตรฐาน ซึ่งเป็นการสร้างแบบทดสอบเพื่อใช้เชิงพาณิชย์ และนำมาจัดสอบได้หลายครั้ง หรือจัดสอบได้ในหลายรูปแบบ จึงมีความคุ้มค่าและสมควรดำเนินการมากกว่า

สำหรับแบบทดสอบที่ให้คำตอบในรูปแบบอื่น เช่น แบบทดสอบที่ต้องการคำตอบอย่างละเอียด การประเมินแบบทดสอบในรูปแบบอื่นๆ จึงมีความสำคัญ หลักการของความสะดวกใช้ และความเชื่อมั่นมีบทบาทสำคัญมากควบคู่ไปกับแนวคิดของความยากง่ายของแบบทดสอบ ประเด็นด้านความสะดวกใช้ของแบบทดสอบเกี่ยวข้องกับความสะดวกของคำสั่ง เวลาในการทำแบบทดสอบ ความสะดวกในการจัดสอบ และระยะเวลาที่ใช้ในการตรวจแบบทดสอบ ความเชื่อมั่นมีบทบาทสำคัญในกรณีที่แบบทดสอบส่วนนั้นมีผู้ตรวจให้คะแนนมากกว่า 1 คน และในกรณีนี้อาจจะมีความสำคัญน้อยกว่าเมื่อผู้ตรวจให้คะแนนคนเดียวต้องตรวจแบบทดสอบเป็นเวลานาน ที่อาจจะนำไปสู่ความไม่มั่นคงในมาตรฐานของการตรวจ ความยากง่าย (Facility) มีความสำคัญต่อความตรงและความสำเร็จในข้อคำถามแต่ละชนิด เช่น คำสั่งที่ไม่ชัดเจน ภาษาที่ซับซ้อน หัวข้อที่คลุมเครือ และข้อมูลที่แสดงถึงความลำเอียงด้านวัฒนธรรม ทั้งหมดนี้อาจนำไปสู่ระดับความยากที่สูงขึ้นที่ไม่เป็นที่ต้องการ

5) การระบุกระบวนการให้คะแนนและรูปแบบการรายงานผล ข้อคำถามที่ได้รับการออกแบบอย่างเป็นระบบ พร้อมทั้งผ่านการคัดเลือกและจัดเรียงลำดับอย่างเหมาะสม และถูกประเมินคุณภาพเพื่อให้อยู่ในระดับความยากที่ต้องการ ควรเป็นแบบทดสอบที่สามารถตรวจให้คะแนนได้อย่างถูกต้อง และรายงานคะแนนกลับไปยังผู้เข้าสอบและสถาบันได้อย่างมีประสิทธิภาพ

6) การศึกษาเพื่อประเมินความตรงของแบบทดสอบ เป็นที่ทราบกันดีว่าไม่มีเครื่องมือวัดมาตรฐานใดที่จะถูกนำมาใช้ซ้ำโดยไม่มีโครงการต่อเนื่องในการประเมินความตรงเชิงโครงสร้าง ไม่ว่าจะแบบทดสอบมาตรฐานชนิดใด เมื่อพัฒนาขึ้นมาแล้ว ต้องมีการตรวจสอบประสิทธิภาพของการนำไปใช้เป็นระยะ ๆ อย่างมีระบบแบบแผน และต้องมีขั้นตอนต่าง ๆ ที่จะปรับปรุงให้ดียิ่งขึ้น ข้อกำหนดนี้มีความจำเป็นสำหรับแบบทดสอบที่ถูกพัฒนาเป็นชุดแบบทดสอบที่คู่ขนานกันหลายชุด นั่นคือแบบทดสอบชุดต่าง ๆ ต้องเชื่อถือได้เท่ากัน โดยที่คะแนนจากการสอบแบบทดสอบชุดหนึ่ง มีความตรงและความสามารถในการตีความได้เหมือนกับแบบทดสอบชุดที่เป็นต้นฉบับ

สรุปขั้นตอนการพัฒนาแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัยได้มาจากแนวคิดของBachman and Lyle F. & Palmer (1996), Genesee & Upshur (1996) Genesee and Fred & Upshur (1996) และ H. D. Brown (2004) ประกอบด้วยขั้นตอนหลัก 3

ขั้น คือ ขั้นตอนออกแบบ (Design) ขั้นปฏิบัติการ (Operationalization) และขั้นบริหารการสอบ (Administration) โดยในแต่ละขั้นจะประกอบด้วยกิจกรรมย่อยที่เกี่ยวข้อง เช่น

ขั้นการออกแบบ (Design) มีกิจกรรมย่อยที่เกี่ยวข้อง คือ การกำหนดจุดประสงค์ของแบบทดสอบ การกำหนดและบรรยายภาระงานในบริบทของการใช้ภาษาเป้าหมาย การบรรยายลักษณะของผู้ใช้แบบทดสอบหรือผู้ทำแบบทดสอบ การนิยามโครงสร้างภาษาที่ต้องการจะวัด การกำหนดแผนการประเมินคุณภาพแบบทดสอบ และการวางแผนและจัดสรรทรัพยากรที่เกี่ยวข้อง

ขั้นปฏิบัติการ (Operationalization) มีกิจกรรมที่เกี่ยวข้อง คือ การจัดทำโครงสร้างลักษณะแบบทดสอบ การสร้างชุดคำถามตามโครงสร้างลักษณะแบบทดสอบ (ชุด A และ ชุด B) การเขียนคำสั่ง การกำหนดวิธีตรวจให้คะแนนแบบทดสอบ การสร้างแบบประเมินคุณภาพแบบทดสอบโดยผู้เชี่ยวชาญ การสร้างแบบสอบถามความคิดเห็นผู้ใช้แบบทดสอบ การสร้างแบบสัมภาษณ์ผู้เข้าสอบ และการสร้างคู่มือการจัดการสอบ

และขั้นบริหารการสอบ (Administration) มีกิจกรรมหลัก 2 อย่าง คือ การทดลองใช้แบบทดสอบ และการใช้แบบทดสอบจริง โดยทั้ง 2 กิจกรรม มีกระบวนการย่อยส่วนใหญเหมือนกัน คือ การเตรียมความพร้อมด้านสถานที่และบุคลากร (การอบรมบุคลากร การจัดเตรียมอุปกรณ์และสถานที่) การทดสอบ (การใช้แบบทดสอบที่พัฒนา การสอบถามความคิดเห็น การสัมภาษณ์ การใช้แบบทดสอบมาตรฐาน) การวิเคราะห์ข้อมูล (ค่าสถิติพื้นฐาน ค่าความยากง่าย ค่าอำนาจจำแนก ค่าความเชื่อมั่นของแบบทดสอบแต่ละชุด ค่าความสัมพันธ์ระหว่างแบบทดสอบ 2 ชุด ค่าความตรงร่วมสมัยของแบบทดสอบแต่ละชุดกับแบบทดสอบมาตรฐาน ค่าระดับคะแนนที่ใช้เป็นเกณฑ์การประเมินผลการสอบ และค่าเฉลี่ยความคิดเห็นของผู้ใช้แบบทดสอบ) การรายงานผล (รายงานค่าสถิติพื้นฐานของคะแนนสอบ ค่าความยากง่าย ค่าอำนาจจำแนก ค่าความเชื่อมั่นของแบบทดสอบแต่ละชุด ค่าความสัมพันธ์ระหว่างคะแนนของแบบทดสอบที่สร้างขึ้นกับแบบทดสอบมาตรฐาน ระดับคะแนนที่ใช้เป็นเกณฑ์การประเมินผลการสอบ และค่าเฉลี่ยความคิดเห็นของผู้ใช้แบบทดสอบ) การแก้ไข ปรับปรุง และคัดเลือกแบบทดสอบ (กรณีการทดลองใช้แบบทดสอบ) และการคัดเลือกแบบทดสอบเข้าธนาคารแบบทดสอบ (กรณีการใช้แบบทดสอบจริง)

2.2.1.2 ตารางกำหนดลักษณะแบบทดสอบ (Table of Test Specifications)

นักวิชาการด้านการทดสอบและประเมินผลภาษาอังกฤษ เช่น Bachman and Lyle F. & Palmer (1996) Alderson, Clapham, and Wall (1995) และ (Hughes, 1989) ได้เสนอแนวคิดเกี่ยวกับหัวข้อที่ควรบรรจุไว้ในตารางกำหนดลักษณะแบบทดสอบหรือภาระงานสอบไว้แตกต่างกัน โดยมีรายละเอียดในแต่ละแนวคิด ดังนี้

Bachman and Lyle F. & Palmer (1996) กล่าวว่า ในกระบวนการของขั้นปฏิบัติการ นักพัฒนาแบบทดสอบเริ่มต้นด้วยการกำหนดภาระงานการใช้ภาษาเป้าหมายชนิดต่าง ๆ และปรับเปลี่ยนลักษณะภาระงานตามความจำเป็นที่เป็นผลมาจากความพอใจในคุณภาพด้านผลประโยชน์ของแบบทดสอบ นอกจากนี้ นักพัฒนาแบบทดสอบอาจจะเพิ่มจุดประสงค์เฉพาะ คำนียามเชิง

โครงสร้างสำหรับภาระงานสอบแต่ละชนิด ผลที่ได้จากการกำหนดลักษณะภาระงานทำให้ได้ตารางกำหนดลักษณะภาระงานสอบ (Test Task Specifications) ที่จะถูกใช้เป็นแม่แบบสำหรับการพัฒนาภาระงานต่าง ๆ ในแบบทดสอบจริง ไม่ว่าจะเป็ภาระงานสอบชนิดใด การกำหนดลักษณะภาระงานจะครอบคลุมสิ่งต่อไปนี้

1) จุดประสงค์ของภาระงานสอบ ส่วนนี้จะเหมือนกับหนึ่งในหลาย ๆ จุดประสงค์ที่ถูกระบุไว้ในประโยค Design Statement สำหรับแบบทดสอบทั้งชุด เหตุผลที่ต้องใส่จุดประสงค์ในส่วนนี้คือเพื่อให้ นักเขียนแบบทดสอบ ผู้ซึ่งจะใช้การกำหนดลักษณะภาระงานสอบเป็นตัวชี้ นำในการเขียนภาระงานต่าง ๆ ในแบบทดสอบจะได้ไม่หลงประเด็นจากจุดประสงค์เฉพาะที่ภาระงานแต่ละอย่างถูกกำหนดให้ไว้

2) คำนิยามของโครงสร้างที่จะวัด สิ่งนี้เป็นเรื่องสำคัญ เพราะโครงสร้างต่าง ๆ ถูกใส่รวมเข้ามาใน Design Statement สำหรับแบบทดสอบทั้งชุด และนักพัฒนาแบบทดสอบและนักเขียนแบบทดสอบต้องมีความเข้าใจที่ชัดเจนในสิ่งที่ภาระงานถูกกำหนดให้ไว้ และคำนิยามนี้จะเป็ประโยชน์สำหรับผู้ใช้แบบทดสอบในการตีความผลลัพธ์ของแบบทดสอบด้วย

3) ลักษณะของบริบทของภาระงานสอบ สิ่งเหล่านี้ อาจจะผันแปรไปตามแต่ละภาระงาน แต่จำเป็นต้องระบุให้ชัดในส่วนนี้

4) การจัดสรรเวลา เนื่องจากภาระงานสอบแต่ละชนิด อาจต้องใช้เวลามากหรือน้อยต่างกัน และเนื่องจากนักพัฒนาแบบทดสอบมักต้องการให้มีการคุมเวลาในการให้ผู้เข้าสอบทำแบบทดสอบให้เสร็จทั้งชุด จึงจำเป็นต้องมีการระบุเวลาในแต่ละภาระงานให้ชัดเจน

5) คำแนะนำสำหรับการตอบสนองต่อภาระงาน ภาระงานสอบต่างชนิดกันต้องการให้มีการตอบสนองที่แตกต่างกัน ดังนั้นจึงจำเป็นต้องระบุคำแนะนำสำหรับภาระงานสอบแต่ละชนิด

6) ลักษณะของตัวกระตุ้น การตอบสนอง และความสัมพันธ์ของตัวกระตุ้นต่อการตอบสนอง

7) วิธีตรวจให้คะแนน ภาระงานสอบชนิดต่าง ๆ อาจจะถูกตรวจให้คะแนนต่างกัน ดังนั้นวิธีการให้คะแนนจึงจำเป็นต้องระบุไว้สำหรับภาระงานแต่ละชนิด

Alderson, Clapham, and Wall (1995: 11-13) กล่าวว่า ในตารางกำหนดลักษณะแบบทดสอบควรมีข้อมูลต่อไปนี้

1) จุดประสงค์ของแบบทดสอบคืออะไร แบบทดสอบต่าง ๆ มีแนวโน้มจะเป็นหนึ่งชนิดในชนิดของการแบ่งแบบทดสอบอย่างกว้าง ๆ ต่อไปนี้ คือ แบบทดสอบวัดระดับ (Placement Test) แบบทดสอบวัดความก้าวหน้า (Progress Test) แบบทดสอบวัดผลสัมฤทธิ์ (Achievement Test) แบบทดสอบวัดความสามารถ (Proficiency Test) และแบบทดสอบวินิจฉัยปัญหา (Diagnostic Test)

2) คุณลักษณะของผู้เรียนที่จะใช้แบบทดสอบ ครอบคลุมข้อมูลด้านอายุ เพศ ระดับความสามารถหรือระดับช่วงชั้นของการเรียน ภาษาแม่ ภูมิหลังด้านวัฒนธรรม ประเทศบ้านเกิด ระดับและธรรมชาติของการศึกษา เหตุผลสำหรับการสอบ แนวโน้มความสนใจส่วนตัว ถ้าเป็นไปได้ควรมีข้อมูลความสนใจด้านอาชีพ และระดับความรู้ที่เป็นพื้นฐาน

3) แบบทดสอบมีกี่ตอน กี่ฉบับ ระดับความยากของข้อสอบจะให้เห็นความแตกต่างอย่างไร เช่น แบบทดสอบที่ใช้เวลา 3 ชั่วโมง แบบทดสอบที่ใช้เวลา 2 ชั่วโมง 5 ฉบับ ตอนที่ใช้เวลา 45 นาที

มี 3 ตอน การอ่านได้รับการทดสอบแยกจากไวยากรณ์ การฟังและการเขียนรวมกันเป็นแบบทดสอบฉบับเดียว และอื่น ๆ

4) สถานการณ์ทางภาษา เป้าหมายที่ได้รับการให้ความสำคัญ และสิ่งนี้ได้รับการกระตุ้นด้วยวิธีการต่างๆ ในสาระการสอบและวิธีการ

5) ชนิดของเนื้อความที่ควรเลือกใช้คือรูปแบบใด ภาษาเขียนและ/หรือภาษาพูด แหล่งที่มาของเนื้อความเหล่านี้คืออะไร กลุ่มเป้าหมาย ระดับความยากหรือความยาว หน้าที่ภาษาที่ควรมีในเนื้อความ เช่น การโน้มน้าวใจ การให้คำนิยาม การสรุปความ เป็นต้น และระดับความซับซ้อนของภาษาควรอยู่ในระดับใด

6) ควรทดสอบทักษะภาษาอะไรบ้าง ระบุจุดประสงค์นำทางหรือทักษะย่อยให้ชัดเจน และออกแบบข้อคำถามให้ทดสอบสิ่งเหล่านี้เป็นรายชื่อ หรือในลักษณะการบูรณาการ และมีความแตกต่างระหว่างข้อคำถามที่ทดสอบ ประเด็นสำคัญ รายละเอียดเฉพาะและการลงความเห็นหรือไม่

7) ควรทดสอบประเด็นทางภาษาอะไรบ้าง มีรายการโครงสร้างหรือลักษณะทางไวยากรณ์ที่จะให้รวมไว้ในแบบทดสอบหรือไม่ การระบุคำศัพท์ในบางลักษณะ เช่น รายการความถี่ที่ใช้บ่อย เป็นต้น หรือไม่ หัวเรื่อง และหน้าที่ การใช้ภาษาเฉพาะในบางสถานการณ์หรือลักษณะการใช้ภาษาในทางปฏิบัติถูกระบุหรือไม่

8) ต้องการให้เป็นภาระงานแบบใด ทักษะเดี่ยว ทักษะสัมพันธ์ สถานการณ์จำลอง ที่สามารถวัดได้อย่างเป็นปรนัยหรือไม่

9) แต่ละตอนถูกกำหนดให้มีแบบทดสอบกี่ข้อ การกำหนดน้ำหนักของแบบทดสอบแต่ละข้อเป็นอย่างไร เท่ากันทุกข้อ มีการให้น้ำหนักเพิ่มพิเศษสำหรับข้อคำถามที่ยากกว่าหรือไม่

10) วิธีการทดสอบรูปแบบใดที่จะนำมาใช้ แบบมีคำตอบให้เลือกหลายตัวเลือก แบบเติมคำแบบจับคู่ การแปล แบบให้คำตอบสั้น ๆ การอธิบายรูปภาพ การแสดงบทบาทสมมติตามบัตรภาระงาน เรียงความ การเขียนแบบมีโครงสร้าง

11) คำสั่งรูปแบบใดที่จะนำมาใช้ในการให้คำแนะนำสำหรับผู้เข้าสอบ ตัวอย่างจำเป็นต้องใช้เพื่อช่วยให้ผู้เข้าสอบรู้ว่าอะไรถูกคาดหวัง เกณฑ์การประเมินควรรวมไว้ในคำสั่งด้วยหรือไม่

12) เกณฑ์ใดที่จะใช้สำหรับการประเมินโดยผู้ให้คะแนน ความถูกต้อง ความเหมาะสม การสะกดคำ ความยาวของการพูดหรือการเขียน ฯลฯ สำคัญอย่างไร

Hughes (1989: 49-51) กล่าวว่า ข้อมูลที่ควรปรากฏในตารางกำหนดลักษณะแบบทดสอบคือ เนื้อหา รูปแบบการทดสอบและเวลา ระดับเกณฑ์ความสามารถ และกระบวนการให้คะแนน โดยมีรายละเอียดในแต่ละประเด็น ดังนี้

1) สาระของแบบทดสอบ (Content) ส่วนนี้ไม่ได้หมายถึงสาระของแบบทดสอบฉบับใดฉบับหนึ่ง แต่หมายถึงเนื้อหาโดยรวมทั้งหมดของแบบทดสอบทุกฉบับ ตัวอย่างของเนื้อหาจะปรากฏในแบบทดสอบแต่ละฉบับ วิธีการอธิบายสาระแบบทดสอบจะแตกต่างกันไปตามธรรมชาติของแบบทดสอบ ตัวอย่างเช่น สาระของแบบทดสอบวัดไวยากรณ์ อาจจะแจกแจงรายการของโครงสร้างสำคัญทั้งหมด ในทางกลับกันสาระของแบบทดสอบวัดทักษะภาษา อาจจะถูกระบุแง่มุมต่าง ๆ ต่อไปนี้จะเป็นกรอบแนวทางที่เป็นไปได้สำหรับการเขียนสาระของแบบทดสอบ ที่สามารถนำไป

ดัดแปลงให้เหมาะสมกับการใช้งานได้ แต่สิ่งสำคัญคือ สารแบบทดสอบควรระบุให้ครบถ้วนในสิ่งต่อไปนี้

ภาระงาน (Operations) หมายถึง สิ่งที่ผู้เข้าสอบต้องสามารถทำได้ ตัวอย่างเช่น ในแบบทดสอบวัดการอ่าน ภาระงานอาจเกี่ยวกับการอ่านบทอ่านอย่างละเอียดเพื่อหาข้อมูลเฉพาะอย่าง การเดาความหมายของคำศัพท์ที่ไม่คุ้นเคยจากบริบท เป็นต้น

ชนิดของสาร (Types of Text) หากเป็นแบบทดสอบวัดการเขียน ชนิดของสาร อาจหมายถึง จดหมาย แบบฟอร์ม เรียงความเชิงวิชาการความยาว 3 หน้า เป็นต้น

ผู้รับสาร (Addressees) ส่วนนี้หมายถึงชนิดของบุคคลที่ผู้เข้าสอบต้องเขียนถึงหรือพูดด้วย ตัวอย่างเช่น ผู้พูดเจ้าของภาษาที่มีอายุและสถานะเท่ากัน หรือ บุคคลที่เป็นกลุ่มเป้าหมายในการรับสารที่เป็นบทอ่านหรือบทฟัง ตัวอย่างเช่น นักเรียนระดับมหาวิทยาลัยที่เป็นเจ้าของภาษา

หัวข้อ (Topics) หัวข้อในแบบทดสอบจะถูกคัดเลือกให้เหมาะสมกับผู้เข้าสอบ และชนิดของแบบทดสอบ

2) รูปแบบการทดสอบและเวลา (Formant and Timing) สิ่งนี้เป็นตัวกำหนดโครงสร้างแบบทดสอบ รวมทั้งเวลาที่ใช้ในแต่ละตอน และชนิดสารหรือกระบวนการให้การตั้งคำตอบ พร้อมทั้งให้ตัวอย่าง ส่วนนี้ควรกล่าวด้วยว่าแต่ละตอนในแบบทดสอบ จะถูกให้ค่าน้ำหนักคะแนนเท่าไร ควรกล่าวด้วยว่า บทที่จะให้อ่านหรือฟังมีเรื่องกี่ตอน (ในกรณีที่เป็นการทดสอบการอ่าน หรือการฟัง) หรือต้องการให้เขียนกี่เรื่องกี่ตอน (ในกรณีของการวัดการเขียน) และในแต่ละตอนจะมีแบบทดสอบกี่ข้อ

3) ระดับเกณฑ์ความสามารถ (Criterial Levels of Performance) ระดับความสามารถที่ต้องการสำหรับความสำเร็จในระดับต่าง ๆ ควรถูกกำหนดไว้ สิ่งนี้อาจเกี่ยวข้องกับค่ากล่าวง่าย ๆ ที่จะได้รับ เช่น เพื่อแสดงถึงระดับเชี่ยวชาญ (Mastery) ต้องตอบแบบทดสอบได้ถูกต้อง 80% สำหรับแบบทดสอบอัตนัย การกำหนดเกณฑ์จะซับซ้อนมากกว่า โดยอาจจะมีคำอธิบายที่บอกถึงระดับที่ยอมรับได้ในองค์ประกอบต่าง ๆ เช่น ถ้าเป็นการวัดความสามารถในการเขียน อาจมีการระบุความสามารถในองค์ประกอบที่เกี่ยวกับ ความถูกต้อง (Accuracy) ความเหมาะสม (Appropriacy) ความหลากหลายของโครงสร้างภาษา (Range) ความยืดหยุ่น (Flexibility) และปริมาณของภาษาที่ผลิตออกมา (Size)

4) เทคนิคการให้คะแนน (Scoring Procedures) สิ่งเหล่านี้จะมีความสำคัญมาก ในกรณีของการให้คะแนนแบบทดสอบอัตนัย นักพัฒนาแบบทดสอบควรชัดเจนว่าเราจะทำให้แบบทดสอบมีความเชื่อมั่นจากผู้ให้คะแนนสูงอย่างไร

สรุปได้ว่า รายละเอียดหัวข้อที่ควรปรากฏในตารางกำหนดลักษณะแบบทดสอบ คือ 1) ทักษะ (Skill)/จุดประสงค์การสอบ (Test Purpose) 2) ความสามารถทางภาษาที่ต้องการวัด (Language Ability) 3) ภาระงานสอบย่อย (Test Task) 4) หัวเรื่อง (Topic) 5) ประเภทเนื้อความ (Text Type) 6) วิธีการทดสอบ (Testing Technique) 7) ระดับความรู้ที่วัด (Level of Knowledge) 8) จำนวนข้อ (Number of Questions) 9) เวลา (Time) 10) น้ำหนักคะแนน (Weighting) และ 11) วิธีตรวจให้คะแนน (Scoring Method)

2.2.1.3 หลักการเขียนข้อคำถามในแบบทดสอบภาษาให้มีคุณภาพดี

Brown (1996) ได้ให้ข้อเสนอแนะในการการพัฒนาข้อคำถามทางภาษาที่มีคุณภาพไว้ 3 ประการ คือ 1) ข้อเสนอแนะทั่วไปสำหรับข้อคำถามในทุกรูปแบบ 2) คำแนะนำสำหรับข้อคำถามที่ให้ตอบแบบรับสาร (Receptive Response Item) และ 3) ข้อเสนอแนะสำหรับข้อคำถามที่ให้ตอบโดยการสื่อสาร (Productive Response Item) โดยมีรายละเอียดดังนี้

ข้อเสนอแนะทั่วไปสำหรับข้อคำถามในทุกรูปแบบ

ข้อเสนอแนะทั่วไปสำหรับข้อคำถามในทุกรูปแบบมีประเด็นคำถามที่ใช้ตรวจสอบแบบทดสอบ 10 ข้อ คือ 1) รูปแบบคำถามสอดคล้องเหมาะสมกับจุดประสงค์และสาระของข้อคำถามหรือไม่ 2) มีคำตอบที่ถูกต้องเพียงข้อเดียวใช่หรือไม่ 3) เขียนข้อคำถามเหมาะสมกับระดับความสามารถของผู้เข้าสอบหรือไม่ 4) มีการใช้คำศัพท์หรือภาษาที่เข้าใจยากกำกวมหรือไม่ 5) มีรูปปฏิเสธและปฏิเสธซ้อนปฏิเสธถูกใช้ในข้อคำถามหรือไม่ 6) ตัวข้อคำถามได้บอกไปคำตอบที่สามารถใช้ในการตอบข้อคำถามอื่นหรือไม่ 7) ส่วนต่างๆของข้อคำถามอยู่ในหน้ากระดาษเดียวกันหรือไม่ 8) ตัวข้อคำถามมียึดติดข้อมูลที่สำคัญ มีประโยชน์หรือไม่ 9) มีความลำเอียงที่อาจจะเกี่ยวกับเชื้อชาติ เพศ และสัญชาติหรือไม่ และ 10) ได้ให้เพื่อนร่วมงานอย่างน้อย 1 คน ตรวจสอบแบบทดสอบให้แล้วหรือไม่

คำแนะนำสำหรับข้อคำถามที่ให้ตอบแบบรับสาร (Receptive Response Item)

คำแนะนำสำหรับข้อคำถามที่ให้ตอบแบบรับสาร (Receptive Response Item) แบ่งออกเป็น 3 ชนิด คือ ข้อคำถามแบบให้ระบุ ถูก-ผิด (True-False) ข้อคำถามแบบเลือกตอบ (Multiple-Choice) และข้อคำถามแบบจับคู่ (Matching Items) โดยประเด็นคำถามสำหรับการตรวจสอบข้อคำถามแบบให้ระบุ ถูก-ผิด (True-False) มี 2 ข้อ คือ 1) การใช้คำในประโยค ข้อความเพียงพอที่จะให้พิจารณาหาคำตอบได้โดยไม่กำกวมใช่หรือไม่ และ 2) มีข้อความ กระเด็นที่เป็นตัวชี้หน้าที่เห็นได้ชัดเจนเพื่อให้ตอบถูกหรือไม่ ส่วนประเด็นคำถามสำหรับการตรวจสอบข้อคำถามแบบเลือกตอบ (Multiple-Choice) มี 4 ข้อ คือ 1) มีการให้ข้อความ กระเด็นที่เป็นตัวชี้คำตอบโดยไม่ตั้งใจหรือไม่ 2) ตัวลวงทุกตัวมีความเป็นไปได้ที่จะถูกเลือกหรือไม่ 3) มีข้อความ ข้อมูลฟุ่มเฟือยที่ไม่จำเป็นในตัวเลือกต่างๆหรือไม่ และ 4) การจัดเรียงตัวเลือกต่างๆได้ถูกตรวจสอบวางแผนอย่างดีแล้วหรือไม่ หรือการกำหนดตัวเลือกที่เป็นคำตอบกระจายกันอย่างทั่วถึง เท่ากัน หรือไม่ และประเด็นคำถามสำหรับการตรวจสอบข้อคำถามแบบจับคู่ มี 3 ข้อ คือ 1) มีตัวเลือกมากกว่าส่วนที่เป็นคำถามหรือไม่ 2) ตัวเลือกสั้นกว่าส่วนที่เป็นคำถามเพื่อลดภาระการอ่านหรือไม่ และ 3) รายการตัวเลือกและส่วนที่เป็นคำถามสัมพันธ์กันในหัวเรื่องเดียวกันหรือไม่

ข้อเสนอแนะสำหรับข้อคำถามที่ให้ตอบโดยการสื่อสาร (Productive Response Item)

ข้อแนะนำสำหรับข้อคำถามที่ให้ออกโดยการสื่อสาร (Productive Response Item) แบ่งออกเป็น 3 ชนิด คือ ข้อคำถามแบบเติมคำ (Fill-in) ข้อคำถามที่ให้ออกสั้นๆ (Short Response) และข้อคำถามภาระงาน (Tasks) โดยประเด็นคำถามสำหรับการตรวจสอบข้อคำถามแบบเติมคำ (Fill-in) มี 5 ข้อ คือ 1) คำตอบที่ต้องการกระชับหรือไม่ 2) มีบริบทเพียงพอที่จะช่วยให้ผู้เข้าสอบตอบได้หรือไม่ 3) ช่องว่างมีความยาวเป็นมาตรฐานเดียวกันหรือไม่ 4) ส่วนสำคัญ ประเด็นหลักของข้อคำถามอยู่หน้าช่องว่างหรือไม่ และ 5) มีการลองคิดถึงรายการคำตอบที่ยอมรับได้เป็นไปได้หรือไม่ ส่วนประเด็นคำถามสำหรับการตรวจสอบข้อคำถามที่ให้ออกสั้นๆ (Short Response) มี 2 ข้อ คือ 1) ตัวข้อคำถามถูกกำหนด ออกแบบเพื่อให้ได้คำตอบที่เป็นไปได้เพียงคำตอบเดียวหรือไม่ และ 2) ข้อคำถามถูกวางขอบเขตชัดเจนและเป็นคำถามตรงหรือไม่ และประเด็นคำถามสำหรับการตรวจสอบข้อคำถามภาระงาน (Tasks) มี 6 ข้อ คือ 1) ภาระงาน ภารกิจ ที่ผู้เข้าสอบต้องทำกล่าวไว้ชัดเจนหรือไม่ 2) ภาระงาน ภารกิจจำกัดแคบลง (และหรือกว้างขึ้น) เหมาะสมกับเวลาที่ให้ทำหรือไม่ 3) แผนการตรวจให้คะแนนถูกกำหนดล่วงหน้าสอดคล้องกับวิธีการที่จะใช้หรือไม่ 4) แผนการตรวจให้คะแนนถูกกำหนดล่วงหน้าตามองค์ประกอบด้านภาษาที่จะวัด 5) แผนการตรวจให้คะแนนถูกนิยามล่วงหน้าในแง่ของเกณฑ์การระดับคะแนนในแต่ละองค์ประกอบ และ 6) การตรวจให้คะแนนที่จะทำโดยไม่ต้องรู้ชื่อผู้เข้าสอบใช้หรือไม่

2.2.2 การประเมินคุณภาพแบบทดสอบวัดความสามารถภาษาอังกฤษ

วิธีการประเมินคุณภาพแบบทดสอบวัดความสามารถภาษาอังกฤษหรือแบบทดสอบที่มีจุดประสงค์อื่นมักจะใช้วิธีการประเมินที่เหมือนกัน คือ 1) ความเชื่อมั่น (Reliability) 2) ความตรง (Validity) 3) ความลำเอียงในการทดสอบ (Test Bias) 4) การวิเคราะห์แบบทดสอบรายข้อ (Item Analysis) และ 5) การสอบถามความคิดเห็นผู้เกี่ยวข้องโดยใช้แบบสอบถาม โดยแต่ละวิธีมีรายละเอียดดังต่อไปนี้

2.2.2.1 ความเชื่อมั่น (Reliability)

ความเชื่อมั่น (Reliability) เป็นคุณสมบัติอย่างหนึ่งที่ควรมีอยู่ในแบบทดสอบที่มีคุณภาพทุกชนิด การพัฒนาแบบทดสอบวัดความสามารถภาษาอังกฤษให้ได้มาตรฐานและมีคุณภาพ นักพัฒนาแบบทดสอบจำเป็นต้องเข้าใจความหมายของความเชื่อมั่น ปัจจัยที่ส่งผลต่อความเชื่อมั่น ชนิดของความเชื่อมั่นและการคำนวณหาความเชื่อมั่น ดังที่ผู้เชี่ยวชาญด้านการวัดและประเมินผลภาษาอังกฤษได้กล่าวไว้ ดังนี้

ความหมายของความเชื่อมั่น

ผู้เชี่ยวชาญด้านการวัดและประเมินผลภาษาอังกฤษ เช่น Heaton (1975) (Madsen & Harold S., 1983) (H. D. Brown, 2004) Brown (2004: 20) และ Cambridge English (2013: 27) ได้ให้คำนิยามความหมายของความเชื่อมั่นไว้ไม่แตกต่างกัน ดังนี้

Heaton (1975: 155) กล่าวว่า ความเชื่อมั่น (Reliability) เป็นลักษณะที่จำเป็นของแบบทดสอบที่ดีทุกชนิด เพราะก่อนที่แบบทดสอบจะวัดได้ตรง แบบทดสอบชุดนั้นต้องมีความน่าเชื่อถือที่จะใช้เป็นเครื่องวัด ถ้าแบบทดสอบถูกจัดสอบกับผู้เรียนคนเดิมในโอกาสที่ต่างกัน โดยที่ไม่มีการฝึกฝนภาษาเพิ่มเติมในระหว่างนั้น หลังจากนั้นพบว่าแบบทดสอบให้ผลลัพธ์ที่ต่างกัน อาจกล่าวได้ว่าแบบทดสอบชุดนั้นไม่มีความเชื่อมั่น

Madsen (1983: 179) กล่าวว่า แบบทดสอบที่มีความเชื่อมั่น (Reliability) คือ แบบทดสอบที่ให้ผลลัพธ์อย่างคงที่ในเวลาที่แตกต่างกัน โดยที่สภาวะแวดล้อมอื่นของแบบทดสอบยังคงเหมือนเดิม เช่น ให้ความเวลาในการทำแบบทดสอบแต่ละครั้งเท่ากัน ถ้าเป็นการทดสอบทักษะการฟัง ควรจัดห้องสอบให้ปราศจากเสียงรบกวนเท่ากันในการสอบแต่ละครั้ง เป็นต้น

Brown (2004: 20) กล่าวว่า แบบทดสอบที่น่าเชื่อถือ (A Reliable Test) จะมีความคงที่ของคะแนนและเชื่อถือได้ ถ้านำแบบทดสอบชุดเดียวกันไปใช้สอบกับนักเรียนคนเดิมหรือนักเรียนที่ถูกกำหนดให้สอบได้ใน 2 โอกาสที่ต่างกัน แบบทดสอบชุดนั้นควรแสดงผลลัพธ์ที่คล้ายกัน

Cambridge English (2013: 27) กล่าวว่า ความเชื่อมั่น (Reliability) เกี่ยวข้องกับระดับมากหรือน้อยที่ผลการสอบมีลักษณะคงที่ เทียบ และอิสระจากข้อผิดพลาดที่เกิดจากการวัด

สรุปได้ว่า ความเชื่อมั่น (Reliability) คือ คุณสมบัติอย่างหนึ่งที่แสดงให้เห็นว่า แบบทดสอบมีความชัดเจน เทียบ เชื่อถือได้ และคงที่ในการวัด โดยเมื่อนำแบบทดสอบชุดเดิมไปใช้สอบกับผู้เข้าสอบกลุ่มเดิมเป็นครั้งที่สองหรือมากกว่า คะแนนที่ได้จากการทดสอบซ้ำควรมีลักษณะไม่แตกต่างหรือให้ผลลัพธ์ที่คล้ายกัน และอิสระจากข้อผิดพลาดที่เกิดจากการวัด

วิธีการหาค่าความเชื่อมั่นและการคำนวณ

เนื่องจากความเชื่อมั่นของแบบทดสอบเกิดขึ้นจากปัจจัยที่เกี่ยวข้องกับการสอบหลายชนิด และความพยายามที่จะควบคุมปัจจัยเหล่านี้ให้คงที่เป็นที่มาของการคำนวณหาค่าความเชื่อมั่นด้วยวิธีการต่าง ๆ ดังที่ Henning (1987) ได้กล่าวถึงวิธีการหาค่าความเชื่อมั่นและการคำนวณไว้ดังนี้

Henning (1987) กล่าวว่า ปกติค่าความเชื่อมั่นจะอยู่ระหว่าง 0 ถึง 1 และมักจะใช้แทนสัดส่วนของคะแนนจริงกับความแปรปรวนของคะแนนที่สังเกตได้ การคำนวณหาค่าความเชื่อมั่นของคะแนนสามารถทำได้หลายวิธีการ การเลือกวิธีการคำนวณขึ้นอยู่กับปัจจัยต่าง ๆ เช่น ธรรมชาติของปัจจัยที่เกี่ยวข้องกับความเชื่อมั่น ความง่ายในการคำนวณ และธรรมชาติของแบบทดสอบ ต่อไปนี้จะเป็นการอธิบายวิธีการหาความเชื่อมั่นแบบต่าง ๆ พร้อมแสดงสูตรที่ใช้คำนวณ เพื่อให้เลือกใช้ให้เหมาะสมกับสถานการณ์

วิธีการใช้แบบทดสอบคู่ขนาน (Parallel Forms Method) เป็นการคำนวณความเชื่อมั่นจากการใช้แบบทดสอบ 2 ชุด กับผู้สอบคนเดียวกัน และผลที่ได้จากการสอบถูกนำมาหาความสัมพันธ์โดยใช้ Product Moment Correlation อย่างไรก็ตาม แบบทดสอบที่จะนำมาใช้ด้วยวิธีนี้ต้องมีความเท่ากันในประเด็นที่กำหนด แบบทดสอบเหล่านี้จึงถูกเรียกว่าแบบทดสอบที่มีรูปแบบคู่ขนานกัน (Parallel Forms) รูปแบบที่เท่ากัน (Equivalent Forms) หรือรูปแบบทางเลือก (Alternate Forms) เพื่อแสดงความเท่ากันของแบบทดสอบ แบบทดสอบเหล่านี้ต้อง 1) มีระดับความยากที่เท่ากัน โดยไม่มีความแตกต่างอย่างมีนัยสำคัญในคะแนนเฉลี่ยเมื่อแบบทดสอบถูกนำไปใช้สอบกับผู้เข้าสอบคนเดียวกันและคะแนนเฉลี่ยของผู้เข้าสอบถูกเปรียบเทียบโดยใช้ t-test 2) มีความแปรปรวนที่เท่ากัน เมื่อความแปรปรวนของคะแนนของแบบทดสอบทั้ง 2 ชุด ถูกเปรียบเทียบสำหรับกลุ่มตัวอย่างคนเดิม โดยการใช้ F-Max Test และ 3) แสดงความแปรปรวนร่วมที่เท่ากัน โดยไม่แตกต่างอย่างมีนัยสำคัญ ในค่าสัมประสิทธิ์สหสัมพันธ์ในกลุ่มของรูปแบบแบบทดสอบที่เท่ากัน หรือในกลุ่มค่าสัมประสิทธิ์สหสัมพันธ์ของรูปแบบที่เท่ากันกับเกณฑ์คู่ขนาน ทั้งหมดใช้สอบกับผู้เข้าสอบคนเดิมและเปรียบเทียบโดยการใช้ t-test โดยทั่วไป กระบวนการสำหรับการคำนวณค่าความเชื่อมั่นโดยใช้แบบทดสอบคู่ขนาน คือ การจัดให้ผู้เข้าสอบคนเดิม ทำแบบทดสอบ 2 ชุด ในเวลาเดียวกัน และหาความสัมพันธ์ของผลการสอบตามสูตรต่อไปนี้

$$r_{tt} = r_{A, B}$$

โดยที่ r_{tt} = ค่าสัมประสิทธิ์ความเชื่อมั่น
 $r_{A, B}$ = ความสัมพันธ์ของคะแนนที่ได้จากแบบทดสอบชุด A และที่ได้จากแบบทดสอบชุด B เมื่อจัดสอบกับผู้สอบคนเดียวกันในเวลาเดียวกัน

ความเชื่อมั่นระหว่างผู้ให้คะแนน (Inter-Rater Reliability) เป็นอีกวิธีการหนึ่งสำหรับการคำนวณความเชื่อมั่น วิธีการนี้ถูกนำมาใช้เมื่อคะแนนในแบบทดสอบมาจากผู้ให้คะแนนโดยอิสระจากผู้ตรวจหรือกรรมการ 2 คนหรือมากกว่า ซึ่งจะมีใช้กับแบบทดสอบวัดความสามารถในการเขียนเรียงความหรือการพูด ความเชื่อมั่นในวิธีนี้ถูกคำนวณจากความสัมพันธ์ของการให้คะแนนของกรรมการคนหนึ่งกับการให้คะแนนของกรรมการคนอื่นๆ และผลคะแนนสุดท้ายของผู้เข้าสอบเป็นส่วนผสมของการให้คะแนนที่มาจากกรรมการทุกคน ไม่ว่าจะป็นค่าเฉลี่ย หรือคะแนนที่มาจากการรวมกันแบบง่าย ๆ ระดับความเชื่อมั่นที่ปกติจะขึ้นอยู่กับจำนวนกรรมการ โดยทั่วไป การใช้กรรมการให้คะแนนหลายคน คะแนนในการสอบจะมีความน่าเชื่อถือมากยิ่งขึ้น วิธีการคำนวณหาความเชื่อมั่นระหว่างผู้ให้คะแนนมีข้อควรคำนึง 2 ประการ คือ 1) การให้คะแนนของกรรมการทุกคนต้องนำมาหาความสัมพันธ์รวมกัน ถ้ามีการใช้กรรมการมากกว่า 2 คนในการตรวจแบบทดสอบ ต้องหาค่าเฉลี่ยของสัมประสิทธิ์ความสัมพันธ์จากการให้คะแนนของกรรมการทุกคู่ โดยเฉพาะอย่างยิ่ง เมื่อค่าสัมประสิทธิ์ความสัมพันธ์ห่างกันมาก หรือต่ำกว่า .50 และ 2) ค่าสัมประสิทธิ์ความสัมพันธ์หรือค่าสัมประสิทธิ์ความสัมพันธ์เฉลี่ยถูกปรับโดยการใช้ The Spearman-Brown Prophecy Formula เพื่อให้ค่า

ความเชื่อมั่นขั้นสุดท้ายสะท้อนจำนวนของผู้ให้คะแนนหรือกรรมการที่ร่วมให้คะแนนผู้เข้าสอบ
ขั้นตอนเหล่านี้จะถูกสรุปไว้ในสูตรข้างล่างนี้

$$r_{tt} = \frac{nr_{A,B}}{1 + (n-1)r_{A,B}}$$

โดยที่ r_{tt} = ค่าสัมประสิทธิ์ความเชื่อมั่นระหว่างผู้ให้คะแนน
 n = จำนวนผู้ให้คะแนนที่ถูกนำมาใช้ในการคำนวณคะแนนสุดท้ายของผู้เข้าสอบ
 $r_{A,B}$ = ความสัมพันธ์ระหว่างผู้ให้คะแนน A และ ผู้ให้คะแนน B หรือ ค่าเฉลี่ย
 ความสัมพันธ์
 ระหว่างผู้ให้คะแนนแต่ละคู่ในกรณีที่มีผู้ให้คะแนนมากกว่า 2 คน

Kuder-Richardson Formula 20 ในกรณีที่ผู้ดำเนินการใช้วิธีการแบ่งครึ่งแบบทดสอบ และสามารถแบ่งแบบทดสอบออกเป็นครึ่งชุดแล้วแต่วิธีการแบ่ง และหาความสัมพันธ์ของคะแนนจากทั้ง
 2 ส่วน ทำให้ได้ค่าสัมประสิทธิ์ความสัมพันธ์ชุดหนึ่ง ถ้าผู้ดำเนินการปรับค่าสัมประสิทธิ์เหล่านี้ด้วย
 The Prophecy Formula และหลังจากนั้นหาค่าเฉลี่ยโดยใช้ Fisher Transformations จะได้รับค่า
 ความเชื่อมั่นที่จะควบคุมความผันแปรที่เกิดจากการแบ่งแบบทดสอบได้ วิธีการตามสูตร Kuder-
 Richardson Formula 20 ได้ค่าความเชื่อมั่นเหมือนกันโดยไม่ต้องคำนวณหาค่าความเชื่อมั่นของ
 แบบทดสอบทั้ง 2 ส่วน โดยมีสูตรในการคำนวณได้ดังนี้

$$r_{tt} = \frac{n}{(n-1)} \frac{(S_t^2 - \sum s_i^2)}{S_t^2}$$

โดยที่ r_{tt} = ค่าสัมประสิทธิ์ความเชื่อมั่นที่คำนวณได้โดย วิธี KR 20
 n = จำนวนข้อของแบบทดสอบ
 S_t^2 = ความแปรปรวนของคะแนนสอบทั้งหมด
 $\sum s_i^2$ = ผลรวมความแปรปรวนของคะแนนแต่ละข้อ

สูตรคำนวณตามวิธี KR 20 จะให้ผลลัพธ์เหมือนกันกับวิธีการคำนวณหาความสอดคล้อง
 ภายในของ Cronbach (Cronbach et al, 1963 อ้างถึงใน Henning, 1987: 84)

Kuder-Richardson Formula 21 ความผันแปรของวิธีการก่อนหน้า อาจจะถูกใช้เมื่อ
 เป็นไปไม่ได้หรือไม่ต้องการที่จะคำนวณความแปรปรวนของแต่ละข้อ อย่างไรก็ตาม ควรจำไว้ว่าวิธีถัด
 ของ KR 21 นี้ ถูกต้องน้อยกว่า KR 20 เพราะประมาณค่าความเชื่อมั่นต่ำกว่าจริงเล็กน้อย ทั้ง KR 20

และ KR 21 มีแนวโน้มที่จะให้ค่าประมาณที่เอนเอียง เมื่อแบบทดสอบมีจำนวนข้อไม่มาก วิธีการคำนวณ KR 21 มีสูตรดังนี้

$$r_{tt} = \frac{n}{(n-1)} \left\{ 1 - \frac{\bar{X} - \frac{\bar{X}^2}{n}}{S_t^2} \right\}$$

โดยที่ r_{tt} = ค่าสัมประสิทธิ์ความเชื่อมั่นที่คำนวณได้โดย วิธี KR 21
 n = จำนวนข้อคำถามในแบบทดสอบ
 \bar{X} = ค่าเฉลี่ยของคะแนนจากแบบทดสอบทั้งฉบับ
 S_t^2 = ความแปรปรวนของคะแนนสอบทั้งฉบับ

ความเชื่อมั่นของแบบทดสอบอัตนัย คำนวณโดยใช้วิธี Cronbach Alpha Coefficient ในรูปสัมประสิทธิ์สหสัมพันธ์แอลฟา (Alpha Coefficient) พัฒนามาจากสูตร KR-20 ซึ่งไม่จำเป็นต้องตรวจให้คะแนน แบบถูก 1 ผิด 0 ก็ได้ (Cronbach, 1951) โดยมีสูตรในการคำนวณ ดังนี้

$$\alpha = \frac{n}{n-1} \left\{ 1 - \frac{\sum S_i^2}{S_t^2} \right\}$$

โดยที่

α แทน ค่าความเชื่อมั่น
 n แทน จำนวนข้อของแบบทดสอบ
 S_i^2 แทน ความแปรปรวนรายข้อของแบบทดสอบ
 S_t^2 แทน ความแปรปรวนทั้งฉบับ

สรุปได้ว่า วิธีการคำนวณหาความเชื่อมั่นที่จะนำมาใช้ในการพัฒนาแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย มี 3 วิธี ได้แก่ 1) การหาค่าความเชื่อมั่นของแบบทดสอบปรนัย โดยใช้ Kuder-Richardson Formula 20 และค่าความเชื่อมั่นของคะแนนแบบทดสอบอัตนัยใช้ Cronbach Alpha Coefficients 2) ความเชื่อมั่นของแบบทดสอบอัตนัย โดยการหาความเชื่อมั่นระหว่างผู้ให้คะแนน (Inter-Rater Reliability) ด้วย The Spearman-Brown Prophecy Formula และ 3) หาความสัมพันธ์ของแบบทดสอบที่สร้างขึ้นกับแบบทดสอบมาตรฐานสากล โดยใช้ Pearson's Product Moment Correlation

2.2.2.2 ความตรงและการประเมินความตรง (Validity and Validation)

ความตรงเป็นคุณสมบัติอีกอย่างหนึ่งที่ควรมีอยู่ในแบบทดสอบทุกชนิด เพื่อแสดงให้เห็นว่าแบบทดสอบชนิดนั้นถูกสร้างขึ้นเพื่อให้วัดหรือทดสอบตรงกับจุดประสงค์ที่ได้กำหนดไว้ ในการพัฒนาแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัยให้มีคุณสมบัติด้านความตรง นักพัฒนาแบบทดสอบจำเป็นต้องเข้าใจความหมายของความตรง ชนิดของความตรง และแนวทางในการประเมินหรือวัดความตรงแต่ละชนิด ดังรายละเอียดต่อไปนี้

ความหมายของความตรง

นักวิชาการหรือผู้เชี่ยวชาญด้านการทดสอบภาษาอังกฤษ อาทิ Heaton (1975: 153) Madsen (1983: 178-179) Henning (1987: 89) Hughes (1989: 22) Cohen (1994: 38) Cambridge English (2013: 23) และ อัจฉรา วงศ์โสธร (2539) ดังรายละเอียดต่อไปนี้

Heaton (1975: 153) ความตรงของแบบทดสอบ (Validity) คือ ระดับมากหรือน้อยที่แบบทดสอบวัดในสิ่งที่ควรวัดและไม่ได้วัดสิ่งอื่นนอกเหนือจากนี้ แบบทดสอบทุกชนิดไม่ว่าจะเป็นแบบทดสอบสั้น แบบทดสอบที่ใช้ในห้องเรียนแบบไม่เป็นทางการ หรือแบบทดสอบที่ใช้โดยทั่วไป ควรมีความตรงเท่าที่นักสร้างแบบทดสอบสามารถทำได้ แบบทดสอบต้องมีเป้าหมายเพื่อเป็นเครื่องวัดที่แท้จริงของทักษะเฉพาะอย่างซึ่งแบบทดสอบถูกกำหนดให้วัด ถ้าปริมาณการวัดความรู้ภายนอกและวัดทักษะอื่นมีมาก แบบทดสอบนั้นจะเป็นเครื่องวัดที่ไม่มีความตรง

Madsen (1983: 178-179) กล่าวว่า แบบทดสอบที่มีความตรง (A Valid Test) คือแบบทดสอบที่วัดเนื้อหาตรงกับสิ่งอ้างไว้ว่าจะวัด ข้อทดสอบวัดทักษะการฟังที่มีคำตอบให้เลือกหลายตัวเลือกอาจขาดความตรง ถ้าตัวเลือกต่าง ๆ ที่ถูกพิมพ์เอาไว้ยากที่จะเข้าใจ โดยทำให้แบบทดสอบชนิดนั้นเหมือนกำลังวัดความเข้าใจในการอ่านเท่า ๆ กับการวัดความเข้าใจในการฟัง แบบทดสอบจะมีความตรงน้อยที่สุดสำหรับนักเรียนที่ทำได้ดีในการฟังมากกว่าการอ่านในทำนองเดียวกัน แบบทดสอบในการอ่านจะขาดความตรง ถ้าความสำเร็จในแบบทดสอบขึ้นอยู่กับข้อมูลที่ไม่ได้มีอยู่ในบทอ่าน ตัวอย่างเช่น ความคุ้นเคยในวัฒนธรรมอังกฤษหรืออเมริกัน เป็นต้น

Henning (1987: 89) กล่าวว่า ความตรง (Validity) คือ ความเหมาะสมของแบบทดสอบที่ใช้สอบหรือส่วนหนึ่งส่วนใดของแบบทดสอบที่เป็นเครื่องวัดในสิ่งที่ถูกกำหนดให้วัด

Hughes (1989: 22) กล่าวว่า แบบทดสอบจะได้รับการยอมรับว่าตรง ถ้าแบบทดสอบนั้นวัดอย่างถูกต้องในสิ่งที่ถูกกำหนดให้วัด

Cohen (1994: 38) ให้ความหมายของความตรง (Validity) ว่าเป็นสิ่งที่เกิดขึ้นเมื่อเครื่องมือได้วัดในสิ่งที่ได้ระบุหรือกำหนดไว้

Cambridge English (2013: 23) ให้ความหมายของความตรง (Validity) ว่า เป็นระดับที่เครื่องมือประเมินสามารถแสดงให้เห็นว่าคะแนนหรือผลการสอบเป็นตัวสะท้อนที่ถูกต้องของระดับความสามารถที่แท้จริงของผู้เข้าสอบ

อัจฉรา วงศ์โสธร (2539) กล่าวว่า แบบทดสอบที่มีความตรง ต้องมีความเที่ยงเป็นพื้นฐาน และสามารถวัดได้ตรงตามวัตถุประสงค์ของการวัด ถ้าเปรียบกับการยิงลูกศร นอกจากว่าลูกศรทุกลูกจะไปสู่จุดเดียวกันแล้ว ยังจะต้องเข้าเป้าที่กำหนดด้วย ค่าความตรงนั้นใช้เพื่อนำผลคะแนนที่ได้จากการทดสอบไปใช้พยากรณ์ความสามารถของผู้เรียน และช่วยให้การตีความหมายของคะแนน ผลการทดสอบถูกต้องตามความมุ่งหมายของการทดสอบที่กำหนดไว้

สรุปได้ว่า ความตรง คือ คุณสมบัติอย่างหนึ่งของแบบทดสอบที่แสดงให้เห็นว่าแบบทดสอบวัดได้ตรงตามจุดประสงค์ที่ได้กำหนดไว้ อีกทั้งผลคะแนนที่ได้ยังสามารถช่วยพยากรณ์ความสามารถของผู้เรียน และช่วยให้การตีความในความหมายของคะแนนผลการสอบถูกต้องตามความจริงที่ควรจะเป็น

ชนิดของความตรงและการประเมินความตรง

นักวิชาการด้านการทดสอบภาษาอังกฤษแบ่งชนิดของความตรงเป็นหมวดหมู่ที่แตกต่างกัน เช่น Cohen (1994: 38-42) แบ่งความตรงออกเป็น 5 ชนิด ในขณะที่ Alderson et al. (1995: 172-186) และ Genesee & Upshur (1996: 250-252) แบ่งความตรงออกเป็น 3 ชนิด ดังรายละเอียดต่อไปนี้

Cohen (1994: 38-42) กล่าวว่า ความตรง หมายถึง สภาวะที่เครื่องมือวัดได้ตรงในสิ่งที่ถูกกำหนดให้วัด และแบบทดสอบต้องเชื่อถือได้ก่อนถึงจะมีความตรง การประเมินความตรงของแบบทดสอบสามารถทำได้หลายวิธี เช่น การประเมินความตรงเชิงเนื้อหา (Content Validity) ความตรงที่สัมพันธ์กับเกณฑ์ (Criterion-related Validity) ความตรงเชิงโครงสร้าง (Construct Validity) ความตรงเชิงระบบ (Systemic Validity) และความตรงร่วมสมัย (Face Validity) โดยมีรายละเอียดในแต่ละวิธี ดังนี้

ความตรงเชิงเนื้อหา (Content Validity) ประเมินได้จากการตรวจสอบความเพียงพอหรือครอบคลุมที่แบบทดสอบได้บรรจุเนื้อหาหรือวัตถุประสงค์ของรายวิชาหรือประเด็นที่ได้รับการประเมินในแบบทดสอบเชิงพาณิชย์ (Commercial Tests) ที่พัฒนาโดยผู้เชี่ยวชาญ บางครั้งยังไม่สามารถระบุได้ว่าครอบคลุมวัตถุประสงค์ใดบ้าง และแบบทดสอบข้อใดที่วัดตรงตามแต่ละวัตถุประสงค์ ด้วยเหตุนี้ครูผู้สอน จึงควรได้รับการกระตุ้นให้ตรวจสอบอย่างรอบคอบว่าแบบทดสอบแต่ละข้อวัดเนื้อหาหรือจุดประสงค์ใด โดยปกติแล้ว การประเมินความตรงเชิงเนื้อหา (Content Validity) สามารถทำได้โดยการตรวจสอบสำเนาของแบบทดสอบข้อต่อข้อ ตรวจสอบตารางกำหนดลักษณะแบบทดสอบ และทบทวนรายการเนื้อหาสาระที่จะใช้สอบ ภาระงานที่เกี่ยวข้องคือการย้อนกลับไปดูคานียามเชิงปฏิบัติการของโครงสร้างปริเฉทความสามารถที่ต้องการวัด และลักษณะของวิธีการทดสอบที่ใช้ (Bachman, 1990)

ความตรงที่สัมพันธ์กับเกณฑ์ (Criterion-related Validity) ใช้สำหรับการตรวจสอบว่าความสามารถของผู้เข้าสอบในวัตถุประสงค์กลุ่มใดกลุ่มหนึ่งตามเครื่องมือวัดที่ได้สอบไปแล้ว สัมพันธ์หรือสอดคล้องกับความสามารถของผู้เข้าสอบจากการใช้เครื่องมือวัดอื่นหรือเกณฑ์อื่นอย่างไร ซึ่งใช้ในการวัดกิจกรรมเดียวกันหรือใกล้เคียงกัน ถ้าเครื่องมือวัดที่กำลังทดสอบสัมพันธ์หรือสอดคล้องกับเครื่องมือวัดที่ใช้เกณฑ์อื่นในเวลาใกล้เคียงกัน วิธีการนี้จะถูกเรียกว่าการประเมินความตรงร่วมสมัยหรือ

ความตรงร่วมสมัย (Concurrent Validation) แต่หากความสัมพันธ์นั้นเกิดขึ้น ณ เวลาใดเวลาหนึ่งในอนาคต กระบวนการนั้นจะถูกเรียกว่า การประเมินความตรงเชิงพยากรณ์ (Predictive Validation) ตัวอย่างของการประเมินความตรงร่วมสมัย เช่น ครูอาจจะอยากรู้ว่าผลการสอบของนักเรียนจากแบบทดสอบที่ครูสร้างขึ้น เปรียบเทียบกับความสามารถของนักเรียนโดยการใช้เกณฑ์อื่นที่ใช้วัดการอ่านจากการใช้แบบทดสอบเชิงพหุวิธี แม้ว่า จะมีความสัมพันธ์กันในระดับสูง เกณฑ์ภายนอกที่ใช้ อาจจะยังไม่ใช่ตัวบ่งชี้ความสามารถที่ต้องการวัด สิ่งหนึ่งที่ต้องพิจารณาคือ แบบทดสอบอื่น จำเป็นต้องมีความตรงเชิงโครงสร้าง (Construct Validity) เพื่อให้ความสัมพันธ์นี้มีคุณค่า นอกจากนั้น ความสัมพันธ์ระดับสูงระหว่างแบบทดสอบ 2 ชนิด ไม่ได้ไปตัวบอกว่าแบบทดสอบฉบับใดเป็นที่ต้องการมากกว่า ถ้าแบบทดสอบฉบับนั้นไม่ตรงกับวัตถุประสงค์การใช้ หรือสามารถใช้แทนแบบทดสอบฉบับอื่นได้ และตัวอย่างสำหรับการประเมินความตรงเชิงพยากรณ์ (Predictive Validation) เช่น แบบทดสอบวัดความถนัดทางภาษา อาจจะใช้เทียบเคียงกับแบบทดสอบวัดผลสัมฤทธิ์ของนักเรียน ในด้านภาษาที่ใช้ในห้องเรียน ซึ่งนักเรียนถูกจัดเข้าชั้นโดยอาศัยความถนัดทางภาษา ปัญหาในประเด็นนี้ คือ แบบทดสอบจัดเข้าชั้นหรือแบบทดสอบเชิงพยากรณ์ อาจจะเป็นเครื่องวัดความสามารถในวิชาเฉพาะที่ไม่เหมาะสม Bachman (1990) กล่าวว่า ยังมีปัจจัยต่าง ๆ จำนวนมากที่เกี่ยวข้องกับการพยากรณ์ผลสัมฤทธิ์ในห้องเรียนภาษา

ความตรงเชิงโครงสร้าง (Construct Validity) คือ ระดับที่คะแนนจากการสอนส่งผลต่อการอ้างอิงถึงคุณสมบัติ (ภาพลักษณ์) ที่ต้องการ (เป็นพื้นฐาน) หรือพูดได้ว่า เป็นตัวตรวจสอบว่าเครื่องมือได้สะท้อนอย่างแท้จริงถึงทฤษฎีที่รองรับคุณสมบัติ/ภาพลักษณ์ที่กำลังถูกวัด การประเมินความตรงเชิงโครงสร้างเป็นการทดสอบความสัมพันธ์เชิงสมมติฐานและระหว่างคะแนนสอบและความสามารถ เป็นกิจกรรมที่ต้องการทั้งการวิเคราะห์อย่างมีเหตุผล เช่น การให้คำนิยามเชิงโครงสร้างตามแนวทฤษฎีและการปฏิบัติ และการสำรวจอย่างมีหลักฐาน และด้วยความสำคัญของการประเมินความตรงเชิงโครงสร้าง ครูหรือนักสร้างแบบทดสอบ จำเป็นต้องถามตนเอง เมื่อจะสร้างเครื่องมือวัดประเมินว่า ข้อคำถามและกระบวนการในการประเมินสะท้อนสิ่งที่บ่งบอกถึงการรู้ภาษาหรือไม่ ให้ตรงข้ามกับภาระงานที่วัดบางอย่างที่แตกต่างจากความรู้ด้านภาษาจริง ๆ

ความตรงเชิงระบบ (Systemic Validity) เกิดจากการขยายความเพิ่มเติมในแง่มุมของความตรงเชิงโครงสร้าง โดยให้ความสำคัญกับผลกระทบของการเรียนการสอนที่เกิดขึ้นจากการนำแบบทดสอบไปสู่ระบบการศึกษา จึงเรียกความตรงในลักษณะนี้ว่า ความตรงเชิงระบบ (Systemic Validity) ตามความคิดเห็นของ Frederiksen and Collins (1989:27 อ้างถึงใน Cohen (1994: 41) ที่กล่าวว่า “แบบทดสอบที่มีความตรงเชิงระบบ คือ แบบทดสอบที่นำมาใช้ในวงจรของระบบการศึกษาและการเปลี่ยนแปลงด้านการเรียนการสอน และส่งเสริมให้มีการพัฒนาทักษะทางปัญญาที่แบบทดสอบต้องการจะวัด หลักฐานสำหรับความตรงเชิงระบบจะเป็นการปรับปรุงในทักษะเหล่านั้น หลังจากที่ใช้แบบทดสอบได้ถูกนำสู่ระบบการศึกษามาสักระยะหนึ่งแล้ว” คุณภาพของผลกระทบไม่ว่าจะเป็นด้านบวกหรือด้านลบ อาจจะเป็นอิสระจากคุณภาพของแบบทดสอบ และแนะนำว่า การพิจารณาผลกระทบของแบบทดสอบที่มีต่อการสอนและ/หรือการเรียนจริง ควรมีการพิจารณาในเรื่อง 1) ผลกระทบของแบบทดสอบต่อเนื้อหาสาระในการเรียนการสอน 2) กระบวนการที่เนื้อหาสาระถูกสอนออกไป 3) อัตราที่เนื้อหาได้รับการสอนอย่างครอบคลุม 4) ธรรมชาติของการจัดลำดับเนื้อหา

และ 5) ปริมาณของเนื้อหาที่นักเรียนได้เรียนจริงและความลึกที่นักเรียนได้เรียน และปริมาณข้อคำถามที่ใช้สอบหรือวัตถุประสงค์

ความตรงเชิงประจักษ์ (Face Validity) หมายถึง จากการดูแบบทดสอบแล้วพบว่าแบบทดสอบได้วัดสิ่งที่ถูกสร้างให้วัดหรือไม่ ตัวอย่างเช่น แบบทดสอบที่วัดการออกเสียงภาษาอังกฤษของตัวผู้สอบเอง โดยประเมินการให้คะแนนการออกเสียงภาษาอังกฤษของผู้สอบคนอื่น อาจจะไม่ได้รับการยอมรับว่าเป็นการวัดที่มีความตรง เช่นเดียวกับการเติมคำในช่องว่างของแบบทดสอบแบบ Cloze Test เพื่อประเมินทักษะในการอ่าน การประเมินเช่นนั้นดูเหมือนว่าจะอ้อมเกินไป การประเมินแบบอ้อมอาจจะทำให้ผู้เข้าสอบสับสนและไม่ดึงดูดใจ

Alderson et al (1995: 172-186) แบ่งความตรงออกเป็น 3 ชนิดหลัก คือ ความตรงภายใน (Internal Validity) ความตรงภายนอก (External Validity) และความตรงเชิงทฤษฎี (Construct Validity) ซึ่งความตรงแต่ละชนิดมีความแตกต่างกันทั้งในด้านความหมายและวิธีการประเมิน ดังรายละเอียดต่อไปนี้

ความตรงภายใน (Internal Validity) เป็นความตรงชนิดที่รู้จักกันมากกว่าชนิดอื่น ๆ การประเมินความตรงภายในสามารถทำได้หลายวิธี แต่วิธีที่ใช้กันมากที่สุดมีอยู่ 3 วิธี คือ การประเมินความตรงเชิงประจักษ์ (Face Validation) ซึ่งผู้ที่ไม่ใช่ผู้ทดสอบ เช่น นักเรียน และผู้บริหาร แสดงความเห็นต่อคุณค่าของแบบทดสอบ การประเมินความตรงเชิงเนื้อหา (Content Validation) ที่นักทดสอบหรือนักภาษาศาสตร์ หรือผู้เชี่ยวชาญเฉพาะด้านประเมินแบบทดสอบ และการประเมินความตรงในการตอบสนอง (Response Validation) ซึ่งมีการใช้วิธีการเชิงคุณภาพจำนวนมาก เช่น การรายงานตนเอง หรือการสังเกตตนเองในส่วนของผู้เข้าร่วม ถูกใช้เพื่อให้เข้าใจว่าพวกเขาจะตอบแบบทดสอบอย่างไรและด้วยเหตุผลอะไร โดยการประเมินความตรงภายในแต่ละวิธี มีรายละเอียดเพิ่มเติมดังต่อไปนี้

ความตรงเชิงประจักษ์ (Face Validity) หมายถึง ความเชื่อถือโดยผิวเผิน หรือการยอมรับโดยทั่วไป และมักจะถูกกละเลยโดยนักทดสอบ เพราะไม่เกี่ยวข้องกับวิธีการทางวิทยาศาสตร์และไม่สำคัญ โดยส่วนใหญ่แล้วความตรงเชิงประจักษ์เกี่ยวข้องกับการตัดสินใจโดยไม่รู้ตัวเกี่ยวกับสาระแบบทดสอบโดยบุคคลที่คำตัดสินของพวกเขาไม่ได้เป็นผู้เชี่ยวชาญ บุคคลเหล่านั้น อาจจะเป็นผู้บริหาร ผู้ใช้แบบทดสอบที่ไม่ใช่ผู้เชี่ยวชาญและนักเรียน คำตัดสินมักจะใช้เป็นในภาพรวม อ้างถึงแบบทดสอบทั้งหมด แม้ว่าความสนใจอาจจะเน้นตรงที่ข้อที่ไม่ดีโดยเฉพาะ คำสั่งที่ไม่ชัดเจน หรือการจำกัดเวลาที่ไม่สมจริง เป็นเหมือนวิธีการปรับเปลี่ยนคำตัดสินในภาพรวมเกี่ยวกับแบบทดสอบ บางครั้งข้อมูลความตรงเชิงประจักษ์อาจจะถูกรวบรวมได้โดยการสัมภาษณ์นักเรียน หรือขอให้พวกเขาตอบแบบสอบถามเกี่ยวกับทัศนคติ หรือปฏิกิริยาต่อ และความรู้สึกเกี่ยวกับแบบทดสอบที่พวกเขาเพิ่งได้ทำไปหรือได้เห็น ผลที่ได้สามารถสรุปเป็นสถิติ ถ้าต้องการ และการยอมรับข้อคำถาม และส่วนต่าง ๆ ของแบบทดสอบสามารถถูกตัดสินใจได้

ความตรงเชิงเนื้อหา (Content Validity) คือ ความเป็นตัวแทน หรือ ความเพียงพอของการเป็นกลุ่มตัวอย่างของเนื้อหา – สาระ ประเด็น หัวเรื่อง – ของเครื่องวัด การประเมินความตรงเชิงเนื้อหาเกี่ยวข้องกับการเก็บรวบรวมคำตัดสินของผู้เชี่ยวชาญ บุคคลที่คำตัดสินของพวกเขามีคนไว้ใจหรือเชื่อถือ แม้จะขัดแย้งกับคำตัดสินของตนเอง (การเก็บข้อมูลจากคนที่เราไว้ใจ หรือ เป็น

ผู้เชี่ยวชาญ) โดยปกติแล้ว การประเมินความตรงเชิงเนื้อหาเกี่ยวข้องกับผู้เชี่ยวชาญทำการตัดสินด้วยวิธีการอย่างเป็นระบบ วิธีการที่พบบ่อย คือ ให้พวกเขาวิเคราะห์สาระของแบบทดสอบ และให้เปรียบเทียบคำกล่าวหรือคำอธิบายที่บอกว่า เนื้อหาควรเป็นอะไร คำกล่าวเกี่ยวกับสาระนั้นนี้อาจจะเป็นการกำหนดโครงสร้างแบบทดสอบ อาจจะเป็นประมวลการสอนอย่างเป็นทางการ หรือ หลักสูตร หรือมันอาจจะเป็นโครงสร้างปริเฉทของแบบทดสอบ

ความตรงตามการตอบสนอง (Response Validity) ตามที่ได้กล่าวไว้แล้วข้างต้น แง่มุมของการประเมินที่มีเพิ่มมากขึ้นบ่อยในระยาะนี้ คือ การเก็บข้อมูลเกี่ยวกับว่าผู้เข้าสอบแต่ละคนตอบแบบทดสอบอย่างไร กระบวนการที่ผู้เข้าสอบใช้ เหตุผลที่พวกเขาใช้ขณะกำลังทำแบบทดสอบ เป็นตัวบ่งชี้ที่สำคัญของสิ่งที่แบบทดสอบกำลังทดสอบ อย่างน้อยที่สุดคือบุคคลเหล่านั้น ดังนั้น มีความสนใจปัจจุบันจำนวนมากในการเก็บข้อมูลจากผู้เรียนหรือผู้เข้าสอบเกี่ยวกับพฤติกรรมและความคิดในการทำแบบทดสอบของพวกเขา

การเก็บข้อมูลแบบ Introspective ในระหว่างการทำแบบทดสอบไม่ใช่เรื่องง่าย เนื่องจาก การเก็บข้อมูลสามารถไปรบกวนกระบวนการที่กำลังถูกสำรวจอยู่ อย่างไรก็ตาม งานวิจัยได้เปิดเผยมุมมองที่น่าสนใจในผลการสอบโดยผ่านวิธีการที่เน้นผู้เรียนเป็นศูนย์กลางเช่นนั้น ตัวอย่างเช่น การสะท้อนคิด (Introspection) ในภาระงานแบบ Cloze แสดงว่านักเรียนต้องตอบคำถามแต่ละข้อโดยใช้ทักษะการอ่านที่นักออกแบบแบบทดสอบตั้งใจไว้หรือไม่ หรือสิ่งที่ต้องการทั้งหมดคือความรู้ด้านโครงสร้างไวยากรณ์ของวลีที่ปรากฏในแบบทดสอบหรือไม่ ในทำนองเดียวกัน Introspection ในระหว่างการอ่านที่วัดความเข้าใจในการอ่าน อาจจะบ่งชี้จุดอ่อนในข้อคำถาม และอาจจะสร้างกรณีต่าง ๆ ที่นักเรียนตอบแบบทดสอบผิดแม้ว่าจะเข้าใจในบทอ่าน หรืออาจจะตอบถูกแม้ว่าไม่เข้าใจบทอ่าน

การเก็บข้อมูลการสะท้อนคิดของนักเรียนควรดำเนินการอย่างไร วิธีง่ายที่สุด คือ เก็บข้อมูลหลังจากสอบ (ภายหลัง) หลังจากที่ผู้เข้าสอบได้ทำแบบทดสอบเสร็จแล้ว หรือ ส่วนประกอบของแบบทดสอบ พวกเขาควรได้รับการสัมภาษณ์เกี่ยวกับเหตุผลว่าทำไมพวกเขาตอบแบบนั้น ตัวกระตุ้นสำคัญสำหรับความทรงจำของพวกเขาในคำตอบ คือ ข้อความในแบบทดสอบในกรณีของการอ่าน การเขียนและแบบทดสอบฟังก็เป็นไปได้ อย่างไรก็ตาม ในกรณีของการสอบพูด ข้อมูลที่มีประโยชน์ควรถูกเก็บได้ดีที่สุด โดยให้ผู้เข้าสอบพูดผ่านวิดีโอ หรือเครื่องบันทึกเสียง ในผลการสอบของพวกเขา ผู้ถามคำถามควรที่จะไม่ชี้แนะที่จะเป็นไปได้ ถามคำถามทั่ว ๆ ไป เช่น ทำไมคุณใส่คำตอบนั้น ทำไมคุณตอบแบบนั้น และควรให้ตัวกระตุ้นเฉพาะเมื่อคำตอบไม่ชัดเจนหรือไม่เจาะจงเพียงพอ สำหรับข้อดีของการสะท้อนความคิดหลังการสอบ คือว่า ผู้เข้าสอบอาจจำไม่ได้ว่าพวกเขาตอบแบบนั้นเพราะอะไร ทางเลือกคือการค้นหาวิธีที่คู่ขนานกัน ที่ให้ผู้เข้าสอบคิดตั้ง ๆ ขณะกำลังตอบ ให้ตัวกระตุ้นในช่วงเวลาที่เจียบโดยผู้สังเกตการณ์ที่เจียบ

ความตรงภายนอก (External Validation) ความตรงภายนอกชนิดที่รู้จักกันมากที่สุด คือ ความตรงร่วมสมัย (Concurrent Validity) และความตรงเชิงพยากรณ์ (Predictive Validity) และสถิติที่ใช้บ่อยมากที่สุด คือ การหาค่าสัมประสิทธิ์สหสัมพันธ์ แนวคิดของการหาความตรงภายนอกบางที่อาจเข้าใจได้ง่ายที่สุดโดยการอภิปรายผ่านความตรงร่วมสมัย

ความตรงร่วมสมัย (Concurrent Validity) การประเมินความตรงร่วมสมัย (Concurrent Validation) เกี่ยวข้องกับการเปรียบเทียบคะแนนสอบกับเครื่องวัดอย่างอื่นสำหรับผู้เข้าสอบคนเดิมที่สอบในเวลาใกล้เคียงกัน เครื่องวัดชนิดอื่นนี้อาจจะเป็นแบบทดสอบฉบับคู่ขนานของแบบทดสอบชุดเดียวกัน หรือแบบทดสอบชนิดอื่น การประเมินตนเองของผู้เข้าสอบด้านความสามารถทางภาษาของพวกเขา หรือการจัดอันดับของผู้เข้าสอบตามเกณฑ์หรือมิติต่าง ๆ โดยครู ผู้เชี่ยวชาญเฉพาะด้าน หรือผู้ให้ข้อมูลอื่น สิ่งสำคัญ คือ เครื่องวัดชนิดนี้สามารถแสดงผลเป็นตัวเลขได้ เหมือนกับการให้คะแนนตามมาตรฐานประมาณค่า และไม่ต้องสัมพันธ์กับการสอบใด ๆ ผลของการเปรียบเทียบมักจะแสดงออกมาในรูปค่าสัมประสิทธิ์ความสัมพันธ์ มีค่าระหว่าง -1.0 ถึง +1.0 ค่าสัมประสิทธิ์ความตรงร่วมสมัยส่วนใหญ่มีค่าตั้งแต่ +.5 ถึง +.7 เพราะค่าความสัมพันธ์สูงพบได้สำหรับแบบทดสอบที่สัมพันธ์กันอย่างใกล้ชิดและเป็นแบบทดสอบที่เชื่อถือได้ แต่มักจะไม่ใช้การวัดที่มาจาก การประเมินตนเองหรือการประเมินโดยครู

ความตรงเชิงพยากรณ์ (Predictive Validity) ความตรงชนิดนี้แตกต่างจากความตรงร่วมสมัยในประเด็นที่ว่า การที่จะเก็บข้อมูลที่ได้จากเครื่องวัดภายนอกในเวลาเดียวกันกับการใช้แบบทดสอบที่อยู่ระหว่างการทดสอบ เครื่องวัดภายนอกจะถูกใช้เก็บข้อมูลในช่วงเวลาอื่นหลังจากการใช้แบบทดสอบที่อยู่ระหว่างการทดสอบ การประเมินความตรงเชิงพยากรณ์ พบบ่อยมากที่สุดกับแบบทดสอบวัดความสามารถโดยทั่วไป แบบทดสอบที่ถูกใช้เพื่อทำนายว่าผู้เข้าสอบจะทำได้ดีเพียงใดในอนาคต วิธีการที่ง่ายที่สุดของการประเมินความตรงเชิงพยากรณ์ คือ การให้นักเรียนทำแบบทดสอบ และเมื่อถึงเวลาที่เหมาะสมในอนาคต ให้นักเรียนทำแบบทดสอบอีกชุดหนึ่งที่วัดความสามารถที่แบบทดสอบชุดแรกต้องการจะทำนาย

ความตรงเชิงโครงสร้าง (Construct Validity) Ebel and Frisbie (1991 อ้างถึงใน Alderson et al, 1995: 183) กล่าวว่า คำว่า “โครงสร้าง” หมายถึง โครงสร้างทางจิตวิทยา แนวคิดทางทฤษฎีเกี่ยวกับแง่มุมของพฤติกรรมมนุษย์ที่ไม่สามารถวัดได้หรือสังเกตได้โดยตรง ตัวอย่างของโครงสร้าง คือ ความฉลาด แรงจูงใจต่อผลสำเร็จ ความกังวล ผลสัมฤทธิ์ ทศนคติ ความโดดเด่น และความเข้าใจในการอ่าน การประเมินความตรงเชิงทฤษฎี (Construct Validation) คือกระบวนการหาหลักฐานเพื่อสนับสนุนความขัดแย้งที่ว่า แบบทดสอบที่ใช้จริง ๆ แล้ววัดโครงสร้างทางจิตวิทยาที่ผู้สร้างต้องการให้แบบทดสอบวัด เป้าหมายคือเพื่อตัดสินความหมายของคะแนนจากการสอบ เพื่อให้มั่นใจว่า คะแนนสื่อความหมายถึงสิ่งที่เราคาดหวังให้มันหมายถึง การประเมินความตรงเชิงทฤษฎีสามารถทำได้หลายวิธี เช่น การเปรียบเทียบกับทฤษฎี (Comparison with Theory) การหาความสัมพันธ์ภายใน (Internal Correlations) การเปรียบเทียบกับข้อมูลประวัติของนักเรียน และลักษณะทางจิตวิทยา (Comparison with Students' Biodata and Psychological Characteristics) การวิเคราะห์แบบพหุคูณ-พหุวิธี และการประเมินความตรงแบบลู่เข้า-แบบกระจายออก (Multitrait-Multi Method Analysis and Convergent-Divergent Validation) และการวิเคราะห์ปัจจัย (Factor Analysis) โดยมีรายละเอียดในแต่ละวิธีดังนี้

การเปรียบเทียบกับทฤษฎี (Comparison with Theory) สำหรับนักทฤษฎีการทดสอบบางคน ความตรงเชิงทฤษฎี คือ รูปแบบหนึ่งของการประเมินความตรงของแบบทดสอบซึ่งจำเป็นต้องเกี่ยวข้องกับการหาระดับมากน้อยที่แบบทดสอบวัดทฤษฎีพื้นฐานของมันอย่างประสบความสำเร็จ

ที่สามารถจัดการได้มากขึ้นด้วยวิธีการทางสถิติ ผลที่ได้จากการลดความซับซ้อนนั้น ปกติทำให้ได้ปัจจัยที่มีจำนวนน้อยลง แบบทดสอบต่าง ๆ ที่ถูกเปรียบเทียบกันในระบบเมตริกซ์ความสัมพันธ์แบบเดิม จะสัมพันธ์กันในวิธีต่าง ๆ ตามปัจจัยหลายชนิด การวิเคราะห์ปัจจัยมี 2 ชนิดใหญ่ ๆ คือ ชนิดหนึ่งเรียกว่าการวิเคราะห์ปัจจัยแบบสำรวจ (Exploratory Factor Analysis- EFA) และอีกชนิดหนึ่งเรียกว่าการวิเคราะห์ปัจจัยแบบยืนยัน (Confirmatory Factor Analysis- CFA) สำหรับในวิธีแรก (EFA) ผู้จัดทำสำรวจข้อมูลที่จะทดลองใช้ และทำความเข้าใจกับปัจจัยที่ปรากฏ สิ่งนี้มักถูกทำโดยการตรวจดูว่าแบบทดสอบใดสัมพันธ์ใกล้ชิดกับปัจจัยใดมากที่สุดและบันทึกปัจจัยเหล่านั้นเอาไว้ ส่วนในวิธีที่สอง (CFA) ผู้วิจัยทำนายว่าแบบทดสอบใดหรือตอนใด จะสัมพันธ์กับแบบทดสอบอื่นหรือตอนอื่นและอย่างไร และหลังจากนั้น ดำเนินการทดสอบความเหมาะสมของคำทำนายกับข้อมูล เหตุผลที่วิธีการวิเคราะห์ปัจจัย 2 วิธีนี้ อยู่ในกระบวนการประเมินความตรงเชิงทฤษฎี คือว่า ปัจจัยต่าง ๆ ที่ปรากฏถูกอธิบายหรือทำนายโดยอาศัยพื้นฐานทฤษฎี ในกรณีนี้คือ ทฤษฎีด้านภาษาศาสตร์ประยุกต์ของผู้เชี่ยวชาญที่บอกว่าอะไรสัมพันธ์กับอะไร

Genesee & Upshur (1996: 250-252) เสนอแนวคิดเกี่ยวกับการประเมินและรายงานความตรง (Assessing and Reporting Validity) ว่า ผู้จัดพิมพ์แบบทดสอบมาตรฐานรับผิดชอบสำหรับการประเมินความตรงของแบบทดสอบและสำหรับการรายงานข้อมูลนี้ในคู่มือแบบทดสอบ ข้อมูลอื่น ๆ เกี่ยวกับความตรงอาจพบได้ในวารสารหรือรายงานการวิจัย ผู้ใช้ไม่ควรนำแบบทดสอบมาตรฐานมาใช้เพื่อการตัดสินใจสำคัญ โดยปราศจากการพิจารณาข้อมูลที่มีอยู่เกี่ยวกับความตรงของแบบทดสอบ ถ้าไม่มีการแสดงข้อมูลความตรงสำหรับจุดประสงค์การใช้ที่กำหนดไว้ การใช้แบบทดสอบนั้นต้องพิจารณานำมาทดลองใช้ก่อน

การประเมินความตรงสามารถทำได้หลายวิธี เช่น การประเมินความสัมพันธ์ของเนื้อหา (Content Relevance) ความสัมพันธ์ของเกณฑ์ (Criterion Relatedness) และการวิจัยประเมินโครงสร้าง (Construct Validation Research) นอกจากนี้มีความประทับใจหรือความรู้สึกแบบไม่รู้ตัวที่แบบทดสอบวัดตรง หลักฐานของความตรงประเภทหลังมักถูกเรียกว่า ความตรงเชิงประจักษ์ (Face Validity) มีความอ่อนมากและควรพิจารณาให้ดีในการใช้เป็นหลักฐานของการยอมรับแบบทดสอบต่อสาธารณชน ผู้จัดพิมพ์แบบทดสอบมาตรฐานมักจะให้ข้อมูลเกี่ยวกับความสัมพันธ์กับเกณฑ์บ่อยมากที่สุด น้อยครั้งมากหรือไม่ค่อยพบบ่อยที่พวกเขาจะรายงานความสัมพันธ์กับเนื้อหาและการศึกษาหาประเมินความตรงเชิงโครงสร้าง ดังมีรายละเอียดของแต่ละวิธี ดังนี้

ความสัมพันธ์กับเกณฑ์ (Criterion Relatedness) ความสัมพันธ์กับเกณฑ์ ถูกแสดงโดยความสัมพันธ์ระหว่างคะแนนสอบและเครื่องวัดที่ใช้เป็นเกณฑ์ เครื่องวัดที่ใช้เป็นเกณฑ์อาจจะเป็นแบบทดสอบอื่นที่วัดความสามารถเดียวกันซึ่งได้รับการประเมินว่ามีความตรงดีอยู่แล้ว ตัวอย่างเช่น เมื่อมีการนำเสนอรูปแบบใหม่ของแบบทดสอบเดิมที่มีอยู่แล้ว ผู้ผลิตจะต้องแสดงให้เห็นว่าแบบทดสอบฉบับใหม่สัมพันธ์หรือสอดคล้องกับแบบทดสอบฉบับเดิมอย่างไร ซึ่งเป็นเครื่องวัดที่ใช้เป็นเกณฑ์ ในอีกตัวอย่างหนึ่ง แบบทดสอบใหม่ที่ถูกออกแบบด้วยประสิทธิภาพที่มากกว่า ประหยัดกว่า หรือบริหารจัดการง่ายกว่าสัมพันธ์กับแบบทดสอบที่มีอยู่แล้ว เช่น การสัมภาษณ์สดเพื่อวัดความสามารถในการพูดเป็นการเตรียมการทดสอบที่ทำได้ยาก ผู้สัมภาษณ์ที่ได้รับการฝึกอบรมมาอาจจะหาได้ยากและจัดตารางเวลาให้ตรงกันได้ยากเช่นกัน แบบทดสอบแบบวัดกึ่งตรง (A

Semidirect Test) ที่ออกแบบให้ผู้เข้าสอบตอบกลับต่อคำถามของเสียงเทปที่ได้บันทึกไว้ล่วงหน้า แล้วมีประโยชน์ในการนำมาใช้มาก คะแนนจากการสอบที่วัดแบบกึ่งตรงสามารถนำมาหาความสัมพันธ์กับคะแนนที่ได้จากการสัมภาษณ์สดที่เป็นเครื่องวัดที่ใช้เป็นเกณฑ์ เพื่อแสดงความสัมพันธ์กับเกณฑ์ของแบบทดสอบรูปแบบใหม่

ความสัมพันธ์กับเนื้อหา (Content Relevance) ความสัมพันธ์กับเนื้อหา คือ ระดับมากน้อยที่แบบทดสอบแสดงให้เห็นถึงความเป็นตัวแทนที่เพียงพอ (ความครอบคลุม) ของปริบทภาษาที่คาดว่าจะทดสอบ การหาความสอดคล้องของภาระงานกับวัตถุประสงค์และกิจกรรมการเรียนการสอน ในขณะที่กำลังสร้างแบบทดสอบเป็นวิธียืนยันหรือหาความสัมพันธ์กับเนื้อหาที่นิยมปฏิบัติกัน เป็นการทำให้มั่นใจว่าของสอบที่สร้างขึ้นเป็นตัวแทนที่เพียงพอของปริบท เช่น เนื้อหาในวิชาที่แบบทดสอบตั้งใจจะวัด เพราะแบบทดสอบมาตรฐานไม่ได้ถูกออกแบบเพื่อวัดผลสัมฤทธิ์ในการเรียนวิชาใดวิชาหนึ่ง จึงควรมีการใช้วิธีการที่แตกต่างไป ความสามารถที่แบบทดสอบถูกออกแบบให้วัดควรถูกอธิบายก่อน หลังจากนั้นให้กรรมการผู้เชี่ยวชาญในสาขานั้นพิจารณาหรือตัดสินความสัมพันธ์ของสาระแบบทดสอบกับแนวคิดที่เป็นปัจจุบันที่สุดของความสามารถนั้น ๆ ความสัมพันธ์ของเนื้อหาถูกประเมินอย่างมีเหตุผล ไม่มีวิธีการทางสถิติที่จะต้องนำมาใช้ตัดสิน

ความตรงเชิงโครงสร้าง (Construct Validation) การประเมินความตรงเชิงโครงสร้างจะพบได้จากโครงการวิจัยที่นิยามโครงสร้างโดยแสดงให้เห็นว่าแบบทดสอบที่มีโครงสร้างต่างกันคืออะไรและไม่เกี่ยวข้องหรือสัมพันธ์กับอะไร คำตัดสิน (Judgment) ถูกนำมาใช้ในการตัดสินใจว่าการวิจัยถูกทำได้ดีหรือไม่และความรู้หรือทักษะภาษาที่ถูกสำรวจถูกนำมาใช้อย่างเหมาะสมหรือไม่ การฝึกฝนและทักษะบางอย่างในวิธีวิจัยมีความจำเป็นในการประเมินกระบวนการประเมินความตรงเชิงโครงสร้าง การประเมินความตรงเชิงโครงสร้างมีความสำคัญโดยเฉพาะอย่างยิ่งเมื่อสาระที่ชัดเจนของทักษะที่กำลังถูกวัดโดยแบบทดสอบไม่สามารถระบุได้อย่างชัดเจน หรือเมื่อไม่สามารถระบุเกณฑ์ใด ๆ ที่จะใช้ประเมินแบบทดสอบได้ (หรือระบุได้ยาก) ลองพิจารณาตัวอย่างของความถนัดในภาษาที่สองตามแนวคิดทฤษฎี ผู้เรียนที่มีความถนัดในภาษาที่สองอยู่ในระดับสูงมีแนวโน้มที่จะมีความสามารถในภาษาที่สองสูงกว่าผู้ที่มีความถนัดต่ำ ความถนัดถูกคิดว่าเป็นปัจจัยที่สร้างความต่าง ไม่เกี่ยวข้องกับความฉลาดโดยทั่วไป แรงจูงใจ บุคลิกภาพ และอื่น ๆ แบบทดสอบวัดความถนัดในภาษาที่สองหลายชนิด ได้รับการพัฒนาเพื่อวัดคุณสมบัติเหล่านี้ ความถนัดในภาษาที่สองเป็นโครงสร้างทางทฤษฎีที่ไม่สามารถสังเกตเห็นได้โดยตรง และไม่มีพฤติกรรมใด ๆ ที่สามารถใช้เพื่อพิสูจน์ว่าความถนัดนี้มีอยู่

อย่างไรก็ตาม ถ้าโครงสร้างของความถนัดในภาษาที่สองมีความตรง ควรเป็นไปได้ที่จะพบหลักฐานว่าแบบทดสอบนั้นมีความจริงทางจิตวิทยาอยู่บ้างบางอย่าง สิ่งนี้สามารถทำได้โดยการทำวิจัยที่แสดงว่า โครงสร้างสัมพันธ์กับความสามารถในแบบทดสอบอื่น ๆ ที่มันควรสัมพันธ์กัน ตัวอย่างเช่น ความถนัดในภาษาที่สองควรสัมพันธ์กับผลสัมฤทธิ์ในการเรียนภาษาที่สอง ตามที่ทฤษฎีของความถนัดทำนายไว้ ในขณะที่เดียวกัน เป็นสิ่งสำคัญที่แสดงว่าโครงสร้างของความถนัดไม่เหมือนโครงสร้างอย่างอื่น ตัวอย่างเช่น เป็นสิ่งสำคัญที่แสดงว่าความถนัดไม่ใช่สิ่งเดียวกันกับความฉลาด ถ้าความสามารถในแบบทดสอบวัดความถนัดสัมพันธ์กับความสามารถในการสอบวัดความฉลาด แสดงว่าความถนัดและความฉลาดเป็นสิ่งเดียวกัน ดังนั้น แนวคิดด้านความถนัดอาจไม่จำเป็น อาจจะ

ซ้ำซ้อนกับแนวคิดด้านความฉลาด เนื่องจากเรามีแบบทดสอบวัดความฉลาดอยู่แล้ว เราไม่จำเป็นต้องมีแบบทดสอบวัดความถนัดอีก

แนวคิดของความตรงเชิงโครงสร้าง เริ่มมีความสำคัญมากขึ้นในการทดสอบภาษา เนื่องจากแนวคิดของแบบทดสอบวัดความสามารถทั่วไปทางภาษาที่เป็นนามธรรม และซับซ้อนถูกใช้เพิ่มมากขึ้น แนวคิดของความสามารถทางภาษาศาสตร์เชิงสังคม ตัวอย่างเช่น ต้องการกระบวนการของการประเมินความตรงเชิงโครงสร้างเนื่องจากไม่มีปริมาณเนื้อหาหรือเกณฑ์ใด ๆ หรือที่ง่าย ๆ ที่สามารถใช้ในการประเมินความตรงเชิงโครงสร้างได้เลย เช่นเดียวกับทักษะภาษาทั้งหมด และแม้กระทั่งชนิดที่ค่อนข้างง่าย เช่น คำศัพท์

สรุปได้ว่า การประเมินคุณภาพด้านความตรงที่สามารถนำมาใช้ในการพัฒนาคุณภาพแบบทดสอบ มี 3 ประเภท คือ ความตรงภายใน (Internal Validity) ความตรงภายนอก (External Validity) และความตรงเชิงโครงสร้าง (Construct Validity) โดยความตรงภายใน (Internal Validity) แบ่งย่อยเป็น 3 ชนิด คือ ความตรงเชิงประจักษ์ (Face Validity) ความตรงเชิงเนื้อหา (Content Validity) และความตรงที่เกิดจากการตอบสนอง (Response Validation) ความตรงภายนอก (External Validity) มี 2 ชนิด คือ ความตรงร่วมสมัย (Concurrent Validity) และความตรงเชิงพยากรณ์ (Predictive Validity) และความตรงเชิงโครงสร้าง (Construct Validity) แบ่งได้เป็น 5 ชนิด คือ การเปรียบเทียบกับทฤษฎี (Comparison with Theory) ความสัมพันธ์ภายใน (Internal Correlations) การเปรียบเทียบกับชีวประวัติและลักษณะทางจิตวิทยาของนักเรียน (Comparison with Students' Biodata and Psychological Characteristics) การวิเคราะห์แบบพหุคุณ-พหุวิธี และการประเมินความตรงแบบลู่เข้า-แบบกระจายออก (Multitrait-Multi Method Analysis and Convergent-Divergent Validation) และการวิเคราะห์ปัจจัย (Factor Analysis) และในการวิจัยครั้งนี้ ผู้วิจัยจะใช้ความตรงเชิงเนื้อหา (Content Validity) และความตรงร่วมสมัย (Concurrent Validity) ในการวิเคราะห์คุณภาพของแบบทดสอบที่สร้างขึ้น

การหาดัชนีความสอดคล้องระหว่างข้อคำถามและวัตถุประสงค์ (Index of Item-Objective Congruence - IOC)

การหาค่า IOC คือ การหาค่าความสอดคล้องระหว่างข้อคำถามและวัตถุประสงค์ (Index of Item-Objective Congruence) ของแบบทดสอบอิงเกณฑ์ ตามแนวคิดของ Rovinelli และ Hambleton (1976) โดยเป็นการสอบถามความคิดเห็นจากผู้เชี่ยวชาญในสาขาวิชาที่ผู้วิจัยสร้างแบบทดสอบ ว่าข้อคำถามที่สร้างขึ้นแต่ละข้อสอดคล้องกับวัตถุประสงค์ที่จะวัดหรือไม่ มากน้อยเพียงใด แล้วนำมาคำนวณหาดัชนีความสอดคล้องระหว่างวัตถุประสงค์กับข้อคำถาม โดยใช้สูตร ดังนี้ (Rovinelli and Hambleton (1976) คือ

$$I_{ik} = \frac{(N-1) \sum_{j=1}^n X_{ijk} - \sum_{i=1}^N \sum_{j=1}^n X_{ijk} + \sum_{j=1}^n X_{ijk}}{2(N-1)n}$$

เมื่อ I_{ik} = ดัชนีความสอดคล้องระหว่างข้อคำถามที่ k กับวัตถุประสงค์ที่ i

($i = 1, 2, \dots, N$)

N = จำนวนวัตถุประสงค์ในการสอบ

n = จำนวนผู้ประเมิน

X_{ijk} = ผลการประเมินข้อคำถามที่ k เทียบกับวัตถุประสงค์ที่ i

โดยผู้ประเมินคนที่ j ($j = 1, 2, \dots, n$)

+1 = เห็นว่าสอดคล้องกัน

= ยังสงสัย ไม่แน่ใจ หรือตัดสินใจไม่ได้

-1 = เห็นว่าไม่สอดคล้องกัน

กระบวนการในการประเมินคุณภาพ (Validation)

Bachman (2004) กล่าวว่า การประเมินคุณภาพ (Validation) หมายถึง กระบวนการในการค้นหาคุณภาพของการใช้คะแนน ในลักษณะของการกล่าวอ้างเชิงตีความ และการเก็บหลักฐานเพื่อสนับสนุนการกล่าวอ้างนั้น กระบวนการในการค้นหาความตรงประกอบด้วยกิจกรรมมีสัมพันธ์ซึ่งกันและกัน 2 กิจกรรม คือ 1) การกำหนดค่ากล่าวอ้างเชิงตีความที่ทำให้ได้กรอบแนวคิดพื้นฐานเชื่อมโยงผลการสอบเข้ากับการตีความและการใช้ที่ต้องการ และ 2) การเก็บรวบรวมหลักฐานที่สำคัญเพื่อสนับสนุนการตีความและการใช้ที่ต้องการ ดังมีรายละเอียดในการดำเนินแต่ละกิจกรรมดังต่อไปนี้

การอ้างอิงเหตุผลด้านความตรง (Formulating the Validation Argument)

การอ้างอิงเหตุผลด้านความตรงกระทำได้โดยอาศัยผลคะแนนจากการสอบด้านภาษา มักจะมีอยู่ 2 ชนิด คือ การทำนายเกี่ยวกับความสามารถที่อยู่นอกเหนือแบบทดสอบ และ การอ้างอิงเกี่ยวกับความรู้หรือความสามารถที่ผู้เข้าสอบมี การอ้างอิงทั้ง 2 ชนิดนี้แตกต่างกันมาก และต้องการหลักฐานสนับสนุนที่แตกต่างกัน ดังตัวอย่างต่อไปนี้

การทำนายเกี่ยวกับความสามารถในอนาคต (สิ่งที่ผู้เข้าสอบสามารถทำได้)

การอ้างอิงชนิดหนึ่ง คือ การทำนายเกี่ยวกับความสามารถในอนาคตที่อยู่นอกเหนือแบบทดสอบ การอ้างอิงชนิดนี้ส่งผลให้เกิดการกล่าวอ้างได้ว่า ความสามารถในภาระงานสอบเป็นตัวกำหนดพื้นฐานสำหรับการทำนายว่าผู้เข้าสอบสามารถทำอะไรได้บ้างในปริเขตของการใช้ภาษาเป้าหมายนอกเหนือจากแบบทดสอบ คำกล่าวอ้างชนิดนี้ไม่จำเป็นต้องสื่อความถึงการกล่าวอ้างเกี่ยวกับความรู้หรือความสะกดที่ผู้เข้าสอบมี แต่คำกล่าวอ้างนี้อาศัยพื้นฐานที่เป็นคำอธิบายปริเขตการใช้ภาษาเป้าหมายเท่านั้น (The TLU Domain) ของลักษณะภาระงานที่อยู่ในปริเขตนั้นและจากลักษณะของภาระงานสอบนั้น ๆ

การอ้างอิงอีกชนิดหนึ่งเกี่ยวกับความรู้หรือความสามารถที่ผู้เข้าสอบหรือมี การอ้างอิงชนิดนี้สื่อความได้ว่า ความสามารถในภาระงานสอบเป็นพื้นฐานสำหรับการอ้างอิงเกี่ยวกับสิ่งที่ผู้เข้าสอบรู้หรือเกี่ยวกับความสามารถที่พวกเขามี การกล่าวอ้างเช่นนี้มักจะอาศัยคำนิยามความสามารถที่นักพัฒนาแบบทดสอบสร้างไว้สำหรับจุดประสงค์ของการวัด คำนิยามในโครงสร้างนี้เป็นพื้นฐานสำหรับการตีความผลจากคะแนนสอบ นอกจากนี้ นักพัฒนาแบบทดสอบยังตั้งใจที่จะอ้างอิงเกี่ยวกับความรู้หรือความสามารถที่จะอ้างอิงผลนอกเหนือแบบทดสอบ ดังนั้น คำกล่าวอ้างด้านความตรงช่วยสื่อถึงคำกล่าวอ้างเกี่ยวกับประสิทธิภาพที่ใช้ภาษาที่นักพัฒนาแบบทดสอบได้อ้างว่าสอดคล้องกับความรู้หรือความสามารถที่น่าเสนอ คำกล่าวอ้างนั้นจะเกี่ยวข้องกับคำอธิบายประสิทธิภาพที่ใช้ภาษาเป้าหมาย ลักษณะภาระงานต่างๆ ในปริศานั้น ๆ และลักษณะของภาระงานการสอบแต่ละชนิด

การอ้างสิทธิแย้ง (Counterclaims)

ดังที่ได้กล่าวไว้แล้วข้างต้น การกล่าวอ้างอิงด้านความตรงยังจำเป็นต้องครอบคลุม การอ้างสิทธิแย้ง (Counterclaims) จากการประเมินความตรงที่เป็นไปได้ เกี่ยวกับการตีความที่ไม่ได้ตั้งใจสำหรับการทำนายเกี่ยวกับความสามารถในอนาคต การอ้างสิทธิแย้ง (Counterclaims) ที่เป็นไปได้อย่างหนึ่งคือ ภาระงานที่บรรจุในแบบทดสอบไม่ได้เป็นตัวแทนที่ตัวอย่างเพียงพอของภาระงานในปริศการใช้ภาษาเป้าหมายที่กำหนดไว้โดยนักพัฒนาแบบทดสอบ ดังนั้น คะแนนสอบจึงไม่มีประโยชน์สำหรับจุดประสงค์นี้ ถ้าการอ้างสิทธิแย้ง (Counterclaims) รูปแบบนี้เกิดขึ้นในแบบทดสอบที่สร้างขึ้นและถ้าไม่มีการตัดออกตั้งแต่ต้น อาจจะทำให้เกิดปัญหาในการประเมินความตรงโดยใช้คะแนนที่เราต้องการให้เกิดขึ้นจากการใช้แบบทดสอบชุดนี้

สำหรับการอ้างอิงถึงความรู้หรือความสามารถ การอ้างสิทธิแย้ง (Counterclaims) ที่อาจเกิดขึ้นได้ คือ ความสามารถในการทำภาระงานการสอบที่ไม่เกี่ยวข้องหรือไม่ได้ใช้ความรู้ความสามารถที่นักพัฒนาแบบทดสอบได้กำหนดไว้ และนั่นคือความสามารถในแบบทดสอบได้รับอิทธิพลโดยส่วนใหญ่จากคุณสมบัติอื่น ๆ ของผู้เข้าสอบหรือกระบวนการอื่นนอกเหนือสิ่งที่ต้องการวัด การอ้างสิทธิแย้ง (Counterclaims) เหล่านี้ถ้ามีอยู่ในแบบทดสอบและไม่ได้ถูกตัดออกจะมีผลต่อความตรงในการอ้างอิงใด ๆ และการใช้งานของคะแนนสอบที่นักพัฒนาแบบทดสอบต้องการ การอ้างสิทธิแย้ง (Counterclaims) อีกอย่างหนึ่งคล้ายกับที่กล่าวมาแล้วข้างต้น คือ เกี่ยวกับภาระงานสอบที่ไม่เป็นตัวแทนที่ดีของภาระงานในปริศการใช้ภาษาเป้าหมาย ถ้า การอ้างสิทธิแย้ง (Counterclaims) เหล่านี้ไม่ถูกตัดออกจะจำกัดความสามารถของนักพัฒนาแบบทดสอบที่จะอ้างอิงผลออกนอกเหนือแบบทดสอบ

การเก็บข้อมูลที่เกี่ยวข้อง (Collecting Relevant Evidence)

เพื่อเก็บหลักฐานเชิงปริมาณที่เกี่ยวข้องในการใช้สนับสนุนคำกล่าวอ้างและปฏิเสธ counterclaim ในคำกล่าวอ้างด้านความตรงของเรา เราได้ออกแบบการศึกษาเพื่อประเมินความตรงที่ประกอบด้วย 5 ขั้นตอน คือ

1) พัฒนาระบวนการในการปฏิบัติ เช่น การวัด การสังเกต สำหรับการเก็บข้อมูลที่สำคัญต่อการกล่าวอ้างและ counterclaim แต่ละอย่าง กระบวนการนี้รวมถึงภารกิจสองอย่าง คือ การเก็บคะแนนที่เป็นผลการสอบ และการตรวจให้คะแนนที่เป็นผลการสอบ

2) กำหนดสมมติฐานเกี่ยวกับรูปแบบความสัมพันธ์ที่คาดหวังหรือความแตกต่างในกลุ่มหรือตัวแปรต่าง ๆ ที่สังเกตได้

ก. สำหรับการสนับสนุนการกล่าวอ้าง เราตั้งสมมติฐานที่ตัวแปรควรสัมพันธ์ซึ่งกันและกัน หรือซึ่งกลุ่มต่าง ๆ ควรแสดงความสามารถที่ต้องการวัดในแบบทดสอบอย่างแตกต่างกัน

ข. สำหรับการปฏิเสธ counterclaims เราควรตั้งสมมติฐานที่ตัวแปรต่างๆ ไม่ควรสัมพันธ์กัน หรือกลุ่มต่างๆ ควรแสดงความสามารถที่เหมือนกันในแบบทดสอบที่วัดความสามารถที่ต้องการ

3) ออกแบบการศึกษาเพื่อเก็บข้อมูลที่เกี่ยวข้อง

รูปแบบการศึกษาที่ใช้บ่อยที่สุด คือ สำหรับการประเมินความตรงเชิงโครงสร้างที่เก็บข้อมูลเชิงปริมาณ คือ การศึกษาความสัมพันธ์ ซึ่งเกี่ยวข้องกับการให้ผู้เข้าสอบจำนวนมากพอสมควร ได้สอบแบบทดสอบต่างชนิดกัน อีกรูปแบบหนึ่งซึ่งมีคุณภาพพอกันแต่ได้รับความนิยมน้อยกว่า คือ การออกแบบการศึกษาเชิงทดลองกับกลุ่มผู้เข้าสอบที่มีคุณสมบัติเท่าเทียมกัน (Equivalent Groups) ซึ่งได้รับการสอนที่แตกต่างกันด้วย และชนิดที่สาม คือ การจัดสอบกับกลุ่มผู้เข้าสอบที่สามารถเปรียบเทียบกันได้ (Comparison Groups) ที่ผู้เข้าสอบแต่ละคนมีระดับความสามารถแตกต่างกัน โดยเป็นที่คาดหวังว่าผู้เข้าสอบจะแสดงความสามารถในการทำแบบทดสอบที่แตกต่างกันด้วย

4) เก็บข้อมูล

5) ทดสอบสมมติฐาน

2.2.2.3 ความลำเอียงและความยุติธรรมในการทดสอบ

ความลำเอียงและความยุติธรรมเป็นคุณสมบัติที่สำคัญของแบบทดสอบวัดความสามารถภาษาอังกฤษที่มีคุณภาพและมาตรฐาน นักพัฒนาแบบทดสอบจำเป็นจะต้องมีความรู้ความเข้าใจเกี่ยวกับ 1) ความหมายของความลำเอียงและความยุติธรรมในการทดสอบ 2) แหล่งที่มาของความลำเอียงในการทดสอบ 3) แนวทางการลดความลำเอียงในการทดสอบ และ 4) การวิเคราะห์ความลำเอียงของแบบทดสอบ เพื่อให้สามารถสร้างและพัฒนาแบบทดสอบวัดความสามารถภาษาอังกฤษให้มีมาตรฐานระดับสากล ดังได้แสดงรายละเอียดในแต่ละประเด็น ดังต่อไปนี้

ความหมายของความลำเอียงและความยุติธรรมในการทดสอบ

ความลำเอียงและความยุติธรรมเป็นคำที่มักจะใช้คู่กัน ในการพัฒนาแบบทดสอบ นักพัฒนาแบบทดสอบต้องสร้างข้อคำถามให้มีความลำเอียงต่อผู้เข้าสอบกลุ่มใดกลุ่มหนึ่งต่ำที่สุด จึงจะกล่าวได้ว่าแบบทดสอบชุดนั้นมีความยุติธรรม อย่างไรก็ตาม นักวิจัยและนักพัฒนาแบบทดสอบได้ให้ความหมายของความลำเอียงและความยุติธรรมในการทดสอบไว้ดังนี้

นักวิจัยด้านการพัฒนาแบบทดสอบหลายคนได้ให้ความหมายของความลำเอียงไว้ค่อนข้างแตกต่างกัน ได้แก่ ALTE (1990 อ้างถึงใน Brown, 2005: 246) Khoii & Shamsi (2012) Peterson (2012) และ Whiting (2009) ดังรายละเอียดต่อไปนี้

ALTE (1990 อ้างถึงใน Brown, 2005: 246) กล่าวว่า แบบทดสอบหรือข้อคำถามสามารถถูกพิจารณาได้ว่ามีความลำเอียงถ้าผู้เข้าสอบกลุ่มใดกลุ่มหนึ่งมีการได้เปรียบหรือเสียเปรียบในการได้คะแนนสอบอันมีผลมาจากลักษณะบางประการของแบบทดสอบหรือข้อคำถาม ซึ่งไม่มีส่วนเกี่ยวข้องกับเนื้อหาสาระที่กำลังถูกทดสอบ

Khoii (2012) กล่าวว่า ความลำเอียง เป็นปัจจัยหนึ่งที่สามารถเป็นอุปสรรคของความยุติธรรมในการสอบ แบบทดสอบที่มีความลำเอียง คือ แบบทดสอบที่มีความแตกต่างกันอย่างเป็นระบบในความหมายของคะแนนสอบที่สัมพันธ์กับความเป็นสมาชิกของกลุ่ม โดยอาจหมายความว่า คน 2 กลุ่มที่ได้คะแนนจากการสังเกตครั้งเดียวกัน อาจมีความเข้าใจในเรื่องที่ออกสอบต่างกัน

Peterson Doug (2012) กล่าวว่า ความลำเอียง หมายถึง การแสดงความชอบในกลุ่มหนึ่งกลุ่มใดมากกว่ากลุ่มอื่น ๆ ในกระบวนการพัฒนาแบบทดสอบ มีหลายวิธีที่อาจจะทำให้การเขียนแบบทดสอบเกิดความลำเอียงได้ เช่น ถ้าผู้เขียนแบบทดสอบใช้ภาษาที่คุ้นเคยกับผู้เข้าสอบกลุ่มหนึ่งกลุ่มใด ที่อาศัยอยู่ในที่ตั้งทางภูมิศาสตร์เฉพาะถิ่น อาจเป็นการสร้างความไม่ยุติธรรมให้เกิดขึ้นในกลุ่มผู้เข้าสอบที่มาจากส่วนอื่น ๆ ของโลกได้ เป็นต้น

Whiting (2009) กล่าวว่า ความลำเอียงเกี่ยวข้องกับผลการสอบที่แตกต่างทั้งในกลุ่มและระหว่างกลุ่ม สาเหตุที่ทำให้กลุ่มหนึ่งได้คะแนนแตกต่างจากอีกกลุ่มหนึ่ง เช่น การเป็นกลุ่มชนผิวดำกับกลุ่มชนผิวขาว การอยู่ในกลุ่มเพศหญิงกับกลุ่มเพศชาย กลุ่มที่มีรายได้สูงกับกลุ่มที่มีรายได้ต่ำ เมื่อมีระดับความลำเอียงที่คงที่

สรุปได้ว่า ความลำเอียงในการทดสอบเกิดขึ้นเมื่อความแตกต่างในผลคะแนนสอบมีผลมาจากความได้เปรียบหรือเสียเปรียบของผู้เข้าสอบกลุ่มใดกลุ่มหนึ่ง เช่น กลุ่มเพศชายหรือกลุ่มเพศหญิง กลุ่มผู้มีรายได้สูงหรือกลุ่มผู้มีรายได้น้อย กลุ่มผิวขาวหรือกลุ่มผิวดำ เป็นต้น โดยลักษณะเฉพาะของแต่ละกลุ่มที่ไม่เกี่ยวข้องกับโครงสร้างความสามารถที่ถูกทดสอบ มีส่วนช่วยสนับสนุนให้ทำคะแนนได้ดีกว่าผู้เข้าสอบอีกกลุ่มหนึ่ง

ในด้านความยุติธรรมในการสอบ นักวิจัยด้านการพัฒนาแบบทดสอบส่วนใหญ่มีมุมมองด้านความหมายของความยุติธรรมที่คล้ายกัน คือ เห็นว่าความยุติธรรมเป็นเรื่องของความเสมอภาคหรือเท่าเทียมกันที่ผู้เข้าสอบทุกคนควรจะได้รับในการทำแบบทดสอบ นอกจากนี้ นักวิจัยแต่ละคนได้กล่าวถึงความยุติธรรมไว้ดังนี้

Brown (2004: 111) กล่าวว่า ความยุติธรรมในการสอบอาจจะถูกชักนำโดยความลำเอียงในการสอบ อาจจะอยู่ในรูปแบบของภาษา วัฒนธรรม เชื้อชาติ เพศ และรูปแบบการเรียนรู้ (Learning Styles) สำหรับความลำเอียงด้านวัฒนธรรมนั้นควรระวังเป็นพิเศษโดยการไม่ใช่ข้อคำถามหรือประเด็นที่นำมาออกสอบที่ทำให้ผู้เข้าสอบเสียเปรียบด้วยการมีวัฒนธรรมท้องถิ่นหรือวัฒนธรรมดั้งเดิมที่แตกต่าง

Weyant, Katie, Chisholm, and Amanda (2014) กล่าวว่า ความยุติธรรม คือ การไม่มีความลำเอียง การปฏิบัติที่เท่าเทียมกันในกระบวนการทดสอบ การเท่าเทียมกันของผลลัพธ์ของการสอบ และความยุติธรรมในโอกาสที่จะเรียนรู้

Djiwandono and Patrisius Istiarto (2006) กล่าวว่า แบบทดสอบที่จัดให้กับกลุ่มผู้เข้าสอบที่มีความแตกต่างกันด้านวัฒนธรรม ชาติพันธุ์ และเพศ สามารถเป็นแบบทดสอบที่ยุติธรรมได้ถ้าแบบทดสอบนั้นมีความเท่าเทียมกันในการวัด การอธิบายโดยใช้คำพูดง่าย ๆ ก็คือ แบบทดสอบยุติธรรมถ้าคะแนนสอบสะท้อนความสามารถในระดับต่าง ๆ ของผู้เข้าสอบในทักษะที่แบบทดสอบวัดและไม่ใช้ความต่างในแง่มุมอื่นใด เช่น แบบทดสอบที่มีเป้าหมายเพื่อวัดความเข้าใจในการฟัง อาจจะไม่ยุติธรรม และมีความลำเอียง ถ้าคะแนนสะท้อนถึงความแตกต่างในกลุ่มผู้เข้าสอบในแง่ของความคุ้นเคยกับวัฒนธรรมเฉพาะอย่าง ผู้เข้าสอบที่ไม่คุ้นเคยกับวัฒนธรรม อาจจะทำได้ต่ำไม่ใช่เฉพาะพวกเขาทักษะการฟังที่แย่ แต่เป็นเพราะขาดความรู้ที่เกี่ยวกับแง่มุมของวัฒนธรรมเฉพาะอย่างที่ถูกระบุไว้ในข้อคำถาม

Jin (2011) เสนอ มาตรฐานกำหนดกรอบความยุติธรรมเป็นคุณภาพของแบบทดสอบที่สอดคล้องโดยตรงกับความตรงของแบบทดสอบ ครอบคลุม 4 ประเด็นหลัก คือ การปฏิบัติที่ยุติธรรมต่อผู้เข้าสอบทุกคน ความอิสระจากความลำเอียง ความยุติธรรมจากผลการสอบ และ ความยุติธรรมในการเรียนรู้สาระที่สอบ

สรุปได้ว่า ความยุติธรรมในการสอบ คือ ลักษณะที่ปราศจากความลำเอียงของแบบทดสอบ ส่งผลให้ผู้เข้าสอบทุกคนได้รับความเสมอภาคในการทำแบบทดสอบ การเรียนรู้สาระแบบทดสอบ และการได้รับผลการสอบ โดยผลการสอบไม่ขึ้นอยู่กับความแตกต่างด้านเพศ สัญชาติ ศาสนา ภาษา วัฒนธรรม และรูปแบบการเรียนรู้ เป็นต้น

แหล่งที่มาของความลำเอียงในการทดสอบ

ความลำเอียงในการทดสอบอาจมีที่มาจากหลายแหล่งที่ต่างกัน ดังที่ Khoii & Shamsi (2012) Whiting (2009) และ Reynolds (1998 อ้างถึงใน Whiting (2009) ได้กล่าวไว้ต่อไปนี้

Khoii & Shamsi (2012) กล่าวว่า แหล่งที่มาที่เป็นไปได้ของความลำเอียงที่สามารถลดความยุติธรรมได้เป็นอย่างมาก ประกอบด้วย การบริหารการสอบที่ไม่ได้ประสิทธิภาพ (Poor Test Administration) การเลือกใช้แบบทดสอบที่ไม่เหมาะสม (Inappropriate Test Choice) และความคาดหวังของผู้เข้าสอบ (Test Taker's Expectation) หลักการนี้คล้ายกับมุมมองของ Society for Industrial and Organizational Psychology or SIOP ที่เห็นว่า ความยุติธรรมเป็นการปฏิบัติที่เสมอภาคแก่ผู้เข้าสอบทุกคนเกี่ยวข้องกับสภาวะแวดล้อมในการทดสอบ (Testing Conditions) การ

เข้าถึงวัสดุในการฝึก (Access to Practice Materials) การให้ข้อมูลย้อนกลับของผลการสอบ (Performance Feedback) โอกาสในการทดสอบซ้ำ (Retest Opportunities) และลักษณะอื่น ๆ ของการบริหารการสอบ (Other Features of Test Administration)

Whiting (2009) นำเสนอความลำเอียงในการทดสอบ 4 ชนิด คือ 1) ความลำเอียงด้านความตรงเชิงเนื้อหา (Bias in Content Validity) 2) ความลำเอียงด้านความตรงเชิงเนื้อหา (Bias in Content Validity) 3) ความลำเอียงในการเลือกข้อคำถาม (Bias in Item Selection) และ 4) ความลำเอียงด้านความตรงเชิงทำนายหรือเกณฑ์สัมพันธ์ (Bias in Predictive or Criterion-Related Validity) โดยมีรายละเอียดในแต่ละชนิด ดังนี้

ความลำเอียงด้านความตรงเชิงโครงสร้าง (Bias in Construct Validity) ความลำเอียงด้านความตรงเชิงโครงสร้างเกิดขึ้นเมื่อแบบทดสอบวัดโครงสร้างทางทฤษฎีแตกต่างกันในผู้เข้าสอบกลุ่มหนึ่งมากกว่าอีกกลุ่มหนึ่ง และความลำเอียงชนิดนี้ยังเกิดขึ้นได้เมื่อแบบทดสอบวัดโครงสร้างความสามารถเดียวกันสำหรับทุกกลุ่ม แต่มีระดับความถูกต้องที่แตกต่างกัน สถิติที่มักจะใช้สำหรับวิธีการนี้คือการวิเคราะห์ปัจจัย โดยแบบทดสอบที่มีความลำเอียงจะแสดงโครงสร้างปัจจัยต่าง ๆ ที่เกิดขึ้นในกลุ่มต่าง ๆ จะมีระดับความเหมือนกันของโครงสร้างปัจจัยที่ต่ำกว่า และแสดงผลตามการจัดอันดับ หรือความยากของข้อคำถามในแต่ละกลุ่ม คำถามพื้นฐานที่มักจะถูกถาม คือ ข้อคำถามหรือแบบทดสอบวัดในสิ่งที่กำหนดว่าจะวัดหรือไม่ ผลลัพธ์ที่สำคัญเชื่อมโยงกับภาษา การทดสอบภาษาอังกฤษในกลุ่มผู้เรียนที่ยังไม่ชำนาญด้านภาษาเป็นสิ่งที่พบปัญหาได้ การทดสอบทางสติปัญญาจะกลายเป็นการทดสอบทางภาษา นักเรียนบางคนหรือบางกลุ่มอาจจะมีความรู้หรือประสบการณ์ในการตอบข้อคำถามได้ถูกต้อง แต่ไม่สามารถทำได้เพราะไม่เข้าใจคำถาม อันเนื่องมาจากอุปสรรคด้านภาษา

ความลำเอียงด้านความตรงเชิงเนื้อหา (Bias in Content Validity) ความลำเอียงด้านความตรงเชิงเนื้อหาเกิดขึ้นเมื่อคำถามหรือประเด็นย่อยค่อนข้างยากกว่าระดับความรู้ของสมาชิกกลุ่มใดกลุ่มหนึ่งมากกว่ากลุ่มอื่น หลังจากที่สามารถโดยทั่วไปของทั้ง 2 กลุ่มถูกทำให้คงที่ ตัวอย่างเช่น ถ้าคำถามถามว่า Soccer และ Football ต่างกันอย่างไร นักเรียนคนหนึ่งหรือกลุ่มหนึ่งที่ไม่เคยเล่นหรือดูหรือพูดคุยเกี่ยวกับ Soccer อยู่ในตำแหน่งที่เสียเปรียบ การขาดความรู้หรือประสบการณ์ทำให้พวกเขาเสียเปรียบ

Reynolds (1998 อ้างถึงใน Whiting (2009) นิยามความลำเอียงเชิงเนื้อหาว่า ข้อคำถามหรือประเด็นย่อยของแบบทดสอบถูกพิจารณาว่ามีความลำเอียงเมื่อแบบทดสอบหรือข้อคำถามนั้นแสดงให้เห็นว่าค่อนข้างยากสำหรับสมาชิกกลุ่มหนึ่งมากกว่าอีกกลุ่มหนึ่ง เมื่อความสามารถทั่วไปของทั้ง 2 กลุ่มคงที่ และไม่มีการอธิบายเชิงทฤษฎีใด ๆ จะใช้อธิบายความแตกต่างของกลุ่มในข้อคำถามหรือประเด็นย่อยที่อยู่ในความสงสัย Reynolds ยังได้แจกตัวอย่างความลำเอียงเชิงโครงสร้างไว้ 3 ประเด็น คือ 1) ข้อคำถามถามข้อมูลที่นักเรียนจำนวนน้อยมีโอกาสไม่เท่ากันในการเรียนรู้ 2) การตรวจให้คะแนนข้อคำถามไม่เหมาะสม เพราะผู้เขียนหรือผู้พัฒนาแบบทดสอบได้ตัดสินใจให้มีคำตอบที่ถูกเพียงคำตอบเดียวอย่างไม่มีกฎเกณฑ์ และนักเรียนจำนวนน้อยอาจถูกทำโทษอย่างไม่เหมาะสมโดยที่ตัวเลือกของคำตอบอื่นที่กำหนดให้อาจจะถูกต้องในอีกวัฒนธรรมวัฒนธรรมหนึ่ง และ 3) โจทย์

คำถามไม่เป็นที่คุ้นเคยของนักเรียนทั่วไปและอาจจะมีนักเรียนกลุ่มหนึ่งที่รู้คำตอบแต่ไม่สามารถตอบได้เพราะไม่เข้าใจคำถามและ/หรือไม่คุ้นเคยกับรูปแบบคำถาม

ความลำเอียงในการเลือกข้อคำถาม (Bias in Item Selection) ความลำเอียงในการเลือกข้อคำถามเกิดขึ้นเมื่อข้อคำถามหรือภาระงานที่ถูกเลือกอยู่บนพื้นฐานของประสบการณ์ด้านการเรียนและภาษาของสมาชิกส่วนใหญ่ ความลำเอียงชนิดนี้มีลักษณะคล้ายกับความลำเอียงเชิงเนื้อหา แต่สื่อให้เห็นความเกี่ยวข้องโดยตรงที่เกิดจากความเหมาะสมของคำถามแต่ละข้อมากกว่า ในภาพรวมแบบทดสอบอาจจะไม่มีความลำเอียงในทางสถิติ แต่อาจจะมี 2-3 ข้อคำถามที่มีความลำเอียงได้ โดยเฉพาะอย่างยิ่ง อาจจะมีประเด็นคำถามว่าข้อคำถามข้อหนึ่งที่ถูกนำเข้ามาอยู่ในแบบทดสอบได้อย่างไร ในขณะที่ข้ออื่น ๆ ไม่ถูกเลือก

ความลำเอียงด้านความตรงเชิงทำนายหรือเกณฑ์สัมพันธ์ (Bias in Predictive or Criterion-Related Validity) ความลำเอียงด้านความตรงเชิงทำนายหรือเกณฑ์สัมพันธ์เกิดขึ้นเมื่อการลงความเห็นจากคะแนนสอบไม่ได้อาศัยความผิดพลาดในการสุ่มที่เป็นไปได้ที่มีโอกาสน้อยที่สุด หรือเมื่อมีความผิดพลาดเกิดขึ้นคงที่ในการลงความเห็นหรือทำนายตามหน้าที่ในความเป็นสมาชิกในกลุ่มเฉพาะกลุ่มใดกลุ่มหนึ่ง คำถามที่ครอบคลุมประเด็นนี้คือ คะแนนสอบสามารถใช้ทำนายได้อย่างถูกต้องหรือไม่ว่านักเรียนหรือกลุ่มผู้เข้าสอบแสดงความสามารถในภาระงานอย่างไร โดยทั่วไป มักจะสันนิษฐานว่าคะแนนด้านสติปัญญาที่สูงใช้ทำนายเกรดเฉลี่ยสูงและความสำเร็จในมหาวิทยาลัยและในการทำงาน และสำหรับฝ่ายที่ไม่เห็นด้วยกับแนวคิดนี้ คือ เห็นว่าแบบทดสอบวัดสติปัญญาถูกให้ความสำคัญมากเกินไป และถ้านักเรียนทำคะแนนได้ต่ำในแบบทดสอบวัดสติปัญญา ความเป็นไปได้สูงที่พวกเขาจะไม่ได้มีโอกาสคัดเลือกเข้าสู่โปรแกรมหรือบริการ เพราะพวกเขาถูกคาดหวังไว้ต่ำ ในทำนองเดียวกัน แบบทดสอบจะถูกพิจารณาว่าไม่ลำเอียง ถ้าผลของกลุ่มย่อยที่เป็นตัวแทนของกลุ่มประชากรทั้งหมดกระจายอยู่รอบ ๆ เส้นถดถอยเดียวกัน แบบทดสอบที่ไม่ลำเอียงทำนายความสามารถเท่าเทียมกันสำหรับทุกกลุ่ม แม้ว่าค่าเฉลี่ยอาจจะแตกต่างกัน

แนวทางการลดความลำเอียงในการทดสอบ

ในเมื่อการสร้างแบบทดสอบที่ปราศจากความลำเอียงเป็นสิ่งที่เป็นไปได้ยาก นักพัฒนาแบบทดสอบจึงจำเป็นต้องหาแนวทางในการลดความลำเอียงของแบบทดสอบให้เหลือน้อยที่สุด ดังเช่น นักพัฒนาแบบทดสอบหลายคน อาทิ Jin (2011) Khoii (2012) Weyant et al. (2014) และ Munson Robin and Parton Cinda (2013) ได้เสนอแนวทางในการลดความลำเอียงไว้ ดังนี้

Jin (2011) เสนอแนวทางในการลดความลำเอียงในการทดสอบโดยให้นักพัฒนาแบบทดสอบด้านภาษาและผู้ใช้แบบทดสอบได้รับข้อเสนอแนะที่เป็นประโยชน์และปฏิบัติได้ เป็นวิธีที่ช่วยรักษามาตรฐานที่สามารถปรับเปลี่ยนได้อย่างมีเหตุผล และช่วยให้ทำการอ้างอิงผลคะแนนสอบได้อย่างมีความหมาย และส่งเสริมการใช้แบบทดสอบที่เหมาะสมตามหลักจริยธรรม

นอกจากนี้ Jin (2011) ยังได้ยกตัวอย่างมาตรการในการกำหนดความตรงและความยุติธรรมของแบบทดสอบ CET ของ NCETC (The National College English Testing Committee) 3 มาตรการ คือ 1) การทบทวนสาระและรูปแบบของทดสอบ 2) การรักษามาตรฐานของ

กระบวนการพัฒนาแบบทดสอบ และ 3) การพัฒนาแบบทดสอบวัดการพูดของ CET และแบบทดสอบ CET ที่จัดสอบด้วยระบบอินเทอร์เน็ต โดยมีรายละเอียด ดังนี้

การทบทวนสาระและรูปแบบของแบบทดสอบ เป็นการพัฒนาแบบทดสอบให้สอดคล้องกับทฤษฎีภาษาศาสตร์และข้อกำหนดในหลักสูตรการสอนภาษาอังกฤษระดับมหาวิทยาลัยฉบับเก่าและฉบับปัจจุบัน The NCETC ได้ทบทวน/แก้ไขสาระและรูปแบบของแบบทดสอบ CET ให้มีความตรงสูงและมีผลกระทบที่ดีต่อการเรียนการสอนที่มากขึ้น

การรักษามาตรฐานของกระบวนการพัฒนาแบบทดสอบ ตั้งแต่เริ่มใช้มาตั้งแต่ปี 1987 แบบทดสอบ CET ได้กำหนดและปฏิบัติตามมาตรฐานของกระบวนการในการพัฒนาทุกแง่มุมของกระบวนการทดสอบ ได้แก่ การสร้างข้อคำถาม (Item Construction) การบริหารการสอบ (Test Administration) การเทียบแบบทดสอบระหว่างชุด (Test-form Equation) การตรวจให้คะแนนและการรายงานผล (Scoring and Score Reporting) ดังเช่นตัวอย่างของการสร้างข้อคำถาม จะมีขั้นตอนในการกำหนดมาตรฐานของแบบทดสอบดังนี้ คือ 1) การฝึกอบรมนักเขียนแบบทดสอบอย่างสม่ำเสมอ 2) กรรมการทบทวนข้อคำถาม ทำการตรวจสอบข้อคำถามที่ส่งมาจากนักเขียนแบบทดสอบในด้านความเหมาะสมของแหล่งข้อมูลและคุณภาพของข้อคำถาม 3) การทดลองใช้ข้อคำถามแต่ละข้อกับกลุ่มตัวแทนที่มีลักษณะคล้ายกันกับกลุ่มตัวอย่างเป้าหมาย 4) การวิเคราะห์ค่าสถิติรายข้อเพื่อตรวจสอบคุณภาพของข้อคำถามและตรวจจับความลำเอียงที่อาจเกิดขึ้นกับกลุ่มผู้เข้าสอบที่มีพื้นฐานแตกต่างกัน 5) การปรับแก้เพิ่มเติมสำหรับข้อคำถามที่ได้รับการยอมรับจากการตรวจสอบของผู้ทบทวนข้อคำถามที่เป็นผู้เชี่ยวชาญ 6) การสร้างชุดแบบทดสอบโดยอาศัยรายละเอียดตามตารางกำหนดลักษณะแบบทดสอบ (Test Specifications) 7) การทบทวนชุดแบบทดสอบฉบับร่างโดยเจ้าของภาษาเพื่อตรวจสอบความถูกต้องสมบูรณ์ของภาษาที่ใช้ และ 8) สมาชิกของ NCETC ตรวจสอบ/ทบทวนชุดแบบทดสอบฉบับล่าสุด

การพัฒนาแบบทดสอบวัดการพูดของ CET และ แบบทดสอบ CET ที่จัดสอบด้วยระบบอินเทอร์เน็ต ในการพัฒนาแบบทดสอบการพูดของแบบทดสอบ CET มีการรักษามาตรฐานและสร้างความยุติธรรมโดยกำหนดวิธีการให้คะแนนแบบวิเคราะห์องค์ประกอบในการพูดของผู้เข้าสอบ 3 ด้าน คือ 1) คุณภาพด้านภาษา (ความถูกต้องและปริมาณภาษาที่ใช้) 2) การมีปฏิสัมพันธ์ ความต่อเนื่องและความสัมพันธ์ของเนื้อความที่ผลิต (ขนาดและการจัดการเนื้อความ) และ 3) ความยืดหยุ่นในการจัดการกับหัวเรื่องต่าง ๆ และการใช้กลยุทธ์ในการสื่อสาร (ความยืดหยุ่นและความเหมาะสม) ส่วนการพัฒนาแบบทดสอบ CET เพื่อให้ใช้สอบผ่านระบบอินเทอร์เน็ตได้ มีการคำนึงถึงด้านความยุติธรรมของการจัดสอบแบบทดสอบ CET ผ่านระบบอินเทอร์เน็ต คือ การหาความสัมพันธ์ระหว่างคะแนนสอบที่มาจากการสอบแบบ Paper-Based และการสอบแบบ Internet-Based โดยตั้งข้อสังเกตว่าความไม่คุ้นเคยกับการใช้คอมพิวเตอร์ ความวิตกกังวลด้วยวิธีการที่แปลกใหม่อาจจะมีผลต่อคะแนนสอบ ผู้ดำเนินการจึงได้ทำการศึกษาเพื่อหาความสัมพันธ์ของวิธีการดังกล่าว และดูเหมือนจะพบว่าผู้เข้าสอบที่ชำนาญในการใช้คอมพิวเตอร์บอกว่าไม่ได้รับผลกระทบ โดยวิธีการสอบแบบนี้ และจากการวิเคราะห์ข้อความที่ผู้เข้าสอบผลิตได้จากการสอบแบบ Paper-Based และการทดสอบแบบ Internet-Based กลับพบว่าผู้เข้าสอบสามารถผลิตข้อความได้มากกว่า และมีโครงสร้างทางภาษาซับซ้อนกว่าด้วยการสอบผ่านอินเทอร์เน็ต การวิเคราะห์ด้านกระบวนการเขียนยังยืนยันอีกว่า การมี

ความคุ้นเคยกับการใช้คอมพิวเตอร์ที่เพิ่มขึ้น การใช้กลยุทธ์ด้านปัญญา (Cognitive) ที่เกี่ยวข้องในการเขียนเรียงความดีขึ้นด้วย

Whiting and Gilman (2009) กล่าวว่า ไม่มีแบบทดสอบใดที่ปราศจากความลำเอียง (A bias-free test) จะมีเพียงแต่แบบทดสอบที่ถูกลดความลำเอียงให้น้อยลง (Bias-reduced test) โดยสามารถใช้มาตรการเหล่านี้ คือ 1) แปลแบบทดสอบให้เป็นภาษาของผู้เข้าสอบ 2) ใช้ล่ามแปลข้อคำถามในแบบทดสอบสำหรับผู้เข้าสอบ 3) ตรวจสอบข้อคำถามหรือภาระงานเพื่อดูว่าแต่ละกลุ่มทำคะแนนได้แตกต่างกันหรือไม่ และตัดข้อคำถามหรือภาระงานเหล่านั้นออกไป 4) ตัดข้อคำถามที่จะส่งผลให้เกิดความลำเอียงกับผู้เข้าสอบ 5) เมื่อตีความคะแนนสอบ ให้พิจารณาประสบการณ์เดิมของผู้เข้าสอบด้วยเสมอ 6) อย่าสนับสนุนสมมติฐานของการมีประสบการณ์ร่วมกันหรือโอกาสเท่าเทียมกันในการเรียนรู้ กลุ่มต่าง ๆ มีพื้นฐานและประสบการณ์ที่แตกต่างที่ส่งผลต่อความสามารถในการสอบ 7) อย่ายึดการตัดสินใจอยู่บนพื้นฐานของแบบทดสอบฉบับเดียวหรือคะแนนเดียว ข้อมูลชิ้นเดียวหรือคะแนนอย่างเดียว ไม่มีประโยชน์มากนักที่จะใช้ประกอบการตัดสินใจที่มีประสิทธิภาพและเหมาะสม 8) อย่าตีความคะแนนสอบโดยใช้ข้อมูลแหล่งเดียว ควรเก็บข้อมูลจากหลายแหล่งและใช้วิธีการที่เข้าใจได้ง่าย 9) หากมีบุคคลหรือกลุ่มหนึ่งได้คะแนนต่ำให้ตั้งข้อสังเกตว่าแบบทดสอบอาจจะมีปัญหาแบบทดสอบอาจจะไม่เหมาะสมและควรถูกตัดออก 10) ในแบบทดสอบวัดสติปัญญา ถ้ากลุ่มหนึ่งกลุ่มใดทำคะแนนได้ไม่ดีนัก ให้ตรวจสอบปัจจัยที่อาจจะส่งผลและปริมาณที่อาจจะส่งผลหรือมีประโยชน์ต่อกลุ่มนั้น 11) ทบทวนข้อมูลที่ใช้เป็นหลักฐานหรือขนาดกลุ่มตัวอย่าง ขณะเดียวกัน กลุ่มที่แตกต่างสามารถเป็นตัวแทนได้ตามสัดส่วนในการหากกลุ่มตัวอย่างมาตรฐาน จำนวนที่กำหนดไว้ว่าจะน้อยเกินไปที่จะเป็นตัวแทน จึงทำให้ไม่สามารถนำผลไปใช้อ้างอิงในวงกว้าง (Generalizability) ได้ และ 12) เพิ่มความยุติธรรมทางวัฒนธรรมหรือแบบทดสอบที่ลดเนื้อหาต้านวัฒนธรรมในการวัดหรือกระบวนการตัดสินใจ แบบทดสอบเหล่านี้ถูกออกแบบให้ลดอิทธิพลของการเรียนรู้ด้านวัฒนธรรมและบรรยากาศเชิงสังคมที่ไม่เกี่ยวข้อง และดังนั้นทำการแยกความสามารถหรือการแสดงออกจากโอกาสในการเรียนรู้ให้ชัดเจนขึ้น

Khooi (2012) นำเสนอรายการกิจกรรมที่ส่งเสริมความยุติธรรมของแบบทดสอบ 5 ประการ คือ 1) ความเกี่ยวข้องของนักการศึกษา กลุ่มเล็กและคณะกรรมการที่เป็นตัวแทนในทุกขั้นตอนของการพัฒนาและการตรวจให้คะแนน 2) การประเมินความยุติธรรมที่หลากหลายวิธีโดยผู้ทบทวนที่ได้รับการฝึกอบรม 3) การวิเคราะห์ข้อคำถามอย่างสม่ำเสมอ เพื่อให้มั่นใจว่าคำถามต่าง ๆ ไม่ก่อให้เกิดความแตกต่างของกลุ่มอย่างไม่ยุติธรรม 4) การอบรมหรือฝึกอบรมอย่างจริงจังสำหรับทุกคนที่เกี่ยวข้องในการพัฒนาหรือการตรวจให้คะแนนแบบทดสอบเพื่อให้มั่นใจว่าผู้เข้าสอบทุกคนมีโอกาสเท่าเทียมกันในการแสดงทักษะและความสามารถของพวกเขา และ 5) การอำนวยความสะดวกที่เหมาะสม เช่น รูปแบบที่เป็นทางเลือก การให้เวลาเพิ่มเติมสำหรับผู้เข้าสอบบางคนที่ต้องการ

Weyant et al. (2014) นำเสนอ หลักการ 6 ประการ ของ The CaMLA's Fairness Principles เพื่อช่วยลดความลำเอียงของแบบทดสอบ โดยมีรายละเอียด คือ

1) แสดงการเคารพบุคคล (Demonstrate Respect for People) ได้แก่ 1) การให้ความสนใจเกี่ยวกับแหล่งกำเนิดของบุคคลด้วยความระมัดระวัง 2) หลีกเลี่ยงการกล่าวถึงความเหนือกว่า/ดีกว่า ในด้านเพศ เชื้อชาติ ประเทศ หรือชนชั้นทางสังคม 3) หลีกเลี่ยงการทำให้มีการจดจำหรือ

ยอมรับหลักการโดยทั่วไปของบุคคลเฉพาะกลุ่ม 4) ถ้าเป็นไปได้ ควรใช้สัญลักษณ์ และสรรพนามบอกเพศที่เป็นกลาง และ 5) รักษาสมดุลการเป็นตัวแทนที่ปรับเปลี่ยนได้ของบุคคลที่อยู่ในตำแหน่งที่มีอำนาจตลอดแบบทดสอบทั้งชุด

2) แสดงถึงการเคารพความเชื่อมั่น และความเชื่อส่วนบุคคล (Demonstrate Respect for Personal Convictions and Beliefs) รวมทั้งความเชื่อและวันหยุดทางศาสนา

3) แสดงถึงความไวต่อการรับรู้ในความแตกต่างของประชากรและความรู้ต่าง ๆ ในโลก (Demonstrate Sensivity to Population Differences and World Knowledge) โดยการ 1) ไม่สันนิษฐานเองว่ามีเพียงแง่มุมเดียวหรือการตีความของข้อเท็จจริงอันเดียว 2) ไม่สันนิษฐานเอาเองว่ามีความคุ้นเคยกับเทคโนโลยีเฉพาะด้าน 3) ไม่ทักท้วงเอาเองว่ามีความคุ้นเคยกับประเพณีของอเมริกาเหนือ 4) ไม่สันนิษฐานเอาเองว่า อเมริกา หมายถึง สหรัฐอเมริกาเพียงเท่านั้น และ 5) การหลีกเลี่ยงการอธิบายภาษาและหัวเรื่องต่าง ๆ

4) หลีกเลี่ยงการคัดค้านที่เลยเถิดไม่เหมาะสม (Avoid Undue Negativity) เช่น ความล้มเหลวทางวิชาการ (Academic Failure) ช่องว่างความสำเร็จ (Achievement Gaps) และสถานการณ์การสอบที่ยุ่งยาก (Difficult Testing-Situations) ความตระหนี่ที่ไม่จำเป็น (Unnecessary Meanness) พฤติกรรมของการเสพติดและเป็นอันตราย (Addictive and Harmful Behaviors) อาชญากรรมและการกักขังควบคุม

5) หลีกเลี่ยงหัวเรื่องทีก่อให้เกิดการโต้แย้งอย่างรุนแรงหรือก่อให้เกิดความวุ่นใจ (Avoid Unduly Controversial or Upsetting Topics) เช่น การกระตุ้นให้ผู้เข้าสอบวิจารณ์หรือสรรเสริญการปฏิบัติเพื่อชาติ นักการเมืองบางคน หรือ การโต้แย้งทางการเมืองที่รุนแรง สงคราม ความรุนแรง การฆ่าตัวตาย ความยากจน ความแตกต่างกันด้านรายได้ การเจ็บป่วยจากภาวะข้าวยากหมากแพง เชื้อโรค ความตาย ภาวะใกล้ตาย หรือระยะสุดท้าย กระบวนการทางการแพทย์ในลักษณะกราฟฟิก หรือสร้างความวุ่นใจ ความชอบที่จะหลีกเลี่ยงการแก้ปัญหาที่คุ้นเคย

6) หลีกเลี่ยงการทดสอบความรู้ที่ไม่สัมพันธ์กับโครงสร้าง (Avoid Testing Construct-Irrelevant Knowledge) เช่น คณิตศาสตร์หรือวิทยาศาสตร์ เป็นต้น นอกจากนี้ Weyant & Chisholm (2014) ได้กล่าวถึง หัวเรื่องที่ไม่เหมาะสมสำหรับการสอบของ Cambridge English Language Assessment Tests ทุกชนิด และถูกระบุไว้ในเอกสารสำหรับผู้เขียนแบบทดสอบของ CaMLA ประกอบด้วย 1) แอลกอฮอล์ 2) บุหรี่ (Cigarettes) 3) ยาเสพติด (Drugs) 4) การทดสอบ การสอบผ่านและการสอบตก (Examinations, Passing and Failing) 5) เรื่องราวหรือการอ้างอิงเชิงประวัติศาสตร์ ที่น่าจะให้ต่อต้านบางประเทศหรือคนบางกลุ่ม 6) จุดเด่นของบางประเทศ (โดยเฉพาะอย่างยิ่งสิ่งที่ปฏิบัติในประเทศใดประเทศหนึ่งอาจจะถูกข่มขู่ในด้านลบโดยคนอื่น ๆ) 7) การเมือง 8) หัวเรื่องที่ทำให้เกิดความไม่สบายใจ เช่น ความตาย การเจ็บป่วยที่กำลังจะตาย ปัญหาครอบครัว/สังคมที่รุนแรง หายนะทางธรรมชาติและสิ่งที่ก่อให้เกิดความกลัวโดยทั่วไป เช่น แมงมุม หรืองู 9) ศาสนา (รวมถึงหัวเรื่องที่เกี่ยวข้อง เช่น หัวเรื่องที่ไม่เป็นที่ยอมรับในบางศาสนา) 10) เพศ หรือการบ่งชี้เพศ และหน้าที่ต่าง ๆ ทางร่างกาย 11) ลักษณะเฉพาะ รวมถึงลัทธิชนชาตินิยม ลัทธิเพศนิยม ความคิดเก่าแก่ทางวัฒนธรรม และทัศนคติซึ่งสามารถตีความได้ว่าเป็นการ

อุปการะประเทศ วัฒนธรรม ความเชื่อหรือบุคคลอื่น ๆ และ 12) สงคราม โดยข้อมูลในหัวข้อข้างต้นนี้ มีได้บางส่วนแต่ไม่ได้เป็นประเด็นหลักของเนื้อหาที่จะนำมาออกสอบ

Munson Robin and Parton Cinda (2013) กล่าวว่า ผลการสอบไม่ควรได้รับอิทธิพลจากความทุพพลภาพหรือไร้ความสามารถของนักเรียน ชาติพันธุ์ เพศ หรือความสามารถทางภาษาอังกฤษ นักเรียนกับแบบทดสอบมีพันธะผูกพันกันที่จะต้องทำให้แบบทดสอบไม่มีความลำเอียงและมีความยุติธรรม โดยมีการตรวจสอบและควบคุมจากการออกแบบ การพัฒนา การทดลองสอบ และการใช้

แบบทดสอบจริง นักพัฒนาแบบทดสอบต้องทำให้แบบทดสอบปราศจากความลำเอียงและยุติธรรมสำหรับนักเรียนที่มีภูมิหลังอันแตกต่างกันมาก โดย Munson & Parton (2013) ได้นำเสนอกิจกรรมต่าง ๆ ตลอดกระบวนการพัฒนานาแบบทดสอบ เพื่อช่วยในการลดความลำเอียงของแบบทดสอบ ดังนี้

กิจกรรมการเขียนแบบทดสอบ ประกอบด้วย 1) การฝึกอบรมภายในเกี่ยวกับความลำเอียงและการรับรู้ถึงความลำเอียงให้กับผู้เขียนแบบทดสอบแต่ละคน 2) รวบรวมรายการเกี่ยวกับความลำเอียง และการรับรู้ถึงความลำเอียงที่จะให้ผู้เขียนแบบทดสอบใช้ในระหว่างการเขียนแบบทดสอบ และ 3) สำหรับภาระงานการปฏิบัติ ต้องให้นักเรียนอ่านและเข้าใจข้อมูลตามบริบท ซึ่งมักจะมีคำศัพท์เฉพาะด้านที่ไม่เกี่ยวข้องกับเนื้อหา อีกทั้งยังพัฒนากิจกรรมในห้องเรียนสำหรับภาระงานภาคปฏิบัติแต่ละอย่าง กิจกรรมในห้องเรียน อธิบายข้อมูลตามบริบท และทบทวนคำศัพท์ที่ไม่เกี่ยวข้องกับโครงสร้างเนื้อหา กิจกรรมที่จะทำให้ข้อมูลตามบริบทและคำศัพท์เป็นที่เข้าใจของนักเรียนทั้งหมด

กิจกรรมการทบทวนความลำเอียงและความยุติธรรม ประกอบด้วย 1) มีการทบทวนภายในของข้อคำถามแต่ละข้อ เพื่อตรวจสอบความลำเอียงและการรับรู้ถึงความลำเอียงของแบบทดสอบ โดยเฉพาะ 2) ส่งตัวอย่างข้อคำถามให้กรรมการภายนอกตรวจสอบความลำเอียงและการรับรู้ต่อความลำเอียง อาจจะเป็นนักการศึกษาและบุคลากรภาควิชาการศึกษาและภาควิชาหลักสูตรและการสอน ซึ่งเป็นผู้ที่มีความเชี่ยวชาญด้านความลำเอียงและความยุติธรรมในการทดสอบเข้าร่วมในการทบทวนด้วย ผู้ทบทวนจะได้รับการอบรมและใช้แบบตรวจสอบรายการ (Checklist) เกี่ยวกับความลำเอียงและการรับรู้ด้านความลำเอียง และ 3) ผู้เขียนแบบทดสอบทำการประเมินข้อเสนอแนะเพื่อค้นหารูปแบบประเด็นที่จะต้องกล่าวถึงหรือพิจารณาเป็นพิเศษ

บทบาทของการทดลองใช้แบบทดสอบ คือ 1) การทดลองใช้แบบทดสอบกับนักเรียนที่เป็นตัวแทนของกลุ่มตัวอย่างที่มีความต่างกันในด้านฐานะทางสังคมและเศรษฐกิจ ความสามารถทางภาษาอังกฤษ ภูมิสำเนา เพศ และความทุพพลภาพ การทดลองใช้แบบทดสอบมีเป้าหมายเพื่อตรวจสอบรูปแบบ หน้าที่ และลักษณะของข้อคำถาม 2) การทดลองกลุ่มเล็กเพื่อหาความสัมพันธ์ระหว่างการตรวจให้คะแนนโดยเครื่องและการตรวจมือ คำตอบของนักเรียนที่นำมาใช้เป็นตัวแทนของกลุ่มประชากรนักเรียนที่แตกต่างกัน 3) การสำรวจความคิดเห็นของนักเรียนหลังการสอบซึ่งมีข้อคำถามเกี่ยวกับการมีส่วนร่วมเกี่ยวข้องในการทดสอบ และ 4) การสำรวจความคิดเห็นของผู้บริหารการสอบและนักเรียน เพื่อสอบถามการมีปฏิสัมพันธ์ของพวกเขาเกี่ยวกับภาระงานภาคปฏิบัติและกิจกรรมในห้องเรียนที่สัมพันธ์สอดคล้องกับภาระงานภาคปฏิบัติ

บทบาทของการทดสอบภาคสนาม คือ การทดสอบภาคสนาม เป็นการศึกษเพื่อประมวผลของการใช้ข้อคำถามชนิดต่าง ๆ รวมทั้งการวิเคราะห์ความสามารถของนักเรียน มีการใช้สิ่งอำนวยความสะดวกและการสนับสนุนตามที่กำหนดไว้จำนวนมาก

นอกจากนี้ Munson & Parton (2013) ได้นำเสนอข้อคำถามเพื่อใช้ในการตรวจสอบความลำเอียงในแบบทดสอบ แบ่งออกเป็น 7 ด้าน คือ 1) ด้านเชื้อชาติ/ชาติพันธุ์/วัฒนธรรม 2) ด้านเพศ 3) ด้านศาสนา 4) ด้านอายุ 5) ด้านความทุพพลภาพ 6) ด้านสังคม-เศรษฐกิจ และ 7) ด้านอื่น ๆ โดยมีรายละเอียดในแต่ละด้าน ดังนี้

1) การพิจารณาด้านเชื้อชาติ/ชาติพันธุ์/วัฒนธรรม มีคำถามในการตรวจสอบ 5 ข้อ คือ 1) แบบทดสอบได้ตอบสนองต่อผู้เข้าสอบสัญชาติใดหรือกลุ่มชาติพันธุ์ใดมากกว่ากลุ่มอื่น ๆ หรือไม่ 2) แบบทดสอบได้แสดงให้เห็นลักษณะที่เป็นลบหรือเฉพาะหรือการปฏิบัติต่อกลุ่มใดอย่างไม่เหมาะสมอย่างน้อยหนึ่งชนิดที่เกี่ยวกับชาติพันธุ์หรือวัฒนธรรมในหรือไม่ 3) แบบทดสอบมีการใช้ภาษา เนื้อหาหรือบริบทที่ไม่สามารถเข้าถึงหรือไม่เป็นที่รู้จักอย่างแพร่หลายต่อกลุ่มชาติพันธุ์ สัญชาติ หรือวัฒนธรรมใดวัฒนธรรมหนึ่งหรือไม่ 4) แบบทดสอบมีภาษาหรือสาระที่ขัดกับค่านิยมของกลุ่มวัฒนธรรมใดโดยเฉพาะหรือไม่ และ 5) แบบทดสอบปฏิบัติอย่างไม่เหมาะสมเกี่ยวกับประสบการณ์ของมนุษย์ที่สำคัญหรือน่าสลดใจหรือไม่

2) การพิจารณาด้านเพศ มีคำถามในการตรวจสอบ 5 ข้อ คือ 1) แบบทดสอบมีการใช้ภาษา สาระ หรือบริบทที่โจมตีต่อเพศชายหรือเพศหญิงหรือไม่ 2) แบบทดสอบมีการใช้ภาษา สาระ หรือบริบทที่ไม่สามารถเข้าถึงหรือไม่เป็นที่รู้จักอย่างกว้างขวางต่อเพศชายหรือเพศหญิงหรือไม่ 3) แบบทดสอบเป็นตัวแทนในด้านลบหรือลักษณะเฉพาะของบุคคลตามเพศหรือความชอบด้านเพศหรือไม่ 4) ความสำเร็จในการทำแบบทดสอบของผู้เข้าสอบต้องอาศัยประสบการณ์ ชนิดของโครงสร้างครอบครัวชนิดใดชนิดหนึ่งโดยเฉพาะหรือไม่ และ 5) แบบทดสอบนำเสนอการประชดประชัน เสียดสีด้านเพศหรือไม่

3) การพิจารณาด้านศาสนา มีคำถามในการตรวจสอบ 6 ข้อ คือ 1) แบบทดสอบสนับสนุนศาสนาใดศาสนาหนึ่งมากกว่าศาสนาอื่น ๆ หรือไม่ 2) แบบทดสอบนำเสนอเรื่องราวของผู้นำศาสนาในลักษณะที่เป็นลบหรือลักษณะเด่นเฉพาะอย่างน้อย 1 ศาสนาหรือไม่ 3) แบบทดสอบใช้ภาษา สาระ หรือบริบทที่ไม่เหมาะสมต่อศาสนาใดศาสนาหนึ่งหรือไม่ ตัวอย่างเช่น การอ้างอิงถึงการใช้เวทมนตร์ คาถา วิชาการใช้เวทมนตร์ และมายากล 4) แบบทดสอบใช้ภาษา สาระ หรือบริบทเกี่ยวกับศาสนาที่ไม่สามารถเข้าใจได้โดยทั่วไปโดยสมาชิกในทุกกลุ่มศาสนาหรือไม่ 5) แบบทดสอบต้องการให้ผู้ปกครอง ครู หรือผู้เข้าสอบสนับสนุนตำแหน่งที่อยู่ตรงข้ามมาความเชื่อด้านศาสนาของพวกเขาหรือไม่ และ 6) แบบทดสอบต้องการความรู้ของบุคคล เหตุการณ์ หรือกลุ่มต่าง ๆ ที่ไม่เป็นที่คุ้นเคยของนักเรียนทั้งหมดหรือไม่

4) การพิจารณาด้านอายุ มีคำถามในการตรวจสอบ 4 ข้อ คือ 1) แบบทดสอบตอบสนองกลุ่มอายุใดกลุ่มหนึ่งมากกว่ากลุ่มอื่น ยกเว้นในบริบทหรือประสบการณ์หรือการมีอายุมากกว่ามีความจำเป็นหรือเป็นประโยชน์หรือไม่ 2) แบบทดสอบแสดงให้เห็นถึงด้านลบหรือลักษณะเด่นเฉพาะของกลุ่มอายุใดหรือไม่ 3) แบบทดสอบมีการใช้ภาษา สาระ หรือบริบทที่ตรงข้ามหรือไม่เหมาะสมกับกลุ่ม

อายุใดกลุ่มหนึ่งมากกว่ากลุ่มอื่น ๆ หรือไม่ และ 4) แบบทดสอบมีการใช้ภาษา สาระ หรือบริบทที่ไม่สามารถเข้าถึงต่อกลุ่มอายุใด หรือ กลุ่มอายุอีกหลายกลุ่มที่ถูกทดสอบหรือไม่

5) การพิจารณาด้านความพึงพอใจ มีคำถามในการตรวจสอบ 4 ข้อ คือ 1) แบบทดสอบมีการลดระดับบุคคลตามพื้นฐานของลักษณะทางร่างกายหรือความสามารถด้านร่างกาย สมอง หรืออารมณ์หรือไม่ 2) แบบทดสอบมุ่งประเด็นไปที่ความพึงพอใจของบุคคลมากกว่าการนำเสนอเกี่ยวกับบุคคลในภาพรวมหรือไม่ 3) แบบทดสอบมีการใช้ภาษา สาระ หรือบริบทที่ไม่เหมาะสมต่อบุคคลที่พึงพอใจหรือไม่ และ 4) แบบทดสอบสื่อถึงด้านลบหรือไม่ดี หรือลักษณะเด่นเฉพาะของบุคคลที่มีความพึงพอใจอย่างน้อย 1 คนหรือไม่

6) การพิจารณาด้านสังคม-เศรษฐกิจ มีคำถามในการตรวจสอบ 6 ข้อ คือ 1) แบบทดสอบแนะนำว่า ความร่ำรวยมีสัมพันธภาพกับคุณความดีหรือสติปัญญาหรือไม่ 2) แบบทดสอบใช้ภาษา สาระ หรือบริบทที่ไม่เหมาะสมต่อบุคคลที่มีลักษณะทางเศรษฐกิจเฉพาะกลุ่มหรือไม่ 3) แบบทดสอบแสดงความชื่นชอบ ชื่นชม ยกย่องกลุ่มสังคม-เศรษฐกิจกลุ่มใดกลุ่มหนึ่งมากกว่ากลุ่มอื่น ๆ หรือไม่ 4) กลุ่มเฉพาะทางด้านชาติพันธุ์ วัฒนธรรม หรือศาสนาถูกทำให้เด่น เพราะมีสถานภาพทางสังคม เศรษฐกิจที่เฉพาะหรือไม่ 5) แบบทดสอบมีส่วนดึงดูดใจหรือลดคุณค่าบุคคลโดยอาศัยสถานะทางสังคม-เศรษฐกิจหรือไม่ และ 6) แบบทดสอบมีการชี้แนะหรือไม่ว่า ความยากจนสัมพันธ์กับพฤติกรรมด้านลบที่เพิ่มขึ้นในสังคม

7) การพิจารณาด้านอื่น ๆ มีคำถามในการตรวจสอบ 8 ข้อ 1) แบบทดสอบต้องทำให้นักเรียนขั้นสุดตำแหน่งโดยต้องอาศัย อำนาจ หรือบารมีของพ่อแม่หรือไม่ 2) แบบทดสอบนำเสนอ สงคราม หรือความรุนแรงผ่านตัวอักษร หรือ สิ่งที่เป็นตัวอย่างรวมมากเกินไป 3) แบบทดสอบนำเสนอหัวข้อเรื่องที่ยิวโตะ หรือก่อให้เกิดความขัดแย้ง เช่น ความตาย สงคราม การทำแท้ง หรือ การฆ่าหรือช่วยให้บุคคลที่ทุกข์ทรมานจากโรคหรืออาการบาดเจ็บที่รักษาไม่หายตายอย่างไม่เจ็บปวด ยกเว้นเป็นเรื่องที่ต้องทำให้สอดคล้องกับมาตรฐานของรัฐหรือไม่ 4) แบบทดสอบคาดหวังให้เกิดค่านิยมที่ไม่ได้เป็นของผู้เข้าสอบทุกคนหรือไม่ 5) แบบทดสอบลดระดับบุคคลหรือวัฒนธรรมจากภูมิภาคต่าง ๆ ของประเทศ หรือรัฐ หรือไม่ 6) แบบทดสอบยอมรับหรือไม่เห็นด้วยในการยกเลิกสัญญา พฤติกรรมทางด้านอาชญากรรมผิดกฎหมาย หรือเป็นอันตรายหรือไม่ 7) แบบทดสอบต้องทำให้ผู้เข้าสอบเปิดเผยค่านิยมที่พวกเขาต้องยึดถือไว้ว่าเป็นความลับหรือไม่ และ 8) แบบทดสอบมีการใช้ภาษาที่ทำให้เกิดความตระหนักถึงภาษาที่น่าละอายหรือภาษาของพวกเร้าร่วมเพศ

สรุปได้ว่า แนวทางการลดความลำเอียงของแบบทดสอบที่จะนำมาใช้ในการวิจัยครั้งนี้ ประกอบด้วย 2 ประการ คือ 1) จัดทำรายการหัวข้อเรื่องที่ไม่เหมาะสมในการนำมาออกสอบให้นักพัฒนาแบบทดสอบได้ศึกษาและใช้ตรวจสอบเนื้อหาที่เหมาะสมก่อนดำเนินการเขียนข้อคำถาม และ 2) จัดทำรายการข้อคำถามเพื่อใช้ตรวจสอบเนื้อหาและข้อคำถามอีกครั้ง โดยให้นักเขียนข้อคำถามประเมินตนเองหลังจากเขียนข้อคำถามเสร็จ และให้ผู้เชี่ยวชาญด้านการทดสอบภาษา ตรวจสอบหาความลำเอียงของแบบทดสอบอีกครั้งก่อนนำแบบทดสอบไปทดลองใช้

การวิเคราะห์ความลำเอียงของแบบทดสอบ

หลังจากการทดลองใช้แบบทดสอบกับกลุ่มทดลองหรือการใช้แบบทดสอบฉบับจริงกับกลุ่มตัวอย่าง นักพัฒนาแบบทดสอบควรมีการวิเคราะห์ความลำเอียงของข้อคำถามต่าง ๆ เพื่อให้มั่นใจว่าเมื่อมีการรายงานผลการสอบโดยใช้ค่าสถิติแล้ว แบบทดสอบที่พัฒนาขึ้นนั้นมีข้อใดที่ควรตัดออกหรือควรปรับปรุงเนื่องจากยังมีค่าความลำเอียงอยู่ในระดับที่ก่อให้เกิดความไม่ยุติธรรมกับผู้เข้าสอบทุกคน วิธีการวิเคราะห์ความลำเอียงที่นักพัฒนาแบบทดสอบสามารถนำมาใช้ได้ เช่น 1) วิธีการแปลงค่าความยาก (Transformed Item Difficulty Methods) 2) วิธีใช้ค่าอำนาจจำแนก (Item Discrimination Procedures) และ 3) วิธีใช้ตารางความสอดคล้อง (Contingency Table Approaches) (วารสารณ์ ประจวบเกษรชษฐ์, 2561) โดยมีรายละเอียดแต่ละวิธีดังนี้

วิธีการแปลงค่าความยาก (Transformed Item Difficulty Methods)

Angoff and Ford (1973) ได้เสนอวิธีการแปลงค่าความยากในการตรวจสอบความลำเอียงของข้อสอบ โดยนิยามข้อสอบที่มีความลำเอียงว่าเป็นข้อสอบที่เบี่ยงเบนไปจากเส้นแกนหลัก โดยการวิเคราะห์นั้นจะใช้ค่าความยากของข้อสอบแต่ละข้อ จากกลุ่มผู้สอบทั้ง 2 กลุ่ม แล้วนำมาแปลงให้เป็นค่าความยากมาตรฐานที่มีค่าเฉลี่ย 13 และส่วนเบี่ยงเบนมาตรฐาน 4 ด้วยสูตร $\Delta = 13 + 4Z$ จากนั้นนำค่าความยากมาตรฐานของทั้ง 2 กลุ่มมาสร้างเป็นแผนภาพกระจายจุดกระจาย ถ้าทั้ง 2 กลุ่มมีความสามารถกลุ่มมีความสามารถเท่าเทียมกันแล้ว จุดต่าง ๆ จะต้องมิลักษณะเป็นวงรีล้อมรอบเส้นแกนหลัก 45 องศาจากจุด 0 ถ้ากลุ่มผู้สอบมีความสามารถแตกต่างกัน จุดต่าง ๆ จะกระจายจุดกระจายออกไปจากเส้นแกนหลัก 45 องศาจากจุด 0 ข้อสอบที่มีความลำเอียงคือข้อสอบที่มีค่าลำดับอยู่ห่างจากเส้นแกนหลักมาก ๆ ซึ่งระยะห่างระหว่างจุดคู่ลำดับกับเส้นแกนหลักคำนวณได้โดยใช้สูตร

$$d_g = \frac{(a\Delta_{g1} - \Delta_{g2} + b)}{\sqrt{a^2 + 1}}$$

$$\text{เมื่อ } a = \frac{(s_2^2 - s_1^2) + \sqrt{(s_2^2 - s_1^2)^2 + 4r_{12}^2 s_1^2 s_2^2}}{2r_{12} s_1 s_2}$$

$$b = \bar{x}_2 - a\bar{x}_1$$

d_g = ระยะห่างระหว่างข้อสอบข้อ g ไปยังเส้นแกนหลัก

Δ_{g1} = ค่าความยากมาตรฐานของข้อสอบข้อ g ในกลุ่ม 1

Δ_{g2} = ค่าความยากมาตรฐานของข้อสอบข้อ g ในกลุ่ม 2

s_1 = ส่วนเบี่ยงเบนมาตรฐานของค่าความยากมาตรฐานกลุ่ม 1

s_2 = ส่วนเบี่ยงเบนมาตรฐานของค่าความยากมาตรฐานกลุ่ม 2

\bar{x}_1 = ค่าเฉลี่ยของค่าความยากมาตรฐานในกลุ่ม 1

$$\begin{aligned} \bar{x}_2 &= \text{ค่าเฉลี่ยของค่าความยากมาตรฐานในกลุ่ม 2} \\ r_{12} &= \text{สหสัมพันธ์ระหว่างค่าความยากมาตรฐานของกลุ่ม 1 และ 2} \end{aligned}$$

ค่า d_g ของข้อใดมีค่ามากแสดงว่ามีความลำเอียงมาก แต่เกณฑ์ของค่า d_g นี้ยังไม่พบว่ามีเอกสารหรือตำราเล่มใดระบุไว้ว่าค่าเท่าใดจึงจะถือว่าข้อสอบนั้นลำเอียง แต่ในงานวิจัยที่ศึกษาเกี่ยวกับความลำเอียงด้วยวิธีแปลงค่าความยาก จะใช้เกณฑ์ 0.75 หรือก็คือ $d_g < -0.75$ และ $d_g > +0.75$ จึงจะถือว่าข้อสอบนั้นมีความลำเอียง

วิธีใช้ค่าอำนาจจำแนก (Item Discrimination Procedures)

วิธีนี้เหมือนกับวิธีการแปลงค่าความยาก โดยการคำนวณหาค่าอำนาจจำแนกของข้อสอบของทั้ง 2 กลุ่ม แล้วนำมาสร้างเป็นแผนภาพกระจัดกระจาย แล้วตรวจสอบดูว่า ค่าลำดับของข้อสอบข้อใดที่กระจายออกจากเส้นแกนหลัก 45 องศา มาก ถ้าว่าข้อสอบข้อนั้นมีความลำเอียง ส่วนกรีนและดราเปอร์ (Green and Draper) เสนอให้ใช้สหสัมพันธ์พอยท์ไบเซเรียล (Point-biserial correlation) ในการคำนวณหาค่าอำนาจจำแนก แล้วพิจารณาว่าข้อสอบข้อใดที่จำแนกได้ดีในกลุ่มหนึ่ง แต่จำแนกได้ไม่ดีในอีกกลุ่มหนึ่ง จะถือว่าข้อสอบข้อนั้นมีความลำเอียง

วิธีใช้ตารางความสอดคล้อง (Contingency Table Approaches)

วิธีการใช้ตารางความสอดคล้องได้เสนอแนะครั้งแรกโดยชูนีแมน (Scheuneman) บนพื้นฐานที่ว่า บุคคลที่มีความสามารถเท่ากัน ย่อมมีความน่าจะเป็นในการตอบถูกเท่ากัน ดังนั้นจึงเกิดการสร้างตารางความสอดคล้องระหว่างกลุ่มและความสามารถ ซึ่งจะนิยามเป็นพิสัยของคะแนนในช่วงต่าง ๆ เทคนิคที่อยู่ในการใช้ตารางความสอดคล้องนี้ คือ วิธีโคสแควร์ โดยชูนีแมน และคามิลลี (Scheuneman and Camilli) ใช้เทคนิคนี้และนิยามข้อสอบที่มีความยุติธรรมไว้ว่า ในแต่ละกลุ่มของผู้สอบที่อยู่ในช่วงคะแนนเดียวกัน สัดส่วนในการตอบถูกของผู้สอบแต่ละกลุ่มจะต้องเท่ากัน เทคนิคโคสแควร์จะแบ่งคะแนนที่สังเกตได้ออกเป็นช่วง ในแต่ละช่วงจะแสดงการตอบถูกในแต่ละกลุ่ม ถ้าสัดส่วนในแต่ละช่วงของแต่ละกลุ่มมีความแตกต่างกันมาก เป็นหลักฐานบ่งชี้ถึงความลำเอียงของข้อสอบ

ในการวิจัยครั้งนี้จะวิเคราะห์ความลำเอียงของแบบทดสอบในชั้นทดลองใช้แบบทดสอบโดยใช้วิธีการแปลงค่าความยาก (Transformed Item Difficulty Methods) ตามสูตรและวิธีที่ได้กล่าวไว้ข้างต้น

สรุปแนวทางการลดความลำเอียงที่จะนำมาใช้ในการวิจัย

แนวทางการลดความลำเอียงที่จะนำมาใช้ในการวิจัยครั้งนี้ ประกอบด้วย 3 แนวทาง คือ 1) การตรวจสอบรายการหัวข้อที่ไม่เหมาะสมสำหรับการนำมาทดสอบ โดยให้ผู้ออกแบบทดสอบเป็นผู้ตรวจสอบและพิจารณาความเหมาะสม 2) การประเมินความยุติธรรมของหัวเรื่องหรือเนื้อหาต่าง ๆ ที่นำมาใช้ออกข้อสอบโดยผู้เชี่ยวชาญด้านการทดสอบภาษาอังกฤษเป็นผู้ประเมินก่อนที่จะมีการนำแบบทดสอบไปทดลองใช้กับกลุ่มทดลอง และ 3) การประเมินความลำเอียงของแบบทดสอบโดยใช้วิธีการแปลงค่าความยาก (Transformed Item Difficulty Methods) ในขั้นทดลองใช้แบบทดสอบ โดยใช้เกณฑ์ 0.75 หรือก็คือ $d_g < -0.75$ และ $d_g > +0.75$ จึงจะถือว่าข้อสอบนั้นมีความลำเอียง

2.2.2.4 การวิเคราะห์แบบทดสอบรายข้อ (Item Analysis)

Brown (1995) กล่าวว่า ในการวิเคราะห์ผลการสอบรายข้อ ส่วนใหญ่จะใช้สถิติ 3 ชนิด คือ 1) สถิติเชิงบรรยาย (Descriptive Statistics) 2) สถิติสำหรับการคำนวณค่าความยากง่าย และ 3) สถิติสำหรับการคำนวณค่าอำนาจจำแนก โดยมีรายละเอียด ดังนี้

สถิติเชิงบรรยาย (Descriptive Statistics)

สถิติเชิงบรรยาย (Descriptive Statistics) ใช้ในการรายงานผลการสอบรายข้อและทั้งหมดในรูปของค่าสูงสุด-ต่ำสุด ร้อยละ ค่าเฉลี่ย และ ส่วนเบี่ยงเบนมาตรฐาน

สถิติสำหรับการคำนวณค่าความยากง่าย

แบบทดสอบปรนัย: ความยากง่ายของแบบทดสอบรายข้อ (Item facility / Item difficulty / Item easiness) หรือ P เป็นค่าสถิติที่ใช้ตรวจสอบร้อยละของจำนวนผู้เข้าสอบที่ทำแบบทดสอบถูกต้อง วิธีการคำนวณค่า P ให้บวกจำนวนผู้เข้าสอบที่ตอบแบบทดสอบถูกต้องของแบบทดสอบแต่ละข้อ และหารด้วยจำนวนรวมของผู้เข้าสอบทั้งหมด โดยมีสูตรในการคำนวณ ดังนี้

$$P = \frac{N_{\text{correct}}}{N_{\text{total}}}$$

โดยที่

N_{correct} = จำนวนผู้สอบที่ตอบถูก

N_{total} = จำนวนผู้สอบที่เข้าสอบทั้งหมด

แบบทดสอบอัตนัย: การวิเคราะห์ค่าความยากง่ายของแบบทดสอบอัตนัย (Item Difficulty) มีสูตรในการคำนวณ (Scannell and Tracy, 1975) ดังนี้

$$P = \frac{S_U + S_L + (N_T)X_{min}}{N_T (X_{max} + X_{min})}$$

เมื่อ	S_U	=	คะแนนรวมของกลุ่มที่ได้คะแนนสูง
	S_L	=	คะแนนรวมของกลุ่มที่ได้คะแนนต่ำ
	N_T	=	จำนวนผู้เข้าสอบทั้งหมดที่ใช้ในการวิเคราะห์
	X_{max}	=	คะแนนสูงสุด
	X_{min}	=	คะแนนต่ำสุด

การตีความค่าความยากง่าย (Ebel, 1978) ตามเกณฑ์นี้ คือ 0.40 – 0.60 หมายถึง ดีมาก, 0.31 – 0.39, 0.61 – 0.69 หมายถึง ดี, 0.20 – 0.29, 0.70 – 0.80 หมายถึง พอใช้, ควรปรับปรุงง่ายเกินไป, และ ≥ 0.81 และ $0.19 \leq$ ควรปรับปรุงยากเกินไป

สถิติสำหรับการคำนวณค่าอำนาจจำแนก

การหาค่าอำนาจจำแนก แบ่งได้เป็น 2 ชนิด คือ ค่าอำนาจจำแนกของแบบทดสอบปรนัย และค่าอำนาจจำแนกของแบบทดสอบอัตนัย โดยมีรายละเอียดและสูตรการคำนวณ ดังนี้ การวิเคราะห์หาค่าอำนาจจำแนก (Item Discrimination) ของแบบทดสอบปรนัยทั้งรายข้อ ตอน ทักษะ และทั้งฉบับ มีสูตรในการคำนวณ (Brown, 2005: 69) ดังนี้

$$R = IF_{upper} - IF_{lower}$$

โดยที่

R = ค่าอำนาจจำแนกของแบบทดสอบแต่ละข้อ

IF_{upper} = ค่าความยากสำหรับกลุ่มที่ได้คะแนนสูงในการทำแบบทดสอบทั้งชุด

IF_{lower} = ค่าความยากสำหรับกลุ่มที่ได้คะแนนต่ำในการทำแบบทดสอบทั้งชุด

การวิเคราะห์ค่าอำนาจจำแนกของแบบทดสอบอัตนัย มีสูตรที่ใช้ในการคำนวณ (Scannell and Tracy, 1975) ดังนี้

$$R = \frac{S_U + S_L}{N_U (X_{max} - X_{min})}$$

เมื่อ	R	=	ค่าอำนาจจำแนก
	S_U	=	คะแนนรวมของกลุ่มที่ได้คะแนนสูง

S_L	=	คะแนนรวมของกลุ่มที่ได้คะแนนต่ำ
N_U	=	จำนวนผู้เข้าสอบในกลุ่มที่ได้คะแนนสูง
X_{max}	=	คะแนนสูงสุด
X_{min}	=	คะแนนต่ำสุด

การตีความค่าอำนาจจำแนก (Ebel, 1978) ตามเกณฑ์ คือ 0.60 ขึ้นไป หมายถึง ดีมาก, 0.35 – 0.59 หมายถึง ดี, 0.20 – 0.34 หมายถึง พอใช้, ต่ำกว่า 0.20 หมายถึง ต่ำ

2.2.2.5 การสอบถามความคิดเห็นผู้เกี่ยวข้องโดยใช้แบบสอบถาม

ในการสอบถามความคิดเห็นผู้เกี่ยวข้องเกี่ยวกับคุณภาพของแบบทดสอบโดยใช้แบบสอบถาม ผู้เชี่ยวชาญด้านการทดสอบภาษาอังกฤษหลายคน เช่น Bachman & Palmer (1996: 139-155) Genesee & Upshur (1996: 241-244) Brown (2004: 30-38) และ อัจฉรา วงศ์โสธร (2544: 263-269) ได้ให้แนวคิดที่เป็นประโยชน์และตัวอย่างข้อคำถามที่นักพัฒนาแบบทดสอบสามารถนำมาปรับใช้และสร้างเป็นแบบสอบถามเพื่อประเมินคุณภาพด้านต่าง ๆ ของแบบทดสอบที่พัฒนาได้ ดังมีรายละเอียดต่อไปนี้

Bachman & Palmer (1996: 139-155) นำเสนอข้อคำถามสำหรับการตรวจสอบคุณภาพของแบบทดสอบใน 6 ด้าน คือ ความเชื่อมั่น ความตรงเชิงโครงสร้าง ความสมจริง การมีปฏิสัมพันธ์ ผลกระทบ และความสะดวกใช้ โดยแต่ละด้านมีรายละเอียดดังนี้

ด้านความเชื่อมั่น (Reliability) มีคำถามสำหรับการประเมิน 5 ข้อ คือ 1) ลักษณะของบริบทแวดล้อมในการสอบแต่ละครั้งแตกต่างกันมากน้อยเพียงใด 2) ลักษณะของคำสั่งในแบบทดสอบผันแปรไปในทิศทางที่จะไม่จูงใจ จากแบบทดสอบตอนหนึ่งไปสู่แบบทดสอบตอนอื่น ๆ หรือต่อแบบทดสอบที่มีรูปแบบต่างกันมากน้อยเพียงใด 3) ลักษณะของสารรับเข้าในแบบทดสอบผันแปรไปในทิศทางที่จะไม่จูงใจ จากแบบทดสอบตอนหนึ่งไปยังแบบทดสอบตอนอื่น ๆ จากภาระงานหนึ่งไปสู่ภาระงานอื่น ๆ และต่อแบบทดสอบที่มีรูปแบบแตกต่างกันมากน้อยเพียงใด 4) ลักษณะของการตอบสนองต่อแบบทดสอบที่คาดหวังผันแปรไปในทิศทางที่จะไม่จูงใจ จากตอนหนึ่งของแบบทดสอบสู่ตอนอื่น ๆ หรือต่อแบบทดสอบที่มีรูปแบบแตกต่างกันมากน้อยเพียงใด และ 5) ลักษณะของความสัมพันธ์ระหว่างสารรับเข้าและการตอบกลับผันแปรไปในทิศทางที่จะไม่จูงใจ จากตอนหนึ่งของแบบทดสอบไปยังตอนอื่น ๆ หรือในแบบทดสอบที่มีรูปแบบแตกต่างกันมากน้อยเพียงใด

ด้านความตรงเชิงโครงสร้าง (Construct Validity) มีคำถามสำหรับการประเมิน 5 ข้อ คือ 1) กรอบความสามารถทางภาษาสำหรับแบบทดสอบชุดนี้ชัดเจนและไม่คลุมเครือหรือไม่ 2) กรอบความสามารถทางภาษาสำหรับแบบทดสอบชุดนี้สัมพันธ์กับจุดประสงค์ของแบบทดสอบ 3) ภาระงานของแบบทดสอบสะท้อนค่านิยมเชิงโครงสร้างมากน้อยเพียงใด 4) กระบวนการตรวจให้คะแนนสะท้อนค่านิยมเชิงโครงสร้างมากน้อยเพียงใด และ 5) คะแนนที่ได้จากการสอบนำไปสู่การตีความความสามารถทางภาษาของผู้เข้าสอบตรงตามที่ตั้งใจไว้หรือไม่

ด้านความสมจริง (Authenticity) มีคำถามสำหรับประเมิน 2 ข้อ คือ 1) คำอธิบายของภาระงานในปริศนาการใช้ภาษาในอนาคต ครอบคลุมข้อมูลเกี่ยวกับบริบท สารรับเข้า พฤติกรรมการตอบ และความสัมพันธ์ระหว่างสารรับเข้าและพฤติกรรมการตอบมากน้อยเพียงใด และ 2) ลักษณะของภาระงานการทดสอบสอดคล้องกับภาระงานที่ได้กล่าวไว้ในปริศนาการใช้ภาษาในอนาคตมากน้อยเพียงใด

ด้านการมีปฏิสัมพันธ์ของแบบทดสอบ (Interactiveness) มีคำถามสำหรับการประเมินที่แบ่งออกเป็น 6 ชนิด คือ การมีปฏิสัมพันธ์ที่เกี่ยวข้องกับความรู้ในหัวเรื่องของผู้เข้าสอบ จำนวน 1 ข้อ คือ ภาระงานแสดงถึงระดับของความรู้ในหัวเรื่องที่เหมาะสมมากน้อยเพียงใด และนักพัฒนาแบบทดสอบสามารถคาดหวังให้ผู้เข้าสอบมีระดับความรู้ในหัวเรื่องนี้ ได้มากน้อยเพียงใด การมีปฏิสัมพันธ์ที่เกี่ยวข้องกับความเหมาะสมของภาระงานในแบบทดสอบต่อคุณลักษณะส่วนตัวของผู้เข้าสอบ จำนวน 2 ข้อ คือ 1) มีการบรรยายลักษณะส่วนตัวของผู้เข้าสอบไว้ในคำกล่าวถึงการออกแบบ (Design Statement) ของแบบทดสอบมากน้อยเพียงใด และ 2) ลักษณะของภาระงานในแบบทดสอบเหมาะสมกับผู้เข้าสอบที่มีลักษณะส่วนตัวที่ได้ระบุไว้แล้วมากน้อยเพียงใด การมีปฏิสัมพันธ์ที่เกี่ยวข้องกับความรู้ด้านภาษาของผู้เข้าสอบ จำนวน 1 ข้อ คือ การทำภาระงานสอบให้สำเร็จเกี่ยวข้องกับประเด็นความรู้ด้านภาษาของผู้เข้าสอบมากน้อยเพียงใด การมีปฏิสัมพันธ์ที่เกี่ยวข้องกับทักษะภาษาในภาระงานสอบ จำนวน 1 ข้อ คือ นอกเหนือจากทักษะการสาธิตแบบง่ายๆ ของความสามารถทางภาษามีทักษะอะไรบางอย่างที่เกี่ยวข้องในการทำความเข้าใจสารรับเข้าและสร้างสรรค์คำตอบ การมีปฏิสัมพันธ์ที่เกี่ยวข้องกับกลยุทธ์ในการอภิธานของผู้เข้าสอบ จำนวน 2 ข้อ คือ 1) ภาระงานในแบบทดสอบสัมพันธ์ซึ่งกันและกันมากน้อยเพียงใด และ 2) มีการจัดโอกาสสำหรับการใช้กลยุทธ์ที่เกี่ยวข้องมากน้อยเพียงใด และ การมีปฏิสัมพันธ์ที่เกี่ยวข้องกับประสบการณ์เดิมของผู้เข้าสอบในการตอบสนองต่อภาระงานในแบบทดสอบ มี 1 ข้อ คือ ภาระงานในแบบทดสอบช่วยกระตุ้นการตอบสนองที่เกิดจากประสบการณ์ที่ทำให้ค่อนข้างง่ายหรือยากสำหรับผู้เข้าสอบในการทำแบบทดสอบได้เต็มความสามารถ

ด้านผลกระทบ (Impact) มีคำถามสำหรับการประเมินที่แบ่งออกเป็น 3 ชนิด คือ ผลกระทบต่อบุคคลที่เป็นผู้เข้าสอบ มีจำนวน 7 ข้อ คือ 1) ประสบการณ์ในการทำแบบทดสอบหรือข้อมูลย้อนกลับที่ได้รับอาจจะส่งผลต่อลักษณะของผู้เข้าสอบที่สัมพันธ์กับการใช้ภาษามากน้อยเพียงใด (เช่น ความรู้ในหัวเรื่อง การรับรู้ในสถานการณ์การใช้ภาษาเป้าหมาย ประเด็นของความรู้ด้านภาษา และการใช้กลยุทธ์) 2) มีการเตรียมการอะไรบางอย่างในการทำให้ผู้เข้าสอบเกี่ยวข้องโดยตรงหรือสำหรับการรวบรวม หรือใช้ข้อมูลย้อนกลับจากผู้เข้าสอบในขั้นการออกแบบและการพัฒนาแบบทดสอบ 3) ข้อมูลย้อนกลับที่จัดให้แก่ผู้เข้าสอบมีความจำเป็น สมบูรณ์ และมีความหมายอย่างไร 4) กระบวนการตัดสินใจและเกณฑ์ที่ประยุกต์ใช้ได้เหมือนกันสำหรับผู้เข้าสอบทุกกลุ่มหรือไม่ 5) คะแนนสอบสัมพันธ์และเหมาะสมกับการตัดสินใจที่ได้ทำไว้แล้วอย่างไร 6) ผู้เข้าสอบได้รับการแจ้งเกี่ยวกับกระบวนการและเกณฑ์ที่จะใช้ในการตัดสินใจโดยตลอดทุกเรื่องราวหรือไม่ และ 7) กระบวนการและเกณฑ์เหล่านี้ถูกปฏิบัติตามในการทำการตัดสินใจหรือไม่ ผลกระทบต่อบุคคลที่เป็นครู มีจำนวน 3 ข้อ คือ 1) ประเด็นความสามารถทางภาษาที่จะวัดสอดคล้องกับส่วนที่กล่าวไว้ในสื่อประกอบการสอนอย่างไร 2) ลักษณะของแบบทดสอบและภาระงานในแบบทดสอบสอดคล้องกับลักษณะของกิจกรรมการเรียน

การสอนอย่างไร และ 3) จุดประสงค์ของการสอบสอดคล้องกับคุณค่าและเป้าหมายของครูและของโปรแกรมการเรียนการสอน และคำถามสำหรับประเมินผลกระทบสังคมและระบบการศึกษา มีจำนวน 5 ข้อ คือ 1) การตีความจากผลการสอบ สอดคล้องกับค่านิยมและเป้าหมายของสังคมและระบบการศึกษา 2) ค่านิยมและเป้าหมายของนักพัฒนาแบบทดสอบตรงกันหรือขัดแย้งกับค่านิยมและเป้าหมายของสังคมและระบบการศึกษา 3) ผลที่อาจตามมาทั้งในเชิงบวกและลบ สำหรับสังคมและระบบการศึกษา จากการใช้แบบทดสอบในทิศทางนี้คืออะไร 4) ผลด้านบวกที่พึงปรารถนาที่สุดหรือสิ่งที่ดีที่สุดที่อาจจะเกิดขึ้น อันเป็นผลพวงของการใช้แบบทดสอบแบบนี้คืออะไร และสิ่งนี้อาจจะเกิดขึ้นอย่างไร และ 5) ผลด้านลบที่พึงปรารถนาลดน้อยที่สุด หรือสิ่งที่แย่ที่สุดที่อาจจะเกิดขึ้น อันเป็นผลพวงของการใช้แบบทดสอบแบบนี้คืออะไร หรือสิ่งนี้อาจจะเกิดขึ้นได้อย่างไร

ด้านความสะดวกใช้ (Practicality) มีคำถามสำหรับประเมินจำนวน 2 ข้อ คือ 1) มีความต้องการทรัพยากรอะไร ปริมาณเพียงใดสำหรับ (ก) ขั้นตอนการออกแบบ (ข) ขั้นตอนการพัฒนา (ค) ขั้นตอนนำไปใช้ และ 2) ทรัพยากรอะไรที่มีไว้พร้อมแล้วสำหรับการดำเนินการในข้อ (ก) (ข) และ (ค)

Genesee & Upshur (1996: 241-244) นำเสนอเกณฑ์ในการประเมินแบบทดสอบมาตรฐาน 6 ด้าน คือ 1) สาระแบบทดสอบและจุดประสงค์ (Test Content and Purpose) 2) ความเหมาะสม (Appropriateness) 3) ความสะดวกใช้ (Practicality) 4) คุณภาพต่อผู้ใช้ (User Qualities) 5) ความเชื่อมั่น (Reliability) และ 6) ความตรง (Validity) โดยมีรายละเอียดในแต่ละด้าน ดังนี้

สาระแบบทดสอบและจุดประสงค์ (Test Content and Purpose) เมื่อประเมินแบบทดสอบมาตรฐานให้พิจารณา 1) ทักษะภาษาหรือทักษะย่อยที่นักพัฒนาแบบทดสอบสนใจจะทดสอบ 2) ชนิดของการตัดสินใจที่นักพัฒนาแบบทดสอบต้องการ โดยธรรมชาตินักพัฒนาแบบทดสอบจะต้องการพิจารณาอย่างละเอียด เฉพาะแบบทดสอบมาตรฐานเหล่านั้นที่ทดสอบทักษะภาษาหรือความรู้ชนิดเดียวกันที่ต้องการประเมิน เป็นเรื่องสำคัญที่นักพัฒนาแบบทดสอบจะต้องตรวจสอบรายละเอียดของทักษะที่อยู่ในแบบทดสอบ เพื่อให้แน่ใจว่าสอดคล้องกับทักษะที่ตนสนใจสามารถทำได้โดยการตรวจสอบคำอธิบายสาระแบบทดสอบที่อยู่ในคู่มือแบบทดสอบ และที่สำคัญมากกว่านั้น คือ การตรวจสอบสาระของแบบทดสอบฉบับจริง เป็นสิ่งสำคัญมากที่ต้องตรวจสอบสาระแบบทดสอบอย่างละเอียดเอง เพราะชื่อแบบทดสอบหรือคู่มืออาจให้ข้อมูลไม่สมบูรณ์ ไม่เพียงพอและในบางกรณีคู่มืออาจให้คำอธิบายที่ไม่ถูกต้องเกี่ยวกับสิ่งที่ได้บรรจุไว้

ความเหมาะสม (Appropriateness) ความเหมาะสม ครอบคลุมถึง การเปรียบเทียบระหว่างนักเรียนที่จะถูกทดสอบและนักเรียนที่เป็นกลุ่มเป้าหมายของแบบทดสอบ ความเหมาะสมควรระบุประเด็นต่อไปนี้ คือ ก) อายุหรือระดับชั้นปี โดยพิจารณาว่า แบบทดสอบเหมาะสมกับอายุหรือระดับชั้นปีของนักเรียนที่คุณต้องการจะทดสอบหรือไม่ ข) พื้นฐานการศึกษา โดยพิจารณาว่าแบบทดสอบเหมาะสมสำหรับใช้กับนักเรียนที่มีพื้นฐานการศึกษาในแบบที่นักพัฒนาแบบทดสอบต้องการหรือไม่ ตัวอย่างเช่น ถ้านักพัฒนาแบบทดสอบวางแผนที่จะทดสอบนักเรียนที่ได้รับการศึกษาแบบเป็นทางการเพียงเล็กน้อย เขาควรเลือกแบบทดสอบที่ไม่เกี่ยวข้องกับการงานด้านภาษา เพราะนักเรียนของเขาอาจจะไม่คุ้นเคยกับสิ่งนั้น ค) ระดับความสามารถ โดยพิจารณาว่า ระดับความสามารถโดยทั่วไปที่แบบทดสอบต้องการจะทดสอบเหมาะสมกับนักเรียนที่จะถูกทดสอบหรือไม่

ความเหมาะสมยังรวมถึงคำถามที่ว่า แบบทดสอบเป็นที่ยอมรับของนักเรียน ครู ผู้บริหารโรงเรียน และแม้กระทั่งพ่อแม่ หรือสมาชิกอื่นของชุมชนที่ได้รับผลจากแบบทดสอบ

ความสะดวกในการใช้ (Practicality) แบบทดสอบใช้ได้สะดวกหรือไม่ จากการพิจารณาด้านงบประมาณ เวลา พื้นที่ การบริหาร การตรวจ และวัสดุหรืออุปกรณ์พิเศษ คำถามจำนวนหนึ่งที่จะใช้ถามเกี่ยวกับความสะดวกในการใช้ คือ ก) มีงบประมาณเพียงพอในการผลิตแบบทดสอบหรือไม่ ข) มีเวลาเพียงพอที่จะใช้ในการจัดสอบตามข้อแนะนำของนักพัฒนาแบบทดสอบหรือไม่ ค) มีสถานที่ตามที่ต้องการที่จะใช้เป็นสถานที่สอบหรือไม่ เช่น ห้องเรียนที่จัดที่นั่งสอบได้ 25 ที่นั่ง โดยที่โต๊ะเก้าอี้วางห่าง ๆ กัน ห้องที่เก็บเสียง เป็นต้น ง) การบริหารการสอบต้องมีการอบรมพิเศษแก่ผู้ดำเนินการสอบหรือไม่ ตัวอย่างเช่น กระบวนการสอบสัมภาษณ์บางอย่างไม่สามารถดำเนินการโดยบุคคลที่ไม่ได้รับการฝึกอบรม ผู้ดำเนินการสัมภาษณ์ต้องได้รับการฝึกพิเศษแบบเข้มข้น เป็นต้น เป็นไปได้หรือไม่ที่จะจัดให้มีการอบรมเช่นนั้น จ) แบบทดสอบต้องการกระบวนการให้คะแนนแบบพิเศษหรือไม่ ตัวอย่างเช่น แบบทดสอบมาตรฐานบางชนิดไม่สามารถให้ตรวจได้โดยครูหรือผู้ดูแลการสอบแบบทดสอบต้องถูกส่งไปตรวจโดยผู้เชี่ยวชาญ บางครั้งอาจจะมีค่าใช้จ่ายเพิ่มเติม สิ่งนี้เป็นไปได้หรือไม่ ถ้าต้องการผลการสอบโดยทันทีหรือไม่มีกองทุนเพิ่มเติมสำหรับจ่ายค่าบริการตรวจนั้น แบบทดสอบนี้ไม่สามารถนำมาใช้ได้ใช่หรือไม่ และ ฉ) ต้องมีการใช้วัสดุหรืออุปกรณ์พิเศษในการจัดสอบหรือไม่ เช่น เครื่องบันทึกเทป เครื่องบันทึกวิดีโอ เครื่องฉายสไลด์ เป็นต้น มีอุปกรณ์เหล่านี้อยู่แล้วหรือไม่

คุณภาพต่อผู้ใช้ (User Qualities) เป็นเรื่องสำคัญที่ต้องตรวจสอบมาตรฐานจากมุมมองของทั้งครูและนักเรียนในฐานะผู้ใช้แบบทดสอบ โดยข้อคำถามที่เหมาะสมสำหรับใช้ถามนักเรียนในฐานะผู้ใช้แบบทดสอบ ได้แก่ ก) แบบทดสอบได้อธิบายจุดประสงค์ ความตั้งใจ หรือการใช้ที่ถูกให้คำแนะนำแก่นักเรียนด้วยวิธีการที่ซื่อสัตย์และตรงไปตรงมาหรือไม่ ข) คำสั่งของแบบทดสอบ ชัดเจน และเฉพาะเจาะจง ทัวถึงหรือไม่ ค) มีแบบทดสอบที่เป็นตัวอย่างหรือไม่ มีให้อย่างเพียงพอหรือไม่ ง) สมุดสอบ และกระดาษคำตอบมีการจัดพิมพ์ที่ช่วยส่งเสริมความเข้าใจและการตอบแบบทดสอบหรือไม่ พิจารณาจากประเด็นต่อไปนี้ คือ 1) ตัวอักษรที่พิมพ์ชัดเจนและอ่านออกหรือไม่ 2) ในกรณีของข้อคำถามแบบให้เลือกคำตอบ สามารถแยกแยะคำถามออกจากคำตอบได้ง่ายหรือไม่ 3) การใช้ตัวเลขหรือตัวอักษรแสดงข้อต่าง ๆ ชัดเจนหรือไม่ และ 4) รูปภาพหรือส่วนที่มีลักษณะกราฟฟิกอื่น ๆ ชัดเจน ดีความได้ง่ายหรือไม่ และ จ) ข้อคำถามไม่มีความลำเอียง เสียดีสี กระทบกระเทียบต่อบุคคลเหมาะสมกับวัฒนธรรมและน่าสนใจต่อนักเรียนหรือไม่

ข้อคำถามที่ใช้ถามครูในฐานะผู้ใช้แบบทดสอบ ได้แก่ ก) กระบวนการบริหารแบบทดสอบถูกอธิบายไว้อย่างชัดเจนและเข้าใจง่ายในคู่มือแบบทดสอบหรือไม่ ใครคือผู้บริหารแบบทดสอบ ต้องจัดสรรเวลาอย่างไร นักเรียนต้องตอบแบบทดสอบทุกข้อหรือไม่ และถ้าไม่ต้องตอบทุกข้อ ข้อใดบ้างที่นักเรียนตอบ ข) คู่มือแบบทดสอบแสดงคำสั่งที่ต้องแจ้งให้ผู้เข้าสอบทราบไว้อย่างชัดเจนและกระชับหรือไม่ ค) กระบวนการให้คะแนนอธิบายไว้อย่างชัดเจนและเข้าใจได้ง่ายหรือไม่ ง) มีการให้บรรทัดฐาน (Norms) ไว้หรือไม่ บรรทัดฐาน (Norms) ถูกให้ไว้สำหรับผู้เข้าสอบกลุ่มย่อยต่าง ๆ หรือไม่ คะแนนอิงกลุ่มหรือคะแนนที่ถูกแปลงอธิบายไว้อย่างชัดเจนและเข้าใจง่ายหรือไม่ จ) มีคำแนะนำสำหรับการแปลงคะแนนดิบเป็นคะแนนตามเกณฑ์หรือไม่ ฉ) มีคำแนะนำเกี่ยวกับการตีความคะแนน

สอบให้ไว้หรือไม่ ข) คู่มือสอบให้คำแนะนำเกี่ยวกับการตัดสินใจชนิดต่าง ๆ ที่สามารถทำได้ และสามารถทำได้อย่างไร โดยใช้คะแนนสอบนี้หรือไม่ และ ข) มีคู่มือด้านเทคนิคหรือไม่ ถ้ามี คู่มือได้ให้ คำอธิบายที่สมบูรณ์และชัดเจนเกี่ยวกับตัวอย่างมาตรฐาน กระบวนการคัดเลือกแบบทดสอบ ความ เชื่อมั่น และความตรงหรือไม่

Brown (2004: 30-38) ประยุกต์ใช้หลักการสำหรับการประเมินคุณภาพแบบทดสอบใน 6 ด้าน คือ 1) ความสะดวกใช้ 2) ความเชื่อมั่น 3) ความตรงเชิงเนื้อหา 4) ความตรงเชิงประจักษ์ 5) ความสมจริง และ 6) ผลกระทบที่เป็นประโยชน์ต่อผู้เรียน โดยมีรายละเอียดในแต่ละด้าน ดังนี้

1) กระบวนการของแบบทดสอบสามารถนำไปใช้ได้อย่างสะดวกหรือไม่ การนำไปใช้ถูกตัดสินใจโดยการจำกัดเรื่องของเวลาของครู (ของนักเรียน) ค่าใช้จ่าย และรายละเอียดในการบริหาร และอาจจะเกี่ยวข้องกับสิ่งที่เกิดขึ้นก่อนและหลังการสอบ เพื่อตัดสินว่าแบบทดสอบเหมาะที่จะใช้กับ ความต้องการหรือไม่ โดยสามารถตรวจสอบได้จากรายการต่อไปนี้ คือ ก) ข้อมูลรายละเอียดเกี่ยวข้อง กับการบริหารแบบทดสอบถูกกำหนดไว้อย่างชัดเจนก่อนการสอบหรือไม่ ข) นักเรียนสามารถทำ แบบทดสอบให้เสร็จสมบูรณ์ได้อย่างสมเหตุสมผล (เหมาะสม) ภายในระยะเวลาที่กำหนดไว้หรือไม่ ค) แบบทดสอบสามารถจัดสอบได้อย่างราบรื่น โดยปราศจากความบกพร่องเชิงกระบวนการ (Procedural 'Glitches') หรือไม่ ง) วัสดุและอุปกรณ์ทั้งหมดพร้อมหรือไม่ จ) งบประมาณในการ พัฒนาแบบทดสอบอยู่ภายในวงเงินที่กำหนดไว้หรือไม่ ฉ) ระบบการตรวจให้คะแนนหรือประเมินผล เสร็จทันในกรอบเวลาที่มีอยู่หรือไม่ และ ช) วิธีการรายงานผลการสอบได้รับการวางแผนล่วงหน้า แล้วหรือไม่

2) แบบทดสอบเชื่อถือได้หรือไม่ ความเชื่อมั่นใช้ได้ทั้งกับแบบทดสอบและครู และอย่างน้อย ที่สุดต้องมีการป้องกันแหล่งที่มาของความไม่น่าเชื่อถือของแบบทดสอบจาก 4 แหล่ง คือ ความ แปรปรวนที่เกิดจากนักเรียน การตรวจให้คะแนน การจัดสอบ และตัวแบบทดสอบ ความเชื่อมั่นที่มา จากแบบทดสอบและการจัดสอบสามารถทำได้โดยการทำให้มั่นใจว่านักเรียนทุกคนได้รับตัวกระตุ้นที่ มีคุณภาพเดียวกัน ไม่ว่าจะเป็นภาษาเขียนหรือทางเสียง บางส่วนของความเชื่อมั่นในแบบทดสอบ ขึ้นอยู่กับลักษณะทางกายภาพ จึงควรทำให้มั่นใจว่า ก) นักเรียนทุกคนได้รับแบบทดสอบที่ถ่ายสำเนา อย่างสะอาด ข) ระบบเสียงในการฟังได้ยินทั่วถึงเท่ากันเหมือนกันสำหรับผู้เข้าสอบทุกคน ค) ตัวกระตุ้นที่เป็นวิดีโอเห็นภาพชัดเจนเท่ากันสำหรับผู้เข้าสอบทุกคน ง) แสงสว่าง อุณหภูมิ เสียง รบกวน และสภาพห้องเรียนอื่นๆ เท่าเทียมกันสำหรับผู้เข้าสอบทุกคน และ จ) กระบวนการตรวจ แบบทดสอบปรนัยมีปัญหาถกเถียงในเรื่องความถูกต้องของคำตอบน้อยมาก

ความเชื่อมั่นที่มาจาก การตรวจให้คะแนน เป็นอีกประเด็นหนึ่งที่พบบ่อยในการประเมินและ อาจจะดูยากกว่า เนื่องจากแบบทดสอบที่ใช้ในห้องเรียนมักไม่เกี่ยวข้องกับคนตรวจ 2 คน ความ เชื่อมั่นที่เกิดจากผู้ให้คะแนนจึงไม่ได้เป็นปัญหา สิ่งที่จะเป็นปัญหาคือ การตรวจแบบทดสอบโดย ผู้ตรวจคนเดียว และใช้เวลาในการตรวจนาน ซึ่งอาจจะทำให้ผู้ตรวจเหนื่อยล้า หรือไม่มีสมาธิในการ ตรวจที่เป็นมาตรฐานเดียวกันสำหรับผู้เข้าสอบทุกคน ความเชื่อมั่นในลักษณะนี้เกิดจากผู้ตรวจคน เดียว ในการตรวจแบบทดสอบแบบปลายเปิด (Open-Ended Responses) อาจปฏิบัติตาม คำแนะนำต่อไปนี้ คือ ก) ใช้ชุดเกณฑ์การให้คะแนนที่คงที่ในการตรวจ ข) ให้ความสนใจที่เป็นแบบ เดียวกันกับคำตอบทั้งหมดตลอดระยะเวลาในการประเมิน ค) อ่านแบบทดสอบอย่างน้อย 2 รอบ เพื่อ

ตรวจสอบความคงที่ ง) ถ้าผู้ตรวจแบบทดสอบรู้สึกว่าจะต้องเปลี่ยนการให้คะแนนระหว่างการตรวจ และอาจต้องปรับเปลี่ยนในสิ่งที่เป็นการตอบที่ถูกบ้าง ต้องกลับไปตรวจแบบทดสอบส่วนที่ตรวจมาแล้วใหม่ทั้งหมดและใช้เกณฑ์เดียวกัน และ จ) หลีกเลี่ยงการเหนื่อยล้าโดยอาการอ่านในหลาย ๆ สถานที่ ถ้าต้องใช้เวลาในการอ่านแบบทดสอบหลายชั่วโมง

3) กระบวนการสอบแสดงถึงความตรงเชิงเนื้อหาหรือไม่ แหล่งที่มาของความตรงที่สำคัญของแบบทดสอบที่ใช้ในห้องเรียนคือ ความตรงเชิงเนื้อหา ความมากหรือน้อยที่การประเมินให้นักเรียนได้ ทำภาระงานที่เคยได้เจอในชั้นเรียน และสิ่งนั้นเป็นตัวแทนของวัตถุประสงค์ต่าง ๆ ของเนื้อหาในบทต่าง ๆ ที่การประเมินยึดเป็นฐาน มีขั้นตอน 2 ขั้นตอนในการประเมินความตรงเชิงเนื้อหาของแบบทดสอบที่ใช้ในห้องเรียน คือ 1) วัตถุประสงค์การเรียนรู้ได้ถูกระบุ และกำหนดกรอบอย่างเหมาะสมหรือไม่ภายใต้แบบทดสอบที่ใช้ในห้องเรียน คือ วัตถุประสงค์ของบทเรียน หน่วย หรือเนื้อเรื่องในวิชาที่สอน ดังนั้น การประเมินขั้นแรกของแบบทดสอบที่ใช้ในห้องเรียนที่มีประสิทธิภาพ คือ การระบุวัตถุประสงค์ และ 2) วัตถุประสงค์ถูกนำมาใช้ในรูปของโครงสร้างแบบทดสอบหรือไม่

4) ความตรงเชิงประจักษ์ที่พบได้ในกระบวนการมีความตรงและก่อให้เกิดความลำเอียงหรือไม่ คำถามนี้รวมแนวคิดของความตรงเชิงประจักษ์เข้ากับความสำคัญของการวัดโครงสร้าง กระบวนการประเมินเพื่อวัดความสามารถทางเลือกของนักเรียน โดยทั่วไปนักเรียนจะตัดสินใจว่าแบบทดสอบมีความตรง ถ้าคำสั่งชัดเจน โครงสร้างแบบทดสอบถูกจัดวางอย่างสมเหตุผล ระดับความยากถูกกำหนดอย่างเหมาะสม แบบทดสอบไม่ได้สร้างความแปลกใจ และเวลาที่ใช้ในการทำแบบทดสอบเหมาะสม

5) ภาระงานสอบสมจริงเท่าที่จะเป็นไปได้หรือไม่ ประเมินระดับความสมจริงของแบบทดสอบโดยการถามคำถามต่อไปนี้ คือ ก) ภาษาที่ใช้ในแบบทดสอบเป็นธรรมชาติเท่าที่จะเป็นไปได้หรือไม่ ข) ข้อคำถามต่าง ๆ มีการใช้บริบทเท่าที่จะเป็นไปได้มากกว่าที่จะให้อยู่เดี่ยว ๆ ใช่หรือไม่ ค) หัวข้อและสถานการณ์ น่าสนใจ สนุกสนาน และ/หรือ ช้าช้าใช่หรือไม่ และ ง) ภาระงานเป็นตัวแทนหรือใกล้เคียงกับภาระงานในโลกแห่งความเป็นจริงใช่หรือไม่

6) แบบทดสอบให้ผลกระทบที่เป็นประโยชน์ต่อผู้เรียนหรือไม่ หลักฐานอื่นของผลกระทบอาจจะเห็นได้น้อยจากตัวแบบทดสอบ สิ่งที่เกิดขึ้นก่อนและหลังการสอบมีความสำคัญมาก เวลาในการเตรียมตัวก่อนสอบสามารถให้ผลกระทบ เนื่องจากผู้เรียนได้ทบทวนและเน้นย้ำประเด็นต่างๆ ที่อยู่ในจุดประสงค์อย่างกว้างขวาง โดยการใช้เวลาของนักเรียนหลังจากการสอบทบทวนเนื้อหา นักเรียนค้นพบจุดอ่อนและจุดแข็งของตนเอง

อัจฉรา วงศ์โสธร (2544: 263-269) นำเสนอเกณฑ์มาตรฐานสำหรับการประเมินแบบทดสอบอิงวัตถุประสงค์เชิงอุดมและแบบทดสอบความสามารถสูงสุด โดยมีรายละเอียดดังต่อไปนี้

รายการตรวจสอบเพื่อยืนยันมาตรฐานของการพัฒนาแบบทดสอบอิงวัตถุประสงค์แบบอุดม ซึ่งเป็นการประเมินกระบวนการดำเนินงาน ประกอบด้วยเกณฑ์ต่อไปนี้ คือ 1) มีการหาค่าดัชนีความง่ายและดัชนีความไว 2) มีการหาค่าความเที่ยงในการประมาณความสามารถและองค์ประกอบที่วัดมีความคลาดเคลื่อนต่ำ 3) มีการหาค่าความตรงตามเกณฑ์มาตรฐาน 4) มีการหาค่าความตรงเชิงเนื้อหา กำหนดโดยการที่ข้อทดสอบสร้างขึ้นได้ตรงตามตารางกำหนดลักษณะแบบทดสอบ 5) มีความตรงเชิงภาวะสันนิษฐาน วัดคุณลักษณะทางจิตวิทยาที่ถูกต้องตามทฤษฎี และสมมติฐานที่ตั้งไว้ สามารถ

ยืนยันได้โดยสถิติ 6) มีการจัดสอบเป็นมาตรฐาน โดยดำเนินการตามคำสั่งที่ได้ให้ไว้ เพราะคำสั่งชัดเจนทำให้ผู้ดำเนินการสอบและผู้เข้าสอบเข้าใจอย่างละเอียด 7) การรายงานผลการสอบให้คะแนนรวมและคะแนนแต่ละส่วน พร้อมทั้งคำอธิบายผลตามหลักการที่ใช้ในการสร้างข้อทดสอบ มีการนำเสนอรายงานที่ชัดเจน ช่วยให้สถาบันการศึกษา ครู นักเรียน ใช้ผลเพื่อการประเมินได้อย่างมีประสิทธิภาพ 8) เนื้อหาของข้อทดสอบจะต้องไม่ลำเอียง และไม่ขัดต่อศีลธรรมอันดี 9) ต้องมีความสะดวกในการใช้ กระบวนการสอบต้องไม่ยุ่งยาก ซับซ้อน และ 10) แบบทดสอบมีประโยชน์ทางด้านการพัฒนาผู้เข้าสอบและโปรแกรมการเรียนการสอน

รายการตรวจสอบเพื่อประเมินคุณสมบัติของแบบทดสอบความสามารถสูงสุด ประกอบด้วย การประเมินด้านคำสั่ง เนื้อหา รูปแบบและการจัดลำดับข้อความในแบบทดสอบ และการใช้ผลการสอบ โดยมีรายละเอียดของเกณฑ์ในแต่ละด้านดังนี้

ด้านคำสั่งประกอบด้วยเกณฑ์การประเมิน 13 ข้อ คือ 1) คำสั่งรวมและคำสั่งสำหรับแต่ละตอนของแบบทดสอบให้ไว้อย่างชัดเจนหรือไม่ 2) ควรให้ตัวอย่างการตอบไว้ในคำสั่งหรือไม่ ถ้าใช้วิธีการวัดซับซ้อน 3) มีการระบุน้ำหนักคะแนนของแบบทดสอบแต่ละตอนหรือไม่ 4) มีการกำหนดเวลาในการทำแบบทดสอบแต่ละตอนหรือไม่ 5) แบบทดสอบทุกข้อมีการปฏิบัติที่สอดคล้องกับคำสั่งในการทำแบบทดสอบหรือไม่ 6) มีการใช้ภาษาที่ยากเกินไปจนผู้เข้าสอบไม่เข้าใจคำสั่งหรือไม่ 7) ภาษาที่ใช้ในคำสั่งมีระดับความยากง่ายเดียวกันหรือใกล้เคียงกันกับภาษาของแบบทดสอบหรือไม่ 8) ในการวัดแบบผู้เข้าสอบให้คำตอบเอง เช่น การเขียนตอบได้มีการกำหนดความยาวของคำตอบให้ไว้หรือไม่ (โดยกำหนดเป็นจำนวนคำหรือจำนวนประโยค) 9) การกำหนดเวลาในการทำสอบสมจริงสอดคล้องกับน้ำหนักและความยากง่ายของแบบทดสอบหรือไม่ 10) ผู้เข้าสอบมีความเข้าใจเกี่ยวกับน้ำหนักหรือความสำคัญของแบบทดสอบทั้งฉบับและของแบบทดสอบในแต่ละตอน และเวลาที่ให้ในการทำสอบหรือไม่ 11) ผู้เข้าสอบเข้าใจล่วงหน้าถึงขอบเขตการวัด และน้ำหนักคะแนนในการสอบหรือไม่ 12) ผู้เข้าสอบได้รับคำแนะนำการทำสอบอย่างชัดเจน ครอบคลุม และเพียงพอหรือไม่ และ 13) วิธีการดำเนินการสอบมีระบุไว้ชัดเจน ช่วยให้ผู้คุมสอบใด ๆ สามารถดำเนินการโดยวิธีเดียวกันหรือไม่

ด้านเนื้อหาประกอบด้วยเกณฑ์การประเมิน 11 ข้อ คือ 1) เนื้อหาสอดคล้องกับมิติสัมพันธ์ของการสื่อสารทางภาษาตามกรอบกำหนดของวัตถุประสงค์แบบอุดมอย่างเหมาะสมและเพียงพอหรือไม่ 2) ในการวัดผลสัมฤทธิ์ในการเรียน แบบทดสอบที่ใช้สามารถวัดและประเมินวัตถุประสงค์ของการเรียนการสอนได้อย่างเหมาะสมและเพียงพอหรือไม่ 3) แบบทดสอบวัดทักษะทางภาษาทั้งด้านการรับสารและส่งสาร ตามวัตถุประสงค์ครบถ้วนหรือไม่ 4) มีการทดสอบทักษะทางภาษาทั้งแบบทักษะเดี่ยวและทักษะสัมพันธ์หรือไม่ 5) เนื้อหาของแบบทดสอบใช้วัดทักษะทางการสื่อสารได้ครอบคลุมหรือไม่ (วัดทั้งระเบียบวิธี/กฎเกณฑ์ทางภาษา และการใช้ภาษาอย่างสมจริง 6) การทดสอบสอดคล้องกับวัตถุประสงค์ โดยใช้วิธีการวัดที่ประสงค์หรือไม่ 7) ระดับภาษาที่ใช้เหมาะสม มีลีลาที่สอดคล้องกับสภาพการใช้ภาษาอย่างสมจริง เช่น การใช้ภาษาเป็นกันเองในการสนทนา และการใช้ภาษาเป็นทางการในการอ่าน และการบรรยาย ฯลฯ หรือไม่ 8) แบบทดสอบใช้ภาษาที่มีลีลาสุภาพ ไม่ขัดต่อจริยธรรมและศีลธรรมอันดีหรือไม่ 9) ชื่อของแบบทดสอบสอดคล้องกับเนื้อหาที่วัด ไม่ทำให้ผู้เข้าสอบไขว้เขวหรือเข้าใจผิดได้ 10) แบบทดสอบไม่ลำเอียงต่อเชื้อชาติ พื้นภูมิลัทธิ เพศ และสภาพทาง

ร่างกายของผู้เข้าสอบหรือไม่ และ 11) มีการใช้ภาษาโดยไม่คำนึงถึงความรู้สึกนึกคิด จิตใจของผู้เข้าสอบบางกลุ่มหรือไม่

ด้านรูปแบบและการจัดลำดับข้อความในแบบทดสอบ ประกอบด้วยเกณฑ์การประเมิน 14 ข้อ คือ 1) แบบทดสอบยาวหรือสั้นไปหรือไม่ 2) มีการวัดวัตถุประสงค์บางข้อมากเกินไปหรือน้อยเกินไปหรือไม่ 3) แบบทดสอบมีการชี้แนะคำตอบซึ่งกันและกันหรือไม่ 4) มีการใช้ภาษาที่ชี้แนะคำตอบที่ถูกหรือไม่ 5) สำหรับแบบทดสอบแบบเลือกตอบ ผู้เข้าสอบทำสอบโดยใช้ความรู้ความสามารถในการเลือกคำตอบที่ถูก และไม่ใช้การเดาหรือการคัดคำตอบที่คิดว่าไม่ใช่คำตอบออกหรือไม่ 6) การทดสอบแบบเลือกตอบแบบจริง/เท็จ (True/False) ถูก/ผิด และแบบมีตัวเลือกมากกว่า 2 ขึ้นไป มีการกระจายคำตอบที่ถูกเพื่อหลีกเลี่ยงการเดาหรือไม่ โดยแบบจริง/เท็จ หรือถูก/ผิด ควรมีคำตอบ “เท็จ” และ “ผิด” มากกว่า และแบบมีตัวเลือกมากกว่า 2 ขึ้นไป ไม่ควรกำหนดให้ข้อ “b” และ “c” เป็นคำตอบที่ถูกมากเกินไป 7) มีการกำหนดเวลาที่เหมาะสมโดยให้ผู้เข้าสอบที่อ่อนสามารถทำได้ด้วยบ้างหรือไม่ 8) มีการใช้วิธีการวัดเหมาะสมกับคำตอบที่ต้องการหรือไม่ เช่น แบบจับคู่ มีการใช้ภาษาในตัวแบบทดสอบและคำตอบในระดับความยากง่ายเดียวกัน 9) มีการใช้วิธีการวัดที่หลากหลาย เพื่อวัดครอบคลุมทุกทักษะ หรือวัดทักษะย่อยที่ประสงค์หรือไม่ 10) มีการใช้ตัวกระตุ้น/ตัวแบบทดสอบ เพื่อให้ได้คำตอบที่ต้องการอย่างสมจริง เช่น มีการใช้เทปบันทึกเสียง วิดีโอทัศน์ ภาพประกอบ ฯลฯ หรือไม่ 11) การวางรูปแบบแบบทดสอบชัดเจน สะดวกต่อการทำสอบ ไม่ทำให้ผู้เข้าสอบพลาดพลั้ง เช่น การเรียงข้อ และตัวเลือก a b c d ชัดเจน และถ้าเป็นแบบโคลงหรือแบบเติมคำ มีการวางตัวเลือกในหน้าเดียวกันกับตัวแบบทดสอบ เพื่อให้ผู้เข้าสอบไม่ต้องสับสน หรือเสียเวลาพลิกหน้ากระดาษหรือไม่ 12) การเว้นช่องไฟและช่องว่างระหว่างบรรทัดเหมาะสม ชัดเจนหรือไม่

การใช้ผลการสอบ ประกอบด้วยเกณฑ์การประเมิน 2 ข้อ คือ 1) มีแผนการใช้ผลของการทดสอบเพื่อการศึกษาหรือไม่ และ 2) ผลของการทดสอบจะมีประโยชน์ต่อการพัฒนาและปรับปรุงหรือไม่

สรุปได้ว่าในการพัฒนาแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test) จะมีการประเมินคุณภาพแบบทดสอบโดยใช้แบบสอบถาม 2 ระยะ คือ 1) การประเมินคุณภาพแบบทดสอบก่อนนำไปใช้ และ 2) การประเมินคุณภาพแบบทดสอบหลังใช้ โดยในการประเมินคุณภาพแบบทดสอบก่อนนำไปใช้ จะมีการใช้แบบประเมิน 2 ชุด คือ 1) แบบประเมินค่า IOC โดยผู้เชี่ยวชาญ เพื่อตรวจสอบความสอดคล้องของจุดประสงค์การสอบและข้อคำถามทั้งชุด ทุกตอน และรายข้อ และ 2) แบบประเมินคุณประโยชน์ของแบบทดสอบ (Test Usefulness) ซึ่งข้อคำถามในแบบประเมินจะเป็นข้อคำถามที่ใช้ในการประเมินคุณประโยชน์ของแบบทดสอบตามแนวคิดของ Bachman and Palmer (1996) โดยในการวิจัยครั้งนี้ผู้วิจัยจะสร้างแบบประเมิน 7 ด้าน คือ ความตรงเชิงโครงสร้าง (Construct Validity) ความเชื่อมั่น (Reliability) ความสมจริง (Authenticity) การมีปฏิสัมพันธ์ (Interactiveness) ผลกระทบ (Impact) ความสะดวกใช้ (Practicality) และความเป็นธรรม (Fairness) ส่วนการประเมินคุณภาพแบบทดสอบหลังใช้ จะมีการใช้แบบประเมิน 1 ชุด คือ แบบประเมินความพึงพอใจของผู้เกี่ยวข้องต่อแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency

Test) ตามแนวคิดของ อัจฉรา วงศ์โสธร (2544) Genesee & Upshur (1996) และ Brown (2004) โดยจะสอบถามความคิดเห็นของผู้เชี่ยวชาญ อาจารย์ผู้สอนภาษาอังกฤษระดับมหาวิทยาลัยและผู้เข้าสอบ ใน 6 ด้าน คือ คำสั่ง เนื้อหา รูปแบบและการจัดลำดับความ วิธีการตรวจให้คะแนน การใช้ผล การสอบ และสภาพแวดล้อมในการสอบ

2.3 งานวิจัยที่เกี่ยวข้องกับการพัฒนาแบบทดสอบและการประเมินคุณภาพแบบทดสอบวัดความสามารถภาษาอังกฤษ

งานวิจัยที่เกี่ยวข้องกับการพัฒนาแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย แบ่งออกได้เป็น 2 ชนิด คือ งานวิจัยที่เกี่ยวข้องกับการพัฒนาแบบทดสอบวัดความสามารถภาษาอังกฤษ และงานวิจัยที่เกี่ยวข้องกับการประเมินคุณภาพแบบทดสอบภาษาอังกฤษ โดยมีรายละเอียดต่อไปนี้

2.3.1 งานวิจัยที่เกี่ยวข้องกับการพัฒนาแบบทดสอบวัดความสามารถภาษาอังกฤษ

งานวิจัยที่เกี่ยวข้องกับการพัฒนาแบบทดสอบวัดความสามารถภาษาอังกฤษ มีจำนวน 6 เรื่อง ผู้ดำเนินการวิจัย คือ Wullur (2011), Samaie (2014), Ivanova (2011), Urmston (2013), อัจฉรา วงศ์โสธร (2532) และ อัจฉรา วงศ์โสธร และ ศิริพร พงษ์สุรพิพัฒน์ (2544) โดยมีรายละเอียดในแต่ละงานวิจัยดังนี้

Wullur (2011) ศึกษาเพื่อพัฒนาและหาคุณภาพแบบทดสอบวัดความสามารถภาษาอังกฤษ ที่ถูกออกแบบสำหรับใช้เป็นเครื่องมือวัดในการจัดนักศึกษาใหม่เข้าชั้นเรียนในมหาวิทยาลัยต่าง ๆ ในประเทศอินโดนีเซีย การวิจัยมีจุดประสงค์เพื่อออกแบบแบบทดสอบวัดความสามารถที่มีความตรงในการใช้ประเมินความสามารถของนักศึกษาใหม่ในอินโดนีเซีย โดยมุ่งหวังที่จะตอบคำถามการวิจัย 6 ข้อ คือ 1) มีการกำหนดความตรงเพื่อใช้ในการประเมินความสามารถอย่างไร 2) มีการกำหนดความเชื่อมั่นในการตรวจให้คะแนนของการประเมินความสามารถอย่างไร 3) ระบบการอธิบายคะแนนเหมาะสมกับกระบวนการในการวัดความสามารถเพื่อจัดนักศึกษาเข้าชั้นเรียนอย่างไร 4) แบบทดสอบวัดความสามารถแต่ละตอนสัมพันธ์ซึ่งกันและกัน และสัมพันธ์กับแบบทดสอบทั้งชุดอย่างไร 5) จุดแข็งของความสัมพันธ์ระหว่างคะแนนจากแบบทดสอบวัดความสามารถเมื่อนำมาเทียบกับคะแนนของแบบทดสอบวัดระดับเข้าชั้นเรียนที่มีอยู่แล้วคืออะไร และ 6) ผู้ตรวจให้คะแนนตีความระบบการให้คะแนนอย่างไร

การศึกษาครั้งนี้ใช้รูปแบบการวิจัยแบบผสมผสานกันของวิธีการตรวจสอบคุณภาพแบบทดสอบก่อนและหลังการทดลองใช้แบบทดสอบ (Priori and Post Priori Approach for Validation Study) วิธีการตรวจสอบคุณภาพก่อนเน้นคุณภาพด้านเนื้อหา ขณะที่วิธีการตรวจสอบ

คุณภาพหลังเน้นการตรวจสอบคุณภาพเชิงประจักษ์หลังจากการทดลองใช้แบบทดสอบ และ การศึกษาครั้งนี้ใช้การวิเคราะห์คุณภาพแบบทดสอบ 6 ชนิด ทั้งก่อน ระหว่าง และหลังการทดลองใช้ แบบทดสอบ คือ ความตรงตามเนื้อหา (Content Validation) ความตรงเชิงประจักษ์ (Face Validation) ความตรงร่วมสมัย (Concurrent Validation) ความตรงตามโครงสร้าง (Construct Validation) ความคงที่ภายใน (Internal Consistency) และความสะดวกใช้ของแบบทดสอบ (Practicality)

กระบวนการเก็บข้อมูลมี 4 ขั้นตอน คือ 1) การสร้างข้อคำถาม (Item Generating) 2) การศึกษานำร่อง 3) การลดจำนวนข้อ และ 4) การตรวจสอบคุณภาพภาคสนาม โดยมีรายละเอียด กระบวนการ และวิธีการวิเคราะห์ในแต่ละขั้น คือ 1) การสร้างข้อคำถามประกอบด้วย การประเมิน ความต้องการ การกำหนดโครงสร้างแบบทดสอบ การสร้างข้อทดสอบ การหาคุณภาพตามเนื้อหาและ ข้อมูลย้อนกลับ วิธีการวิเคราะห์ในขั้นนี้คือ การให้คะแนนโดยผู้เชี่ยวชาญและประวัติการศึกษาของ กรรมการ 2) การศึกษานำร่อง การอบรมผู้ให้คะแนน (Rater Training) การทดลองใช้แบบทดสอบ กับนักศึกษาที่มีลักษณะคล้ายกลุ่มตัวอย่าง และการตรวจให้คะแนนนักศึกษากลุ่มทดลอง 3) การลด จำนวนแบบทดสอบ ประกอบด้วย การวิเคราะห์คุณภาพแบบทดสอบตามโครงสร้างและการปรับปรุง แบบทดสอบ การวิเคราะห์ที่ใช้ คือ การวิเคราะห์แบบทดสอบ (ความยากง่ายและอำนาจจำแนก) สำหรับแบบทดสอบปรนัย และการวิเคราะห์ Rasch Analysis สำหรับแบบทดสอบอัตนัย และ 4) การตรวจสอบคุณภาพภาคสนาม ประกอบด้วย การอบรมผู้ให้คะแนน การใช้แบบทดสอบกับกลุ่ม ตัวอย่าง การให้คะแนนกลุ่มตัวอย่าง ทดสอบความเชื่อมั่น คุณภาพแบบทดสอบเชิงประจักษ์และความ สะดวกในการใช้ คุณภาพแบบคู่ขนาน และการทบทวน การวิเคราะห์หรือสถิติที่ใช้ คือ Kendull's tau-b and Kendull's w test, spearman's rho test และ Gamma test

วิธีการประเมินคุณภาพแบบทดสอบที่ใช้ คือ 1) ความตรงเชิงประจักษ์ (Face Validity) ซึ่ง ประเมินโดยการเปรียบเทียบคะแนนของชาวอินโดนีเซียที่เรียนในประเทศฟิลิปปินส์ และกลุ่มที่เรียน ในอินโดนีเซีย ผลการศึกษาพบว่า คะแนนสัมพันธ์กันอย่างสูงที่ Spearman $p = .85$ 2) ความตรง ตามเนื้อหา (The Content Validation) ที่ได้มาจากการประเมินโดยผู้เชี่ยวชาญ 4 คน และ กรรมการผู้ประเมินซึ่งเป็นอาจารย์ 10 คน ผลการวิจัยพบว่า การครอบคลุมเนื้อหาและความสำคัญ ของแบบทดสอบอยู่ในระดับที่น่าพอใจอย่างสูง 3) ความตรงร่วมสมัย (Concurrent Validation) ที่ ได้มาจากคะแนนของแบบทดสอบ Placement Test ของ Universitas Klabat และพบว่า แบบทดสอบทั้ง 2 ชนิด สัมพันธ์กันอย่างสูง คือ .964 โดยการใช้ gamma test และ 4) ความตรง ตามโครงสร้าง (The Construct Validation) ซึ่งดำเนินการ 2 แบบ คือ การวิเคราะห์รายข้อ สำหรับ แบบทดสอบปรนัย และใช้ Rasch Analysis สำหรับแบบทดสอบอัตนัย ผลของการวิเคราะห์รายข้อ พบว่า มีเพียง 3 ข้อเท่านั้นอยู่ในระดับพอใช้ ส่วนที่เหลืออยู่ในระดับดี (Good Items) และ ดีมาก (Very Good Items) ผลของการวิเคราะห์ Rasch แสดงว่า ผู้ให้คะแนนสามารถตีความระบบการให้ คะแนน และแบบทดสอบมีความเหมาะสมดีและถูกต้อง โดยทั่วไปแล้วแบบทดสอบที่พัฒนาขึ้นมี คุณภาพตรง

สรุปคุณภาพของแบบทดสอบ คือ 1) ความตรงเชิงประจักษ์ (Face Validation) หากจาก เปอร์เซ็นต์การกระจายของคะแนน ผลลัพธ์ที่ได้คือ ค่า Spearman rho .85 เพื่อหาการกระจายของ

คะแนนที่เทียบกันได้ระหว่างผู้สอบที่เป็นกลุ่มทดลองชาวอินโดนีเซีย ดีความได้ว่าแบบทดสอบใช้ประเมินความสามารถทางด้านวิชาการได้ 2) ความตรงเชิงเนื้อหา (Content Validation) ประเมินโดยใช้แบบประเมินมาตราส่วนประมาณค่า 1-5 ผลที่ได้คือค่าเฉลี่ยของแบบทดสอบในภาพรวม คือ 4.08 สารแบบทดสอบได้รับความพอใจในระดับสูง ดีความได้ว่า การครอบคลุมเนื้อหาและความสำคัญของสอบมีความตรง 3) ความตรงร่วมสมัย (Concurrent Validation) หาโดยใช้ Gamma Test ค่าที่ได้คือความสัมพันธ์เท่ากับ .964 พบว่าแบบทดสอบที่พัฒนาสัมพันธ์ในระดับสูงกับแบบทดสอบที่มีอยู่ ดีความได้ว่า แบบทดสอบทั้ง 2 ชนิดวัดโครงสร้างเดียวกัน 4) ความตรงเชิงโครงสร้าง (Construct Validation) แยกเป็น 2 ชนิด คือ การวิเคราะห์แบบทดสอบรายข้อสำหรับแบบทดสอบปรนัย และการวิเคราะห์แบบ Rasch Analysis สำหรับแบบทดสอบอัตนัย การวิเคราะห์แบบทดสอบรายข้อ พบว่า 58% เป็นแบบทดสอบที่ดีมาก 30% เป็นแบบทดสอบที่ดี และ 12% เป็นแบบทดสอบที่อยู่ในระดับพอใช้ ดีความได้ว่า โครงสร้างของแบบทดสอบส่วนปรนัยมีความตรง และในส่วนของ Rasch Analysis ผลลัพธ์ที่ได้คือ 1) ค่าความยากง่ายอยู่ระหว่าง -1.426 ถึง 2.129 หมายถึง ความคงเส้นคงวาของผู้ตรวจและความยากของแบบทดสอบมีความคงที่ และดีความได้ว่า ผู้ตรวจตีความเกณฑ์การให้คะแนนได้ดี 2) ค่า Q-index อยู่ระหว่าง .000 ถึง .025 หมายถึงความผิดพลาดในการวัดต่ำมาก ดีความได้ว่า แบบทดสอบมีความถูกต้องในการวัด 3) ค่า Z_q index กระจาย หมายถึง ไม่มีแบบทดสอบที่เหมาะสมเกินไป ดีความได้ว่า โครงสร้างถูกวัดอย่างเป็นอิสระ 4) ค่า P index อยู่ระหว่าง .097 ถึง .094 หมายถึง ไม่มีแบบทดสอบที่ไม่เหมาะสม ดีความได้ว่า แบบทดสอบวัดได้เหมาะสมดี และ 5) ค่าความเชื่อมั่นของแบบทดสอบถูกกำหนดโดยใช้ค่าดัชนีความคงที่ภายใน ผลที่ได้คือ ความคงที่ภายในของแบบทดสอบแต่ละตอนคงที่ระดับสูง โดยมีค่า Kendall's Coefficient of Concordance เท่ากับ .927 ความสัมพันธ์ระหว่างตอนต่าง ๆ กับคะแนนรวมถูกคำนวณด้วย และผลที่ได้เปิดเผยว่า แบบทดสอบตอนต่าง ๆ สัมพันธ์กันในระดับสูง โดยมีค่า Kendall's tau-b correlation เท่ากับ .942 (Formal Letter) .934 (Application Form) .917 (Making Appointment) และ .862 (Oral Presentation) ความสะดวกใช้ของแบบทดสอบคำนวณโดยการวิเคราะห์เปอร์เซ็นต์การกระจายของคะแนนในการทดลองใช้แบบทดสอบ และการใช้แบบทดสอบจริง ผลที่พบคือ คะแนนกระจายดีทั้งใน 4 ระดับ และรูปแบบการกระจายสัมพันธ์กันที่ระดับ $p = .85$ ดังนั้น ระบบการให้คะแนนสามารถนำมาใช้ได้และทำนายความสามารถได้

Samaie and Mahmoud (2014) ทำการศึกษาเรื่อง Achievement Test Development and Validation: A Measure of Reading Comprehension Strategies for Iranian Learners of English เพื่อหาคุณภาพของแบบทดสอบวัดผลสัมฤทธิ์ที่ใช้เป็นเครื่องวัดกลยุทธ์ความเข้าใจการอ่านของผู้เรียนภาษาอังกฤษในฐานะภาษาต่างประเทศชาวอิหร่านในระดับกลางต่ำ (Pre-intermediate) ณ Ilam University กลยุทธ์การอ่านที่นำมาใช้ในการพัฒนาความเข้าใจของผู้เรียน มี 4 ชนิด คือ Making Connections, Visualizing, Inferencing (inferring) และ Questioning the author คำถามการวิจัย คือ แบบทดสอบวัดผลสัมฤทธิ์ที่พัฒนาขึ้นเป็นเครื่องมือวัดที่ตรงสำหรับการประเมินกลยุทธ์ความเข้าใจในการอ่านที่กำหนดสำหรับผู้เรียนภาษาอังกฤษในฐานะภาษาต่างประเทศชาวอิหร่านหรือไม่ ผู้เข้าร่วมในการศึกษาครั้งนี้ คือ นักศึกษาชาวอิหร่านที่มีความสามารถในระดับกลางต่ำ (Pre-intermediate) จำนวน 26 คน ที่เรียนวรรณกรรมภาษาอังกฤษ ณ Ilam

University ทั้งเพศชายและหญิง อายุระหว่าง 18-25 ปี มีความสามารถใน 2 ระดับ คือ ระดับสูงและระดับต่ำ นักศึกษาในกลุ่มสูง เรียกว่า Mastery Group เป็นกลุ่มที่ได้รับการสอนให้ใช้กลยุทธ์ความเข้าใจในการอ่าน และนักศึกษาในกลุ่มต่ำ เรียกว่า Non-Mastery Group ซึ่งเป็นกลุ่มที่ไม่ได้รับการสอนกลยุทธ์ที่เกี่ยวข้อง เครื่องมือที่ใช้สำหรับการวิจัย คือแบบทดสอบวัดผลสัมฤทธิ์ความเข้าใจในการอ่านที่พัฒนาโดยผู้วิจัย ซึ่งประกอบด้วย บทอ่าน 5 บท แต่ละบทมีคำถามวัดความเข้าใจ 8 คำถาม เป็นข้อคำถามแบบให้เลือกคำตอบทั้งหมด และมีความพยายามที่จะเลือกบทอ่านที่มีความยากเท่ากับตัวบทอ่านที่นักศึกษาได้รับระหว่างเรียน

การพัฒนาแบบทดสอบดำเนินการตามกระบวนการในการพัฒนาแบบทดสอบ 3 ขั้นตอนของ Bachman and Palmer (1996) คือ ขั้นตอนออกแบบ (Design) ขั้นตอนปฏิบัติการ (Operationalization) และขั้นบริหารแบบทดสอบ (Administration) โดยมีรายละเอียด คือ ขั้นตอนออกแบบ (Design Stage) เป็นการเตรียมคำอธิบายของส่วนประกอบของการออกแบบอย่างละเอียด เพื่อให้มั่นใจว่าความสามารถในการทำภาระงานสอบตรงหรือสอดคล้องกับการใช้ภาษา และคะแนนสอบเป็นประโยชน์อย่างมากต่อจุดประสงค์ที่กำหนด Bachman and Palmer (1996) แบ่งขั้นนี้ออกเป็น 5 ส่วน คือ 1) คำบรรยายจุดประสงค์การสอบ ในการศึกษาครั้งนี้ จุดประสงค์การสอบ คือ เพื่อประเมินประสิทธิภาพของกลยุทธ์ความเข้าใจในการอ่านในการส่งเสริมความสามารถในการเข้าใจการอ่านของผู้เรียนภาษาอังกฤษในฐานะภาษาต่างประเทศชาวอิหร่าน 2) คำอธิบายขอบเขตการใช้ภาษาเป้าหมาย (Target Language Use Domain) และชนิดของภาระงาน โดยสมมติว่า ผลของการศึกษาครั้งนี้สามารถอ้างอิงถึงขอบเขตการใช้ภาษาเป้าหมายในอนาคต เพราะนักศึกษาทุกคนอาจจะเจอภาระงานด้านความเข้าใจการอ่านในชีวิตประจำวัน เช่น ในการอ่านหนังสือพิมพ์ เป็นต้น 3) คำอธิบายเกี่ยวกับประชากรผู้ทำแบบทดสอบ โดยแบบทดสอบชุดนี้เตรียมให้กับผู้เรียนภาษาอังกฤษในฐานะภาษาต่างประเทศชาวอิหร่าน ในห้องเรียนภาษาอังกฤษ ณ Ilam University ทั้งเพศชาย และเพศหญิง อายุระหว่าง 18-25 ปี 4) คำนิยามโครงสร้างที่จะวัด ในที่นี้โครงสร้างที่จะวัด คือ กลยุทธ์ความเข้าใจในการอ้างอิงของนักศึกษา 4 ด้าน คือ Visualizing, Making Connections, Inferencing และ Questioning the Author และ 5) แผนสำหรับการประเมินคุณภาพด้านคุณประโยชน์ของแบบทดสอบ ยึดตามแนวคิดการประเมินด้านคุณประโยชน์ของแบบทดสอบของ Bachman and Palmer (1996) โดยนำมาใช้เฉพาะความตรงตามทฤษฎี (Construct Validity) ส่วนรายการทรัพยากรที่ต้องการและมีอยู่แล้ว และแผนสำหรับการจัดสรรและจัดการทรัพยากรที่ต้องการสำหรับแบบทดสอบนี้ คือ ทรัพยากรบุคคล ได้แก่ นักพัฒนาแบบทดสอบ นักบริหารแบบทดสอบ และผู้ตรวจแบบทดสอบ ซึ่งเป็นผู้วิจัยเอง ทรัพยากรด้านวัสดุ ได้แก่ ห้องเรียนภาษาอังกฤษที่ Ilam University สำหรับการทดสอบสื่อวัสดุที่เกี่ยวข้อง วัสดุสอบ ได้แก่ บทอ่านที่เกี่ยวข้อง กระดาษและปากกา และทรัพยากรด้านเวลา ได้แก่ เวลาในการพัฒนา บริหาร และตรวจแบบทดสอบ

ขั้นตอนปฏิบัติการ (Operationalization Stage) เกี่ยวข้องกับการกำหนดสถานที่ (Setting) คำสั่ง (Rubric) ตัวกระตุ้น (Input) คำตอบที่คาดหวัง (Expected Response) และความสัมพันธ์ระหว่างตัวกระตุ้นและคำตอบ (Relationship between Input and Response) โดยมีรายละเอียด คือ สถานที่ (Setting) ในการใช้แบบทดสอบ คือ แบบทดสอบที่พัฒนาจะถูกใช้สอบในตอนท้ายของ

การเรียนในห้องเรียนของผู้เข้าสอบเอง เพื่อป้องกันปัญหาที่เกี่ยวข้องกับการไม่คุ้นเคยกับสถานที่ ในด้านคำสั่ง (Rubric) เนื่องจากผู้เข้าสอบอยู่ในระดับ Pre-Intermediate จึงสันนิษฐานว่ามีความคุ้นเคยกับการสอบประเภทนี้ ภาระงานต่างๆ ถูกตรวจสอบอย่างละเอียดโดยผู้สอน ส่วนตัวกระตุ้นที่ใช้ (Input) เป็นบทอ่านในรูปแบบพิมพ์ (Printed) ที่เกี่ยวข้อง ในด้านคำตอบที่คาดหวัง (Expected Response) นักศึกษาถูกคาดหวังให้ตอบคำถามแบบมีตัวเลือกของคำตอบ โดยการวางคำตอบที่ถูกต้องในกระดาษคำตอบของตนเอง และความสัมพันธ์ระหว่างตัวกระตุ้นและคำตอบ (Relationship between Input and Response) มีความสัมพันธ์กันโดยตรงระหว่างตัวกระตุ้นและคำตอบที่ต้องการ คำตอบจะถูกสร้างโดยตรงจากข้อมูลที่ให้ไว้ในตัวกระตุ้น เช่น ผู้เข้าสอบอ่านตัวกระตุ้น และตอบคำถามโดยใช้ข้อมูลที่อยู่ในบทอ่าน

ขั้นบริหารแบบทดสอบ (Test Administration Stage) เกี่ยวกับการจัดสอบกับกลุ่มเป้าหมาย เก็บข้อมูลและวิเคราะห์ข้อมูลสำหรับจุดประสงค์ต่างๆ ในการประเมินคุณประโยชน์ของแบบทดสอบ และทำการอ้างอิงหรือตัดสินใจในสิ่งที่แบบทดสอบได้ถูกกำหนดไว้ ในการศึกษาครั้งนี้ แบบทดสอบที่พัฒนาขึ้นถูกจัดสอบในคาบเรียนสุดท้ายในการเรียนของกลุ่มเป้าหมาย เพื่อหลีกเลี่ยงความไม่คุ้นเคยกับสถานที่ สำหรับจุดประสงค์ของแบบทดสอบถูกพัฒนาออกเป็น 5 ตอน แต่ละตอนมี 8 ข้อ ทั้งหมดเป็นข้อคำถามแบบให้เลือกคำตอบ มีส่วนที่เป็นคำถาม (Stem) และตัวเลือก 4 ตัวเลือก และนักศึกษาถูกกำหนดให้ตอบคำถามโดยวงคำตอบในกระดาษคำตอบ ทุกข้อมีคะแนนเท่ากัน คือ ข้อละ 1 คะแนน และมีเวลาในการสอบ 40 นาที ข้อละ 1 นาที เนื่องจากเป็นแบบทดสอบแบบเลือกตอบ แบบทดสอบจะถูกตรวจอย่างเป็นปรนัยโดยผู้สอนเอง

การประเมินคุณภาพของแบบทดสอบ (Evaluating the Usefulness of the Test) ใช้หลักการประเมินคุณประโยชน์ของแบบทดสอบของ Bachman and Palmer(1996) ซึ่งประเมินใน 6 ด้าน คือ ความตรงตามโครงสร้าง (Construct Validity) ความเชื่อมั่น (Reliability) ความสมจริง (Authenticity) การมีปฏิสัมพันธ์ (Interactiveness) การมีผลกระทบที่ดี (Impact) และความสะดวกใช้ (Practicality) สำหรับการศึกษาครั้งนี้ประเมินเฉพาะสิ่งที่สำคัญที่สุด คือ ความตรงตามโครงสร้าง (Construct Validity)

จากผลการศึกษาและข้อเสนอแนะ ระบุว่าในการใช้กลุ่มที่แตกต่างกันในการหาความตรงตามโครงสร้าง พบว่า คะแนนเฉลี่ยจากการสอบของกลุ่มเก่ง (Master Group) สูงกว่าคะแนนของกลุ่มอ่อน และผลของการเปรียบเทียบโดยการจับคู่ (A paired samples t-test) โดยการใช้ Statistical Package for the Social Sciences (SPSS) พบว่า มีความแตกต่างกันระหว่างค่าเฉลี่ยคะแนนของ 2 กลุ่ม และสนับสนุนคำกล่าวที่ว่า แบบทดสอบเป็นเครื่องมือที่ตรงสำหรับกลยุทธ์ที่ต้องการในทักษะความเข้าใจการอ่านของผู้เรียนภาษาอังกฤษในฐานะภาษาต่างประเทศระดับ Pre-Intermediate อย่างไรก็ตาม ผู้วิจัยกล่าวว่าการศึกษานี้มีข้อจำกัด 2 อย่าง คือ 1) ข้อจำกัดในการกล่าวอ้างผลการวิจัย เนื่องจากจำนวนผู้เข้าสอบมีจำนวนจำกัด 2) การประเมินคุณภาพแบบทดสอบประเมินเพียงประเด็นเดียว คือ ความตรงตามโครงสร้าง ถ้าวัดไม่เพียงพอ ดังนั้น ข้อเสนอแนะสำหรับการวิจัยครั้งต่อไป คือ ควรพัฒนาและประเมินแบบทดสอบให้มีคุณภาพสูงกวานี้ เพื่อใช้ประเมินผลกระทบของกลยุทธ์ที่กล่าวมาในการช่วยปรับปรุงทักษะความเข้าใจในการอ่านของผู้เรียนที่เรียนภาษาอังกฤษในฐานะภาษาต่างประเทศ

Ivanova (2011) อธิบายกระบวนการพัฒนาและประเมินแบบทดสอบวัดผลสัมฤทธิ์ ภาษาอังกฤษที่มีรูปแบบข้อคำถามแบบเลือกตอบ เพื่อใช้กับนักศึกษาที่ The Faculty of Mathematics and Informatics วิชาเอก Business Information Technologies (BIT) โดยมี ขั้นตอนในการพัฒนาแบบทดสอบ คือ การวางแผนโดยรวม (Overall Plan) การนิยามสาระหรือ เนื้อหา (Content Definition) และการกำหนดโครงสร้างแบบทดสอบ ดังรายละเอียด ต่อไปนี้

ในการวางแผนโดยรวม (Overall Plan) ผู้สร้างแบบทดสอบจำเป็นต้องพิจารณาอย่างละเอียด เกี่ยวกับแบบทดสอบว่าวัดทักษะอะไร (Which is its manifestations are) มีปัจจัยแวดล้อม อะไรบ้างที่ส่งผลต่อผลของการวัด การนิยามจุดประสงค์การสอบที่ชัดเจน การตัดสินใจเลือกรูปแบบ การสอบ (Test Format) ที่เหมาะสมกับจุดประสงค์การสอบมากที่สุด ในกรณีนี้เลือกใช้ Multiple-choice Format เพื่อวัดความสามารถในการใช้ไวยากรณ์และคำศัพท์ที่สอดคล้องกับภาษาในระดับ B1 โดยข้อดีของรูปแบบคำถามแบบเลือกตอบ (Multiple-choice Questions) คือ แม้จะสร้างได้ ยาก ใช้เวลานาน แต่ตรวจง่าย นักศึกษาทำได้รวดเร็ว และน่าเชื่อถือมากกว่าคำถามแบบเปิด เพราะ ความเชื่อมั่นของแบบทดสอบสัมพันธ์โดยตรงกับจุดประสงค์ของการวัด (Its Objectivity) ในแต่ละข้อ และจำนวนข้อในชุดแบบทดสอบ

ในการนิยามสาระหรือเนื้อหา (Content Definition) ผู้สร้างแบบทดสอบต้องตัดสินใจว่าเนื้อหาอะไร ที่จะถูกทดสอบจากสาระหรือเนื้อหาในตำราเรียน แจงรายการโครงสร้างไวยากรณ์ และ Tense และ คำศัพท์ภาษาอังกฤษที่ฝึกในวิชาเรียน ซึ่งต้องกล่าวไว้ในตารางกำหนดโครงสร้างแบบทดสอบในขั้น ต่อไป

โครงสร้างแบบทดสอบเปรียบเสมือนแผนที่ของแบบทดสอบ เป็นเอกสารที่มีรายละเอียดเพื่อ บอกว่าแบบทดสอบมีเนื้อหาเกี่ยวกับอะไร และจะวัดความสามารถในด้านใด โดยมีส่วนประกอบ คือ

- 1) จุดประสงค์การสอบ ไม่ว่าจะเป็นการทดสอบเพื่อจัดผู้เรียนเข้าชั้น (Placement Test) วัดผลสัมฤทธิ์ (Achievement Test) วัดความสามารถ (Proficiency Test) และตรวจสอบวินิจฉัย (Diagnostic Test)
- 2) รายละเอียดที่เกี่ยวกับผู้เข้าสอบ ได้แก่ อายุ เพศ ระดับความสามารถ ภาษา แรก ประเทศถิ่นเกิด ระดับการศึกษา เหตุผลสำหรับการสอบ
- 3) จำนวนตอนของแบบทดสอบ เวลาที่ใช้ในการทดสอบ ความแตกต่างจากแบบทดสอบอื่น ๆ
- 4) ประเภทของเนื้อความที่เลือกมาออก แบบทดสอบ เช่น เขียน และ/หรือ พูด มาจากแหล่งใด หัวข้อ และระดับของความสมจริง ความ ซับซ้อนของภาษา
- 5) ควรทดสอบทักษะภาษาอะไรบ้าง มีการแยกความต่างระหว่างแบบทดสอบที่วัด ความคิดหลัก รายละเอียดเฉพาะ การลงความเห็น หรือไม่
- 6) ควรเพิ่มทักษะภาษาใด และควรมี รายการโครงสร้างไวยากรณ์ และคำศัพท์หรือไม่
- 7) ต้องการภาระงานแบบไหน วัดได้แบบ ตรงไปตรงมา (Objectively Assessible) บูรณาการ (Integrative) สถานการณ์จำลองแบบสมจริง (Simulated Authentic) เป็นต้น
- 8) แต่ละตอนควรมีกี่ข้อ และน้ำหนักคะแนนที่ให้ อยากให้มี น้ำหนักคะแนนเท่ากันหรือมากกว่าข้ออื่น
- 9) ใช้เทคนิคการสอบแบบไหน แบบเลือกคำตอบ (Multiple-choice) เติมช่องว่าง (Gap Filling) จับคู่ (Matching) ถ่ายโอนข้อมูล (Transformations) การบรรยายภาพ (Picture Descriptions) การเขียนเรียงความ (Essay Writing) เป็นต้น
- 10) ใช้การให้คำสั่งแบบใด จะยกตัวอย่างการตอบ หรือแจ้งเกณฑ์การให้คะแนน

ด้วย และ 11) ใช้เกณฑ์การประเมินในองค์ประกอบใดบ้าง ความสำคัญระหว่างความถูกต้อง ตัวสะกด ความยาวของข้อความที่เขียน เป็นต้น

รายละเอียดในการพัฒนาแบบทดสอบ

1) การพัฒนาข้อคำถาม (Item Development) ในขั้นเริ่มต้น ข้อคำถามควรมีมากกว่าที่ ต้องการใช้ หลังจากทดลองใช้ บางส่วนอาจจะถูกตัดออกไป อัตราส่วนที่แนะนำคือ 3:2 (สร้าง: เก็บไว้) จำนวนแบบทดสอบขึ้นอยู่กับหลายปัจจัย เช่น ระยะเวลา อายุของผู้สอบ ชนิดของข้อคำถาม แบบทดสอบที่มีจำนวนข้อมาก มีความเชื่อมั่นสูง ถ้าแบบทดสอบน้อยกว่า 10 ข้อ จะไม่ได้รับค่าความ เชื่อมั่นตามที่ต้องการ

แบบทดสอบแบบเลือกตอบ ประกอบด้วย ส่วนคำถาม (Stem) ซึ่งนำเสนอปัญหาหรือคำถาม ตัวลวง (Distractors) ที่เสนอทางเลือกในการแก้ปัญหา และคำตอบ (The Key) หลักในการออกแบบ ทดสอบ Multiple-choice คือ 1) สร้างข้อคำถามที่ชัดเจนและกระชับ 2) ไม่ใช่คำซ้ำในตัวเลือกแต่ละ ตัว 3) ตัวลวงมีความสมเหตุสมผล 4) ไม่แสดงนัยหรือชี้นำข้อมูลเพื่อตอบคำถามถัดไป 5) หลีกเลี่ยงการใช้ประโยคในรูปปฏิเสธ 6) ตรวจสอบคำตอบที่ถูกต้องอีกครั้ง และให้กระจายอย่างเท่า ๆ กันในทั้ง 4 ตัวเลือก 7) ต้องมั่นใจว่าทุกคำถามมีจุดประสงค์ชัดเจน 8) มั่นใจว่าคำตอบที่ถูก ถูกจริงหรือดีที่สุด 9) มั่นใจว่าคำตอบผิดมีความเป็นไปได้ และ 10) ทำให้คำตอบมีความยาวที่หลากหลาย เพื่อหลีกเลี่ยง การคิดว่าข้อยาวคือคำตอบ

2) การออกแบบและการทดลองใช้ (Test Design and Assembly) ในการสร้างข้อคำถาม และจัดเรียงให้เหมาะสม มีข้อควรพิจารณา 2 ประการ คือ ความตรงของแบบทดสอบและรูปแบบ นั้นคือ 1) ต้องมั่นใจว่าสาระของแบบทดสอบแต่ละข้อสอดคล้องกับสาระในโครงสร้างแบบทดสอบ และ 2) แบบทดสอบควรมีรูปแบบที่อ่านง่ายและประหยัดเวลาในการทำ

3) การผลิตแบบทดสอบ (Test Production) ในขั้นตอนนี้เกี่ยวข้องกับการผลิต การพิมพ์ หรือการสำเนาแบบทดสอบ สิ่งสำคัญคือ ความปลอดภัยของแบบทดสอบ เพราะจะมีคนเข้าถึง แบบทดสอบมากกว่าในขั้นการพัฒนาแบบทดสอบ นอกจากนั้น ควรตรวจสอบคุณภาพและความ ถูกต้องของภาษาและตัวสะกดของแบบทดสอบฉบับล่าสุด เพื่อให้มีความตรงของแบบทดสอบด้วย

4) การบริหารการสอบ (Test Administration) ต้องมั่นใจว่าการกำหนดเวลาและการคุม สอบมีมาตรฐาน เพื่อไม่ให้เกิดความผิดปกติในระหว่างการสอบ และควรจัดสภาพแวดล้อมต่าง ๆ ที่ เหมาะสมและเท่าเทียมกันสำหรับผู้สอบทุกคน

5) การตรวจคำตอบ (Scoring Test Responses) ต้องใช้คำตอบที่ถูกต้อง (เฉลย) ในการ ตรวจ และตรวจความสอดคล้องของกฎการให้คะแนนกับจุดประสงค์ของแบบทดสอบอีกครั้ง และ ควรมีการวิเคราะห์แบบทดสอบ ควรทำให้สมบูรณ์ และทบทวนอีกครั้ง โดยตรวจดู 1) คะแนนเฉลี่ย ของคะแนนดิบ (The raw score mean) 2) ส่วนเบี่ยงเบนมาตรฐาน (Standard deviation) 3) ค่า ความยากเฉลี่ย (The mean item difficulty) 4) ค่าอำนาจจำแนกเฉลี่ย (The mean item discrimination) 5) ช่วงของคะแนนดิบ (Range of raw score) 6) ค่าดัชนีบางอย่างของคุณภาพ แบบทดสอบในภาพรวม (Some indices of overall test quality) และ 7) การวิเคราะห์ตัวลวง (Distractor Analysis)

6) การกำหนดเกณฑ์ผ่าน (Establishing passing scores) เนื่องจากแบบทดสอบแต่ละข้อมีน้ำหนักคะแนนเท่ากัน คะแนนของแต่ละข้อรวมกันเป็นคะแนนรวมของนักศึกษาแต่ละคน คะแนนรวมของนักศึกษาถูกแปลงเป็นเปอร์เซ็นต์ และจัดระดับตามเกณฑ์การตัดเกรดของแบบทดสอบมาตรฐาน คือ 0 – 49 % = ตก (Fail), 50 – 59 % = เป็นที่พอใจ (Satisfactory), 60 – 69 % = ดี (Good), 70 – 79 % = ดีมาก (Very Good), 80 – 100 % = ยอดเยี่ยม (Excellent)

7) รายงานผลการสอบ (Reporting test results) เมื่อจะรายงานผลให้นักศึกษาทราบ ควรพิจารณา 1) ความถูกต้องของคะแนน 2) เนื้อหาและรูปแบบของคะแนนที่จะรายงานว่าจะให้นักศึกษาทราบคะแนนที่เป็นเปอร์เซ็นต์จากคะแนนเต็ม หรือให้เป็นเกรดที่ได้ 3) ความเหมาะสมของคะแนนสอบ 4) ควรหลีกเลี่ยงความเข้าใจผิด และการใช้คะแนนที่ไม่ถูกต้อง และ 5) เรื่องการสอบใหม่สำหรับนักศึกษาที่สอบไม่ผ่าน

8. ธนาคารแบบทดสอบ (Item Banking) ธนาคารแบบทดสอบ คือ กระบวนการจัดเก็บแบบทดสอบอย่างปลอดภัย สำหรับการเรียกใช้ในอนาคต แบบทดสอบที่มีประสิทธิภาพและเป็นแบบทดสอบที่ดีสร้างได้ยาก เป็นความคิดที่ดีที่จะเก็บคำถามเหล่านั้นไว้ด้วยกัน เพื่อนำมาใช้อีกครั้งในอนาคต

9. รายงานด้านเทคนิคของแบบทดสอบ (Test technical Report) แบบทดสอบทุกชนิดควรมีการจัดเก็บและสรุปอย่างเป็นระบบในรูปแบบรายงาน ในการรายงานประเด็นสำคัญต่าง ๆ ของการพัฒนา การจัดสอบ การตรวจ การรายงาน และการวิเคราะห์และประเมินแบบทดสอบ ทำหน้าที่เป็นเหมือนประวัติของแบบทดสอบ เช่นเดียวกับข้อแนะนำสำหรับการพัฒนาแบบทดสอบในอนาคต รายงานยังมีความสำคัญ สำหรับครู ผู้เตรียมนักศึกษาสำหรับการสอบ เช่นเดียวกับผู้บริหารอื่น ๆ และผู้เกี่ยวข้อง ผู้ที่ต้องการอยากรู้ว่าสถาบันแปลงหลักสูตรสู่การปฏิบัติอย่างไร

บทสรุป: 1) ข้อมูลย้อนกลับจากนักศึกษบอกว่าแบบทดสอบไม่ยากมากแต่ยาวเกิน นักศึกษาคิดว่าแบบทดสอบแบบเลือกตอบง่ายกว่าเมื่อเทียบกับแบบทดสอบปลายเปิด แต่ตัวลงที่มีมากขึ้น มีความเป็นไปได้ที่ให้นักศึกษาตอบผิด 2) คะแนนที่นักศึกษาได้รับ สอดคล้องกันอย่างคร่าว ๆ กับคะแนนที่นักศึกษาได้จากการทำแบบทดสอบวัดความก้าวหน้าก่อนหน้า และ 3) เนื่องจากจำนวนแบบทดสอบที่มากขึ้น ช่วยเพิ่มความเชื่อมั่นของแบบทดสอบ การเปลี่ยนแปลงที่เป็นไปได้ของแบบทดสอบคือ การตัดแบบทดสอบที่ไม่ได้คุณภาพออกไป หรือเพิ่มเวลาในการทำแบบทดสอบ หรือแยกแบบทดสอบออกเป็น 2 ส่วน แบบคู่ขนานกัน

Urmston Alan et al. (2013) ทำวิจัยเรื่องการทดสอบเพื่อการวินิจฉัยความสามารถทางภาษาของนักศึกษามหาวิทยาลัยในประเทศฮ่องกง: การพัฒนาและประเมินคุณภาพแบบทดสอบ DELTA (The Diagnostic English Language Tracking Assessment) โดยมีวัตถุประสงค์เพื่อพัฒนาแบบทดสอบที่สอบผ่านเว็บและใช้ตรวจสอบติดตามความสามารถของนักศึกษาในมหาวิทยาลัย 3 แห่งในประเทศฮ่องกง การพัฒนาแบบทดสอบ DELTA อาศัยโครงสร้างความสามารถทางภาษาของ Bachman & Palmer (1996) และเป็นแบบทดสอบประเภทเลือกตอบอย่างเดียว วัดความสามารถทั้งด้านการฟัง คำศัพท์ การอ่าน และไวยากรณ์ (โดยแบบทดสอบการเขียน และพูดอยู่ในระหว่างการพัฒนา) เหมือนกับแบบทดสอบชนิดอื่น ๆ ที่ใช้โครงสร้างความสามารถนี้ แบบทดสอบการอ่าน การฟัง และไวยากรณ์ อาศัยการใช้ข้อความประกอบ (Text-based) และมีเป้าหมายเพื่อวัดความสามารถ

ทางภาษาอังกฤษโดยทั่วไปและเชิงวิชาการ ขณะที่แบบทดสอบด้านคำศัพท์เป็นประเภทวัดแยกส่วน และต้องการวัดความรู้คำศัพท์ของนักศึกษา แม้ว่าจะรู้ดีว่าแบบทดสอบแบบเลือกตอบมีข้อจำกัดในการประเมินความสามารถในการสื่อสารของผู้เข้าสอบ การเลือกใช้วิธีนี้เพราะต้องการให้ได้รับผลจาก การตรวจด้วยคอมพิวเตอร์โดยทันที รายละเอียดของแบบทดสอบแต่ละส่วนมีดังนี้

แบบทดสอบในส่วนการอ่านทดสอบความสามารถของนักศึกษาในการอ่านและเข้าใจภาษาเขียนชนิดต่าง ๆ ที่นักศึกษาจะได้อ่านสำหรับการเรียนรู้ภาษาอังกฤษและการศึกษาในระดับมหาวิทยาลัยโดยทั่วไป บทความที่นักศึกษาจะได้อ่านมีหลายชนิด คือ บทความวิชาการ (Academic Articles) บทความจากข่าว (News Articles) และบทความพิเศษ (Feature Articles) บทความเหล่านี้เป็นชนิดที่นักศึกษาระดับมหาวิทยาลัยในประเทศฮ่องกงถูกคาดหวังว่าจะได้เจอในสภาพแวดล้อมการใช้ภาษาอังกฤษ และแต่ละบทความเขียนด้วยภาษาอังกฤษมาตรฐาน ทักษะย่อยในการอ่านของแบบทดสอบ DELTA ประกอบด้วย 1) การระบุข้อมูลเฉพาะอย่าง 2) การตีความคำศัพท์หรือวลีที่ใช้โดยผู้เขียน 3) ความเข้าใจความคิดหลักและความคิดสนับสนุน 4) ความเข้าใจข้อมูลและการทำ การอ้างอิง 5) การอนุมานเหตุผลของผู้เขียน 6) การตีความทัศนคติหรือเจตนาของผู้เขียน 7) ความเข้าใจความสัมพันธ์เชิงไวยากรณ์ของคำหรือวลีต่าง ๆ ภายในตัวบทความ และ 8) การระบุประเภทของบทความ

แบบทดสอบในส่วนการฟังทดสอบความสามารถของนักศึกษาในการเข้าใจเสียงพูดชนิดต่าง ๆ ในบริบทมหาวิทยาลัย บทฟังที่นำมาใช้มีหลายชนิด เช่น การโต้เถียง (Debates) การอภิปราย (Discussion) และบทสนทนา (Conversation) บทฟังเหล่านี้ เป็นเสียงพูดที่คาดหวังว่านักศึกษาระดับมหาวิทยาลัยจะเจอในสภาพแวดล้อมการใช้ภาษาอังกฤษ บทฟังแต่ละบทใช้เสียงพูดมาตรฐานที่มีสำเนียงต่าง ๆ สมดุลกัน ทั้งสำเนียงอังกฤษ อเมริกัน ออสเตรเลีย และแคนาดา และใช้ผู้พูดที่เป็นเพศชาย-หญิงในปริมาณที่เท่า ๆ กัน ผู้พูดใช้ความเร็วในการพูดด้วยอัตราปกติ ทักษะย่อยในการฟังของแบบทดสอบ DELTA ประกอบด้วย 1) การระบุข้อมูลเฉพาะอย่าง 2) การตีความคำศัพท์หรือวลีที่ใช้โดยผู้พูด 3) ความเข้าใจความคิดหลักและความคิดสนับสนุน 4) ความเข้าใจข้อมูลและการทำ การอ้างอิง 5) การอนุมานเหตุผลของผู้พูด 6) การตีความทัศนคติหรือเจตนาของผู้พูด

แบบทดสอบส่วนไวยากรณ์วัดความสามารถของนักศึกษาในการระบุและแก้ไขข้อผิดพลาดทางไวยากรณ์ที่นักศึกษามหาวิทยาลัยในฮ่องกงเป็นผู้ทำ แบบทดสอบไวยากรณ์ประกอบด้วย 2 ส่วน คือ ส่วนที่เป็นบทความ และส่วนที่เป็นตัวเลือกในการตอบ ในการทำแบบทดสอบนักศึกษาต้องอ่านบทความก่อน เมื่อเจอคำที่ขีดเส้นใต้ ให้มาเลือกจากตัวเลือกกว่าคำที่ขีดเส้นใต้ควรแก้ไขให้ถูกต้องอย่างไร และจะมีตัวเลือกข้อ D ซึ่งจะคงเดิมไม่เปลี่ยนแปลงเป็นการยืนยันว่าคำที่ขีดเส้นใต้ถูกต้องแล้ว บทความที่เลือกใช้มาจากบทความจริงที่เขียนโดยนักศึกษามหาวิทยาลัยในฮ่องกง เป็นบทความประเภทต่าง ๆ เช่น เชิงพรรณนา (Narrative) เชิงอธิบาย (Expository) และเชิงโต้แย้ง (Argumentative) บทความจะถูกดัดแปลงเพื่อไม่ให้มีข้อผิดพลาดอื่น ๆ หลงเหลืออยู่ เพื่อไม่ให้นักศึกษาเกิดความสับสนในการตอบ ตัวอย่างของทักษะย่อยด้านไวยากรณ์ที่นำมาใช้ในแบบทดสอบ DELTA เช่น Adjective: John is (friend) FRIENDLY to his new roommates.; Adverb: The man ran more (quick) QUICKLY than Jill.; Determiner: She had a long holiday in (the Asia) ASIA last year.;

Gerund: I look forward to (meet) MEETING you next week.; Infinitive: It is easy (making) TO MAKE mistakes if you are not careful.

แบบทดสอบคำศัพท์ต้องการทดสอบความสามารถของนักศึกษาในการอ่านและเข้าใจคำและวลีต่าง ๆ ในภาษาอังกฤษที่นักศึกษาอาจพบได้ในการเรียนรู้ภาษาอังกฤษ และการศึกษาในระดับมหาวิทยาลัยโดยทั่วไป นั่นคือ นักศึกษาจะถูกวัดความสามารถในการเลือกคำศัพท์มาใช้ให้เหมาะสมกับบริบทที่กำหนดให้ ในการสอบนักศึกษาจะได้รับประโยคที่มีช่องว่างเว้นไว้ และให้เลือกคำศัพท์จากตัวเลือกมาเติมให้ได้ใจความที่ถูกต้องเหมาะสม คำศัพท์ที่นำมาใช้ในแบบทดสอบ DELTA มาจาก Coxhead's (2000) Academic Word List (AWL)

การวิเคราะห์ข้อมูลจากการสอบใช้ Rasch Measurement เพื่อตรวจสอบข้อคำถามต่าง ๆ และแปลผลออกมาเป็นรายงานการวินิจฉัย (Diagnostic Reports) ที่มีข้อมูลความยากรายข้อ (Item Difficulty) ทักษะย่อยทางภาษา (Language Subskills) ประเภทบทอ่าน (Text Type) (สำหรับการอ่านและการฟัง) และหัวข้อเรื่องหลัก (Theme) (สำหรับการอ่านและการฟัง)

สำหรับความท้าทายที่นักพัฒนาแบบทดสอบอยากให้มีในอนาคต คือ แบบทดสอบ DELTA สามารถใช้ทำนายความก้าวหน้าในการเรียนภาษาอังกฤษของนักศึกษาเพื่อให้พวกเขาทราบว่าจำเป็นต้องเรียนรู้อะไรเพิ่มเติม เพื่อให้ไปถึงระดับความสามารถที่ต้องการเมื่อจบการศึกษา และการที่นักศึกษาทำแบบทดสอบ DELTA บ่อย ๆ จะช่วยทำให้เกิดการเรียนรู้ภาษาเพิ่มมากขึ้นขณะที่อยู่ในมหาวิทยาลัย

อัจฉรา วงศ์โสธร (2532) ทำการวิจัยเรื่องการวิจัยเพื่อพัฒนาแบบทดสอบอิงปริเฉทภาษาอังกฤษสำหรับใช้กับนักศึกษาไทยในระดับต่าง ๆ โดยมีวัตถุประสงค์เพื่อ 1) พัฒนาข้อทดสอบอิงปริเฉทให้เป็นมาตรฐานมีคุณสมบัติตามเกณฑ์ โดยกำหนดปริเฉทของการใช้ภาษาในทักษะมาลา (Modality) ทางเสียงและทางตัวอักษร ได้แก่ ฟัง พูด อ่าน เขียน ตามหลักสูตรของแต่ละระดับชั้นเป็นแบบทดสอบอิงปริเฉท (Domain-referenced Test) สำหรับมัธยมศึกษาตอนต้น ตอนปลาย และอุดมศึกษา 2) กำหนดระดับความสามารถในการใช้ภาษาอิงปริเฉทอันเป็นเกณฑ์กลางของนักศึกษาในแต่ละระดับ และ 3) ศึกษา วิเคราะห์ และกำหนดปัจจัยของความสามารถสื่อความหมายทางภาษาอังกฤษของนักศึกษาไทยในระดับมัธยมศึกษาตอนต้น มัธยมศึกษาตอนปลาย และอุดมศึกษา

การวิจัยมีขั้นตอนดังนี้ คือ 1) พัฒนาแบบทดสอบอิงปริเฉทขึ้น 3 ฉบับ เพื่อใช้ในแต่ละระดับการศึกษา โดยใช้ผลการวิเคราะห์หลักสูตรและบทเรียนประกอบกับวัตถุประสงค์ของการศึกษาในแต่ละระดับ สร้างวัตถุประสงค์แบบอุดม ตารางวิเคราะห์ข้อทดสอบและสร้างแบบทดสอบ ได้ทดลองแบบทดสอบที่ใช้ ปรับปรุงตามเกณฑ์มาตรฐานแบบอิงปริเฉท แล้วจึงนำมาใช้กับกลุ่มตัวอย่างประชากร 2) สุ่มกลุ่มตัวอย่างประชากรของการวิจัยในระดับมัธยมศึกษาแบบหลายขั้นตอนอย่างมีชั้นภูมิ (Multi-stage Stratified Random Sampling) โดยสุ่มตามเกณฑ์มาตรฐานห้องเรียน ซึ่งกำหนดโดยกระทรวงศึกษาธิการ เป็น 4 กลุ่ม คือ ขนาดเล็ก ขนาดกลาง ขนาดใหญ่ และขนาดใหญ่พิเศษ ในระดับมัธยมศึกษาปีที่ 3 มีจำนวน 697 คน มัธยมศึกษาปีที่ 6 มีจำนวน 525 คน กลุ่มตัวอย่างประชากรในระดับอุดมศึกษาเป็นนิสิตซึ่งกำลังศึกษาวิชาภาษาอังกฤษตามหลักสูตรของสถาบันภาษา ในปีที่ 1 ปีที่ 2 และปีที่ 3 ได้มาโดยการสุ่มอย่างมีชั้นภูมิ (Stratified Random Sampling) จำนวน

493 คน 3) การวิเคราะห์ข้อทดสอบใช้การวิเคราะห์คุณภาพรายข้อแบบ Rasch Model และแบบอิงปริเฉท โดยหาดัชนีความง่ายและดัชนีความไว ส่วนการวิเคราะห์คุณภาพทั้งฉบับใช้วิธีวิเคราะห์แบบอิงปริเฉท โดยหาค่าความเชื่อถือได้และสัมประสิทธิ์ความเที่ยงตรง นอกจากนี้ยังใช้ t-test และ ANOVA พิสูจน์ความเที่ยงตรงแบบอิงเกณฑ์ของแบบทดสอบ การวิเคราะห์ความสามารถของกลุ่มตัวอย่างประชากรใช้สถิติแบบพาราเมตริก โดยหาค่า X, SD, CV, Kurtosis และ Skewness โดยกำหนดระดับความสามารถเป็น 5 ระดับ คือ อ่อนมากต้องแก้ไขปรับปรุงอีกมาก ค่อนข้างอ่อนยังต้องแก้ไขปรับปรุง ปานกลางยังต้องแก้ไขปรับปรุง ค่อนข้างดีมีต้องแก้ไขปรับปรุงบ้าง และดีถึงดีมากไม่ต้องแก้ไขปรับปรุง การศึกษาเกี่ยวกับปัจจัยความสามารถในการสื่อความหมายทางภาษา ใช้การวิเคราะห์ค่าสหสัมพันธ์แบบ Pearson Product Moment และการวิเคราะห์ปัจจัยแบบเสาะแสวงหา (Exploratory Factor Analysis) โดยวิธี Principal Component Analysis และหมุนแกนแบบ Varimax

ผลการวิจัยสรุปว่า 1) ข้อทดสอบที่พัฒนาขึ้น จำนวน 220 ข้อ มีความเหมาะสมกับระดับความสามารถของนักศึกษาและสอดคล้องกับโครงสร้างของแบบทดสอบ (IOC) และมีดัชนีความง่ายกับดัชนีความไวในเกณฑ์มาตรฐาน เป็นปริมาณร้อยละ 95 ของแบบทดสอบทั้งหมด 2) คุณภาพของแบบทดสอบวิเคราะห์หาค่าความเชื่อถือได้แบบอิงปริเฉท พบว่ามีค่าความเชื่อถือได้ในแต่ละระดับ คือ ฉบับมัธยมศึกษาตอนต้น .954 ฉบับมัธยมศึกษาตอนปลาย .848 และฉบับอุดมศึกษา จุฬาลงกรณ์มหาวิทยาลัย .782 สัมประสิทธิ์ความเที่ยงตรงของแบบทดสอบทั้ง 3 ฉบับ วิเคราะห์จากผลของการศึกษานำ 2 ครั้ง มีค่า .512 .537 และ .585 ตามลำดับ นอกจากนี้ แบบทดสอบทักษะมาลาทางเสียงและแบบทดสอบทักษะมาลาทางตัวอักษรในระดับมัธยมศึกษาตอนต้นและตอนปลาย และแบบทดสอบทักษะมาลาทางตัวอักษรในระดับอุดมศึกษา จุฬาลงกรณ์มหาวิทยาลัย มีความเที่ยงตรงเชิงเกณฑ์สัมพันธ์ (Criterion - Related Validity) จากการวิเคราะห์โดย t-test และ F-test สำหรับแบบทดสอบอัตโนมัติมีค่าความเชื่อถือได้แบบความเท่ากันของการให้คะแนนและความคงที่ของการประเมินอยู่ในระดับสูงมีค่า r_{xy} เท่ากับ .84 - .93 ค่าความเที่ยงตรงเชิงเนื้อหาแบบค่าความสอดคล้องของข้อทดสอบกับวัตถุประสงค์ของหลักสูตร (Item Objective Congruence) อยู่ในระดับสูง มีค่า .84-.95 3) การศึกษาระดับความสามารถของนักเรียนในการใช้อุปกรณ์ประกอบทั้งทางด้านเนื้อหาและทักษะทางภาษา พบว่านักเรียนในระดับมัธยมศึกษาปีที่ 3 และปีที่ 6 มีความสามารถรวมทุกทักษะในระดับเกือบปานกลางต้องปรับปรุงแก้ไข ส่วนในระดับอุดมศึกษา นิสิตจุฬาลงกรณ์มหาวิทยาลัยมีความสามารถในระดับปานกลาง ต้องปรับปรุงแก้ไข และเมื่อพิจารณาความสามารถแล้วพบว่า นักศึกษาทั้ง 3 ระดับ มีความสามารถใช้ทักษะมาลาทางเสียงและเนื้อหาภาษาในระดับปานกลาง ส่วนความสามารถในการใช้ทักษะมาลาทางตัวอักษรอยู่ในระดับเกือบปานกลาง โดยที่ทักษะอ่านอยู่ในระดับปานกลาง เขียนในระดับค่อนข้างอ่อนต้องแก้ไขปรับปรุง และทักษะสัมพันธ์อ่าน-เขียนอยู่ในระดับเกือบปานกลาง ในด้านการกระจายของคะแนนพบว่า นักเรียนในระดับมัธยมศึกษาปีที่ 3 มีการกระจายของคะแนนสูงที่สุด (6 ระดับ) ส่วนนักศึกษาในระดับอุดมศึกษามีการกระจายของคะแนนน้อยที่สุด (4 ระดับ)

ข้อเสนอแนะทางการวิจัย คือ การศึกษาเพื่อพิสูจน์ทฤษฎีและสมมติฐานเกี่ยวกับปัจจัยของความสามารถในการสื่อสารและรับสารทางภาษา ควรวิจัยในแนวทางต่อไปนี้ 1) ความสามารถในการ

สื่อสารทางภาษาทางการพูดและการเขียนซึ่งเป็นทักษะมาลาต่างกัน มีคุณลักษณะย่อยในระบบการจำแนกประเภทแบบเป็นขั้นตอน (Taxonomy) ว่ามีระดับขั้นตอน (Stage) ของการเรียนรู้ เป็นปัจจัยซึ่งมีลักษณะโดยรวมเข้ากลุ่มสัมพันธ์ และควรวิจัยเกี่ยวกับคุณลักษณะของทักษะรับสาร คือ การฟัง การอ่านว่ามีกลุ่มสัมพันธ์รูปแบบใด 2) การเรียนรู้ภาษาในระดับต้น (Beginning Level) เช่น ในระดับประถมศึกษาและระดับมัธยมศึกษาตอนต้นนั้น วิธีการเรียนการสอนแบบทักษะเดียวหรือทักษะสัมพันธ์ จะสอดคล้องกับธรรมชาติของการเรียนรู้ภาษามากกว่ากันและวิธีใดจะมีประสิทธิภาพมากกว่ากัน 3) ในระดับกลาง กลางขั้นต้น กลางขั้นสูง และระดับสูง (Intermediate, Lower Intermediate, Upper Intermediate และ Advanced) การประสานสัมพันธ์ระหว่างทักษะมาลาทางเสียงและทางตัวอักษร ควรเป็นไปในรูปแบบใด คือ ประสานโดยผ่านสื่อแบบใดจึงจะมีประสิทธิภาพมากที่สุด เช่น โดยการใช้บทอ่าน - บทฟัง โดยเทปบันทึกเสียง เทปโทรทัศน์ และมีหน้าที่ภาษาหัวข้อเรื่องหรือสถานการณ์ใด 4) ในการจัดการเรียนการสอนภาษาในทุกระดับ เนื้อหาทางภาษาได้แก่ เสียง ศัพท์ และไวยากรณ์ ควรประสานสัมพันธ์กับทักษะมาลาทางเสียง ได้แก่ การฟัง พูด ฟัง - พูด ฟัง - เขียน ฟัง - อ่าน - เขียน และทักษะมาลาทางตัวอักษร ได้แก่ การอ่าน เขียน การอ่าน - เขียน เป็นปริมาณเท่าใด และในรูปแบบใด โดยศึกษาจากประเภทของโรงเรียน ระดับการศึกษา วัย ความถนัด วุฒิภาวะ ความพร้อมและพื้นฐานความสามารถทางภาษาของผู้เรียน และ 5) ปัจจัยของทักษะมาลาสื่อสารและรับสารของนักศึกษาในเขตการศึกษาอื่นๆ และในกลุ่มอุดมศึกษาทางภาคเอกชนและรัฐบาล ที่ไม่ได้อยู่ในการวิจัยครั้งนี้ จะมีคุณลักษณะซึ่งมีกลุ่มสัมพันธ์ร่วมกัน เหมือนหรือต่างกับกลุ่มตัวอย่างประชากรในการวิจัยครั้งนี้หรือไม่ เพื่อให้ผลของการวิเคราะห์ที่ละเอียดมากขึ้น

อัจฉรา วงศ์โสธร และ ศิริพร พงษ์สุรพิพัฒน์ (2544) ได้วิจัยด้านการสร้างและพัฒนาแบบทดสอบภาษาอังกฤษ โดยมีวัตถุประสงค์เพื่อพัฒนาแบบทดสอบสมิทธิภาพทั่วไปทางภาษาอังกฤษด้านการฟัง การอ่าน และการเขียน เพิ่มมากขึ้นจากแบบทดสอบที่ได้จากการสร้างขึ้นในการวิจัยระยะปีที่ 1 และให้ได้แบบทดสอบ 4 ฉบับให้คู่ขนานกันในเนื้อหาและสถิติ เพื่อวิเคราะห์และปรับปรุงแบบทดสอบที่สร้างขึ้นทั้งฉบับและเป็นรายข้อ และเพื่อพัฒนาแบบทดสอบให้เป็นมาตรฐานและเก็บไว้ในธนาคารแบบทดสอบโดยใช้คอมพิวเตอร์ (Computerized Item Bank) เก็บข้อมูลกับกลุ่มตัวอย่าง 2 กลุ่ม คือ นิสิตจุฬาลงกรณ์มหาวิทยาลัยตั้งแต่ชั้นปีที่ 1 ถึงบัณฑิตศึกษา รวมทั้งบุคคลภายนอกที่สมัครตามประกาศของสถาบันภาษา และเข้าสอบในภาคปลายปีการศึกษา 2533 เมื่อวันที่ 16-31 มกราคม 2534 จำนวน 481 คน และบุคคลภายนอกที่มีพื้นฐานความรู้ระดับปริญญาตรีและปริญญาโท ซึ่งเข้าทดสอบในภาคปลายปีการศึกษา 2534 จำนวน 87 คน ผลการวิจัยสรุปว่าได้แบบทดสอบปรนัยวัดทักษะการฟัง อ่าน และเขียน และแบบอัตนัยวัดทักษะการเขียน จำนวน 4 ฉบับ ที่มีความเที่ยงอยู่ในเกณฑ์ดี - ดีมาก ($KR_{20} = .854 - .910$ แบบ $\Phi = .822 - .896$) และแบบอัตนัยวัดทักษะการเขียนที่มีค่าความเที่ยงอยู่ในเกณฑ์ดี - ดีมาก ($R_{xy} = .594 - .840$) โดยมีค่าอำนาจในการทำนาย $48.00 - 58.50$ ได้ข้อทดสอบแบบปรนัยวัดสมิทธิภาพทั่วไปทางภาษาอังกฤษที่มีคุณภาพดีมาก - พอใช้ จำนวน 471 ข้อ ได้ข้อทดสอบแบบอัตนัยวัดสมิทธิภาพทางการเขียนที่มีคุณภาพดี - พอใช้ได้ จำนวน 12 ข้อ และได้ข้อทดสอบแบบปรนัยวัดสมิทธิภาพทั่วไปทางภาษาอังกฤษซึ่งได้มาตรฐานเข้าเกณฑ์ของ Rasch Model จำนวน 414 ข้อ นอกจากนี้ผู้วิจัยได้ให้ข้อเสนอแนะว่า ควรทำการวิจัยเปรียบเทียบระดับสมิทธิภาพทางการใช้ภาษาอังกฤษของนิสิต

นักศึกษาและบุคคลทั่วไปที่วัดโดยแบบทดสอบของสถาบันภาษาและแบบทดสอบมาตรฐานที่มีการใช้อย่างแพร่หลาย เช่น แบบทดสอบ TOEFL (Test of English as a Foreign Language) ซึ่ง Educational Testing Service แห่ง Princeton, New Jersey พัฒนาขึ้นใช้ และแบบทดสอบ IELTS (International English Language Testing System) ซึ่ง British Council และ University of Cambridge Local Examinations Syndicate พัฒนาขึ้น

2.3.2 งานวิจัยด้านการประเมินคุณภาพแบบทดสอบวัดความสามารถภาษาอังกฤษ

งานวิจัยที่เกี่ยวข้องกับการประเมินคุณภาพแบบทดสอบวัดความสามารถภาษาอังกฤษ มีจำนวน 12 เรื่อง ผู้ดำเนินการวิจัย คือ Chen (2009) Ito (2005), Pizorn, Karmen, and Moe Ali (2012), Cheng (2013) Wu and Mei (2012), LCRT Consulting (2013) Riazi and Mehdi (2014), Zhou Yuemei (2004), Lopez & Janssen (2010), Lopez and Alexis Augusto & Janssen (2010), Wu (2011), Weir, Chan, and Nakatsuhara (2013) และ ศิริพร พงษ์สุรพิพัฒน์, กรองแก้ว วรรณสุด, and และสุภาณี ชินวงศ์ (2544) โดยมีรายละเอียดในแต่ละงานวิจัยดังนี้

Chen (2009) ประเมินคุณประโยชน์ของแบบทดสอบและเปรียบเทียบแบบทดสอบภาษาอังกฤษระดับมหาวิทยาลัยของประเทศจีนระดับ 4 ระหว่างชุดเก่าและชุดใหม่ (CET-4) โดยมีวัตถุประสงค์เพื่อเปรียบเทียบแบบทดสอบชุดเก่าและชุดใหม่ตามคุณประโยชน์ของแบบทดสอบ 6 ประการ ตามแนวทางของ Bachman and Palmer (1996) คือ ความเชื่อมั่น (Test Reliability) ความตรงเชิงโครงสร้าง (Construct Validity) ความสมจริง (Authenticity) การมีปฏิสัมพันธ์ (Interactiveness) ผลกระทบ (Impact) และความสะดวกใช้ (Practicality) และต้องการที่จะตรวจสอบว่าแบบทดสอบด้านไวยากรณ์และคำศัพท์ ได้รับการทดสอบในระบบใหม่อย่างไร มีประโยชน์และมีประสิทธิภาพเพียงใด โดยสามารถสรุปผลการวิจัยได้ดังนี้

ด้านความตรงเชิงโครงสร้าง (Construct Validity) ของแบบทดสอบชุดเก่าและใหม่ พบว่า

1) ตารางแสดงโครงสร้างข้อสอบแบบทดสอบชุดใหม่ ให้รายละเอียดโครงสร้างเนื้อหาและวิธีการวัดที่ละเอียด ชัดเจนกว่ามาก และ 2) แบบทดสอบชุดใหม่มีการทดสอบทักษะต่าง ๆ อย่างหลากหลายมากกว่าชุดเดิม โดยเฉพาะการฟังและการอ่านมีเทคนิคการทดสอบที่หลากหลายมากขึ้น เช่น มีการให้ภาระงานการฟังที่หลากหลาย โดยการเพิ่ม Long Conversations และ Compound Dictations มากกว่าชุดเดิม ซึ่งมีเฉพาะ Short Conversations และ Listening Passages นอกจากนี้ ในส่วนของการอ่านมีการวัดทักษะต่าง ๆ เพิ่มขึ้น ด้านไวยากรณ์และคำศัพท์มีการทดสอบโดยใช้บริบทมากขึ้น และในส่วนของกรายงานผลคะแนนมีการเพิ่มข้อมูลให้ผู้เข้าสอบได้ทราบจุดแข็งและจุดอ่อนของตนเอง

ด้านความเชื่อมั่น (Test Reliability) แบบทดสอบชุดใหม่มีค่าความเชื่อมั่นสูงกว่าชุดเดิม โดยน่าจะมีผลมาจากจำนวนข้อ เวลาที่เพิ่มขึ้นและการวัดด้วยวิธีการที่หลากหลายมากกว่าเดิมในการทดสอบฟัง

ด้านความสมจริง (Authenticity) และการมีปฏิสัมพันธ์ (Interactiveness) ซึ่งประเมินความสอดคล้องระหว่างภาระงานสอบ (Test Tasks) และภาระงานการใช้ภาษาในอนาคต (TLU Tasks) และความสัมพันธ์ระหว่างภาระงานการสอบและผู้เข้าสอบ โดยรวมพบว่า แบบทดสอบทั้ง 2 ชุด มีความสมจริงในระดับหนึ่ง และแบบทดสอบชุดใหม่มีความสมจริงมากกว่าชุดเก่า กล่าวคือ 1) ด้านการฟังเพื่อความเข้าใจในแบบทดสอบชุดใหม่มีการฟังในเชิงวิชาการเพิ่มเข้ามา และมีการวัดทักษะเชิงวิชาการมากกว่าเดิม 2) ด้านการอ่าน แบบทดสอบทั้ง 2 ชุด มีหัวเรื่องที่หลากหลายและเป็นที่น่าสนใจของผู้เข้าสอบ แต่ในแบบทดสอบการอ่านชุดใหม่ มีการเพิ่มทักษะ Skimming และ Scanning และเป็นเทคนิคที่นักศึกษาใช้บ่อยอยู่แล้วทั้งในชีวิตประจำวันและเชิงวิชาการ จึงทำให้แบบทดสอบอ่านชุดใหม่ดูมีความสมจริงและมีปฏิสัมพันธ์กับผู้เข้าสอบมากกว่า อย่างไรก็ตาม ในแบบทดสอบการอ่าน ไม่ปรากฏการทดสอบสาระที่นักศึกษาอาจได้เจอในอนาคต คือ การอ่านที่เกี่ยวกับข้อควรปฏิบัติ (Instructions) การอ่านตารางเวลา (Schedules) และ 3) เนื่องจากไม่มีการปรับเปลี่ยนแบบทดสอบการเขียน จึงไม่สามารถเปรียบเทียบความสมจริงในทักษะเขียนได้ อย่างไรก็ตาม ในด้านทักษะเขียน ยังไม่มีการทดสอบภาระงานเขียนที่นักศึกษาสามารถนำไปใช้ได้ในอนาคต (Writings for practical purposes)

ด้านผลกระทบ (Impact) ของแบบทดสอบต่อผู้เข้าสอบ ระบบการศึกษา และสังคม แบบทดสอบชุดใหม่มีผลกระทบต่อรูปแบบและเนื้อหาในการเรียนรู้ของผู้เข้าสอบ และการสอบกลายเป็นแรงจูงใจ ทำให้นักศึกษาสนใจการเรียนทั้งในเชิงปริมาณและคุณภาพ แต่ทำให้นักศึกษาอยู่ในภาวะกดดันมากขึ้น สำหรับผลกระทบต่ออาจารย์ผู้สอน พบว่า มีการปรับเปลี่ยนเนื้อหาและวิธีการสอน อาจารย์มีแนวโน้มที่จะสอนให้สอดคล้องกับการสอบมากขึ้น เนื่องจากแบบทดสอบชุดใหม่มีการเพิ่มสัดส่วนในการวัดทักษะฟังและพูดมากกว่าเดิม และอาจารย์ได้ปรับการสอนให้ไปทางการสื่อสารมากขึ้น สำหรับผลกระทบต่อระบบการศึกษาและสังคม แบบทดสอบระดับ 4 ได้ส่งผลให้ผู้เข้าสอบคิดถึงความสำเร็จ ความรู้ และความสามารถทั้งของรายบุคคลและสังคมมากกว่าเดิม

ด้านความสะดวกใช้ (Practicality) เนื่องจากมีการใช้แบบทดสอบมาเกือบ 20 ปีแล้ว จึงไม่พบปัญหาด้านการนำแบบทดสอบมาใช้ทั้งในด้านการพัฒนา การบริหาร และการตรวจให้คะแนน

สำหรับด้านการทดสอบไวยากรณ์และคำศัพท์ พบว่ามีการทดสอบทักษะดังกล่าวโดยใช้บริบทมากขึ้น และมีการวัดทักษะไวยากรณ์และคำศัพท์ในทักษะอื่นด้วย เช่น การฟัง การอ่านและการเขียน อย่างไรก็ตาม มีข้อเสนอแนะสำหรับการวัดทักษะไวยากรณ์และคำศัพท์ คือ 1) การคัดเลือกบทอ่านที่ไม่เหมาะสมและการตัดคำออก อาจมีผลกระทบต่อความตรงและความเชื่อมั่น และ 2) การสร้างตัวลวงทำได้ยาก และ 3) อาจจะได้ข้อคำถามที่ด้อยคุณภาพในการวัดความสามารถด้านไวยากรณ์และคำศัพท์

สรุปได้ว่า แบบทดสอบ CET-4 ชุดใหม่ มีระดับของความมีประโยชน์สูงกว่าชุดใหม่ และยังมีประสิทธิภาพและความตรงในการทดสอบไวยากรณ์และคำศัพท์มากกว่าเดิม

ข้อเสนอแนะสำหรับการศึกษาคำศัพท์ คือ ควรมีการศึกษาเพิ่มเติมในการช่วยลดผลกระทบด้านลบ (Negative Impact) และช่วยเพิ่มผลกระทบด้านบวก (Positive Impact) และเนื่องจากมีการปรับแบบทดสอบระดับ 4 และเนื้อหาในหลักสูตรวิชาภาษาอังกฤษระดับมหาวิทยาลัย ทำให้เกิดการ

ปรับเปลี่ยนแนวทางและการวางแผนการสอนของอาจารย์ การศึกษาครั้งต่อไปควรศึกษาวิธีการสอนที่เหมาะสมกับข้อกำหนดใหม่ ๆ ที่เกิดขึ้นในแบบทดสอบชุดใหม่

Ito (2005) ทำการศึกษาเรื่อง A Validation Study on the English Language Test in a Japanese Nationwide University Entrance Examination โดยมีจุดประสงค์เพื่อหาความเชื่อมั่นและความตรงของคำถามภาษาอังกฤษในแบบทดสอบ JFSAT (The Joint First Achievement Test) เป้าหมายคือ เพื่อตรวจสอบว่าแบบทดสอบภาษาอังกฤษในแบบทดสอบ JFSAT เป็นเครื่องมือที่เชื่อถือได้และตรงกับความสามารถภาษาอังกฤษของนักศึกษา ความตรงที่ศึกษาในครั้งนี้คือความตรงที่สัมพันธ์กับเกณฑ์ (Criterion Validity) และความตรงเชิงโครงสร้าง (Construct Validity) ของแบบทดสอบภาษาอังกฤษที่อยู่ในแบบทดสอบ JFSAT คำถามการวิจัยในการศึกษาครั้งนี้ คือ 1) ความเชื่อมั่นและความตรงที่สัมพันธ์กับเกณฑ์ของแบบทดสอบส่วนภาษาอังกฤษในแบบทดสอบ JFSAT อยู่ในระดับใดและแบบทดสอบภาษาอังกฤษวัดความสามารถภาษาอังกฤษของนักศึกษาหรือไม่ และ 2) ไม่มีความสัมพันธ์ระหว่างแบบทดสอบวัดการออกเสียงที่ให้สอบโดยการเขียนและแบบทดสอบวัดความเข้าใจในการฟังใช้หรือไม่ และคำถามการวิจัยจากการศึกษาครั้งนี้ คือ 3) ความตรงเชิงโครงสร้างของแบบทดสอบภาษาอังกฤษที่อยู่ในแบบทดสอบ JFSAT อยู่ในระดับใด

สำหรับคำถามวิจัยที่ 1 เกี่ยวข้องกับการหาค่าสัมประสิทธิ์ความเชื่อมั่นของแบบทดสอบภาษาอังกฤษที่อยู่ในแบบทดสอบ JFSAT เช่น Cronbach's alpha และค่าสัมประสิทธิ์ความสัมพันธ์ของ Pearson (Pearson's Product-Moment Correlation Coefficient) ระหว่างแบบทดสอบภาษาอังกฤษในแบบทดสอบ JFSAT และแบบทดสอบที่เป็นเกณฑ์จากภายนอก เช่น a cloze test และระหว่างแบบทดสอบวัดการออกเสียงที่ให้สอบโดยการเขียนและแบบทดสอบวัดความเข้าใจในการฟัง เช่น แบบทดสอบส่วนการฟังในแบบทดสอบ TOEFL (Test of English as a Foreign Language)

คำถามการวิจัยที่ 2 เกี่ยวข้องกับการหาค่าสัมประสิทธิ์ความสัมพันธ์ของแบบทดสอบตอนต่าง ๆ ของแบบทดสอบภาษาอังกฤษที่อยู่ในแบบทดสอบ JFSAT และตรวจสอบค่าสหสัมพันธ์บนพื้นฐานของเกณฑ์ที่ใช้ในการศึกษาความสัมพันธ์ภายในที่ผ่านมา ในการศึกษาความสัมพันธ์ภายในแบบทดสอบภาษาอังกฤษที่อยู่ในแบบทดสอบ JFSAT ถูกแบ่งออกเป็น 5 ตอน คือ 1) แบบทดสอบวัดการออกเสียง (Pronunciation) 2) แบบทดสอบวัดไวยากรณ์ (Grammar) 3) แบบทดสอบวัดทักษะการพูด (Spoken English) 4) แบบทดสอบวัดทักษะการเขียน (Written English) และ 5) แบบทดสอบวัดความเข้าใจในการอ่าน (Reading Comprehension)

ผู้เข้าร่วมในการศึกษา 100 คน เป็นกลุ่มตัวอย่างของนักศึกษาชั้นปีที่ 1 ผู้ที่ลงทะเบียนเรียนในวิชาภาษาอังกฤษทั่วไปที่ Aichi University of Education ในประเทศญี่ปุ่น กลุ่มตัวอย่างส่วนใหญ่อายุ 18 ปี อายุเฉลี่ย คือ 18 ปี 2 เดือน อัตราส่วนระหว่างเพศชายและหญิง คือ 1 ต่อ 1 ทุกคนได้รับการเรียนภาษาอังกฤษอย่างเป็นทางการมากกว่า 6 ปี ก่อนเข้าร่วมการศึกษาครั้งนี้ วิชาเอกของกลุ่มตัวอย่าง คือ สาขาทางด้านวิทยาศาสตร์ กลุ่มตัวอย่างมีลักษณะคล้ายกัน (Homogeneous) ทั้งด้านสัญชาติ ภูมิหลังทางด้านภาษา การศึกษาและอายุ

เครื่องมือที่ใช้ในการศึกษาครั้งนี้ คือ 1) แบบทดสอบ Cloze Test แบบให้เติมคำเอง จำนวน 70 ข้อ คะแนนเต็ม 70 คะแนน ใช้เวลาในการทำ 30 นาที 2) แบบทดสอบวัดความเข้าใจในการฟัง

ของแบบทดสอบ TOEFL จำนวน 50 ข้อ คะแนนเต็ม 50 คะแนน ใช้เวลาในการทำ 25 นาที 3) แบบทดสอบภาษาอังกฤษ (The JFSAT-English) ฉบับปี 1991 จำนวน 58 ข้อ คะแนนเต็ม 200 คะแนน ใช้เวลาในการทำ 80 นาที และ 4) แบบทดสอบวัดการออกเสียงที่ให้สอบโดยการเขียน ฉบับปี 1989 และ 1992 จำนวน 20 ข้อ ข้อละ 2 คะแนน รวม 40 คะแนน ใช้เวลาในการทำ 10 นาที

ในการศึกษานำร่อง ผู้เข้าร่วมในการศึกษาเป็นนักศึกษาชั้นปีที่ 2 ที่เรียนวิชาภาษาอังกฤษทั่วไป ที่ Aichi University of Education ในประเทศญี่ปุ่น อายุเฉลี่ย คือ 19 ปี 4 เดือน อัตราส่วนเพศชายต่อเพศหญิง คือ 1 ต่อ 1 ทุกคนได้เรียนภาษาอังกฤษอย่างเป็นทางการมากกว่า 7 ปี ก่อนการศึกษาครั้งนี้ วิชาเอกของกลุ่มทดลองคือ การศึกษาภาษาญี่ปุ่น และศิลปศึกษา จึงกล่าวได้ว่ากลุ่มทดลองมีความคล้ายกันทั้งด้านสัญชาติ ภูมิหลังด้านภาษา ระดับการศึกษาและอายุ

ผลการศึกษานำร่องพบว่า แบบทดสอบ Cloze Test มีค่าสัมประสิทธิ์ความเชื่อมั่นค่อนข้างสูง ($r = .840$) และค่าความสัมพันธ์ ($r = .604$) กับแบบทดสอบ TOEFL และผู้เขียนได้ตัดสินใจใช้แบบทดสอบ Cloze Test เป็นเกณฑ์ภายนอกสำหรับวัดภาษาอังกฤษของผู้เข้าร่วม เพราะค่าความเชื่อมั่นของแบบทดสอบ Cloze Test เกินกว่าระดับที่ตั้งเป้าไว้ที่ระดับ .8 ($r = .840$) และแบบทดสอบมีความสัมพันธ์กับแบบทดสอบที่น่าเชื่อถืออย่างแบบทดสอบ TOEFL สูงกว่า .5 ($r = .604, p < .01$) แบบทดสอบวัดความเข้าใจในการฟังของแบบทดสอบ TOEFL ถูกใช้เป็นแบบทดสอบวัดความเข้าใจในการฟัง สำหรับตรวจดูว่าไม่มีความสัมพันธ์กับแบบทดสอบวัดการออกเสียงที่ให้ทำโดยการเขียนจริงหรือไม่ ผู้เขียนไม่ได้ดำเนินการศึกษาความตรงของแบบทดสอบการฟังในแบบทดสอบ TOEFL เนื่องจากเป็นแบบทดสอบที่เชื่อถือได้อยู่แล้ว

สำหรับกระบวนการตรวจให้คะแนนสำหรับแบบทดสอบอื่นๆ ยกเว้นแบบทดสอบ Cloze Test ผู้วิจัยให้นักศึกษาแลกเปลี่ยนกระดาษคำตอบกันและตรวจตามคำสั่งของผู้วิจัย หลังจากนั้นผู้วิจัยเก็บรวบรวมแบบทดสอบหรือคำตอบเพื่อนำไปคำนวณทางสถิติต่อไป

ในด้านซอฟต์แวร์ทางสถิติและระดับนัยสำคัญ การวิเคราะห์ทางสถิติถูกกำหนดโดยโปรแกรมคอมพิวเตอร์ SPSS เวอร์ชัน 7.5 ในห้องสมุดของภาควิชาการศึกษาภาษาอังกฤษ คณะศึกษาศาสตร์ มหาวิทยาลัยโรซามา ประเทศญี่ปุ่น ระดับ alpha สำหรับการตัดสินใจทางสถิติทุกชนิดถูกกำหนดไว้ที่ $\alpha = .01$

ผลการศึกษา โดยตอบคำถามการวิจัยทั้ง 3 ข้อ คือ 1) ค่าความเชื่อถือได้และความตรงที่สัมพันธ์กับเกณฑ์ของแบบทดสอบภาษาอังกฤษ JFSAT สูงเพียงใด แบบทดสอบภาษาอังกฤษ JFSAT วัดความสามารถภาษาอังกฤษของนักศึกษาหรือไม่ ค่าสัมประสิทธิ์ของความเชื่อถือได้ของแบบทดสอบ Cloze Test เท่ากับ .853 เป็นค่าที่ค่อนข้างสูงกว่าในการทดลองใช้ ($r = .841$) ผลที่ได้ชี้ให้เห็นว่าแบบทดสอบ Cloze Test มีความเชื่อถือได้ค่อนข้างสูงอย่างคงที่ ค่าสัมประสิทธิ์ความเชื่อถือได้ของแบบทดสอบภาษาอังกฤษ JFSAT ค่อนข้างสูงเช่นกัน ($r = .817$) อย่างไรก็ตาม ค่าสัมประสิทธิ์ความเชื่อถือได้ของแบบทดสอบวัดการออกเสียงแบบสอบโดยการเขียน ($r = .208$) และแบบทดสอบวัดความเข้าใจในการฟังของแบบทดสอบ TOEFL ($r = .398$) ต่ำมาก เนื่องจากแบบทดสอบ Cloze Test และแบบทดสอบภาษาอังกฤษ JFSAT มีค่าความเชื่อถือได้ค่อนข้างสูง ผู้วิจัยหาความสัมพันธ์ของแบบทดสอบทั้ง 2 ชนิด โดยไม่ได้แก้ไขแบบทดสอบ Cloze Test ค่าสัมประสิทธิ์ความสัมพันธ์ของแบบทดสอบทั้ง 2 ชนิด อยู่ในระดับปานกลาง ($r = .462, r\text{-square}$

= .213) แสดงให้เห็นว่าความแปรปรวนร่วมกันระหว่างแบบทดสอบ 2 ชนิด คือ 21.3 % และแบบทดสอบภาษาอังกฤษ JFSAT เป็นแบบทดสอบที่มีความตรงในการวัดความสามารถภาษาอังกฤษอยู่บ้าง 2) มีความสัมพันธ์ที่ไม่มีนัยสำคัญระหว่างแบบทดสอบวัดการออกเสียงแบบให้สอบโดยการเขียน กับแบบทดสอบวัดความเข้าใจในการฟังหรือไม่ การขาดความสัมพันธ์ระหว่างแบบทดสอบวัดการออกเสียงแบบให้สอบโดยการเขียนและแบบทดสอบวัดความสามารถในการฟังของแบบทดสอบ TOEFL อาจจะมาจากความจริงที่ว่าความสามารถจำแนกเสียงต่าง ๆ แต่ละเสียงในแบบทดสอบวัดการออกเสียงที่ให้สอบโดยการเขียนแตกต่างกันโดยสิ้นเชิงจากความสามารถในการฟังที่มีมากกว่าที่แบบทดสอบวัดความเข้าใจในการฟังของแบบทดสอบ TOEFL พยายามจะทดสอบ ในความหมายนี้ คำตอบสำหรับคำถามการวิจัยข้อ 2 คือ “ใช่” อย่างไรก็ตามพบข้อบกพร่อง (Drawback) ของการศึกษาครั้งนี้ คือ แบบทดสอบวัดความเข้าใจในการฟังของแบบทดสอบ TOEFL มีความเชื่อถือได้ต่ำ ผลการศึกษาต้องการให้มีการศึกษาเพิ่มเติม ดังนั้น จะมีความสัมพันธ์ที่ไม่มีนัยสำคัญระหว่างแบบทดสอบวัดการออกเสียงที่ให้สอบโดยการเขียนและแบบทดสอบวัดความเข้าใจในการฟังของแบบทดสอบ TOEFL หรือไม่ ยังไม่ชัดเจน และ 3) ความตรงเชิงโครงสร้างของแบบทดสอบภาษาอังกฤษในแบบทดสอบ JFSAT สูงระดับใด การศึกษาความสัมพันธ์ภายในนี้ เสนอแนะว่าค่าสัมประสิทธิ์ความสัมพันธ์ระหว่างแบบทดสอบวัดการออกเสียงแบบให้สอบโดยการเขียน และแบบทดสอบตอนย่อยอื่นๆ ($r = .153, n.s. \text{ to } .280, n.s.$) ต่ำเกินไปที่จะสอดคล้องกับเกณฑ์ 3 ด้าน

สรุปได้ว่า ผลของการศึกษาที่ 1 แสดงว่าแบบทดสอบภาษาอังกฤษของแบบทดสอบ JFSAT เป็นเครื่องมือที่เหมาะสมในระดับหนึ่งสำหรับการวัดความสามารถภาษาอังกฤษของผู้เข้าสอบ ในแง่ของความเชื่อถือได้และความตรง ประการที่สอง ไม่มีค่าความสัมพันธ์ที่มีนัยสำคัญทางสถิติระหว่างแบบทดสอบวัดการฟังของแบบทดสอบ TOEFL และแบบทดสอบวัดการออกเสียงที่ให้สอบโดยการเขียน อย่างไรก็ตาม เนื่องจากแบบทดสอบทั้ง 2 ชนิด แสดงถึงความเชื่อถือได้ต่ำ ผลลัพธ์นี้ควรพิจารณาอย่างรอบคอบ การศึกษาที่ 2 คือการศึกษาความสัมพันธ์ภายในเพื่อหาความตรงเชิงโครงสร้างของแบบทดสอบ ผลของการศึกษาที่ 2 แสดงว่าความสัมพันธ์ระหว่างแบบทดสอบวัดการออกเสียงที่ให้สอบโดยการเขียนกับแบบทดสอบตอนย่อยอื่น ๆ ต่ำมาก แสดงถึงความตรงเชิงโครงสร้างของแบบทดสอบที่ต่ำด้วย

ข้อจำกัดและประเด็นที่ยังเป็นปัญหา คือ 1) ผลการวิจัยสามารถอ้างอิงผลสำหรับผู้เรียนภาษาอังกฤษชาวญี่ปุ่นเท่านั้น ผู้ที่ได้เรียนภาษาอังกฤษในบริบทการศึกษาที่เป็นทางการ อย่างไรก็ตาม การเลือกกลุ่มตัวอย่างเพียงสัญชาติเดียวเป็นจุดแข็งที่ต้ออย่างหนึ่งของการศึกษา ทำให้ได้หลักฐานเชิงประจักษ์ 2) การศึกษาปัจจุบันเน้นที่การศึกษาความตรงภายในของแบบทดสอบภาษาอังกฤษในแบบทดสอบ JFSAT โดยการคำนวณสัมประสิทธิ์ความสัมพันธ์ในระหว่างตอนย่อยต่าง ๆ ดังนั้นแม้ว่าแบบทดสอบตอนย่อยทั้งหมดยกเว้นแบบทดสอบวัดการออกเสียงที่ให้สอบโดยการเขียนแสดงถึงการมีความตรงสูงระดับหนึ่ง ในการศึกษาความสัมพันธ์ภายใน ผลที่ได้ไม่สามารถรับประกันได้ว่าแต่ละตอนย่อยกำลังวัดสิ่งที่ตั้งใจจะวัดจริง ๆ ตัวอย่างเช่น บางคนอาจจะมีคำถามว่า ถ้าแบบทดสอบวัดสำนวนภาษาอังกฤษในการพูดสามารถวัดความสามารถในการพูดภาษาอังกฤษของผู้เข้าสอบอย่างอ้อม ๆ ได้หรือไม่ ดังนั้น การวิจัยในอนาคต ควรทำการศึกษาความตรงร่วมสมัย (Concurrent Validation Study) ในแต่ละตอนย่อย เพื่อที่จะรู้เพิ่มเติมเกี่ยวกับปัญหาคุณภาพของข้อคำถามในแบบทดสอบ

ภาษาอังกฤษในแบบทดสอบ JFSAT และ 3) ในอนาคตควรสำรวจอีกครั้งว่ามีการขาดความสัมพันธ์ระหว่างแบบทดสอบวัดการออกเสียงที่ให้สอบโดยการเขียนและแบบทดสอบวัดความเข้าใจในการฟังของแบบทดสอบใดก็ได้ ถ้าพบว่าไม่มีความสัมพันธ์ระหว่างแบบทดสอบ 2 ชนิด เราควรสำรวจต่อว่าแบบทดสอบวัดความเข้าใจในการฟังชนิดใดที่ควรบรรจุไว้ในตอนย่อยของแบบทดสอบส่วนภาษาอังกฤษของแบบทดสอบ JFSAT บางทีประเด็นหนึ่งที่เราควรพิจารณาอาจเป็นจำนวนข้อในแบบทดสอบวัดความเข้าใจในการฟังที่จะทำให้แบบทดสอบใช้งานได้ดี โดยเป็นแบบทดสอบตอนย่อยที่ส่งผลต่อคะแนนรวมของแบบทดสอบอย่างมีนัยสำคัญ ดังนั้น คำถามสำหรับการวิจัยในอนาคตที่ผู้สนใจในวิจัยด้านนี้สามารถนำไปดำเนินต่อได้ คือ 1) มีความสัมพันธ์ที่ไม่มีระดับนัยสำคัญระหว่างแบบทดสอบวัดการออกเสียงแบบให้ทำโดยการเขียนและแบบทดสอบวัดความเข้าใจในการฟังหรือไม่ 2) ความตรงร่วมสมัยของแต่ละตอนย่อยของแบบทดสอบส่วนภาษาอังกฤษในแบบทดสอบ JFSAT สูงในระดับใด และ 3) ควรบรรจุแบบทดสอบวัดความเข้าใจในการฟังจำนวนกี่ข้อไว้ในแบบทดสอบส่วนภาษาอังกฤษของแบบทดสอบ JFSAT ในอนาคต

Pizorn & Moe (2012) ทำการวิจัยเรื่อง A Validation Study of the National Assessment Instrument for Young English Language Learners in Norway and Slovenia การวิจัยมีวัตถุประสงค์เพื่อหาคุณภาพแบบทดสอบภาษาต่างประเทศระดับชาติสำหรับผู้เรียนชั้นพื้นฐาน (11-12 ปี) ของ 2 ประเทศในยุโรป จำนวน 2 ฉบับ โดยการใช้คำแนะนำสำหรับแนวปฏิบัติที่ดีในการทดสอบและประเมินทางภาษา (The European Association for Language Testing and Assessment – EALTA) ของ Alderson and Banerjee (2008, อ้างถึงใน Pizorn & Moe, 2012) เป็นเครื่องมือในการประเมิน โดยต้องการจะแสดงให้เห็นว่า แบบทดสอบเติมเต็มการใช้ที่คาดหวังหรือเป้าหมายไว้ระดับใด และมีผลประโยชน์ต่อผู้เกี่ยวข้อง และกระบวนการเรียนการสอนภาษาต่างประเทศอย่างไร อีกทั้งมีข้อจำกัดอะไรที่จะต้องพิจารณาอย่างรอบคอบและเปิดเผยให้รับทราบ

วิธีการที่ใช้ คือ กระบวนการพัฒนาแบบทดสอบของแบบทดสอบภาษาอังกฤษระดับชาติของประเทศ Norway และ Slovenia ถูกตรวจสอบตามลักษณะ 3 ใน 7 ประการ ที่กำหนดไว้ในแนวปฏิบัติของ The EALTA Guidelines และถูกใช้แล้วในการศึกษาคุณภาพของแบบทดสอบ 2 ครั้ง ลักษณะ 3 ประการนั้น คือ 1) จุดประสงค์การสอบและโครงสร้าง (The Purpose and Specification) 2) การออกแบบแบบทดสอบและการเขียนแบบทดสอบ (Test Design and Item Writing) และ 3) การควบคุมคุณภาพและการวิเคราะห์แบบทดสอบ (Quality Control and Test Analyses) เกณฑ์ 3 ประการนี้ ได้รับการคัดเลือกเพราะว่ามีความสำคัญในการพัฒนาแบบทดสอบสำหรับผู้เรียนภาษาที่เป็นเด็กเล็ก แบบประเมินคุณภาพแบบทดสอบใน 3 ประการ ดังกล่าว ประเมินโดยการตอบคำถามแต่ละคำถาม และการพิจารณาแบบทดสอบแต่ละฉบับ ดังนี้

ด้านจุดประสงค์และโครงสร้างแบบทดสอบ พิจารณาว่า 1) มีการกำหนดจุดประสงค์การสอบไว้อย่างชัดเจนหรือไม่อย่างไร 2) มีการกล่าวถึงการใช้แบบทดสอบผิดพลาดไว้อย่างไร 3) มีการระบุถึงผู้เกี่ยวข้องทั้งหมดไว้โดยเฉพาะหรือไม่ 4) มีโครงสร้างแบบทดสอบหรือไม่ 5) มีการจัดทำโครงสร้างแบบทดสอบสำหรับกลุ่มเป้าหมายต่าง ๆ ไว้อย่างแตกต่างหรือไม่ 6) มีคำอธิบายเกี่ยวกับผู้เข้าสอบหรือไม่ 7) มีการระบุโครงสร้างที่จะวัดอยู่ภายใต้แบบทดสอบชุดใหญ่หรือแบบทดสอบย่อยหรือไม่ 8)

มีการอธิบายวิธีการทดสอบหรือภาระงานและยกตัวอย่างประกอบหรือไม่ 9) มีการบรรยายความหลากหลายในความสามารถของนักเรียนและยกตัวอย่างหรือไม่ 10) มีการบรรยายเกี่ยวกับแผนการตรวจให้คะแนนหรือเกณฑ์การตรวจไว้หรือไม่ และ 11) มีการระบุระดับของแบบทดสอบตามระดับของ CEFR หรือไม่ มีหลักฐานอะไรที่จะใช้สนับสนุนการกล่าวอ้างนี้

ด้านการออกแบบและการเขียนข้อคำถาม พิจารณาว่า 1) นักพัฒนาแบบทดสอบและผู้เขียนแบบทดสอบมีประสบการณ์การสอนที่เกี่ยวข้องตามระดับของแบบทดสอบที่พัฒนาหรือไม่ 2) นักพัฒนาแบบทดสอบและผู้เขียนแบบทดสอบได้รับการฝึกอบรมอะไรบ้าง 3) มีข้อเสนอแนะสำหรับการออกแบบแบบทดสอบและการเขียนแบบทดสอบหรือไม่ 4) มีกระบวนการที่เป็นระบบในการทบทวนปรับปรุงและแก้ไขข้อคำถามและภาระงาน เพื่อให้มั่นใจว่าแบบทดสอบสอดคล้องกับโครงสร้างแบบทดสอบ และการดำเนินงานตามข้อเสนอแนะในการเขียนแบบทดสอบ และ 5) มีข้อมูลป้อนกลับอะไรที่ผู้เขียนแบบทดสอบได้รับในงานของพวกเขา

การควบคุมคุณภาพและการวิเคราะห์แบบทดสอบ พิจารณาว่า 1) แบบทดสอบได้รับการทดลองใช้หรือไม่ ขนาดปกติของกลุ่มตัวอย่างมีจำนวนกี่คน และเทียบกับกลุ่มประชากรในการทดสอบอย่างไร 2) มีการดำเนินการในการเปลี่ยนแปลงแบบทดสอบหลังการวิเคราะห์หลักฐานที่เก็บได้ในชั้นทดลองอย่างไร 3) ถ้ามีแบบทดสอบชุดต่าง ๆ เกิดขึ้น (เช่น แบบทดสอบในแต่ละปี) มีการหาความเท่าเทียมกันของแบบทดสอบอย่างไร 4) วิธีการทางสถิติที่นำมาใช้ในการวิเคราะห์ข้อมูลคืออะไร และ 5) มีกระบวนการหรือขั้นตอนสำหรับผู้เข้าสอบที่ต้องการร้องเรียนหรือขอสอบใหม่อย่างไร

ผลการศึกษารูปได้ว่า ด้านจุดประสงค์การสอบ ดูเหมือนจะกำหนดไว้และเห็นได้อย่างชัดเจน และแสดงออนไลน์ไว้สำหรับผู้เกี่ยวข้องทุกคน แบบทดสอบทั้ง 2 ฉบับ มีเป้าหมายหลักเดียวกัน คือ การให้ข้อมูลเพิ่มเติมเกี่ยวกับผลสัมฤทธิ์ภาษาอังกฤษของนักเรียน ผู้เกี่ยวข้องในทั้งสองบริบทการวัดได้ถูกกำหนด และแจ้งให้ทราบถึงความรับผิดชอบ ผู้จัดสอบในทั้งสองประเทศใช้ความพยายามอย่างมากในการป้องกันความผิดพลาดในการใช้แบบทดสอบ และการตีความผิดในผลการสอบ โดยการรับลงทะเบียนเฉพาะนักเรียนในปีเดียวกัน ให้ข้อมูลอย่างละเอียดเกี่ยวกับแบบทดสอบ และให้ข้อมูลย้อนกลับไปยังนักเรียนและครู ด้านโครงสร้างแบบทดสอบ มีการเผยแพร่ในทั้ง 2 ประเทศ ในรูปแบบของเอกสารออนไลน์ ซึ่งเป็นเอกสารที่มีประโยชน์สำหรับนักออกแบบแบบทดสอบและผู้เกี่ยวข้องอื่น ๆ ในกระบวนการวัดผล มีการปรับเปลี่ยนโครงสร้างแบบทดสอบหลายครั้งตามข้อมูลป้อนกลับหลายแหล่ง ส่งผลให้มีการปรับและแก้ไขประจำปี ด้านความสามารถที่จะวัดของแบบทดสอบทั้ง 2 ฉบับ อยู่บนฐานของความคิดที่เชื่อมโยงภาษาเข้ากับบริบทที่มีการใช้และหน้าที่ในการสื่อสาร ดังนั้นจุดเน้นของภาระงานสอบจึงอยู่บนความสามารถใช้ภาษาเพื่อการสื่อสารของนักเรียน และการบรรยายเกี่ยวกับวิธีการสอบและการทำภาระงานที่ถูกต้องและสมบูรณ์มีไว้สำหรับแบบทดสอบทั้ง 2 ฉบับ และมีการเผยแพร่ภาระงานตัวอย่างในรูปแบบเอกสารออนไลน์

การออกแบบแบบทดสอบและกระบวนการในการควบคุมคุณภาพดูเหมือนว่าจะดำเนินการตามมาตรฐานขั้นสูงที่กำหนดไว้ ด้านการทดสอบและประเมินด้านภาษา มีทีมงานในการทดสอบที่เป็นผู้เชี่ยวชาญด้านการทดสอบและอาจารย์ผู้สอนผู้ที่สามารถเชื่อมโยงความจำเป็นและความสนใจของผู้เข้าสอบได้ ส่วนผู้เขียนแบบทดสอบ/ผู้ออกแบบแบบทดสอบทำงานเป็นทีม และเป็นผู้ที่สามารถอภิปรายแบบทดสอบและภาระงานทั้งหมดอย่างละเอียดก่อนการทดลองใช้

วิธีการทางสถิติหลายชนิดถูกนำมาใช้วิเคราะห์ข้อมูลของแบบทดสอบทั้ง 2 ฉบับ และนำเสนอผลการวิเคราะห์ต่อสาธารณชนและผู้เกี่ยวข้องทั้งหมด หรือผู้ที่สนใจในผลการทดสอบระดับชาติ สิ่งที่สำคัญและต้องปรับปรุงเพิ่มเติม คือ การเผยแพร่ผลของโรงเรียนทางสื่อเฉพาะและความคาดหวังสูงของครูใหญ่และผู้ปกครอง ผู้ที่เพิ่มความกดดันให้กับครูและนักเรียน และในการเปลี่ยนรูปแบบจากแบบทดสอบที่มีความสำคัญน้อยเป็นไปแบบทดสอบที่มีความสำคัญสูง เป้าหมายที่สำคัญของการเรียนการสอนทุกวันนี้คือเพื่อพัฒนาความสามารถทางภาษาของนักเรียน ทั้งด้านการฟัง พูด อ่าน และเขียน แต่ไม่มีแบบทดสอบชนิดใดเลยที่ทดสอบครบทั้ง 4 ทักษะ ดังนั้นแบบทดสอบจึงไม่อาจสื่อภาพรวมในความสามารถทางภาษาของนักเรียนได้ นอกจากนี้ รายงานผลการสอบที่นักเรียนได้รับในสัปดาห์เดียวไม่มีข้อมูลรายละเอียดตามโครงสร้างแบบทดสอบหรือทักษะภาษาที่ถูกประเมิน มีเฉพาะคะแนนรวมเท่านั้น ดังนั้นนักเรียนที่พูดได้ดี แต่อ่านหรือเขียนได้ไม่ดี หรือผู้ที่สามารถใช้โครงสร้างไวยากรณ์อย่างถูกต้องได้น้อย จะไม่สามารถแสดงความเก่งในทักษะภาษาที่สำคัญได้

ช่วงความสามารถด้านการเขียนของนักเรียนมิให้เห็นเฉพาะในสัปดาห์เดียว เนื่องจากแบบทดสอบของนอร์เวย์ไม่มีทักษะการส่งสาร (Productive Skills) มีการแจกเกณฑ์มาตรฐานของสัปดาห์เดียวให้กับครูทุกคนที่เข้าร่วมการประชุมมาตรฐานประจำปี และมีตัวอย่างงานเขียนที่เป็นตัวอย่างคำอธิบายตามเกณฑ์ที่แนบไปพร้อมกับเฉลยคำตอบ (Answer Key) ซึ่งแจกให้ครูทุกคนในสัปดาห์เดียว อย่างไรก็ตาม การมาประชุมเพื่อกำหนดมาตรฐานเป็นความสมัครใจของครู ซึ่งอาจจะมีผลต่อผลของความเชื่อมั่นแบบ Inter-and intra-rater ได้

มีการเทียบมาตรฐานของแบบทดสอบระดับชาติทั้ง 2 ฉบับ กับ CEFR แล้ว ตามกระบวนการที่ชัดเจนที่กำหนดไว้ในคู่มือการเทียบแบบทดสอบทางภาษากับเกณฑ์มาตรฐานการเรียนรู้ทางภาษาของสภายุโรป แต่เนื่องจากยังไม่มีข้อกำหนดคะแนนจุดตัดสำหรับแบบทดสอบปัจจุบัน ผลการสอบของนักเรียนจึงไม่สามารถรายงานโดยใช้ระดับการอ้างอิงของ CEFR ได้ ในนอร์เวย์ การทดลองใช้แบบทดสอบเกี่ยวข้องกับการใช้กลุ่มตัวอย่างจำนวนมากขณะที่กรณีของสัปดาห์เดียวควรใช้กลุ่มทดลองที่สามารถเป็นตัวแทนของนักเรียนที่เป็นกลุ่มประชากร สิ่งนี้อาจเป็นสาเหตุหนึ่งที่ทำให้ค่าดัชนีความยากเปลี่ยนไปทุกปี

การศึกษาคุณภาพแบบทดสอบแสดงว่ามีหลายประเด็นในการพัฒนาและบริหาร และการใช้แบบทดสอบระดับชาติที่ควรศึกษาและปรับปรุงอย่างแท้จริง ข้อเสนอแนะสำหรับนักออกแบบแบบทดสอบในสัปดาห์เดียวให้หาวิธีการที่เหมาะสมที่สุดในการทดลองใช้แบบทดสอบกับกลุ่มตัวอย่างขนาดใหญ่ของประชากรเป้าหมาย ซึ่งสามารถทำให้สำเร็จโดยสร้างระบบการทดลองใช้แบบทดสอบภายในศูนย์การทดสอบแห่งชาติหรือโดยการเปลี่ยนการทดสอบรูปแบบกระดาษเป็นการทดสอบโดยใช้คอมพิวเตอร์

ข้อเสนอแนะอีกประการหนึ่งคือ เกี่ยวกับผู้เกี่ยวข้องทั้งหมดควรตระหนักว่าการเผยแพร่ผลการสอบอย่างเปิดเผยไม่ได้ตอบสนองความต้องการของนักเรียน และแบบทดสอบจำเป็นต้องเทียบกับ CEFR ทุกปี และการเปลี่ยนจากการวัดแบบอิงกลุ่ม ถ้าระบบการสอบแบบอิงเกณฑ์ได้รับการยอมรับโดยผู้เกี่ยวข้องในกระบวนการเรียนการสอนภาษา นักเรียนจะเรียนรู้ที่จะแข่งกับเกณฑ์ (ระดับของ CEFR) และมีการแข่งขันกันระหว่างนักเรียนและ/หรือผู้ปกครองน้อยลง แบบทดสอบระดับชาติทั้ง 2

ฉบับ ควรพัฒนาภาระงานการพูดที่สามารถให้ครูพร้อมกับเกณฑ์มาตรฐานและข้อมูลย้อนกลับที่เหมาะสมสำหรับนักเรียนและผู้ปกครอง ในนอร์เวย์ ภาระงานการเขียนที่พัฒนาอย่างรอบคอบ และเกณฑ์มาตรฐานจะเป็นประโยชน์สำหรับครู นักเรียนและผู้ปกครอง การประชุมเพื่อกำหนดมาตรฐาน ควรถูกจัดและเป็นการบังคับสำหรับครูทั้งหมด แบบทดสอบภาษาอังกฤษระดับชาติสำหรับนักเรียน ป.5 ในนอร์เวย์ การประเมินทักษะการฟังของนักเรียน และครอบคลุมเป้าหมายของหลักสูตรมากกว่านี้

Cheng (2013) ทำการศึกษาเรื่อง Correlating the Gates-MacGinitie Reading Tests with the University Entrance English Examinations Held in Taiwan โดยมีจุดประสงค์เพื่อปูทางสำหรับการศึกษากลุ่มใหญ่ในการหาคุณภาพและสร้างภาพประสิทธิภาพของแบบทดสอบด้านภาษาที่พัฒนาขึ้นในประเทศไต้หวันและเอเชีย กลุ่มตัวอย่างเป็นนักศึกษาชั้นปีที่ 1 จำนวน 224 คน ที่มหาวิทยาลัยทางการแพทย์ในตอนกลางของประเทศไต้หวัน นักศึกษาเหล่านี้มาจากห้องเรียนชั้นปีที่ 1 จำนวน 5 ห้องเรียน (จำนวนรวม คือ 242 คน) ที่สอนโดยผู้วิจัยในภาคเรียนที่ 1 ภาษาอังกฤษสำหรับนักศึกษาปีที่ 1 ในระดับมหาวิทยาลัยเป็นวิชาบังคับ แต่นักศึกษาสามารถที่จะเลือกอาจารย์หรือช่วงเวลาที่เหมาะสมที่สุดกับตารางเรียนของตัวเอง กลุ่มตัวอย่างเหล่านี้ถูกกำหนดรหัสเป็น 2 กลุ่ม คือ GSAT group และ AST group. The AST group ประกอบด้วยนักศึกษา 53 คน คือผู้ที่เข้ามายังมหาวิทยาลัยโดยผ่านแบบทดสอบ AST Exam ที่จัดสอบในเดือนกรกฎาคม 2012 กลุ่ม GSAT group มีนักศึกษา 171 คน เป็นที่เข้ามหาวิทยาลัยโดยแบบทดสอบ GSAT Exam จัดขึ้นในเดือนมกราคม 2012 การคัดเลือกนักศึกษาสำหรับการศึกษาคั้งนี้ไม่รวมนักศึกษาชาวจีนต่างประเทศ นักเรียนต่างชาติ และนักเรียนที่เข้ามหาวิทยาลัยโดยไม่ได้สอบแบบทดสอบ the AST Exam หรือ the GSAT Exam

ในการศึกษาปัจจุบัน ระดับ 7/9 Form S ของ the GMRT ถูกใช้เป็นเครื่องมือ ค่าสัมประสิทธิ์ความเชื่อมั่น อยู่ระหว่าง 0.94 ของเกรด 7 ถึง 0.95 ของเกรด 9 การศึกษาในการเทียบเกรด 7/9 แสดงค่าความสัมพันธ์ 0.90 สำหรับนักเรียนเกรด 7 ผู้ที่สอบทั้งระดับ 6 และระดับ 7/9 ค่าความสัมพันธ์ 0.90 สำหรับนักเรียนเกรด 8 ผู้ที่สอบทั้ง 2 ระดับ ในระดับ 6 และระดับ 7/9 ค่าความสัมพันธ์ 0.90 สำหรับนักเรียนเกรด 9 ผู้ที่สอบทั้งระดับ 6 และระดับ 7/9 และค่าความสัมพันธ์ 0.87 สำหรับนักเรียนเกรด 10 ผู้ที่สอบทั้งสองระดับในเกรด 7/9 และระดับ 10/12 ระดับ 7/9 ดูเหมือนจะเหมาะสมสำหรับการศึกษาคั้งนี้ตามระดับอายุและความสามารถของสาระที่ทำให้เหมาะสมกับนักเรียนเกรด 5 ถึงเกรด 12.9 ในการศึกษาครั้งนี้ the GMRT ถูกจัดสอบให้กับนักเรียนในเวลาเรียนปกติของโรงเรียนในสัปดาห์แรกของภาคเรียนที่ 1 ก่อนการสอน นักศึกษาได้รับการแจ้งเกี่ยวกับจุดประสงค์ของการสอบและบอกว่าคะแนนไม่เกี่ยวข้องใด ๆ กับการรายงานผลการเรียนปลายปี นักเรียนได้รับการกระตุ้นให้ทำดีที่สุด โดยให้ทางเลือกที่จะไม่สอบถ้ารู้สึกไม่สะดวกใจ คำตอบถูกตรวจด้วยมือโดยอาจารย์ 2 คน แต่ละข้อมีค่า 1 คะแนน คะแนนเต็มของแบบทดสอบ คือ 93 คะแนน

ผลการศึกษาและการอภิปรายผล พบว่า จากข้อมูลในสถิติเชิงบรรยายของกลุ่ม AST Exam group ค่าเฉลี่ยของคะแนนภาษาอังกฤษ AST Exam คือ 68.245 (ค่าต่ำสุด = 34, ค่าสูงสุด = 88.5, SD = 12.69) ค่าเฉลี่ยคะแนนดิบของ GMRT คือ 44.226 (ค่าต่ำสุด = 22, ค่าสูงสุด = 70, SD =

10.852) ค่าสัมประสิทธิ์สหสัมพันธ์ของ Pearson ถูกคำนวณเพื่อประเมินความสัมพันธ์ระหว่าง AST Exam English score และ the GMRT score มีความสัมพันธ์กันในเชิงบวกระหว่างตัวแปรทั้งสอง $r = 0.801$, $n = 53$, $p = 0.000$ ในภาพรวมมีความสัมพันธ์ในระดับสูงเชิงบวกระหว่างแบบทดสอบ AST Exam English และ the GMRT

คะแนน GSAT Exam และ the GMRT score ข้อมูลเชิงบรรยายเกี่ยวกับข้อมูลของกลุ่ม GSAT Exam group ระบุว่า จากคะแนนอันดับจาก 1 ถึง 15 คะแนนเฉลี่ยของกลุ่ม GSAT คือ 12.216 (ค่าต่ำสุด คือ 09, ค่าสูงสุด คือ 15, $SD = 01.606$) คะแนนเฉลี่ยคะแนนระดับของคะแนน GMRT คือ 41.55 (ค่าต่ำสุด 23, ค่าสูงสุด 75, $SD = 09.686$) ค่าสัมประสิทธิ์สหสัมพันธ์ถูกคำนวณเพื่อประเมินความสัมพันธ์ระหว่างคะแนนแบบทดสอบ GSAT Exam English และ the GMRT พบว่ามีความสัมพันธ์กันในเชิงบวกระหว่าง 2 ตัวแปร ค่า $r = 0.637$, $n = 171$, $p = 0.000$ สรุปว่าในภาพรวมมีความสัมพันธ์ในเชิงบวกและมีนัยสำคัญระหว่างแบบทดสอบ GSAT Exam English และ the GMRT

การศึกษาแสดงว่า แบบทดสอบภาษาอังกฤษที่พัฒนาขึ้นในไต้หวันสัมพันธ์ระดับสูงกับแบบทดสอบการอ่านที่จัดกลุ่มผู้เรียนมาตรฐานของอเมริกันที่ได้รับความนิยม The GMRT ยังแสดงให้เห็นว่าแบบทดสอบไม่เพียงแต่จัดกลุ่มผู้เรียน แต่ยังช่วยให้การวางแผนจุดเน้นในการเรียนการสอนอีกด้วย ในการจัดกลุ่มนักเรียนที่พร้อมที่จะเรียนในสิ่งที่สูงกว่า และในการเลือกระดับที่เหมาะสมของการสอนให้เหมาะกับนักเรียน เนื่องจาก AST และ GSAT ให้เฉพาะคะแนนหรืออันดับ มหาวิทยาลัยในไต้หวันมีแนวโน้มที่จะสอบนักเรียนใหม่เพื่อจัดกลุ่มความสามารถโดยไม่ได้ตรวจสอบคุณค่าของแบบทดสอบในการส่งเสริมการสอนภาษาอังกฤษสำหรับนักศึกษาใหม่ แบบทดสอบการอ่านที่ดี ควรให้ข้อมูลที่ช่วยให้ครูจัดระดับการอ่านของนักเรียนของตนเองได้ การเรียนรู้จะมีความหมายถ้าความสามารถในการอ่านของนักเรียนสอดคล้องกับความสามารถในการอ่านได้ในหนังสือที่นักเรียนถูกกำหนดให้อ่าน เนื่องจากครูสอนภาษาอังกฤษส่วนมากเรียนวิชาเอกภาษาศาสตร์ภาษาอังกฤษ การสอนภาษาอังกฤษในฐานะภาษาต่างประเทศ และวรรณกรรมภาษาอังกฤษ แบบทดสอบที่ครูใช้บอกระดับภาษาอังกฤษของนักเรียนจึงเป็นแบบทดสอบวัดความสามารถภาษาอังกฤษ เช่น TOEFL, IELTS, GEPT (General English Proficiency Test) ครูไม่มีความรู้เพียงพอเกี่ยวกับหน้าที่ของแบบทดสอบการอ่านที่พัฒนาสำหรับชาวอเมริกันในโรงเรียนเกรดต่างๆ ในอเมริกา ครูควรได้รับความรู้เกี่ยวกับหน้าที่แบบทดสอบภาษาอังกฤษอื่นๆ และวิธีการที่สามารถประยุกต์ใช้ในการสอนของตน และการเรียนรู้ของนักเรียนของตนเอง

Wu (2012) ทำการศึกษาเปรียบเทียบแบบทดสอบ PETST (The Public English Test System) ที่ใช้สอบในจีนแผ่นดินใหญ่ และแบบทดสอบ GEPT (General English Proficiency Test) ใช้สอบในประเทศไต้หวัน โดยใช้วิธีการวิเคราะห์สาระของแบบทดสอบทั้ง 2 ชนิด ในด้าน 1) ระดับของแบบทดสอบ (Test Levels) 2) สาระแบบทดสอบ (Test Contents) และ 3) น้ำหนักของการให้คะแนน (Scoring Weight) เพื่อค้นหาความเหมือนและความต่างของแบบทดสอบ และทฤษฎีที่อยู่เบื้องหลังการออกแบบ สรุปผลได้ว่า แบบทดสอบ GEPT และแบบทดสอบ PETS สามารถเทียบกันได้ ในระดับสูง ทั้ง 2 ชนิดวัดความสามารถภาษาอังกฤษในการสื่อสารโดยทั่วไป ทั้ง 2 ชนิดมี 5 ระดับจากระดับพื้นฐาน (Elementary Level) ไปจนถึงระดับสูง (Superior Level) และทดสอบไม่เฉพาะ

ทักษะรับสาร (Receptive Skills) คือการอ่านและการฟัง แต่ยังทดสอบทักษะส่งสาร (Productive Skills) คือทักษะการพูดและการเขียน

อย่างไรก็ตาม เมื่อศึกษารายละเอียด พบความแตกต่างระหว่างแบบทดสอบ 2 ชนิดนี้ คือ ประการแรก แบบทดสอบ GEPT มีความสามารถมากกว่าในการจำแนกผู้เรียนระดับต้น (Beginner) และระดับกลาง (Intermediate) ออกจากผู้เรียนในระดับสูง (Advanced) และมีความหลากหลายในความยากของแบบทดสอบ ระหว่าง 5 ระดับใน GEPT กว่าในแบบทดสอบ PETS แม้ว่า 3 ระดับแรกของทั้ง 2 แบบทดสอบ วัดความรู้ที่เท่ากันจากเนื้อหาของนักเรียนระดับต้นกับระดับมหาวิทยาลัย สองระดับสุดท้ายของ GEPT ยากกว่าสองระดับหลังของ PETS มาก ประการที่สอง เมื่อเทียบกับ PETS แบบทดสอบ GEPT ให้ความสำคัญอย่างมากกับทักษะผลิตสาร (Productive Skills) คือการพูดและการเขียน และสิ่งที่ขาดหายไปคือการสอนไวยากรณ์แบบดั้งเดิม จุดเน้นในทักษะการพูดและการเขียนของแบบทดสอบ GEPT สามารถพิสูจน์ได้จาก 1) ระยะเวลาในการสอบที่นานกว่าที่ให้กับภาระงานการสอบพูดและเขียน และ 2) ข้อกำหนดที่สูงกว่าสำหรับความสามารถของผู้เข้าสอบในภาระงานการพูดและการเขียน สิ่งที่แตกต่างกันจากแบบทดสอบ GEPT คือ แบบทดสอบ PETS มีการเปลี่ยนน้ำหนักคะแนนสำหรับ 3 ทักษะภาษาอังกฤษ เช่น การฟัง การอ่าน และการเขียน ในทั้ง 5 ระดับ เพื่อเน้นทักษะการผลิตสารที่ระดับสูงกว่าของแบบทดสอบ PETS

โดยทั่วไปแล้ว แบบทดสอบ GEPT สะท้อนทฤษฎีการออกแบบแบบทดสอบโดยอาศัยวิธีการเรียนรู้ภาษาอังกฤษแบบสื่อสาร ขณะที่แบบทดสอบ PETS สะท้อนผลของการสอนไวยากรณ์แบบดั้งเดิม การเปรียบเทียบกันได้ง่ายกว่าแบบทดสอบ PETS แบบทดสอบ 3 ระดับแรกของ GEPT ทำหน้าที่เป็นแบบทดสอบที่ดีเยี่ยมในการกระตุ้นผู้เรียนภาษาอังกฤษขั้นต้นให้เรียนภาษาอังกฤษของเขาต่อไป อย่างไรก็ตาม แบบทดสอบ 2 ระดับสุดท้ายของ GEPT ซึ่งต้องการทักษะการเขียนและการพูดที่ใกล้เคียงเจ้าของภาษา แยกผู้เรียนระดับสูงออกจากผู้เรียนขั้นต้น แบบทดสอบ PETS มีความสามารถน้อยในการแยกผู้เรียนที่มีทักษะภาษาระดับกลางจากผู้เรียนที่มีทักษะภาษาระดับสูง อย่างไรก็ตาม ด้วยรูปแบบแบบทดสอบที่คงที่ในระหว่าง 5 ระดับ และสัดส่วนที่มากกว่าของแบบทดสอบปรนัย แบบทดสอบ PETS จะให้คะแนน และมีประสิทธิภาพมากกว่าในการใช้สอบกับประชากรที่มากกว่าของผู้สอบชาวจีน

LCRT Consulting (2013) ศึกษาความสัมพันธ์ของแบบทดสอบ MELA (The Michener English Language Assessment) กับแบบทดสอบ IELTS (International English Language Testing System) โดยมีวัตถุประสงค์เพื่อสร้างข้อโต้แย้งด้านความตรงที่สนับสนุนการใช้แบบทดสอบ MELA เป็นแบบทดสอบวัดความสามารถภาษาอังกฤษสำหรับการรับใบอนุญาตการประกอบอาชีพและการรับเข้าศึกษาต่อหลังจบชั้นมัธยมศึกษา นอกจากการหาความสัมพันธ์ของคะแนนระหว่างของแบบทดสอบ MELA และแบบทดสอบ IELTS ยังมีการเก็บข้อมูลที่เป็นหลักฐานความสัมพันธ์เพิ่มเติมจากความสัมพันธ์ระหว่างการประเมินตนเองของผู้เข้าสอบกับคะแนน หลักฐานอื่น ๆ ที่เกี่ยวข้องกับวิเคราะห์แบบทดสอบรายข้อและการวิเคราะห์ความเชื่อมั่นซึ่งถูกดำเนินการในแบบทดสอบ MELA แต่ละข้อและภาระงาน และสุดท้ายข้อมูลจากแบบสอบถามที่ใช้เก็บข้อมูลเพื่อศึกษาหลักฐานความตรงด้านเนื้อหาที่ให้โดยผู้เข้าร่วม

การเก็บข้อมูลดำเนินการโดย 1) จัดหมายเชิญถูกพิมพ์บนเว็บไซต์พร้อมด้วยคำแนะนำเกี่ยวกับวิธีลงทะเบียนสำหรับการศึกษา 2) ผู้เข้าสอบได้รับการสมัครเพื่อเข้าร่วมการศึกษา ผู้เข้าร่วมถูกรับเข้ามาโดยผ่านหลักสูตรที่ให้บริการสำหรับผู้เชี่ยวชาญด้านสุขภาพที่ได้รับการฝึกฝนในระดับนานาชาติ 3) ผู้เข้าร่วมลงลายมือชื่อในแบบฟอร์มอนุญาตในสิทธิต่าง ๆ 4) ผู้เข้าร่วมแต่ละคนทำแบบทดสอบ IELTS และ MELA (คนละวันกัน) และ 5) หลังจากทำแบบทดสอบทั้ง 2 ชุด ผู้เข้าสอบแต่ละคนทำแบบสอบถามเกี่ยวกับประวัติของตนเอง การประเมินตนเองเกี่ยวกับความสามารถในทักษะภาษาทั้ง 4 ด้าน และความรู้สึกของพวกเขาต่อแบบทดสอบ MELA

การวิเคราะห์ข้อมูล นำคะแนนจากแบบทดสอบ MELA และ IELTS พร้อมทั้งคำตอบต่าง ๆ ในการสอบถามได้รับการวิเคราะห์โดยวิธีการต่าง ๆ ต่อไปนี้เป็นรายการของการวิเคราะห์ที่อธิบายไว้ในรายงานฉบับนี้ 1) สถิติเชิงพรรณนา สำหรับ แบบทดสอบ MELA และ IELTS และข้อมูลจากการสอบถาม (ทั้งการประเมินตนเอง และความรู้สึกต่อแบบทดสอบ MELA) 2) ความสัมพันธ์ระหว่างคะแนนสอบ MELA อีกฝ่ายหนึ่ง และคะแนนสอบ IELTS และผลจากการประเมินตนเอง อีกฝ่ายหนึ่ง 3) การเปรียบเทียบคะแนน MELA ตามระดับต่าง ๆ ของคะแนนแบบทดสอบ IELTS 4) การเปรียบเทียบคะแนน MELA ในกลุ่มต่าง ๆ ของนักเรียน กลุ่มของผู้เข้าสอบแยกตามปัจจัยภูมิหลังของผู้เข้าสอบ 5) สถิติเชิงพรรณนาสำหรับข้อมูลในแบบสอบถามที่เกี่ยวข้องกับความรู้ของผู้เข้าสอบต่อ MELA และ 6) การวิเคราะห์รายข้อและความเชื่อมั่นของแบบทดสอบแต่ละข้อและภาระงานต่างๆ ในแบบทดสอบ MELA

ผู้เข้าร่วม (Participants) เป็นนักศึกษาจำนวน 56 คน เข้าร่วมในการศึกษา จำนวน 55 คน เข้ารับการอบรมในหลักสูตร An Ontario Bridging Program กลุ่มผู้เข้าสอบเป็นเพศหญิง 36 คน (64%) และเพศชาย 20 คน (36%) ส่วนใหญ่ คือ 38 คน (68%) มีอายุระหว่าง 26 ถึง 40 ปี เป็นผู้พูดเจ้าของภาษา 21 ภาษา และมาจากประเทศต่างๆ 25 ประเทศ ส่วนที่มากที่สุดคือ 40 คน (71%) เรียนจบมหาวิทยาลัย แต่มีเพียง 18 คน (32%) ได้รับการศึกษาโดยใช้ภาษาอังกฤษ ขณะที่ผู้เข้าร่วมการศึกษาที่เหลือได้รับการอบรมโดยใช้ภาษาของตนเอง กลุ่มตัวอย่างเรียนภาษาอังกฤษมาแล้วระหว่าง 1 ถึง 25 ปี (ค่าเฉลี่ย 9 ปี, S.D. = 6.5 ปี) และได้อยู่ในประเทศแคนาดาระหว่าง 1 ถึง 6 ปี (ค่าเฉลี่ย 2.5 ปี, S.D. = 8 ปี) ในขณะช่วงเวลาของการศึกษา มีเพียง 9 คน (16%) ลงทะเบียนในหลักสูตรของ the Michener Institute ในช่วงเวลาเก็บข้อมูล ในด้านกระบวนการลงทะเบียนด้านอาชีพในแคนาดาในช่วงเวลาของการเก็บข้อมูลมีจำนวน 15 คน (27%) ได้ลงทะเบียนตามระเบียบของมหาวิทยาลัย จำนวน 10 คน (19%) บอกว่าพวกเขาได้พบข้อกำหนดด้านความคล่องแคล่วทางภาษาตามระเบียบของมหาวิทยาลัย จำนวน 2 คน (4%) บอกว่าพวกเขาได้ผ่านการสอบเพื่อรับใบอนุญาตในการประกอบอาชีพแล้ว และจำนวน 31 คน (55%) บอกว่าพวกเขาอยู่ในขั้นตอนอื่นๆ ของกระบวนการ

ผลการศึกษา พบว่า ผลจากแบบทดสอบ MELA ในภาพรวม ผู้เข้าสอบได้รับคะแนนสูงกว่าค่าเฉลี่ยในทุกภาระงานและทุกตอนของแบบทดสอบ MELA ตัวอย่างเช่น ผู้เข้าร่วมได้รับคะแนนเฉลี่ย 8.9 (จากคะแนนเต็ม 15) ในตอนการอ่านและได้รับคะแนนเฉลี่ย 9 (จากคะแนนเต็ม 15) ในส่วนการเขียน เป็นต้น และผลจากการสอบ IELTS ผู้เข้าสอบได้รับคะแนนเฉลี่ยประมาณ 6 (จากเต็ม 9) ในตอนต่างๆ ของแบบทดสอบ IELTS ส่วนใหญ่ของผู้เข้าร่วมได้รับคะแนนที่ระดับ 5 ถึง 7 ตัวอย่างเช่น

ค่าเฉลี่ยของคะแนนรวมในแบบทดสอบ IELTS คือ 6.05 (S.D. = .82) แต่ส่วนใหญ่จำนวน 49 คน ได้รับคะแนนรวมอยู่ระหว่าง 5 ถึง 7 ผลเหล่านี้เสนอแนะว่า ผู้เข้าร่วมในการศึกษาครั้งนี้ความสามารถภาษาอังกฤษเหนือหรือมากกว่าค่าเฉลี่ยเล็กน้อย

ความสัมพันธ์ระหว่างคะแนน MELA และ IELTS จากการวิเคราะห์ความสัมพันธ์ระหว่างคะแนนแบบทดสอบ MELA และ IELTS ในแต่ละตอนและคะแนนรวม พบว่า ผู้เข้าร่วมที่ได้คะแนนสูงในแบบทดสอบ IELTS ได้รับคะแนนสูงในแบบทดสอบ MELA ด้วย (และในทางกลับกัน) สำหรับแต่ละทักษะภาษา ตัวอย่างเช่น ความสัมพันธ์ระหว่างคะแนนสอบ MELA ในส่วนการฟัง และคะแนนเต็มของแบบทดสอบ IELTS คือ .78 ความสัมพันธ์ในทักษะภาษาต่าง ๆ อยู่ระหว่าง .67 (สำหรับการพูด) และ .78 (สำหรับการฟัง) ในภาพรวม กลุ่มตัวอย่างที่ได้รับคะแนนสูงในแบบทดสอบ IELTS ได้คะแนนสูงในแบบทดสอบ MELA ด้วย ความสัมพันธ์ของคะแนนในเชิงบวก รวมทั้ง MELA และ IELTS ทำให้ได้หลักฐานด้านความตรงของแบบทดสอบ MELA

ความสัมพันธ์ระหว่างการประเมินตนเองและคะแนนสอบ สเกลในการประเมินตนเองในแบบสอบถาม ประกอบด้วย 25 ข้อ ข้อคำถาม 5 ข้อ เป็นแบบทั่วไป และ 20 ข้อ เป็นแบบเฉพาะด้าน จำนวน 20 ข้อ มีค่าความเชื่อมั่นคงที่ภายใน (An Internal Consistency Reliability) Cronbach alpha เท่ากับ .94 ซึ่งบ่งบอกว่า ผลลัพธ์เหล่านี้มีความสัมพันธ์กัน (วัดโครงสร้างเดียวกัน) ข้อคำถาม 20 ข้อ แบ่งได้ คือ วัดการฟัง 7 ข้อ การอ่าน 6 ข้อ การเขียน 2 ข้อ และการพูด 5 ข้อ

จากข้อมูลในสถิติเชิงบรรยาย พบว่า กลุ่มตัวอย่างในภาพรวมประเมินตนเองสูงใน 4 ทักษะ ตัวอย่างเช่น กลุ่มตัวอย่างให้คะแนนเฉลี่ย 3.5 (จากคะแนนเต็ม 4) ในการประเมินตนเองด้านทักษะการฟังภาษาอังกฤษ ในด้านความสัมพันธ์ระหว่างคะแนนสอบ IELTS และประเมินตนเองพบว่า ค่าความสัมพันธ์โดยทั่วไปอยู่ในระดับต่ำ ส่วนความสัมพันธ์ระหว่างคะแนนสอบ MELA และการประเมินตนเอง พบว่า ค่าความสัมพันธ์อยู่ในระดับปานกลาง และมีรูปแบบเดียวกันเหมือนกับของแบบทดสอบ IELTS และกาประเมินตนเอง โดยทั่วไปแล้ว ผู้ที่รู้สึกว่าคุณมีความสามารถด้านภาษาอังกฤษสูงกว่าได้รับคะแนนจากแบบทดสอบ MELA (และ IELTS) สูง เนื่องจากการประเมินตนเองของกลุ่มตัวอย่างและคะแนนสอบสัมพันธ์แบบคล้าย ๆ กันทั้งในแบบทดสอบ IELTS และ MELA หลักฐานที่จะบอกว่า MELA เป็นเครื่องวัดที่ตรงจึงได้รับการเชื่อถือ

ความรู้สึกรู้สึก/ความคิดเห็นของผู้เข้าสอบต่อ MELA ข้อมูลจากการวิเคราะห์แบบสอบถามความรู้สึกของผู้เข้าสอบต่อแบบทดสอบ MELA พบว่า ผู้เข้าสอบรู้สึกว่าคุณ MELA เป็นเครื่องมือที่ยุติธรรมของความสามารถทางภาษาอังกฤษของพวกเขา และพบว่าภาระงานในแบบทดสอบ MELA สัมพันธ์กับบริบทการใช้ภาษาเป้าหมายของพวกเขา (การทำงานด้านบริการสุขภาพ) ข้อมูลนี้ใช้หลักฐานด้านความตรงด้านเนื้อหาของภาระงานในแบบทดสอบ

ความเชื่อมั่นที่เกิดจากความคงที่ภายใน ความเชื่อมั่นคือความคงที่ของคะแนนในแต่ละภาระงานและผู้ให้คะแนน ค่าสัมประสิทธิ์ความเชื่อมั่นสำหรับแบบทดสอบแต่ละข้อสูงบ่งบอกว่าแบบทดสอบมีความคงที่ ขณะที่ความเชื่อมั่นระหว่างผู้ตรวจให้คะแนนที่สูงกว่าจะบอกว่าผู้เข้าสอบมีแนวโน้มจะได้คะแนนในอันดับเดียวกันไม่ว่าใครจะตรวจหรือให้คะแนนความสามารถของพวกเขา และ ค่าความเชื่อมั่น .07 หรือสูงกว่าเป็นค่าที่ถูกต้องคาดหวัง ค่าความเชื่อมั่นด้านความคงที่ภายใน (Internal Consistency Reliability) สูงในทุกตอนการฟัง ($r = .89$) การอ่าน ($r = .85$) การเขียน (r

= .88) และการพูด ($r = .94$) เป็นหลักฐานที่แสดงว่า แบบทดสอบ MELA มีความตรงด้านความคงที่ภายใน (Internal- Consistency Validity)

ค่าความเชื่อมั่นจากการให้คะแนนของผู้ตรวจ สำหรับภาระงานการพูดและการเขียน ถูกตรวจให้คะแนนโดยผู้ตรวจ 3 คน วิเคราะห์โดยใช้ Spearman rho correlations (เป็นเครื่องวัดความสอดคล้องในการให้คะแนนของกรรมการ) ค่าความเชื่อมั่น .80 หรือสูงกว่า เป็นค่าที่คาดหวัง ความเชื่อมั่นจากการตรวจของกรรมการสำหรับแบบทดสอบการพูดและการเขียน โดยทั่วไปอยู่ในระดับสูง ค่าความสัมพันธ์ระหว่างการให้คะแนนของผู้ให้คะแนนแต่ละคู่อยู่ระหว่าง .73 ถึง .88 สำหรับการพูด และระหว่าง .67 ถึง .94 สำหรับส่วนการเขียน ข้อมูลนี้แสดงว่า MELA มีหลักฐานสำหรับค่าความเชื่อมั่นระหว่างผู้ตรวจ (Inter-rater reliability) สำหรับแบบทดสอบในส่วนของ การพูดและการเขียน

การวิเคราะห์แบบทดสอบรายข้อ (Item Analysis) จากค่าความเชื่อมั่นด้านความคงที่ภายในที่ระบุว่า แต่ละตอนของแบบทดสอบ MELA (การฟัง การพูด การอ่าน และการเขียน) มีค่าความเชื่อมั่นด้านความคงที่ภายในที่สามารถยอมรับได้ (ค่า alpha มากกว่า .70) ข้อมูลนี้สื่อความว่าแบบทดสอบทุกข้อในแต่ละตอนของแบบทดสอบ MELA ผันแปรไปด้วยกันในลักษณะที่เชื่อถือได้หรือคงที่ ภายในแต่ละตอนมีการวิเคราะห์เฉพาะด้าน ได้แก่ ค่าความยากรายข้อ (Item Facility) ค่าดัชนีอำนาจจำแนก (Item Discrimination) การวิเคราะห์ตัวลวง (Distractor Analysis) และค่า alpha ถ้าแบบทดสอบถูกตัดทิ้ง ผลจากการวิเคราะห์นี้ แบบทดสอบบางข้อได้รับการปรับแก้ และใส่รวมกลับไปแบบทดสอบฉบับล่าสุด อย่างไรก็ตาม แม้จะไม่มีการแก้ไขเหล่านี้ การวิเคราะห์ข้อมูลแสดงให้เห็นว่าทุกภาระงานในแต่ละตอนของแบบทดสอบ MELA ส่งผลให้คะแนนคงที่และน่าเชื่อถือไปด้วยกัน

Riazi and Mehdi (2014) ทำการศึกษาเรื่อง Concurrent and Predictive Validity of Pearson Test of Academic (PTE Academic) โดยมีเป้าหมายหลัก 2 ประการ คือ เพื่อศึกษาความตรงร่วมสมัยของแบบทดสอบ PTE Academic และเพื่อศึกษาความตรงเชิงพยากรณ์ โดยเป้าหมาย 2 ประการนี้สามารถแปลงเป็นคำถามการวิจัย 3 ข้อ คือ 1) มีความสัมพันธ์ใดบ้างระหว่างคะแนนจากแบบทดสอบ PTE Academic และคะแนนจากแบบทดสอบ IELTS Academic ของกลุ่มตัวอย่าง ในประเด็นของคะแนนรวมและคะแนนในแต่ละส่วนของการฟัง การอ่าน การพูดและการเขียน 2) แบบทดสอบ PTE Academic สามารถแสดงให้เห็นความแตกต่างระหว่างกลุ่มความสามารถ 2 กลุ่ม (กลุ่มสูงกว่าและต่ำกว่า) ตามผลที่ได้จากคะแนนจากแบบทดสอบ IELTS Academic ของกลุ่มตัวอย่างหรือไม่ และ 3) มีความสัมพันธ์ระหว่างคะแนนของแบบทดสอบ PTE Academic ของผู้เข้าร่วมและผลสัมฤทธิ์ทางวิชาการของพวกเขาที่วัดโดยเกรดเฉลี่ย (GPA) ในภาคเรียนที่ 1 และตามที่เทียบกับกลุ่มปกติที่เป็นผู้พูดเจ้าของภาษา โดยแยกออกเป็น 3 ประเด็นย่อย คือ ก) มีความสัมพันธ์อย่างมีนัยสำคัญระหว่างคะแนนจากแบบทดสอบ PTE Academic ของกลุ่มตัวอย่าง และเกรดเฉลี่ยในภาคเรียนแรกของพวกเขาในแง่ของคะแนนรวมและคะแนนในทักษะการสื่อสารทั้ง 4 หรือไม่ ข) ร้อยละของผู้เข้าร่วมที่มีคะแนน PTE Academic อยู่ในระดับ A2, B1, B2, และ C1 และระดับช่วงที่สูงกว่าของระดับตามกรอบอ้างอิงความสามารถทางภาษาของสภายุโรป (CEFR) และไปถึง Academic Norm เมื่อเทียบกับเกรดเฉลี่ยของกลุ่มปกติ

วิธีการวิจัย ข้อมูลถูกเก็บกับนักศึกษานานาชาติที่ไม่ใช่เจ้าของภาษา 60 คน ระดับปริญญาตรี 37 คน และระดับบัณฑิตศึกษา 23 คน ที่ Macquarie University ในออสเตรเลีย ทั้งเพศชาย (45 คน) และเพศหญิง (15 คน) สัญชาติต่าง ๆ ที่มาจาก 3 วิชาเอก คือ วิทยาศาสตร์ 10 คน ธุรกิจและเศรษฐศาสตร์ 40 คน และศิลปะและมนุษย 10 คน ช่วงอายุของกลุ่มตัวอย่าง 18-36 ปี (เฉลี่ย 23.05 , S.D. = 3.38) กลุ่มตัวอย่างเหล่านี้มีผลการสอบ IELTS Academic อยู่แล้ว เพราะเป็นข้อกำหนดในการสอบเข้ามหาวิทยาลัย นักศึกษากลุ่มอาสาสมัครที่ได้สอบแบบทดสอบมากที่สุด 6 เดือน ก่อนการสอบ PTE Academic ได้รับการคัดเลือก เกณฑ์นี้ถูกกำหนดเพื่ออำนวยความสะดวกในการเก็บข้อมูล เนื่องจากเป็นไปไม่ได้ที่จะคัดเลือกผู้เข้าร่วมทั้งหมด ผู้ที่เพิ่งสอบ IELTS ก่อนที่จะสอบ PTE Academic ช่วงเวลาระหว่างการสอบ PTE Academic และ IELTS Academic คือ 2 ถึง 6 เดือน โดยค่าเฉลี่ย 4.7 เดือน ข้อมูลประกอบด้วยคะแนนสอบ IELTS ของผู้เข้าร่วม คะแนนสอบ PTE Academic เกรดเฉลี่ยภาคเรียนที่ 1 และเกรดเฉลี่ยของกลุ่ม เพื่อที่จะตอบส่วนที่สองของคำถามวิจัยที่ 3 คะแนน PTE Academic ถูกจัดเข้ากลุ่มตามระดับของ CEFR (A2, B1, B2, และ C1 และสูงกว่า) การแบ่งกลุ่มแบบนี้จะถูกใช้เพื่อคำนวณร้อยละของผู้เข้าร่วมในแต่ละระดับผู้ไปถึงหรือไปเกินเกรดเฉลี่ยของกลุ่ม (Norm Group's GPA) ในการวิเคราะห์ข้อมูล Pearson Correlation ถูกใช้ในการตอบคำถามการวิจัยที่ 1 ซึ่งหาความสัมพันธ์ที่เป็นไปได้ระหว่างคะแนน IELTS Academic ของผู้เข้าร่วม และคะแนน PTE Academic นอกจากนี้ Independent t-test ถูกใช้เพื่อค้นหาว่า PTE Academic สามารถแยกความสามารถ 2 กลุ่ม เหมือนการใช้ค่ามัธยฐานของคะแนน IELTS เพื่อตอบคำถามการวิจัยที่ 2 ลำดับสุดท้าย ผู้วิจัยใช้ Pearson correlation เพื่อตอบส่วนแรกของคำถามวิจัยที่ 3 ที่ค้นหาความสัมพันธ์ที่เป็นไปได้ระหว่างคะแนน PTE Academic จากกลุ่มตัวอย่างกับผลสัมฤทธิ์ทางการเรียนของพวกเขา

ผลการวิจัยสรุปได้ว่า คำถามวิจัยที่ 1: มีความสัมพันธ์ในระดับกลางถึงสูงอย่างมีนัยสำคัญระหว่างคะแนนสอบ PTE Academic และ IELTS Academic ของผู้เข้าสอบ ความสัมพันธ์สูงสุดพบได้ระหว่างคะแนนรวมในแบบทดสอบทั้ง 2 ชนิด ($r = 0.824$) ความสัมพันธ์สูงระดับถัดไป พบได้จากการพูด ($r = 0.723$) และการเขียน ($r = 0.686$) ในคำถามการวิจัยที่ 2: แบบทดสอบ PTE Academic สามารถแยกกลุ่มความสามารถสูงและต่ำได้โดยใช้คะแนนสอบ IELTS Academic ทั้งในคะแนนรวม และคะแนนสำหรับการสื่อสารในทักษะทั้ง 4 และคำถามวิจัยที่ 3: ข้อ 1) ความสามารถในการสอบ PTE ทั้งในภาพรวมและในแต่ละทักษะการสื่อสารทั้ง 4 มีความสัมพันธ์อย่างมีนัยสำคัญกับผลสัมฤทธิ์ทางวิชาการ ที่วัดโดยการใช้เกรดเฉลี่ยของกลุ่มตัวอย่าง ความสัมพันธ์สูงสุด ($r = 0.350$) พบได้ระหว่างคะแนนการอ่านกับเกรดเฉลี่ย ตามด้วยคะแนนรวม ($r = 0.344$) การเขียน ($r = 0.292$) การพูด ($r = 0.291$) และการฟัง ($r = 0.286$) ข้อ 2) กลุ่มใหญ่ที่สุดของผู้เข้าร่วม คือ 48.3% อยู่ในระดับ B2 ตามด้วยระดับ B1 (28.3%) ระดับ C1+ (16.7%) และ ระดับ A2 (6.7%) และเป็นที่น่าสนใจว่าความสามารถของผู้เข้าสอบขึ้นสูงจาก A2 ถึง C1+ พวกเขาไปถึงหรือเลยความสามารถทางวิชาการของกลุ่มปกติและมีนักศึกษาทั้งหมด 10 คน ในระดับ C1+ มีความสามารถทางวิชาการเหนือเจ้าของภาษาทั้ง 10 คน

Zhou Yuemei (2004) ทำการศึกษาเรื่อง Comparability Study of Two National EFL Tests (CET-6 and TEM-4) in China เพื่อศึกษาเปรียบเทียบแบบทดสอบภาษาอังกฤษระดับชาติใน

ประเทศจีน 2 ฉบับ คือ แบบทดสอบ CET-6 (The National College English Tests-Band 6) และแบบทดสอบ TEM-4 (The Test for English Majors-Band 4) โดยการวิเคราะห์องค์ประกอบ 3 อย่างของการศึกษาเปรียบเทียบ คือ ผู้เข้าสอบ (หรือกลุ่มตัวอย่าง) คะแนนสอบ และสาระแบบทดสอบ ข้อมูลที่ใช้ในการศึกษามี 3 ชนิด คือ 1) ข้อมูลเกี่ยวกับผู้เข้าสอบ เก็บมาโดยใช้แบบสอบถามหลังจากทำแบบทดสอบ CET-6 เสร็จ 2) คะแนนสอบจากแบบทดสอบทั้ง 2 ฉบับ 3) การจัดอันดับของครูเกี่ยวกับสาระแบบทดสอบ ข้อมูลเหล่านี้ได้รับการตรวจและวิเคราะห์เชิงปริมาณและคุณภาพตามความจำเป็น การวิเคราะห์ในหลาย ๆ ด้าน มีเป้าหมายเพื่อสำรวจความสัมพันธ์ระหว่างแบบทดสอบ 2 ฉบับ ถ้ามีในวิธีการตามการวัดวัตถุประสงค์หรือความเข้าใจ ลำดับสุดท้ายผลการวิจัยจะถูกสรุปด้วยการอ้างอิงถึงทฤษฎีความสามารถในการสื่อสารและ TMF (Test Method Facets)

กลุ่มตัวอย่างได้รับการเลือกจากมหาวิทยาลัย 2 แห่ง ที่มีความแตกต่างกันในเขตเมืองของ Shanghai แห่งหนึ่งเป็นสถาบันชั้นนำด้านการศึกษาภาษาต่างประเทศ และอีกแห่งหนึ่งเป็นมหาวิทยาลัยด้านการเงินและเศรษฐศาสตร์ ทั้ง 2 แห่งรับนักศึกษาเก่ง ผู้ที่มีวิชาเอกต่างกัน นักศึกษาชั้นปีที่ 2 ถูกคัดเลือกจาก 4 ห้องเรียน โดยกลุ่มตัวอย่าง เหตุผลในการเลือกกลุ่มตัวอย่างสะท้อนการพิจารณาดังนี้ 1) กลุ่มตัวอย่างมีลักษณะของกลุ่มประชากรของแบบทดสอบ CET และ TEM ให้มากที่สุดเท่าที่จะเป็นไปได้ในแง่ของวิชาเอก หรือจำนวนปีของการเรียนในมหาวิทยาลัย กลุ่มตัวอย่างจำเป็นต้องผสม 2 สถาบัน หรือแผนกที่แตกต่างกัน 2) ในขณะเดียวกัน กลุ่มตัวอย่างที่เลือก 2 กลุ่มควรมีภูมิหลังที่คล้ายกัน เช่น อายุ การเรียนภาษาอังกฤษก่อนหน้า เพื่อให้สามารถสร้างพื้นฐานที่มีเหตุผลของความต่างและความเหมือนที่มีความหมายในลักษณะของกลุ่มตัวอย่าง (Similarities) 3) เป็นไปได้สำหรับกลุ่มตัวอย่างที่จะสอบแบบทดสอบทั้ง 2 ฉบับ ฉบับหนึ่งหลังจากสอบอีกหนึ่งฉบับภายในระยะเวลาอันสั้น และภายใต้สภาพแวดล้อมที่คล้ายกับการบริหารการสอบปกติเท่าที่จะเป็นไปได้ 4) แต่ละกลุ่มถูกคาดหวังให้คุ้นเคยกับชนิดแบบทดสอบที่ใช้ที่ออกแบบโดยเฉพาะสำหรับพวกเขา หมายถึง ถ้านักศึกษาคนหนึ่งสอบแบบทดสอบ 2 ฉบับ สันนิษฐานได้ว่าเขาจะคุ้นเคยกับแบบทดสอบ 1 ฉบับและไม่คุ้นกับอีก 1 ฉบับ และ 5) สำหรับการเก็บข้อมูลด้วยแบบสอบถามจะครอบคลุมถึงสถานะด้านวิชาการ ขนาดของกลุ่มตัวอย่าง อายุ เพศ คะแนนจากการสอบเอ็นทรานซ์ ทักษะติดต่อห้องเรียนภาษาอังกฤษ และแรงจูงใจสำหรับการเรียนภาษาอังกฤษ

การวิเคราะห์คะแนนและผลการวิจัย พบประเด็นสำคัญ 4 ประเด็น คือ 1) มีความเป็นไปได้ 99 % ที่แบบทดสอบ 2 ฉบับ สัมพันธ์กันอย่างมีนัยสำคัญด้วยค่าความสัมพันธ์ $r = .712$ ดังนั้นจึงตัดสมมติฐานว่างที่ว่า ไม่มีความสัมพันธ์ระหว่างแบบทดสอบ 2 ชุด ออกไป สิ่งที่น่าสนใจอย่างหนึ่งของความสัมพันธ์ คือ ได้จากการเปรียบเทียบครั้งที่ 3 ของค่าเฉลี่ยด้วย MET Test นั่นคือ แบบทดสอบทั้ง 2 ฉบับ สัมพันธ์อย่างมีนัยสำคัญกับแบบทดสอบภาษาอังกฤษที่สอบโดยนักเรียนมัธยมก่อนเข้ามหาวิทยาลัย (MET) โดยมีค่าความสัมพันธ์กับ TEM-4 (.66) สูงกว่าเล็กน้อยกับความสัมพันธ์ของ CET-6 (.527) 2) คะแนนเฉลี่ยของแบบทดสอบค่อนข้างใกล้เคียงกัน ต่างกันเพียง 2 คะแนน แบบทดสอบ CET มีคะแนนเฉลี่ย 67.10; และแบบทดสอบ TEM มีคะแนนเฉลี่ย 69.22 แสดงถึงความคล้ายกันด้านคะแนน อย่างไรก็ตาม ในทางสถิติค่าเฉลี่ยของทั้ง 2 ชนิดต่างกันอย่างมีนัยสำคัญด้วยคะแนนเฉลี่ยของ CET-6 ต่ำกว่าเล็กน้อย และดังนั้นจึงยากกว่าเล็กน้อย สิ่งที่แปลกคือ ค่าระดับ

ความยากของแบบทดสอบทั้ง 2 ชนิด ดูเหมือนจะสับสน โดย TEM-4 (.698) ดูเหมือนจะยากกว่า CET-6 (.707) ความสัมพันธ์ยืนยันด้วยการตอบแบบสอบถามของนักศึกษา ประมาณ 23% คิดว่า CET-6 ยากกว่า ขณะที่ 35% พบว่า TEM-4 ยากกว่า 3) เมื่อนำเฉพาะแบบทดสอบแบบเลือกตอบของแบบทดสอบทั้ง 2 ฉบับ มาพิจารณา จากการวิเคราะห์ Factor Analysis พบว่า ประมาณ 27% ของ CET multiple choice items และประมาณ 77% ของแบบทดสอบแบบเลือกตอบของ TEM เข้าเกณฑ์สำหรับ unidimensionality แสดงถึงความแตกต่างใน latent structure ของ Discrete items ระหว่างแบบทดสอบ 2 ชนิด และประการสุดท้าย การตรวจสอบเปอร์เซ็นต์ไทม์เปิดเผยว่า นักเรียนเก่งมักจะได้คะแนนใน TEM-4 สูงกว่าคะแนนของ CET-6 ขณะที่นักเรียนที่อยู่ในกลุ่มต่ำกว่า ค่าเฉลี่ยจะได้คะแนนใน TEM-4 ต่ำกว่าคะแนนใน CET-6 ด้วยแบบทดสอบ TEM-4 ดูเหมือนจะเป็นตัวจำแนกได้มากกว่าเล็กน้อย

ในการวิเคราะห์สาระและผลการวิจัย การวิเคราะห์สาระโดยอาจารย์สอนภาษาอังกฤษในฐานะภาษาต่างประเทศจำนวน 7 คน หลังจากการฝึกอบรมให้พิจารณาสาระแบบทดสอบโดยใช้มาตราส่วนประมาณค่าที่ดัดแปลงบางส่วนจากทฤษฎีความสามารถใช้ภาษาเพื่อการสื่อสารและ TMF ของ Bachman ผลการวิจัยแยกได้เป็น 2 ส่วน คือ เชิงปริมาณ (Quantity) และเชิงคุณภาพ (Quality) ในเชิงปริมาณ (Quantity) เวลาในการทำ แบบทดสอบ CET-6 ให้เวลาทำแบบทดสอบ 120 นาที แบบทดสอบ TEM-4 ให้เวลาทำแบบทดสอบ 140 นาที ข้อมูลเชิงปริมาณของแบบทดสอบทั้ง 2 ชนิด รวมถึงจำนวนคำในทศวรรษในแบบทดสอบแต่ละชนิด ปริมาณสารป้อนเข้า (Input) สำหรับแต่ละแบบทดสอบ และน้ำหนักคะแนนในแต่ละตอน และจำนวนข้อ ในภาพรวม CET-6 มีปริมาณสารป้อนเข้า (Input) มากกว่าแบบทดสอบ TEM-4 ($5,308 > 5,220$) และใช้เวลาในการประเมินสารป้อนเข้าต่อหน้าที่มากกว่าเล็กน้อย ($44.2 > 37.3$) ในทางตรงกันข้าม แบบทดสอบ TEM มีแบบทดสอบ 4 ใน 7 ส่วนที่มีปริมาณสารป้อนเข้ามากกว่าแบบทดสอบ CET เนื่องจากมีส่วนประกอบของ TEM ในแบบทดสอบมากกว่า CET ภาระงานจึงดูเหมือนที่จะมีความหลากหลายมากกว่าซึ่งอาจจะทำให้ผู้เข้าสอบรู้สึกท้อแท้มากกว่า ในด้านเชิงคุณภาพ (Quality) มีข้อค้นพบอย่างน้อย 3 ประเด็นหลักที่สามารถสรุปได้จากเปอร์เซ็นต์การให้คะแนนและวิเคราะห์สาระของกระดาษแบบทดสอบ 1) จาก 10 ประเด็นหรือ ตัวบ่งชี้ของลักษณะสารป้อนเข้า (Input) ในแง่ของความสามารถใช้ภาษาเพื่อการสื่อสาร และ TMF มีจำนวน 4 ข้อ ของแบบทดสอบ CET และ 6 ข้อ ของแบบทดสอบ TEM แสดงถึงความซับซ้อนที่มากกว่า การค้นพบครั้งนี้จะทำให้วิธีการเชิงคุณภาพที่จะเข้าใจลักษณะของสารป้อนเข้าที่เกี่ยวข้องกับการสร้างแบบทดสอบ ปริมาณหรือความยาว และธรรมชาติของภาษา 2) แบบทดสอบ CET-6 รักษาระดับในการวัดทั้ง Organizational (2.42 บนสเกล 0 ถึง 4) และ Pragmatic (2.41) language abilities ตามทฤษฎีของ CLA ขณะที่แบบทดสอบ TEM-4 เน้นหนักในด้าน Organizational (2.61) abilities มากกว่าในส่วน Pragmatic (2.51) language abilities ในระดับในภาพรวมตามที่ระบุโดยตัวเลขที่บอกอำนาจในการวัดของความสามารถทางภาษาใน TEM-4 สูงกว่า CET-6 เล็กน้อย และประการสุดท้าย ความแตกต่างระหว่างแบบทดสอบ 2 ชนิด อาจจะดำเนินตามทฤษฎี CLA และ TMF 2 ปัจจัยหลักหรือแง่มุมที่ผลกระทบสำคัญต่อผลการสอบในระหว่าง 2 ปัจจัย โดยที่ CLA มีตัวบ่งชี้ 5 ข้อ และ TMF มี 7 ตัวบ่งชี้ และตรวจสอบให้คะแนนในส่วนต่างๆ ของแบบทดสอบโดยอาจารย์ ผลการวิจัยแสดงว่า

แบบทดสอบ 2 ชนิดแตกต่างกันในหน้าที่ของการใช้ภาษาเพื่อการสื่อสาร 28% และในด้านปัจจัยวิธีการทดสอบ 53%

จากผลการวิจัย นำไปสู่การวิจัยในอนาคต ควรมีความพยายามในการรวมแบบทดสอบทั้ง 2 ชุดให้เป็นชุดเดียวกัน หรือสร้างให้แบบทดสอบแต่ละชุดแตกต่างกัน ข้อเสนอแนะที่ 1 อาจจะช่วยประหยัดในการจัดการสอบ ขณะที่ข้อเสนอแนะที่ 2 อาจจะนำความมั่นใจของเอกภษาอังกฤษกลับมา หลังจากที่นักศึกษาใช้เวลาเรียน 1,120 ชั่วโมงในห้องเรียน มากกว่านักศึกษาที่ไม่ใช่เอกภษาอังกฤษ (340 ชั่วโมง) และคะแนนเหมือนกับใน TEM-4 กับ CET-6 น่าจะไม่ใช่เรื่องที่น่ายินดี

Lopez & Janssen (2010) ทำการศึกษาเรื่อง Validation Study of Colombia's ECAES English Exam เพื่อสำรวจความตรงของแบบทดสอบ ECAES English Exam ที่ได้รับการพัฒนาโดยกระทรวงศึกษาธิการของโคลัมเบีย ในช่วงของการปรับปรุงการเรียนการสอนภาษาอังกฤษในโคลัมเบียในปี 2005 โดยมีวัตถุประสงค์เพื่อ 1) สื่อสารกับสาธารณชนเกี่ยวกับประสิทธิภาพของแบบทดสอบ Colombia Bilingue 2) ประเมินโปรแกรมภาษาอังกฤษในมหาวิทยาลัยต่าง ๆ ของโคลัมเบีย และ 3) วัดความสามารถทางภาษาอังกฤษของนักศึกษา และมีคำถามการวิจัย คือ 1) โครงสร้างของแบบทดสอบ ECAES English Exam วัดระดับความสามารถได้อย่างเพียงพอตามค่านิยมของกรอบอ้างอิงความสามารถทางภาษาของสภายุโรป (CEFR) มากน้อยเพียงใด 2) ภาระงานในแบบทดสอบ ECAES English Exam คล้ายกับภาระงานที่นักศึกษาจากมหาวิทยาลัยในโคลัมเบียได้ทำในชีวิตจริงหรือไม่ และ 3) แบบทดสอบ ECAES English Exam มีผลกระทบอะไรบ้างต่อต่อโปรแกรมภาษาอังกฤษในมหาวิทยาลัยของโคลัมเบีย

ในด้านกลุ่มตัวอย่าง ผู้วิจัยเก็บข้อมูลจากผู้สอนโดยใช้วิธีการประเมินสาระแบบทดสอบ (A Content Evaluation of the ECAES English Exam) และเก็บข้อมูลกับนักศึกษาด้วยวิธีการสัมภาษณ์ความคิดภายใน (Think-Aloud Sessions) กลุ่มตัวอย่างเป็นอาจารย์ จำนวน 15 คน ซึ่งสอนภาษาอังกฤษระดับมหาวิทยาลัยที่มาจากมหาวิทยาลัยของรัฐและเอกชน ผู้ให้ข้อมูลเข้าร่วมโดยคัดเลือกจากความสะดวกและเต็มใจที่จะเข้าร่วมในการศึกษาครั้งนี้ ความคุ้นเคยกับกรอบอ้างอิงความสามารถทางภาษาของสภายุโรป (CEFR) และประสบการณ์ในการสอนภาษาอังกฤษในมหาวิทยาลัยต่าง ๆ ในโคลัมเบีย ในการสอบถามรายบุคคล อาจารย์แต่ละคนจะถูกถามให้อธิบายว่า แต่แบบทดสอบแต่ละข้อว่าวัดอะไร ลงรหัสแต่ละข้อตามกรอบอ้างอิงความสามารถทางภาษาของสภายุโรป (CEFR) อธิบายระดับที่แต่ละข้อสอดคล้องกับเป้าหมายการสอนของแต่ละโปรแกรมในมหาวิทยาลัยของตนเอง และแสดงความรู้สึกของตนเองต่อแบบทดสอบ และยังมีกลุ่มตัวอย่างที่เป็นนักศึกษา จำนวน 13 คน มาจากมหาวิทยาลัยต่าง ๆ ใน Bogota ได้รับเลือกโดย 1) เป็นนักศึกษาที่เรียนในชั้นปีสุดท้ายในสาขาวิชาของตนเอง 2) เป็นนักศึกษาที่ได้เรียนรายวิชาภาษาอังกฤษต่าง ๆ ในมหาวิทยาลัยของตนเอง 3) มีความสะดวกและเต็มใจที่จะสอบโดยใช้แบบทดสอบ ECAES English Exam ผู้วิจัยพยายามให้ได้จำนวนนักศึกษามากเท่ากันทั้ง 3 กลุ่ม คือ กลุ่มที่มีความสามารถสูงปานกลาง และต่ำ นักศึกษาแต่ละคนมีหน้าที่ของตนเอง คือ ทำแบบทดสอบและพูดอธิบาย และบอกถึงกลยุทธ์ที่ใช้ในการทำแบบทดสอบแต่ละข้อ

ในด้านการเก็บข้อมูล ผู้วิจัยใช้เครื่องมือในการเก็บรวบรวมข้อมูล 2 ชนิด เพื่อเก็บข้อมูลเกี่ยวกับความตรงของแบบทดสอบ ECAES English Exam คือ การประเมินสาระแบบทดสอบ และ

วิธีการสัมภาษณ์ความคิดภายใน จุดประสงค์ของเครื่องมือทั้ง 2 นี้ คือ เพื่อให้ได้ข้อมูลเกี่ยวกับโครงสร้างของแบบทดสอบ ประโยชน์ของแบบทดสอบ การรับรู้ของอาจารย์และนักศึกษาเกี่ยวกับแบบทดสอบ และผลกระทบของแบบทดสอบ นอกจากนี้ เครื่องมือเหล่านี้ยังได้บันทึกกลยุทธ์ที่ผู้เข้าสอบใช้ในการทำแบบทดสอบอีกด้วย

ในการประเมินสาระแบบทดสอบ ผู้วิจัยสอบถามอาจารย์สอนภาษาอังกฤษระดับมหาวิทยาลัยจำนวน 15 คน เพื่อทำการประเมินเนื้อหาสาระของแบบทดสอบ ECAES English Exam การประเมินเนื้อหาสาระหมายถึง คำตัดสินของผู้เชี่ยวชาญเกี่ยวกับความสามารถที่แบบทดสอบข้อต่าง ๆ วัด การประเมินเนื้อหาสาระในการศึกษาครั้งนี้ ได้ทำขึ้นเป็นรายบุคคล และใช้เวลาประมาณ 2 ชั่วโมงในแต่ละคน ผู้เข้าร่วมทั้งหมดได้รับแบบทดสอบ แบบฟอร์มให้กรอก และรายการคำอธิบายความสามารถระดับต่าง ๆ ตามกรอบอ้างอิงความสามารถทางภาษาของสภายุโรป (CEFR) จุดประสงค์ของการประเมินเนื้อหาสาระนี้คือเพื่อสอบถามอาจารย์ผู้สอนภาษาอังกฤษในฐานะภาษาต่างประเทศในการให้ความเห็นที่เชี่ยวชาญของพวกเขาเกี่ยวกับทักษะต่าง ๆ ที่แบบทดสอบแต่ละข้อในแบบทดสอบ the ECAES English Exam กำลังวัด หรือ ทักษะต่าง ๆ ที่นักศึกษาต้องใช้เพื่อตอบแต่ละภาระงาน และเพื่อให้ข้อเสนอแนะในแต่ละข้อโดยเน้นจุดแข็งและจุดอ่อน อาจารย์ยังได้รับการร้องขอให้โยงความสัมพันธ์แบบทดสอบแต่ละข้อกับกรอบอ้างอิงความสามารถทางภาษาของสภายุโรป (CEFR) และ เป้าหมายที่พบภายในหลักสูตรของมหาวิทยาลัยของตนเอง หลังจากที่อาจารย์ได้ทำส่วนนี้เสร็จสมบูรณ์ พวกเขาได้รับการร้องขอให้ให้ข้อเสนอแนะต่อแบบทดสอบโดยการให้คำตัดสินเกี่ยวกับตัวกระตุ้น (The Stimulus) คำสั่งต่าง ๆ (The Directions) ความคาดหวังของภาระงาน (The Task Expectations)

ในการสัมภาษณ์ความคิดภายใน (Think-Aloud Protocol) นักศึกษาจากสถาบันต่าง ๆ ใน Bogota ได้รับการร้องขอให้เข้าร่วมในการสัมภาษณ์ความคิดภายในขณะที่พวกเขา กำลังทำแบบทดสอบ ECAES English Exam ซึ่งการสัมภาษณ์ความคิดภายในที่ใช้ในการศึกษาครั้งนี้ โดยทั่วไปแล้วนำมาจากกระบวนการที่คล้าย ๆ กันกับที่นำเสนอโดย Ericsson and Simon (1983) การสัมภาษณ์ความคิดภายในมักจะใช้เป็นวิธีการเก็บข้อมูลในการแก้ปัญหา (Problem Solving) การเขียน (Writing) และการวิจัยด้านการอ่าน (Reading Research) การสัมภาษณ์ความคิดภายในที่ใช้ในการศึกษาครั้งนี้ ได้รับการศึกษานำร่องกับนักศึกษาระดับมหาวิทยาลัย 2 คน ซึ่งทำให้ผู้วิจัยมีโอกาสดูการตรวจสอบคุณภาพของคำถามที่ใช้สอบถามข้อมูล ผู้วิจัยยังได้แสดงตัวอย่างของการสัมภาษณ์ความคิดภายในให้นักศึกษาดู โดยการใช้อภาระงานจากแบบทดสอบภาษาอังกฤษในฐานะภาษาต่างประเทศชุดอื่น หลังจากการแสดงตัวอย่าง นักศึกษาได้มีโอกาสฝึกการทำกรสัมภาษณ์ความคิดภายใน ระยะการฝึกใช้เวลาประมาณ 30 นาที

นักศึกษาทุกคนที่เข้าร่วมในการสัมภาษณ์ความคิดภายในได้รับการร้องขอให้เล่าด้วยการใช้ภาษาที่พวกเขาชอบเกี่ยวกับอะไรก็ได้ที่เข้ามาในจิตใจขณะกำลังทำแบบทดสอบ ECAES English Exam การทำกรสัมภาษณ์ความคิดภายในทั้งหมด ถูกดำเนินการเป็นรายบุคคลในห้องส่วนตัวและถูกบันทึกไว้ทั้งหมด เพื่อให้สามารถแปลงความและวิเคราะห์ในเวลาต่อมา ทันทิที่ทำการแบบทดสอบเสร็จในแต่ละตอน ผู้วิจัยได้ทำการสัมภาษณ์หลังจากการทำแบบทดสอบกับนักศึกษาเพื่อเก็บข้อมูลเกี่ยวกับการรับรู้ของพวกเขาต่อตอนต่าง ๆ ของแบบทดสอบและเกี่ยวกับความรู้สึกของ

พวกเขาในการทำแบบทดสอบให้เสร็จ ในแต่ละครั้งของการสัมภาษณ์ความคิดภายในและการสัมภาษณ์แบบย้อนคิด (Think-Aloud Session and Retrospective Interview) ใช้เวลาประมาณ 100 นาที

การวิเคราะห์และตีความของแบบทดสอบทุกข้อในแบบทดสอบได้ให้หลักฐานสำหรับความตรงของแบบทดสอบ ECAES English Exam ผู้เขียนได้แปลงการสัมภาษณ์ความคิดภายในและพิมพ์แบบฟอร์มในการประเมินเนื้อหาสาระทั้งหมด หลังจากนั้น จัดกลุ่มและรวบรวมข้อมูลที่ได้ทั้งหมด โดยรายบุคคล ผู้วิจัยอ่านข้อมูลทั้งหมดหลายรอบ ต่อมาจึงใช้การกำหนดรหัสในการวิเคราะห์ข้อมูลกลุ่มต่าง ๆ เกิดขึ้นโดยเร็วจากข้อมูลที่มีความคิดหรือหัวเรื่องหลักที่ซ้ำกัน เทคนิคนี้ทำให้ผู้วิจัยสามารถใช้การทบทวนแบบสามเส้าโดยผู้วิจัย (Researcher Triangulation) เพื่อหาตรวจความถูกต้องของการลงรหัสและการตีความ กลุ่มต่าง ๆ ที่เกิดขึ้นจากข้อมูลที่ได้คือ หลักฐานที่มีพื้นฐานมาจากโครงสร้างของแบบทดสอบ หลักฐานที่มีพื้นฐานมาจากปฏิสัมพันธ์ของแบบทดสอบ หลักฐานที่มีพื้นฐานมาจากการเป็นสภาพจริงของแบบทดสอบ และหลักฐานที่มีพื้นฐานมาจากผลกระทบของแบบทดสอบ

จากการวิเคราะห์เนื้อหาสาระ ผู้ประเมินพบว่าแบบทดสอบ ECAES English Exam ทดสอบเฉพาะการอ่าน ไวยากรณ์ และทักษะการอ่านอื่น ๆ โดยทั่วไปอาจารย์รู้สึกว่าการสอบ ECAES English Exam ไม่สอดคล้องเพียงพอกับคำอธิบายในระดับต่าง ๆ ของกรอบอ้างอิงความสามารถทางภาษาของสภายุโรป (CEFR) ผู้ประเมินยังเปิดเผยอีกว่าเนื้อหาสาระของแบบทดสอบไม่สอดคล้องโดยสมบูรณ์กับสาระของหลักสูตรภาษาต่างประเทศในมหาวิทยาลัยของตนเอง ดังที่พวกเขาได้บอกว่าหลักสูตรของพวกเขาเป็นหลักสูตรบูรณาการและเน้นการสื่อสาร โดยทั่วไป อาจารย์รู้สึกว่าการสอบ ECAES English Exam ยังไม่สอดคล้องอย่างเพียงพอกับหลักสูตรภาษาอังกฤษในมหาวิทยาลัยของตน การขาดความสอดคล้องส่วนนี้ ทำให้ผู้เขียนสรุปว่ามีความเป็นตัวแทนต่ำในโครงสร้างในแบบทดสอบ นอกจากนี้ในการทำการสัมภาษณ์ความคิดภายในมีความผันแปรที่เกิดจากความไม่จำเป็นในโครงสร้างของแบบทดสอบหลายข้อ เป็นข้อโต้แย้งที่หนักสุดต่อความตรงเชิงโครงสร้างของแบบทดสอบ ตัวแปรที่นอกเหนือบางส่วนในการะงานทำให้แบบทดสอบบางตอนของแบบทดสอบ ECAES English Exam ยากหรือง่ายเกินไป

การมีปฏิสัมพันธ์ระหว่างภาระงานกับผู้เข้าสอบพิจารณาได้ว่า นักศึกษาใช้กลยุทธ์ใดในการทำภาระงานในแบบทดสอบ ไม่ว่าจะเป็นความสามารถเชิงกลยุทธ์ (Strategic Competence) กลยุทธ์เชิงอภิปัญญา (Metacognitive Strategies) และ ความรู้ในหัวเรื่อง (Topical Knowledge) (Bachman & Palmer, 1996) ในด้านความสามารถเชิงกลยุทธ์ (Strategic Competence) ผู้วิจัยพบในการสัมภาษณ์ความคิดภายในว่ามีการเดาเกิดขึ้นมาก อันเนื่องจากรูปแบบแบบทดสอบที่ให้เลือกคำตอบ นอกจากนี้ บางภาระงานในแบบทดสอบต้องการให้นักศึกษาใช้ทักษะภาษาอื่น ๆ ในการให้คำตอบ ตัวอย่างเช่น ตอนที่ 1 ที่ตั้งใจจะวัดทักษะการอ่าน ในขณะที่ความเป็นจริงวัดความรู้ด้านคำศัพท์ ในการสัมภาษณ์นักศึกษา นักศึกษาทั้งหมดบอกว่าเพียงแต่จับคู่คำศัพท์บทสัญลักษณ์กับคำหลักในตัวเลือกเท่านั้น ในด้านความรู้ในหัวเรื่อง (Topical Knowledge) ผู้วิจัยพบว่า ตอนที่ 2 ของแบบทดสอบต้องการให้นักศึกษามีความรู้ในหัวเรื่อง (Topical Knowledge) ในกลุ่มของคำศัพท์ที่กำลังถูกวัด เช่น สิ่งของในบ้าน ส่วนต่าง ๆ ของร่างกาย เป็นต้น ในการสัมภาษณ์ความคิดภายใน

(Think-aloud Sessions) นักศึกษาบางคนกล่าวว่า ภาระงานเป็นสิ่งที่ยาก เพราะแบบทดสอบต้องการให้พวกเขามีความรู้เกี่ยวกับสิ่งของต่าง ๆ ในห้องครัว ในการสัมภาษณ์ความคิดภายใน ผู้วิจัยพบว่า นักศึกษาใช้กลยุทธ์เชิงอภิปัญญา (Metacognitive Strategies) เพียงเล็กน้อยในการทำภาระงานให้สมบูรณ์ ตัวอย่างเช่น ในตอนที่ 7 นักศึกษาเพียงแต่ใช้ความสัมพันธ์ของคำศัพท์ (Word Association) เพื่อเดาความหมายของคำศัพท์ที่ไม่ทราบความหมายเท่านั้น ผู้วิจัยไม่พบหลักฐานใตว่า นักศึกษาใช้กลยุทธ์อื่นๆ เช่น การเดาความหมายจากบริบท (Guessing from Contexts) หรือ การใช้การวิเคราะห์คำ (Using Word Analysis)

ความเป็นสภาพจริงบรรยายระดับที่ผู้เข้าสอบทำภาระงานที่คล้ายกับภาระงานที่พวกเขาจะทำได้ในชีวิตจริง (Bachman & Palmer, 1996) ผลจากการวิเคราะห์และในการสัมภาษณ์ความคิดภายใน ผู้วิจัยพบว่าภาระงานในแบบทดสอบ ECAES English Exam ไม่มีความสมจริงในปริมาณมาก ในการประเมินเนื้อหาสาระ อาจารย์แสดงความคิดเห็นว่าภาระงานในแบบทดสอบไม่เหมือนกับภาระงานที่นักศึกษาทำในชีวิตจริง ดังนั้นผู้วิจัยจึงสรุปได้ว่ามีความสอดคล้องเพียงเล็กน้อยระหว่างผลการสอบและการใช้ภาษาที่ไม่ใช่การสอบ ผู้เขียนรู้สึกว่ นักพัฒนาแบบทดสอบจำเป็นต้องทำการวิเคราะห์ความจำเป็น (Needs Analysis) ในมหาวิทยาลัยต่าง ๆ ในโคลัมเบีย เพื่อระบุภาระงานการใช้ภาษาเป้าหมาย (TLU Tasks) และหลังจากนั้นจึงออกแบบภาระงานที่เหมาะสมและจำเป็นที่มีเปอร์เซ็นต์สูงต่อผู้เข้าสอบชาวโคลัมเบีย

ในด้านผลกระทบรายบุคคล ในระหว่างการสัมภาษณ์ความคิดภายใน นักศึกษาบรรยายความรู้สึ่วว่าแบบทดสอบ ECAES English Exam มีผลกระทบน้อยมาก พวกเขารายงานว่าคะแนนจากการสอบไม่ได้ใช้เพื่อการสมัครงาน การศึกษาในระดับบัณฑิตศึกษา หรือเพื่อการตัดสินใจอื่น ๆ นักศึกษาหลายคนบอกว่าไม่ได้ทำแบบทดสอบนี้อย่างจริงจัง เพราะพวกเขาไม่ได้รับผลอะไร พวกเขาอธิบายอีกว่าไม่ได้ถูกบังคับให้เตรียมการสำหรับการสอบครั้งนี้หรือเน้นย้ำในสาระของแบบทดสอบ และสุดท้าย ผู้วิจัยพบว่าระบบการตรวจให้คะแนนไม่ได้ให้ผลสะท้อนกลับใด ๆ ที่มีความหมายแก่นักศึกษา ในผลของการสอบของพวกเขาที่นอกเหนือคะแนนที่เป็นตัวเลข ในด้านผลกระทบต่อผู้สอน ผู้วิจัยพบว่าแบบทดสอบมีผลกระทบน้อยเช่นกัน ในการประเมินสาระ อาจารย์รายงานว่า พวกเขาไม่ได้สอนเพื่อแบบทดสอบนี้หรือใช้สื่อวัสดุเพื่อเตรียมสำหรับการสอบ พวกเขาแย้งอีกว่า หลักสูตรยังไม่ได้สัมพันธ์กับสาระสำคัญของแบบทดสอบ และยังไม่รู้สึถูกบังคับให้ปรับเนื้อหาสาระในรายวิชาให้คล้ายกับแบบทดสอบ ผู้วิจัยยังได้เรียนรู้อีกว่าอาจารย์บางส่วนไม่ได้รับข่าวสารที่ดีเกี่ยวกับ the ECAES English Exam อาจารย์จำนวน 4 ราย รายงานว่าพวกเขาไม่ทราบเลยว่าแบบทดสอบฉบับใหม่ได้ถูกปรับปรุงในปี 2007 สุดท้าย ผู้เขียนพบว่าผลกระทบเพียงเล็กน้อย ต่อหลักสูตรภาษาอังกฤษของสถาบัน อาจารย์ทั้งหมดรายงานว่ามหาวิทยาลัยของพวกเขาใช้แบบทดสอบ (เช่น TOEFL, IELTS, PET, MELICET หรือแบบทดสอบที่จัดทำเอง) ในระบบความรับผิดชอบของพวกเขาที่จะอธิบายและตรวจสอบความสามารถทางภาษาของนักศึกษา หรือเพื่อจุดประสงค์ในการตัดสินใจบางอย่าง แต่ระบบเหล่านี้ไม่ได้รวมแบบทดสอบ ECAES English Exam เอาไว้ ในทำนองเดียวกัน อาจารย์รายงานว่าพวกเขาไม่ได้หรือมีแผนที่จะหาความสอดคล้องของหลักสูตรภาษาอังกฤษของตนเองกับแบบทดสอบ ECAES English Exam

Wu (2011) ทำการศึกษาเพื่อสร้างกรอบสำหรับการเทียบเคียงและเปรียบเทียบคุณภาพของแบบทดสอบที่อยู่ในระดับเดียวกัน และเพื่อระบุ Parameters ที่เป็นประโยชน์ต่อการอธิบายแบบทดสอบวัดความสามารถในการอ่านที่ต่างระดับกันได้อย่างชัดเจน โดยอาศัยการประเมินอย่างมีวิจรรย์ญาณจากการเกณฑ์การเทียบเคียงของกรอบอ้างอิงความสามารถทางภาษาของสภายุโรป (CEFR) โดยศึกษาเฉพาะระดับ CEFR ระดับ B1 และ B2 เท่านั้น ประยุกต์กรอบการเทียบเคียง Socio-Cognitive ของ Weir's (2005) เพื่อตรวจสอบประเด็นต่าง ๆ ของความตรงของแบบทดสอบ GEPT ระดับต่าง ๆ ในแง่มุมมอง บริบทแวดล้อม (Contextual Parameters) ทักษะกระบวนการทางปัญญา (Cognitive Processing Skills) และผลการสอบ (Test Results) ผู้วิจัยเปรียบเทียบความเหมือนและความต่างระหว่างแบบทดสอบ GEPT และ CEFR ระดับ B1 และ B2 โดยการเทียบกับแบบทดสอบ PET และ FACE ที่พัฒนาโดย Cambridge ESOL ซึ่งได้รับการเทียบคุณภาพกับ CEFR มาแล้ว ผลการวิจัยสนับสนุนความตรงเชิงโครงสร้างของแบบทดสอบ GEPT โดยรวม แต่แสดงว่า ความตรงเชิงปัญญาจำเป็นต้องได้รับการส่งเสริมโดยการเพิ่มแบบทดสอบที่วัดการอ่านอย่างแท้จริงให้มากขึ้นการเทียบแบบทดสอบกับกรอบ CEFR ผลการวิจัยแสดงว่ากระบวนการตามที่คู่มือแนะนำการเทียบแบบทดสอบกับระดับของ CEFR ไม่ได้ให้หลักฐานอย่างเพียงพอที่จะแสดงความเท่าเทียมกันระหว่างแบบทดสอบต่างชนิดที่มุ่งวัดตามกรอบ CEFR ระดับเดียวกัน ผลการวิจัยชี้ให้เห็นว่า the GEPT Intermediate Level และ PET ซึ่งต้องการวัดที่ระดับ B1 มีความเท่าเทียมกัน ขณะที่ the GEPT High-Intermediate Level and FACE ซึ่งมุ่งวัดที่ระดับ B2 มีความแตกต่างกันมากไม่เพียงแต่ผลการสอบแต่ยังเป็นเรื่องของลักษณะบริบทแวดล้อมและกระบวนการทางปัญญาอีกด้วย ผู้วิจัยได้ให้ข้อเสนอแนะว่าการวิจัยครั้งต่อไป ควรเพิ่มจำนวนผู้สอบให้มากขึ้นและให้เป็นผู้สอบที่มีความแตกต่างกันทั้งทางด้านความรู้ภาษาอังกฤษและภูมิหลัง เพื่อดูความสามารถทางปัญญาในการทำแบบทดสอบชนิดต่างๆ กัน เช่น แบบเลือกตอบ หรือให้คำตอบสั้น ๆ

Weir, Chan, and Nakatsuhara (2013) ศึกษาความตรงร่วมสมัยของแบบทดสอบส่วนที่วัดทักษะการอ่านและการเขียนของ the Advanced Level GEPT เพื่อดูหลักฐาน 2 ชนิด คือ ความเท่าเทียมกันของแบบทดสอบและอำนาจในการทำนาย เก็บข้อมูลกับกลุ่มตัวอย่างนักศึกษา 171 คน ที่ไปศึกษาระดับปริญญาตรีหลักสูตรร่วมด้านบริหารธุรกิจ ณ มหาวิทยาลัยแห่งหนึ่งในประเทศอังกฤษ ความเท่าเทียมกันของแบบทดสอบหาได้โดยการวิเคราะห์ความสัมพันธ์ระหว่างคะแนนสอบ the Advanced Level GEPT ในทักษะการอ่านและการเขียน กับ คะแนนสอบ IELTS พบความสัมพันธ์ระดับปานกลางถึงมากระหว่างคะแนนรวมในการทักษะการเขียนของ GEPT และคะแนนการสอบเขียนของ IELTS โดยมีค่าความสัมพันธ์ $r=.432$ ($p<.01$) และระหว่างคะแนนการเขียนของ GEPT กับคะแนนการอ่านและการเขียนของแบบทดสอบ IELTS โดยมีค่าความสัมพันธ์ $r=.503$ ($p<.01$) จากการศึกษาครั้งนี้ยังพบว่า มั่นยากที่กลุ่มตัวอย่างจะผ่านแบบทดสอบวัดทักษะการเขียนของ the GEPT Advanced กว่าการทำงานคะแนน 6.5 (หรือเทียบเท่า ระดับ C1 ของ CEFR) หรือสูงกว่าในการทดสอบทักษะเขียนของแบบทดสอบ IELTS จากคะแนนการสอบเขียนของ GEPT มีค่าความสัมพันธ์ที่มากกว่าที่ระดับ .520 ($p<.01$) ระหว่างคะแนนรวมการสอบอ่านของแบบทดสอบ GEPT (Task 1 and Task 2) และคะแนนการสอบอ่านของ IELTS อาจจะยากกว่าที่จะผ่านแบบทดสอบการอ่านของ the GEPT กว่าการทำงานคะแนน 6.5 หรือสูงกว่าในแบบทดสอบการอ่านของ IELTS ลำดับสุดท้าย

พบว่าสัมพันธที่มากกว่าที่ระดับ .562 ($p < .01$) ระหว่างคะแนนรวมทักษะการอ่านและการเขียนของ GEPT และคะแนนเฉลี่ยของคะแนนการสอบการอ่านและการเขียนของ IELTS ประการที่สอง อำนาจในการทำนายของคะแนนการอ่านและการเขียนของแบบทดสอบ the Advanced Level GEPT หาได้จากความสามารถทางวิชาการของผู้เข้าสอบในชีวิตจริงในการทำภาระงานที่ได้รับมอบหมาย ระหว่างการเรียนและการสอบ ผลการศึกษาแสดงว่า คะแนนการอ่านและการเขียนของแบบทดสอบ GEPT สัมพันธ์กับ ความสามารถทางด้านวิชาการในชีวิตจริงของผู้เข้าสอบที่ระดับ .529 ($p < .01$) มีค่าความแปรปรวนความสามารถในชีวิตจริงเท่ากับ 27.98 % เมื่อศึกษาแยกส่วน คะแนนการอ่านของ GEPT สัมพันธ์กับการแสดงออกในทางวิชาการในชีวิตจริงที่ระดับ .499 ($p < .01$) ความแปรปรวน 24.90% และคะแนนเขียนของ GEPT ที่ระดับ .294 ($p < .01$) อธิบายความแปรปรวนความสามารถในชีวิตจริง 8.64% แบบสอบถามการประเมินตนเองในประสบการณ์ด้านการอ่านและการเขียนทางวิชาการ ใช้เก็บข้อมูลเพื่อหาหลักฐานเชิงปริมาณเพิ่มเติมมาอธิบายว่า the GEPT Advanced Test สามารถสะท้อนการรับรู้ของผู้เข้าสอบในด้านความสามารถการอ่านและการเขียนเชิงวิชาการของตนเองได้ดีเพียงใด ผลการวิจัยแสดงว่า ความยากลำบากในการอ่านและเขียนเชิงวิชาการที่มาจากการประเมินตนเองส่วนใหญ่ เช่น ความเข้าใจประโยคยาวและซับซ้อน การสร้างความหมายเชื่อมโยงระหว่างข้อความ การเขียนเกี่ยวกับเรื่องที่ซับซ้อน การใช้คำศัพท์ที่หลากหลาย การแก้ไขความถูกต้องทางภาษาในงานเขียนของตนเอง ถูกกล่าวไว้ในแบบทดสอบวัดทักษะการอ่านและกาเขียนของ GEPT advanced level

ศิริพร พงษ์สุรพิพัฒน์ และคณะ (2544) จัดโครงการเปรียบเทียบแบบทดสอบ CU-TEP กับแบบทดสอบ TOEFL ซึ่งเป็นการศึกษาความตรงร่วมสมัย เพื่อให้ได้แบบทดสอบที่มีคุณภาพเทียบเคียงกับแบบทดสอบต่างประเทศไว้ให้บริการทดสอบแก่นิสิตและบุคลากรทั่วไป ซึ่งอาจจะใช้เป็นหลักฐานในการสมัครงานหรือศึกษาต่อได้ วัตถุประสงค์ของการวิจัย คือ เพื่อศึกษาหาความตรงร่วมสมัยของแบบทดสอบสมิทธิภาพทั่วไปทางภาษาอังกฤษของจุฬาลงกรณ์มหาวิทยาลัย (CU-TEP) กับแบบทดสอบ TOEFL พลวิจัยในการวิจัยครั้งนี้ คือ ผู้ที่เข้ารับการทดสอบแบบทดสอบสมิทธิภาพทั่วไปทางภาษาอังกฤษ และได้รับการคัดเลือกให้ไปสอบแบบทดสอบ TOEFL จำนวน 155 คน เครื่องมือที่ใช้ในการวิจัยครั้งนี้ ได้แก่ แบบทดสอบ TOEFL และแบบทดสอบวัดสมิทธิภาพทั่วไปทางภาษาอังกฤษที่สร้างขึ้นโดยคณะกรรมการซึ่งมีความชำนาญด้านการสร้างแบบทดสอบภาษาอังกฤษของสถาบันภาษา จุฬาลงกรณ์มหาวิทยาลัย จำนวน 3 ฉบับ ซึ่งเป็นแบบทดสอบคู่ขนาน คณะผู้ออกแบบทดสอบได้ดำเนินการสร้างแบบทดสอบ โดยดำเนินการ 1) ศึกษาตารางกำหนดลักษณะเฉพาะของแบบทดสอบ ซึ่งได้ใช้ในการสร้างแบบทดสอบวัดสมิทธิภาพทั่วไปทางภาษาอังกฤษของสถาบันภาษา จุฬาลงกรณ์มหาวิทยาลัย (CU-TEP) 2) ดำเนินการสร้างแบบทดสอบตามตารางดังกล่าว 3) นำแบบทดสอบที่คณะกรรมการสร้างขึ้นทั้ง 3 ฉบับ มาทดลองใช้กับนิสิต นักศึกษา ข้าราชการ และผู้สนใจทั่วไป หลังจากนั้นนำผลการสอบที่ได้มาวิเคราะห์เพื่อปรับปรุงคุณภาพของแบบทดสอบ 4) นำแบบทดสอบทั้ง 3 ฉบับ ที่ปรับปรุงแล้วไปใช้จริงกับกลุ่มตัวอย่าง และ 5) ส่งผู้เข้าสอบรับการทดสอบแบบทดสอบ CU-TEP และได้รับการคัดเลือกไปสอบแบบทดสอบ TOEFL

การเก็บรวบรวมข้อมูล มีขั้นตอนการดำเนินงาน คือ 1) นำแบบทดสอบทั้ง 3 ฉบับ ไปทำการสอบแก่ผู้สมัครสอบ โดยแบ่งผู้สมัครสอบเป็น 3 กลุ่ม คือ กลุ่มที่ 1 ซึ่งมีผู้สมัครสอบ 273 คน สอบ

แบบทดสอบฉบับที่ 1 กลุ่มที่ 2 ซึ่งมีผู้สมัครสอบ 306 คน สอบแบบทดสอบฉบับที่ 2 กลุ่มที่ 3 ซึ่งมีผู้สมัครสอบ 262 คน สอบแบบทดสอบฉบับที่ 3 2) นำกระดาษคำตอบส่วนที่เป็นแบบทดสอบปรนัย มาตรวจและวิเคราะห์โดยเครื่องคอมพิวเตอร์ของสถาบันภาษา และนำคะแนนที่ได้มาประมวลผลการสอบเพื่อใช้ต่อไป 3) นำกระดาษคำตอบส่วนที่เป็นแบบทดสอบอัตนัยมาทำการตรวจโดยอาจารย์สถาบันภาษา ซึ่งมีความเชี่ยวชาญในการตรวจแบบทดสอบ CU-TEP ฉบับละ 2 คน โดยใช้เกณฑ์การให้คะแนนเกณฑ์เดียวกัน ซึ่งแบ่งออกเป็น 5 ระดับความสามารถ 4) นำคะแนนส่วนที่เป็นแบบทดสอบอัตนัยมาเตรียมข้อมูล โดยการบันทึกข้อมูลลงในแผ่นคอมพิวเตอร์เพื่อใช้ต่อไป 5) สุ่มผู้สมัครสอบจากแต่ละกลุ่ม ๆ ละ ประมาณ 50 คน เพื่อไปทำการสอบแบบทดสอบ TOEFL วิธีการสุ่มนี้ดูจากคะแนนสอบส่วนที่เป็นแบบทดสอบปรนัย โดยคัดเลือกจากผู้ที่มีคะแนนอยู่ในระดับสูง ระดับกลาง และระดับต่ำ จำนวนเท่า ๆ กัน 6) ส่งพลวิจัยไปทำการสอบแบบทดสอบ TOEFL ทั้งส่วนที่เป็นแบบทดสอบปรนัยและแบบทดสอบอัตนัย 7) รวบรวมคะแนนผลการสอบ TOEFL ของพลวิจัยและนำมาบันทึกลงในแผ่นบันทึกข้อมูล และ 8) นำข้อมูลในข้อ 2, 4 และ 7 ไปทำการวิเคราะห์

การวิเคราะห์ข้อมูลดำเนินการวิเคราะห์โดยใช้โปรแกรมสำเร็จรูป SPSS (Statistical Package for the Social Sciences for Personal Computer) และโปรแกรมคอมพิวเตอร์อื่น ๆ ตามขั้นตอน คือ 1) คำนวณค่าสถิติพื้นฐาน ได้แก่ คะแนนสูงสุด คะแนนต่ำสุด คะแนนเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน ค่าสัมประสิทธิ์การกระจายของคะแนนของแบบทดสอบแต่ละฉบับ 2) ทำการวิเคราะห์ข้อทดสอบรายข้อของแบบทดสอบแต่ละฉบับโดยใช้วิธีแบบประเพณีนิยม (Classical Model) ด้วยโปรแกรม CTIA/GRADING (Version 6.30) 3) ทำการวิเคราะห์หาค่าสัมประสิทธิ์สหสัมพันธ์แบบ Pearson Product-Moment Correlation 4) คำนวณหาสมการในการทำนายโดยใช้วิธีการวิเคราะห์การถดถอย (Regression Analysis) 5) คำนวณหาค่าความคลาดเคลื่อนในการทำนาย (Standard error of estimate) และ 6) ทำการวิเคราะห์ค่าความเชื่อมั่นแบบ Inter-rater reliability ของแบบทดสอบอัตนัย

ผลการวิเคราะห์คุณภาพของแบบทดสอบ CU-TEP ทั้ง 3 ฉบับ (ปรนัยและอัตนัย) ในด้านความยากง่ายและค่าอำนาจจำแนก พบว่าแบบทดสอบทั้ง 3 ฉบับมีค่าใกล้เคียงกันมากในส่วนปรนัย โดยมีความยากง่ายอยู่ในระดับปานกลาง ($p = 0.48 - 0.53$) ในส่วนอัตนัยนั้นข้อ 1 และข้อ 3 ของทุกฉบับยากเกินไป ($FV = 0.14 - 0.15$) แต่ข้อ 2 ของทุกฉบับมีค่าความยากง่ายพอใช้ได้ สำหรับค่าอำนาจจำแนกนั้นในส่วนปรนัยมีค่าดีพอใช้ ($r_{bis} 0.34 - 0.38$) ในส่วนอัตนัย ข้อ 2 และ 3 ของฉบับที่ 1 ข้อ 1 และ 3 ของฉบับที่ 2 และข้อ 1, 2, 3 ของฉบับที่ 3 มีค่าอยู่ในเกณฑ์ดี ส่วนข้ออื่นๆ มีค่าอยู่ในเกณฑ์ที่ยอมรับได้ ในด้านค่าความเชื่อมั่น พบว่าแบบทดสอบ CU-TEP ทั้ง 3 ฉบับ มีค่าความเชื่อมั่นสูงทั้งส่วนที่เป็นแบบทดสอบอัตนัยและปรนัย ในส่วนที่เป็นอัตนัยนั้น ฉบับที่ 1 มีค่าความเชื่อมั่นสูงที่สุด (0.94) และฉบับที่ 3 มีค่าความเชื่อมั่นต่ำที่สุด (0.726) แต่ในส่วนที่เป็นปรนัยนั้น ฉบับที่ 3 มีค่าความเชื่อมั่นสูงที่สุด ($KR_{20} = 0.92$) และฉบับที่ 2 มีค่าต่ำที่สุด ($KR_{20} = 0.89$) ในด้านการวิเคราะห์ความสัมพันธ์ของคะแนนจากแบบทดสอบ CU-TEP และแบบทดสอบ TOEFL พบว่า คะแนนจากแบบทดสอบ CU-TEP ฉบับที่ 3 มีความสัมพันธ์กับคะแนนจากแบบทดสอบ TOEFL สูงที่สุดในแต่ละทักษะ ยกเว้นด้านการเขียนแบบปรนัย ซึ่งมีค่าน้อยกว่าฉบับที่ 1 เล็กน้อยแต่ก็ยังมีค่าความสัมพันธ์อยู่ในระดับสูง ($r_{xy} = 0.84$) ในขณะที่คะแนนของแบบทดสอบฉบับที่ 2 มีค่าความสัมพันธ์น้อยกว่าฉบับ

อื่น ๆ ทุกทักษะ นอกจากนี้ผู้วิจัยได้ให้ข้อเสนอแนะสำหรับงานวิจัยทางการทดสอบว่า ควรศึกษาหามาตรการวัดความรู้ความสามารถในการสื่อสารหรือความรู้ความสามารถทางการใช้ภาษาสื่อสารที่เหมาะสม และแสดงออกซึ่งสมรรถภาพทั่วไปทางภาษาอังกฤษอย่างแท้จริง ศึกษาความตรงทางทฤษฎีและความตรงเชิงเนื้อหาเพิ่มเติม และศึกษาความตรงร่วมสมัยของแบบทดสอบ CU-TEP กับแบบทดสอบมาตรฐานระดับนานาชาติอื่น ๆ เพื่อสร้างแบบทดสอบวัดความรู้ความสามารถที่มีประสิทธิภาพ และประสิทธิผลเพิ่มขึ้นตรงตามจุดประสงค์ของผู้ใช้ และสร้างคลังแบบทดสอบหรือธนาคารแบบทดสอบ โดยแยกประเภทของแบบทดสอบ ตามเนื้อหาความยากง่าย อำนาจจำแนก และปัจจัยอื่น ๆ เพื่อสร้างแบบทดสอบด้วยเทคโนโลยีใหม่ ๆ วัดความรู้ความสามารถทางภาษาอังกฤษเฉพาะได้ถูกต้องแน่นอน โดยไม่ใช้เวลานานและการทราบผลก็เป็นไปด้วยความรวดเร็วอีกด้วย

จากการทบทวนงานวิจัย 2 แนวทาง คือ งานวิจัยที่เกี่ยวข้องกับ 1) การพัฒนาแบบทดสอบวัดความสามารถภาษาอังกฤษ และ 2) การประเมินคุณภาพแบบทดสอบ พอจะสรุปได้ว่า งานวิจัยด้านการพัฒนาแบบทดสอบวัดความสามารถภาษาอังกฤษ มีแนวโน้มที่จะพัฒนาแบบทดสอบเพื่อวัดความสามารถในการสื่อสารของผู้เข้าสอบทั้งด้านการฟัง การพูด การอ่าน การเขียน และการใช้ภาษา และแบบทดสอบจะถูกออกแบบให้มีความสำคัญต่อผู้เข้าสอบโดยแท้จริง โดยมักจะพบว่านักพัฒนาแบบทดสอบจะกำหนดเนื้อหาสาระที่ผู้เข้าสอบจะได้พบในสภาพแวดล้อมการใช้ภาษาในชีวิตจริงมาออกสอบ นอกจากนี้ จุดเด่นอย่างหนึ่งที่ได้พบได้ในการพัฒนาแบบทดสอบ คือ การประเมินคุณภาพแบบทดสอบ ที่สามารถทำได้ทั้งก่อนและหลังการนำแบบทดสอบไปใช้ การประเมินคุณภาพก่อนนำแบบทดสอบไปใช้ที่นิยม คือ การตรวจสอบคุณภาพด้านเนื้อหาโดยผู้เชี่ยวชาญ และการประเมินหลังนำแบบทดสอบไปใช้ คือ การหาคุณภาพแบบทดสอบเชิงประจักษ์ ได้แก่ การวิเคราะห์แบบทดสอบรายข้อ (ค่าความเชื่อมั่น ค่าความยากง่าย และค่าอำนาจจำแนก) การเทียบคุณภาพกับแบบทดสอบมาตรฐานหรือเครื่องมือวัดอื่น ๆ และการสอบถามความคิดเห็นของผู้ใช้แบบทดสอบ ส่วนในด้านการสร้างแบบทดสอบ มีขั้นตอนดำเนินงาน 3 ขั้น คือ ขั้นตอนออกแบบ (Design) ขั้นปฏิบัติการ (Operationalization) และขั้นบริหารการสอบ (Administration) โดยขั้นตอนออกแบบ (Design) ประกอบด้วย 1) คำบรรยายจุดประสงค์การสอบ 2) คำอธิบายขอบเขตการใช้ภาษาเป้าหมาย (Target Language Use Domain) และชนิดของภาระงาน 3) คำอธิบายเกี่ยวกับประชากรผู้ทำแบบทดสอบ 4) คำนิยามโครงสร้างที่จะวัด และ 5) แผนสำหรับการประเมินคุณภาพด้านคุณประโยชน์ของแบบทดสอบ ขั้นปฏิบัติการ (Operationalization) ประกอบด้วย 1) กำหนดโครงสร้างแบบทดสอบ 2) สร้างแบบทดสอบหรือพัฒนาข้อคำถาม และ 3) การหาคุณภาพตามเนื้อหา (ประเมินโดยผู้เชี่ยวชาญ) และขั้นบริหารการสอบ (Administration) ซึ่งครอบคลุมทั้งการศึกษานำร่องและการเก็บข้อมูลจริง ประกอบด้วย 1) การเตรียมความพร้อม ได้แก่ การอบรมผู้ให้คะแนน และการผลิตแบบทดสอบ 2) การเก็บข้อมูล ได้แก่ การจัดสอบทั้งกับกลุ่มทดลองและกลุ่มตัวอย่าง การตรวจให้คะแนน การวิเคราะห์คะแนนสอบ ปรับปรุงแบบทดสอบ การลดจำนวนแบบทดสอบ การกำหนดเกณฑ์ผ่าน และการนำแบบทดสอบไปใช้จริงกับกลุ่มตัวอย่าง และ 3) การจัดเก็บแบบทดสอบและรายงานผล ได้แก่ การรายงานผลการสอบ (ทำการอ้างอิงหรือตัดสินใจในสิ่งที่แบบทดสอบได้ถูก

กำหนดไว้) การจัดแบบทดสอบเข้าธนาคารแบบทดสอบ และการจัดทำรายงานด้านเทคนิคของแบบทดสอบ

สำหรับการวิจัยครั้งนี้ ผู้วิจัยจะพัฒนาแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัยที่แตกต่างจากงานวิจัยอื่น ๆ คือ เป็นแบบทดสอบที่วัดความสามารถในการสื่อสารของผู้เข้าสอบครบทุกทักษะ ทั้งการฟัง การพูด การอ่าน การเขียน และการใช้ภาษา อีกทั้งออกแบบเนื้อหาสาระให้

สอดคล้องกับความจำเป็นในการใช้ภาษาในอนาคตของผู้เข้าสอบ โดยเป็นทั้งเนื้อหาสาระที่ผู้เข้าสอบพบได้ในชีวิตประจำวันและเกี่ยวข้องกับการทำงาน และออกแบบให้มีจำนวนข้อคำถามหลายข้อ และใช้เทคนิคการทดสอบที่หลากหลาย ในด้านการประเมินคุณภาพแบบทดสอบ ก่อนนำแบบทดสอบไปใช้ มีการประเมินความสอดคล้องระหว่างจุดประสงค์และข้อคำถาม (IOC) โดยผู้เชี่ยวชาญ และการประเมินคุณประโยชน์ของแบบทดสอบ (Test Usefulness) โดยผู้สอนภาษาอังกฤษระดับมหาวิทยาลัย และหลังการใช้แบบทดสอบ มีการประเมินคุณภาพแบบทดสอบรายข้อและความเชื่อมั่น การตรวจสอบความตรงร่วมสมัย (Concurrent Validity) กับแบบทดสอบมาตรฐาน และการประเมินความพึงพอใจของผู้เกี่ยวข้องต่อแบบทดสอบที่สร้างขึ้น ส่วนการสร้างแบบทดสอบจะดำเนินการใน 3 ขั้นตอน คือ 1) การออกแบบ (Design) ได้แก่ การกำหนดเนื้อหาสาระที่จะสอบและการวางแผนด้านการตรวจสอบคุณภาพ การจัดสรรงบประมาณและทรัพยากร 2) การพัฒนา (Operationalization) ได้แก่ การกำหนดโครงสร้างแบบทดสอบ การพัฒนาแบบทดสอบรายข้อ การประเมินคุณภาพแบบทดสอบก่อนการใช้ และการทดลองใช้แบบทดสอบและปรับปรุง และ 3) การบริหารการสอบ (Administration) ได้แก่ การจัดสอบ และการประเมินคุณภาพแบบทดสอบหลังการใช้



บทที่ 3

วิธีดำเนินการวิจัย

การวิจัยเรื่องการพัฒนาแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test) มีวัตถุประสงค์ 2 ข้อ คือ เพื่อสร้างและพัฒนาแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test) ให้เป็นแบบทดสอบมาตรฐานที่มีลักษณะเฉพาะ คือ

- คุณประโยชน์ของแบบทดสอบเป็นไปตามเกณฑ์ที่กำหนด
- คุณภาพรายข้อและความเชื่อมั่นเป็นไปตามเกณฑ์ที่กำหนด
- ความตรงร่วมสมัย (Concurrent Validity) สูง เมื่อเทียบกับแบบทดสอบมาตรฐาน

เพื่อศึกษาความพึงพอใจของผู้เกี่ยวข้องต่อแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test)

เพื่อให้สามารถดำเนินการวิจัยให้บรรลุวัตถุประสงค์ดังกล่าว ผู้วิจัยได้กำหนดวิธีการวิจัยออกเป็น 5 หัวข้อ คือ 1) ประชากรและกลุ่มตัวอย่าง 2) ตัวแปรที่ศึกษา 3) ระยะเวลา 4) รูปแบบการวิจัย และ 5) การพัฒนาแบบทดสอบ โดยมีรายละเอียดในแต่ละหัวข้อต่อไปนี้

3.1 ประชากรและกลุ่มตัวอย่าง

ประชากรในการศึกษาครั้งนี้ เป็นนิสิตระดับปริญญาตรี ภาคปกติ ชั้นปีที่ 1-4 ที่ลงทะเบียนและกำลังศึกษาอยู่ในภาคเรียนที่ 2 ปีการศึกษา 2561 จำนวน 8,167 คน ประกอบด้วย 5 คณะ คือ 1) คณะมนุษยศาสตร์และสังคมศาสตร์ 2) คณะศึกษาศาสตร์ 3) คณะศิลปกรรมศาสตร์ 4) คณะเศรษฐศาสตร์และบริหารธุรกิจ และ 5) คณะนิติศาสตร์ ดังรายละเอียดในตารางที่ 6

ตารางที่ 6 จำนวนประชากรในการวิจัย

คณะ	ชาย	หญิง	รวม
- มนุษยศาสตร์และสังคมศาสตร์	581	1,838	2,419
- ศึกษาศาสตร์	593	1,558	2,151
- ศิลปกรรมศาสตร์	299	254	553
- เศรษฐศาสตร์และบริหารธุรกิจ	409	1,274	1,683
- นิติศาสตร์	553	808	1,361
รวม	2,435	5,732	8,167

กลุ่มทดลองและกลุ่มตัวอย่างในการศึกษาครั้งนี้ได้มาโดยการรับสมัครนิสิตที่มีความสนใจสอบวัดความสามารถภาษาอังกฤษ กลุ่มทดลอง หมายถึง ผู้เข้าสอบในครั้งที่ 1-2 ของการทดสอบใช้ทดลองใช้แบบทดสอบครั้งนี้ และกลุ่มตัวอย่าง หมายถึง ผู้เข้าสอบในครั้งที่ 3 ซึ่งเป็นการใช้แบบทดสอบฉบับจริงหลังจากได้ปรับปรุงคุณภาพหลังจากการทดลองใช้ทั้ง 2 ครั้ง

นิสิตที่สมัครสอบสามารถเลือกการสอบตามเวลาที่ตนเองสามารถเขาสอบได้เพียง 1 ครั้งเท่านั้น และในแต่ละรอบ ผู้วิจัยสามารถจัดการสอบได้ไม่เกิน 300 คน โดยทั่วไปนิสิตจะให้ความสนใจและสมัครใจที่จะเข้าทดสอบวัดความสามารถภาษาอังกฤษเป็นจำนวนมาก แต่เนื่องจากมีกิจกรรมอย่างอื่นเข้ามาแทรก อีกทั้งมหาวิทยาลัยได้ประกาศให้ช่วงเวลาที่วางแผนจะทำการทดสอบ เป็นการสอนชดเชยของวันหยุด ส่งผลให้จำนวนนิสิตที่เข้าสอบได้จริงน้อยกว่าจำนวนที่สมัครไว้ ดังแสดงในตารางที่ 7





ตารางที่ 7 จำนวนกลุ่มทดลองและกลุ่มตัวอย่าง

กลุ่มอย่าง คณะ	ทดลองครั้งที่ 1			ทดลองครั้งที่ 2			กลุ่มตัวอย่าง			รวม		
	ชาย	หญิง	รวม	ชาย	หญิง	รวม	ชาย	หญิง	รวม	ชาย	หญิง	รวม
มนุษยศาสตร์และสังคมศาสตร์	-	9	9	7	47	54	2	32	34	9	88	97
- ภาษาอังกฤษ	-	9	9	5	34	39	2	21	23	7	64	71
- ภาษาญี่ปุ่น	-	-	-	2	13	15	-	4	4	2	17	19
- ภาษาไทย	-	-	-	-	-	-	-	7	7	-	7	7
ศึกษาศาสตร์	4	7	11	9	22	31	-	3	3	13	32	45
- คณิตศาสตร์	1	-	1	-	-	-	-	-	-	1	-	1
- วิทยาศาสตร์-ชีววิทยา	3	7	10	-	-	-	-	-	-	3	7	10
- วิทยาศาสตร์-ฟิสิกส์	-	-	-	5	10	15	-	-	-	5	10	15
- วิทยาศาสตร์-เคมี	-	-	-	-	-	-	-	2	2	-	2	2
- ภาษาอังกฤษ	-	-	-	3	12	15	-	-	-	3	12	15
- เทคโนโลยีทางการศึกษา	-	-	-	1	-	1	-	-	-	1	-	1
- สังคมศึกษา	-	-	-	-	-	-	-	1	1	-	1	1
เศรษฐศาสตร์และบริหารธุรกิจ	-	8	8	1	7	8	-	-	-	1	15	16
- การบัญชี	-	8	8	1	7	8	-	-	-	1	15	-
นิติศาสตร์	-	-	-	-	1	1	-	8	8	-	9	9
- นิติศาสตร์	-	-	-	-	1	1	-	8	8	-	9	9
รวม	4	24	28	17	77	94	2	43	45	23	144	167

3.2 ตัวแปรที่ศึกษา

ตัวแปรที่จะใช้ในการวิจัย คือ

1) คะแนนที่ได้จากการทดสอบทักษะการฟัง การใช้ภาษา การอ่าน การเขียนและการพูดของแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test) และแบบทดสอบวัดความสามารถภาษาอังกฤษมาตรฐาน (แบบทดสอบปรนัยใช้คะแนนจากแบบทดสอบ TOEFL และแบบทดสอบอัตนัยใช้คะแนนจากแบบทดสอบ IELTS)

2) ความพึงพอใจของผู้เชี่ยวชาญ อาจารย์ผู้สอนภาษาอังกฤษ และผู้เข้าสอบต่อแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test) ที่ประเมินโดยใช้แบบประเมินความพึงพอใจของผู้เกี่ยวข้องต่อแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test)

3.3 ระยะเวลา

การศึกษาครั้งนี้ได้แบ่งช่วงเวลาดำเนินงานออกเป็น 3 ขั้นตอน ตามระยะเวลาในการพัฒนาแบบทดสอบตามแนวคิดของ Bachman and Lyle F. & Palmer (1996), Genesee & Upshur (1996), Genesee and Fred & Upshur (1996) และ Brown (2004) คือ ขั้นตอนออกแบบแบบทดสอบ (Test Design) ขั้นตอนปฏิบัติการออกข้อสอบ (Test Operationalization) และขั้นตอนบริหารการสอบ (Test Administration) โดยมีรายละเอียดในแต่ละขั้น ดังนี้

3.3.1 ขั้นตอนออกแบบแบบทดสอบ (Test Design)

ผู้วิจัยดำเนินกิจกรรมในขั้นตอนออกแบบระหว่างเดือนกันยายน 2558 ถึง เมษายน 2561 ซึ่งมีกิจกรรมหลัก 2 ส่วน คือ การศึกษาข้อมูลพื้นฐานเพื่อนำมากำหนดสาระแบบทดสอบ และการออกแบบแบบทดสอบ โดยการศึกษาข้อมูลพื้นฐานเป็นการศึกษาค้นคว้าหาข้อมูล วิเคราะห์สังเคราะห์ข้อมูลจาก 5 แหล่ง คือ 1) หลักสูตรภาษาอังกฤษระดับอุดมศึกษา (มหาวิทยาลัยทักษิณ, 2554) 2) กรอบอ้างอิงความสามารถทางภาษาของสภายุโรป (CEFR) ระดับ B1-B2 (Council of Europe, 2003) 3) โครงสร้างแบบทดสอบวัดความสามารถภาษาอังกฤษมาตรฐานสากล (Zheng & Cheng, 2008: 408-409; Malaysian Examinations Council, 2006: 12; สถาบันภาษาจุฬาลงกรณ์มหาวิทยาลัย, 2552; Educational Testing Service, 2010: 3; IELTS, 2015: 4; UCLES, 2014: 3; telc GmbH, 2012: 25) 4) หลักการและทฤษฎีการพัฒนาแบบทดสอบวัดความสามารถภาษาอังกฤษ (Bachman & Palmer, 1996; Genesee & Upshur, 1996; Brown, 2004) และ 5) งานวิจัยด้านการพัฒนาแบบทดสอบและการประเมินคุณภาพแบบทดสอบ (อัจฉราวงศ์โสธร, 2532; ศิริพร พงษ์สุรพิพัฒน์ และคณะ, 2544; Ivanova, 2011(Ivanova 2011; Pizorn et al., 2012; Samaie & Mahmoud, 2014; Wu & Rachel Yi-Fen, 2011); Wullur, 2011;

Pizorn & Moe, 2012; Samaie, 2014) สำหรับการออกแบบประกอบด้วย 1) การกำหนดจุดประสงค์ของแบบทดสอบ 2) การกำหนดและบรรยายภาระงานในปริเฉทของการใช้ภาษาเป้าหมายและชนิดของภาระงาน 3) การบรรยายลักษณะของผู้ใช้แบบทดสอบหรือผู้ทำแบบทดสอบ 4) การนิยามสาระหัวเรื่อง โครงสร้างภาษาที่ต้องการจะวัด เทคนิควิธีการวัด การตรวจให้คะแนนและเกณฑ์ 5) การกำหนดแผนการประเมินคุณภาพแบบทดสอบ และ 6) การวางแผนและจัดสรรทรัพยากรที่เกี่ยวข้อง

3.3.2 ขั้นปฏิบัติการออกข้อสอบ (Test Operationalization)

ผู้วิจัยดำเนินการกิจกรรมในขั้นปฏิบัติการออกข้อสอบ (Test Operationalization) ระหว่างเดือนพฤษภาคม ถึง กันยายน 2561 โดยมีกิจกรรมที่เกี่ยวข้อง คือ 1) การจัดทำโครงสร้างลักษณะแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test) ตามรายการหัวข้อที่ควรปรากฏในโครงสร้างลักษณะแบบทดสอบ 2) การตรวจสอบโครงสร้างลักษณะแบบทดสอบกับจุดประสงค์การสอบโดยผู้เชี่ยวชาญ 3) การแก้ไขปรับปรุงโครงสร้างลักษณะแบบทดสอบตามคำแนะนำของผู้เชี่ยวชาญ 4) การเขียนแบบทดสอบหรือสร้างชุดคำถามตามโครงสร้างลักษณะแบบทดสอบ 5) การกำหนดวิธีตรวจให้คะแนนแบบทดสอบ 6) การสร้างแบบประเมินคุณภาพแบบทดสอบโดยผู้เชี่ยวชาญ 7) การสร้างแบบสอบถามความคิดเห็นผู้ใช้แบบทดสอบ 8) การตรวจสอบคุณภาพแบบทดสอบและแบบสอบถามความคิดเห็นผู้ใช้แบบทดสอบโดยผู้เชี่ยวชาญ 9) ปรับปรุงแบบทดสอบ และแบบสอบถามตามความคิดเห็นผู้ใช้แบบทดสอบตามคำแนะนำของผู้เชี่ยวชาญ และ 10) การทดลองใช้แบบทดสอบกับกลุ่มทดลองและปรับปรุงคุณภาพก่อนการใช้จริง

3.3.3 ขั้นบริหารการสอบ (Test Administration)

ผู้วิจัยดำเนินการกิจกรรมในขั้นบริหารการสอบ (Administration) ระหว่างเดือนมกราคม ถึง กุมภาพันธ์ 2562 โดยมีกิจกรรมหลักที่เกี่ยวข้อง 5 ส่วน คือ ก่อนการเก็บข้อมูล ขณะเก็บข้อมูล การตรวจแบบทดสอบ การวิเคราะห์ประสิทธิภาพแบบทดสอบ และการตรวจสอบคุณภาพแบบทดสอบ โดยรวม โดยแต่ละส่วนมีกิจกรรมย่อยที่เกี่ยวข้อง ดังต่อไปนี้

กิจกรรมก่อนการเก็บข้อมูล ประกอบด้วย 1) การปฐมนิเทศผู้เข้าสอบ 2) การจัดคนคุมสอบและผู้ตรวจแบบทดสอบ 3) การประชุมผู้คุมสอบและกับตรวจแบบทดสอบ 4) การอบรมผู้ดำเนินการสอบพูดและผู้ตรวจแบบทดสอบการเขียน 5) การจัดเตรียมอุปกรณ์และห้องสอบ 6) การผลิตแบบทดสอบและการรักษาความปลอดภัยแบบทดสอบ และ 7) การตรวจสอบความพร้อมก่อนก่อนวันสอบ

กิจกรรมขณะเก็บข้อมูล ประกอบด้วย 1) การจัดสอบด้วยแบบทดสอบที่พัฒนาขึ้น 2) การสอบถามความพึงพอใจของผู้เข้าสอบต่อแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test) และ 3) การจัดสอบด้วยแบบทดสอบมาตรฐาน

กิจกรรมในการตรวจแบบทดสอบ ประกอบด้วย 1) การตรวจแบบทดสอบปรนัย (Objective Test) และ 2) การตรวจแบบทดสอบอัตนัย (Subjective Test)

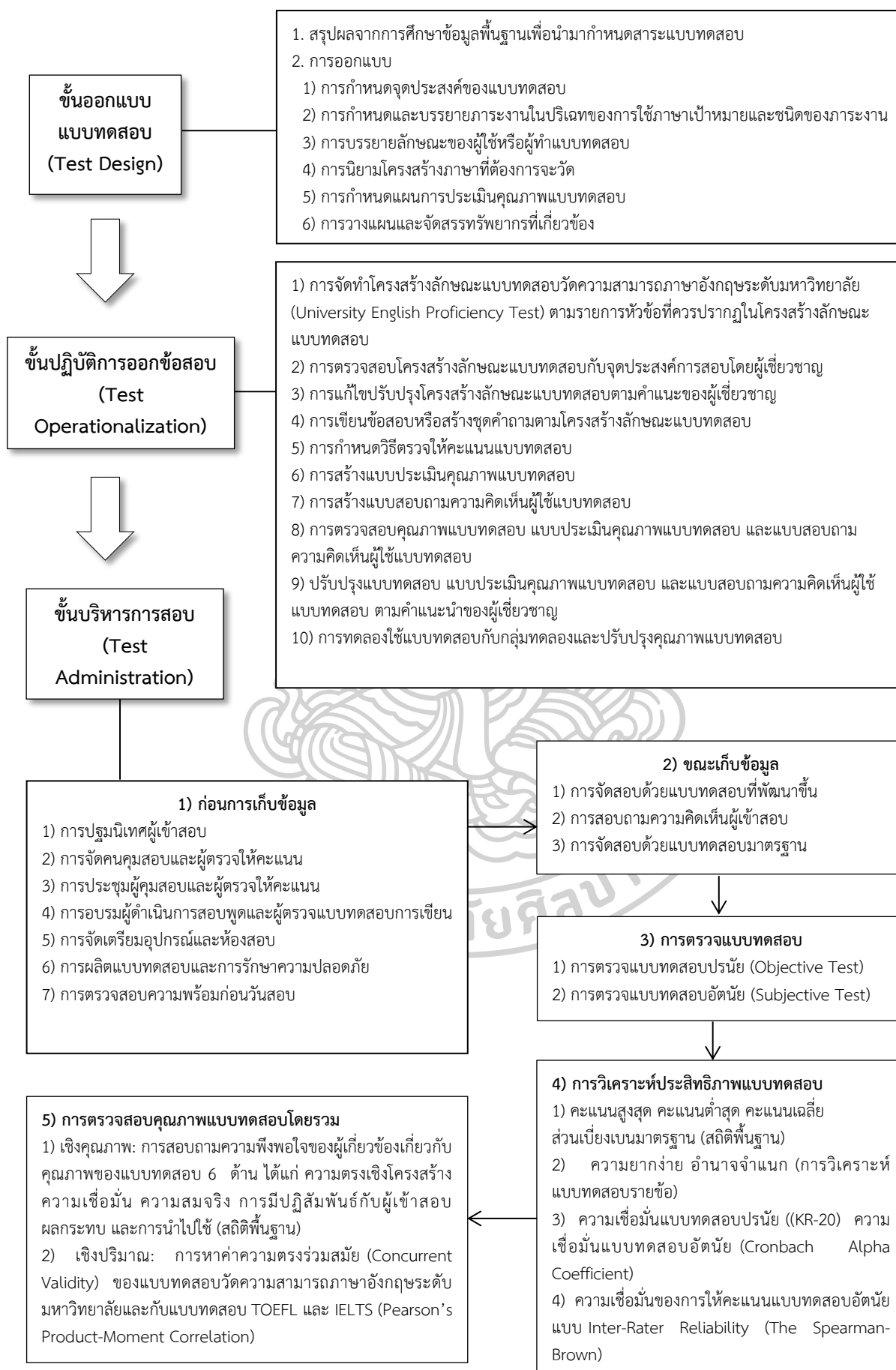
กิจกรรมการวิเคราะห์ประสิทธิภาพแบบทดสอบ ประกอบด้วย 1) คะแนนสูงสุด คะแนนต่ำสุด คะแนนเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน (สถิติพื้นฐาน) 2) ความยากง่าย อำนาจจำแนก (การวิเคราะห์แบบทดสอบรายข้อ) 3) ความเชื่อมั่นแบบทดสอบปรนัย (KR-20) และแบบทดสอบอัตนัย (Cronbach Alpha Coefficient) ซึ่งคำนวณโดยใช้โปรแกรมวิเคราะห์ข้อสอบชื่อ B-Index Version 7.0 (สาครแสงผึ้ง, 2550) และ 4) ความเชื่อมั่นของการให้คะแนนแบบทดสอบอัตนัยแบบ Inter-Rater Reliability (The Spearman-Brown) ซึ่งคำนวณโดยใช้โปรแกรม SPSS ของคณะวิทยาศาสตร์ มหาวิทยาลัยศิลปากร

กิจกรรมการตรวจสอบคุณภาพแบบทดสอบโดยรวม ประกอบด้วย 1) การเก็บข้อมูลเชิงคุณภาพ ได้แก่ การสอบถามความพึงพอใจผู้เชี่ยวชาญ อาจารย์ผู้สอนภาษาอังกฤษ และผู้เข้าสอบ เกี่ยวกับคุณภาพของแบบทดสอบโดยรวม (สถิติพื้นฐาน) และการเก็บข้อมูลเชิงปริมาณ ได้แก่ การหาค่าความสัมพันธ์ระหว่างคะแนนของแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัยกับแบบทดสอบมาตรฐาน (Pearson's Product-Moment Correlation)

3.4 รูปแบบการวิจัย

รูปแบบการวิจัยของการวิจัยครั้งนี้มีลักษณะตามขั้นตอนของการพัฒนาแบบทดสอบมาตรฐานตามแนวคิดของ Bachman & Palmer (1996), Genesee & Upshur (1996) และ Brown (2004) ดังแสดงในแผนผังข้างล่างนี้





การพัฒนาแบบทดสอบแบ่งออกได้เป็น 3 ชั้น คือ 1) ชั้นออกแบบ (Design Stage) 2) ชั้นปฏิบัติการออกข้อสอบ (Test Operationalization) และ 3) ชั้นบริหารการสอบ (Administration Stage) ตาม แนวคิดของ Bachman & Palmer (1996), Genesee & Upshur (1996) และ Brown (2004) โดยมีรายละเอียดในแต่ละชั้นดังนี้

3.5.1 ชั้นออกแบบ (Design Stage)

เพื่อกำหนดโครงสร้าง (Test Construct) ของแบบทดสอบที่จะสร้างขึ้น ผู้วิจัยศึกษาเนื้อหาสาระของแบบทดสอบจาก 1) หลักสูตรภาษาอังกฤษระดับอุดมศึกษา (มหาวิทยาลัยทักษิณ, 2554) 2) กรอบอ้างอิงความสามารถทางภาษาของสภายุโรป (CEFR) (Council of Europe, 2003) และ 3) โครงสร้างแบบทดสอบวัดความสามารถภาษาอังกฤษมาตรฐานสากล (Zheng & Cheng, 2008: 408-409; Malaysian Examinations Council, 2006: 12; สถาบันภาษา จุฬาลงกรณ์มหาวิทยาลัย, 2552; Educational Testing Service, 2010: 3; Educational Testing Service, 2014; IELTS, 2015: 4; UCLES, 2014: 3; telc GmbH, 2012: 25) และสังเคราะห์เป็นเนื้อหาสาระของแบบทดสอบที่จะนำไปจัดทำตารางกำหนดลักษณะแบบทดสอบ (Test Specifications) ในขั้นตอนต่อไป โดยมีการดำเนินการ 2 อย่าง คือ 1) การสรุปกรอบเนื้อหาสาระของแบบทดสอบจากการสังเคราะห์ข้อมูล และ 2) การกำหนดกรอบเนื้อหาสาระของแบบทดสอบที่จะพัฒนา โดยมีรายละเอียดในแต่ละกิจกรรม ดังนี้

3.5.1.1 การสรุปกรอบเนื้อหาสาระของแบบทดสอบจากการสังเคราะห์ข้อมูล

ในการสรุปกรอบเนื้อหาสาระแบบทดสอบที่ได้จากการสังเคราะห์ข้อมูล ผู้วิจัยได้นำเสนอกรอบเนื้อหาสาระของแบบทดสอบในทักษะต่าง ๆ ตามลำดับดังนี้ คือ ทักษะการฟัง ทักษะการใช้ภาษา ทักษะการอ่าน ทักษะการเขียน และทักษะการพูด โดยมีรายละเอียดในการดำเนินการแต่ละขั้นดังนี้

การสรุปผลการศึกษาเนื้อหาสาระของแบบทดสอบด้านการฟัง

แบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test) ด้านการฟัง ควรมีจุดประสงค์หลักเพื่อวัดความสามารถของผู้เข้าสอบในการฟังสาระที่เกี่ยวข้องกับชีวิตประจำวันและเชิงวิชาการ และประกอบด้วย 2 ตอน คือ 1) การฟังในชีวิตประจำวัน และ 2) การฟังเชิงวิชาการ โดยในแต่ละตอนควรมีกรอบเนื้อหาสาระที่เกี่ยวข้องดังนี้

การฟังในชีวิตประจำวัน (Listening for Everyday Life)

1) การฟังในชีวิตประจำวัน (Listening for Everyday Life) มีจุดประสงค์เพื่อวัดความสามารถของผู้เข้าสอบในการฟังสาระที่เกี่ยวข้องกับชีวิตประจำวันและการทำงาน และมีสาระภาษา (Language Knowledge/Subskills) ที่เกี่ยวข้อง คือ ความสามารถในการ 1) จับใจความสำคัญของสาระที่ฟัง (Ability to identify the main idea of the listening texts) 2) เข้าใจรายละเอียดของสาระที่ฟัง (Ability to identify some specific details of the listening texts) 3) ระบุหัวข้อของสาระที่ฟัง (Ability to identify topic of the listening texts) 4) ระบุคำศัพท์สำคัญที่เกี่ยวข้องกับหัวเรื่องของสาระที่ฟัง (Ability to recognize key lexical items related to subject/topic) 5) บอกจุดประสงค์ของผู้พูด (Ability to identify purpose of the speaker) 6) ระบุลำดับเนื้อหาของสาระที่ฟัง (Ability to recognize the sequence of the listening texts) 7) บอกชนิดของสาระที่ฟังว่าเป็นคำเตือน ข้อเสนอแนะ คำแนะนำ หรือวิธีการใช้ เป็นต้น (Ability to recognize types of listening texts (e.g., warning, suggestions, recommendations, instructions) 8) บอกความหมายของคำศัพท์โดยอาศัยบริบทแวดล้อม (Ability to deduce meanings of words from contexts) 9) บอกความคิดเห็นของผู้พูดต่อสาระที่ฟัง (Ability to detect the opinion of the speaker toward subject matter) 10) บอกทัศนคติของผู้พูดต่อสาระที่ฟัง (Ability to detect attitude of speaker toward subject matter) 11) ตีความด้านความสัมพันธ์ เช่น ระหว่างคู่สนทนา เหตุและผล (Ability to infer relationships e.g. between people, cause-effect) 12) ตีความเกี่ยวกับเวลา สถานที่ บุคคล ฯลฯ โดยอาศัยข้อมูลที่ฟัง (Ability to infer about time, place, person, etc. based on the listening texts) 13) การระบุคำแนะนำโดยอาศัยความเข้าใจสาระที่ฟัง (Ability to give suggestions based on understanding the listening texts) 14) ทำนายเหตุการณ์โดยอาศัยความเข้าใจสาระที่ฟัง (Ability to make prediction based on understanding the listening texts) และ 15) สรุปความโดยอาศัยข้อมูลที่ฟัง (Ability to draw conclusion based on the listening texts)

2) การฟังในชีวิตประจำวันกำหนดให้ผู้เข้าสอบฟังหัวเรื่องที่คุ้นเคยทั้งในเรื่องส่วนตัวและเกี่ยวข้องกับงาน โดยใช้บทฟัง 2 ประเภท คือ บทฟังที่มีผู้พูดคนเดียว (Monologues) และบทสนทนา (Conversations) โดยมีรายละเอียด คือ บทฟังที่มีผู้พูดคนเดียว (Monologues) ได้แก่ 1) ข้อความเสียงจากเครื่องรับโทรศัพท์ 2) ข้อความเสียงประกาศในที่สาธารณะ 3) บทสรุปข่าวทางวิทยุ และ 4) คำแนะนำการใช้อุปกรณ์ในชีวิตประจำวัน ส่วนบทสนทนา (Conversations) ได้แก่ 1) บทสนทนาสั้น ๆ จากการพูดคุยในสถานการณ์ประจำวันทั้งในบริบทเรื่องส่วนตัวและเกี่ยวข้องกับงาน 2) บทสนทนาที่ยาวขึ้น ได้แก่ 1) การสัมภาษณ์งาน และ 2) การอภิปราย

3) แหล่งที่มาของบทฟัง ได้มาจากเว็บไซต์ที่ให้ดาวน์โหลดไฟล์เสียงในหัวเรื่องต่าง ๆ ได้ฟรี และผ่านการพิจารณาโดยผู้วิจัยแล้วว่าไม่มีความลำเอียงในด้านหัวเรื่องต่อผู้เข้าสอบทั้ง 3 สาย ได้แก่ สายวิทยาศาสตร์ สายมนุษยศาสตร์ และสายสังคมศาสตร์

- 4) เทคนิคการสอบหรือรูปแบบข้อคำถามที่สามารถนำมาใช้ทดสอบทักษะการฟังของแบบทดสอบวัดความสามารถภาษาอังกฤษควรประกอบด้วย 1) แบบเลือกตอบ 2) แบบจับคู่ 3) แบบถ่ายโอนข้อมูล เช่น การเติมความในแผนผัง แผนที่ ไดอะแกรม) และ 4) แบบเติมประโยคให้สมบูรณ์
- 5) การตรวจให้คะแนนจะเป็นรูปแบบการตอบถูกได้ 1 หรือ ตอบผิด ได้ 0

การฟังเชิงวิชาการ (Listening in Academic Contexts)

1) การฟังเชิงวิชาการ (Listening in Academic Contexts) มีจุดประสงค์เพื่อวัดความสามารถของผู้เข้าสอบในการฟังสาระในบริบทวิชาการ และมีสาระภาษา (Language Knowledge/Subskills) ที่เกี่ยวข้อง คือ ความสามารถในการ 1) จับใจความสำคัญของสาระที่ฟัง (Ability to identify the main idea of the listening texts) 2) เข้าใจรายละเอียดของสาระที่ฟัง (Ability to identify some specific details of the listening texts) 3) ระบุหัวข้อของสาระที่ฟัง (Ability to identify topic of the listening texts) 4) ระบุคำศัพท์สำคัญที่เกี่ยวข้องกับหัวเรื่องของสาระที่ฟัง (Ability to recognize key lexical items related to subject/topic) 5) บอกจุดประสงค์ของผู้พูด (Ability to identify purpose of the speaker) 6) ระบุลำดับเนื้อหาของสาระที่ฟัง (Ability to recognize the sequence of the listening texts) 7) บอกชนิดของสาระที่ฟังว่าเป็นคำเตือน ข้อเสนอแนะ คำแนะนำ หรือวิธีการใช้ เป็นต้น (Ability to recognize types of listening texts (e.g., warning, suggestions, recommendations, instructions) 8) บอกความหมายของคำศัพท์โดยอาศัยบริบทแวดล้อม (Ability to deduce meanings of words from contexts) 9) บอกความคิดเห็นของผู้พูดต่อสาระที่ฟัง (Ability to detect the opinion of the speaker toward subject matter) 10) บอกทัศนคติของผู้พูดต่อสาระที่ฟัง (Ability to detect attitude of speaker toward subject matter) 11) ตีความด้านความสัมพันธ์ เช่น ระหว่างคู่สนทนา เหตุและผล (Ability to infer relationships e.g. between people, cause-effect) 12) ตีความเกี่ยวกับเวลา สถานที่ บุคคล ฯลฯ โดยอาศัยข้อมูลที่ฟัง (Ability to infer about time, place, person, etc. based on the listening texts) 13) การระบุคำแนะนำโดยอาศัยความเข้าใจสาระที่ฟัง (Ability to give suggestions based on understanding the listening texts) 14) ทำนายเหตุการณ์โดยอาศัยความเข้าใจสาระที่ฟัง (Ability to make prediction based on understanding the listening texts) และ 15) สรุปความโดยอาศัยข้อมูลที่ฟัง (Ability to draw conclusion based on the listening texts)

2) การฟังเชิงวิชาการกำหนดให้ผู้เข้าสอบฟังหัวเรื่องที่เกี่ยวข้องกับ สิ่งแวดล้อม การดูแลสุขภาพ บ้านและครอบครัว การศึกษา การสื่อสาร เทคโนโลยี อาชีพและการงาน และประวัติศาสตร์ และบทฟังที่จะใช้สร้างแบบทดสอบเป็นบทฟังที่มีผู้พูดคนเดียว (Monologues) ได้แก่ 1) สารคดีทางวิทยุ 2) การพูดสั้น ๆ และ 3) การบรรยาย

3) แหล่งที่มาของบทฟัง ได้มาจากเว็บไซต์ที่ให้ดาวน์โหลดไฟล์เสียงในหัวเรื่องต่าง ๆ ได้ฟรี และผ่านการพิจารณาโดยผู้วิจัยแล้วว่าไม่มีความลำเอียงในด้านหัวเรื่องต่อผู้เข้าสอบทั้ง 3 สาย ได้แก่ สายวิทยาศาสตร์ สายมนุษยศาสตร์ และสายสังคมศาสตร์

4) เทคนิคการสอบหรือรูปแบบข้อคำถามที่สามารถนำมาใช้ทดสอบทักษะการฟังเชิงวิชาการ ประกอบด้วย 1) แบบเลือกตอบ 2) แบบจับคู่ 3) แบบถ่ายโอนข้อมูล เช่น การเติมความในแผนผังแผนที่ (ไดอะแกรม) และ 4) แบบเติมประโยคให้สมบูรณ์

5) การตรวจให้คะแนนจะเป็นรูปแบบการตอบถูกได้ 1 หรือ ตอบผิด ได้ 0

การสรุปผลการศึกษาเนื้อหาสาระของแบบทดสอบด้านการใช้ภาษา

แบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test) ด้านการใช้ภาษา ควรมีจุดประสงค์หลักเพื่อวัดความสามารถของผู้เข้าสอบในการใช้คำศัพท์และโครงสร้างไวยากรณ์ในการสื่อสารในชีวิตประจำวันและเกี่ยวกับงาน และประกอบด้วยเนื้อหา 2 ตอน คือ 1) ความเข้าใจภาษาเขียนมาตรฐาน (Understanding Standard Written Language) และ 2) การค้นหาข้อผิดพลาดในการเขียน (Detecting Written Mistakes) โดยในแต่ละตอนมีกรอบเนื้อหาสาระที่เกี่ยวข้องดังนี้ ดังนี้

ความเข้าใจภาษาเขียนมาตรฐาน (Understanding Standard Written Language)

1) ความเข้าใจภาษาเขียนมาตรฐาน (Understanding Standard Written Language) มีจุดประสงค์เพื่อวัดความสามารถของผู้เข้าสอบในการใช้ภาษาเขียนมาตรฐานในการสื่อสารประจำวัน และเกี่ยวข้องกับงาน และมีสาระภาษา (Language Knowledge/Subskills) ที่เกี่ยวข้อง คือ ความสามารถในการใช้ประเด็นทางภาษาต่าง ๆ ในการสื่อสารได้ถูกต้อง เช่น idioms, collocations, fixed phrases, complementation, phrasal verbs, semantic precision, tenses, two-word verbs, agreement of subject and verb, clauses, types of sentences, parallel structure, comparatives and superlatives, the form of the verbs, the use of verbs, nouns, pronouns, adjectives and adverbs, articles และ prepositions

2) ความเข้าใจภาษาเขียนมาตรฐาน (Understanding Standard Written Language) กำหนดให้ผู้เข้าสอบอ่านบทอ่านในหัวเรื่องเกี่ยวกับสิ่งแวดล้อม สุขภาพ การสื่อสารและเทคโนโลยี การศึกษา และบ้านและครอบครัว โดยมีรูปแบบของบทอ่าน (Text Form) เป็นบทอ่านประเภทให้ข้อมูลข่าวสาร (Informative Texts)

3) บทอ่านที่นำมาใช้ในการทดสอบ อาจจะนำมาจากวารสาร หนังสือพิมพ์ หรือบทความออนไลน์ที่อนุญาตให้ดาวน์โหลดได้ฟรี และผ่านการพิจารณาขั้นต้นโดยผู้วิจัยแล้วว่า ไม่มีความลำเอียงด้านหัวเรื่องต่อผู้เข้าสอบทั้ง 3 สาย ได้แก่ สายวิทยาศาสตร์ สายมนุษยศาสตร์ และสายสังคมศาสตร์

4) เทคนิคการสอบหรือรูปแบบข้อคำถามที่จะนำมาใช้ในการทดสอบความสามารถในการใช้ภาษาเขียนมาตรฐาน คือ การเติมคำในบทอ่านให้สมบูรณ์ โดยให้เลือกคำตอบได้จากตัวเลือก 4 ตัวเลือก

5) การตรวจให้คะแนนจะเป็นรูปแบบการตอบถูกได้ 1 หรือ ตอบผิด ได้ 0

การค้นหาข้อผิดพลาดในการเขียน (Detecting Written Mistakes)

1) การค้นหาข้อผิดพลาดในการเขียน (Detecting Written Mistakes) มีจุดประสงค์เพื่อให้ผู้เข้าสอบระบุข้อผิดพลาดที่ปรากฏในประโยคและเลือกใช้ประโยคที่เหมาะสมสำหรับการสื่อสารในสถานการณ์ต่าง ๆ และมีสาระภาษา (Language Knowledge/Subskills) ที่เกี่ยวข้อง คือ ความสามารถในการ 1) ระบุข้อผิดพลาดด้านการใช้ภาษาในประโยค และ 2) เลือกใช้ประโยคที่เหมาะสมสำหรับการสื่อสารในสถานการณ์ต่าง ๆ

2) การค้นหาข้อผิดพลาดในการเขียน (Detecting Written Mistakes) กำหนดให้ผู้เข้าสอบทำภาระงานสอบ 2 ชนิด คือ 1) อ่านประโยคที่นำมาจากบทความวิชาการในหัวเรื่องที่คุ้นเคย และหาข้อผิดพลาดในการใช้ภาษาในแต่ละประโยค (Error Recognition) และ 2) เลือกประโยคที่ถูกต้องเหมาะสม (Written Expression) สำหรับใช้สื่อสารในชีวิตประจำวันและเกี่ยวข้องกับงาน

3) แหล่งที่มาของข้อมูลในการทดสอบการค้นหาข้อผิดพลาดในการเขียน (Detecting Written Mistakes) มาจาก 2 แหล่ง คือ 1) วารสารหรือแหล่งทรัพยากรออนไลน์ และ 2) หนังสือเรียนที่ใช้สอนในรายวิชาภาษาอังกฤษทั่วไปในระดับมหาวิทยาลัย

4) เทคนิคการสอบหรือรูปแบบข้อคำถามที่จะนำมาใช้ในการทดสอบความสามารถในการค้นหาข้อผิดพลาดในการใช้ภาษา คือ เลือกส่วนที่ขีดเส้นใต้ซึ่งมีจุดผิดพลาดทางการใช้ภาษาในประโยค 1 ใน 4 ส่วน และเลือกประโยคที่ถูกต้องเหมาะสมเพื่อใช้ในการสื่อสารจาก 1 ใน 4 ประโยค

5) การตรวจให้คะแนนจะเป็นรูปแบบการตอบถูกได้ 1 หรือ ตอบผิด ได้ 0

การสรุปผลการศึกษานี้อาสาของแบบทดสอบด้านการอ่าน

แบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test) ด้านการอ่าน ควรจะมีจุดประสงค์หลักเพื่อวัดความสามารถของผู้เข้าสอบในการอ่านสาระที่เกี่ยวข้องกับชีวิตประจำวันและเชิงวิชาการ และประกอบด้วยเนื้อหา 2 ตอน คือ 1) การอ่านในชีวิตประจำวัน และ 2) การอ่านเชิงวิชาการ โดยในแต่ละตอนควรมีกรอบเนื้อหาสาระที่เกี่ยวข้องดังนี้

การอ่านในชีวิตประจำวัน (Reading for Everyday Life)

1) การอ่านในชีวิตประจำวัน (Reading for Everyday Life) มีจุดประสงค์เพื่อวัดความสามารถของผู้เข้าสอบในการอ่านสาระที่เกี่ยวข้องกับชีวิตประจำวันและการทำงาน และมีสาระภาษา (Language Knowledge/Subskills) ที่เกี่ยวข้อง คือ ความสามารถในการ 1) ระบุประเด็นหลักของบทอ่าน (Ability to identify the main idea in the passage) 2) ระบุประเด็นเฉพาะในบทอ่าน (Ability to identify a specific point in the passage) 3) เลือกคำศัพท์ที่มีความหมายเหมือนกับศัพท์ในบทอ่าน (Ability to choose a general synonym for some specific words in the reading texts) 4) อธิบายความหมายของคำศัพท์ให้สอดคล้องกับบริบท (Ability to explain a word that is specific to the reading passage) 5) ค้นหาคำหรือกลุ่มคำที่ใช้อ้างอิงถึงคำหรือกลุ่มคำอื่น ๆ ใน

บทอ่าน (Ability to identify a word or phrase in the passage that refers to another word or phrase) 6) บอกความหมายของคำศัพท์โดยการพิจารณาจากส่วนประกอบของคำ (Ability to determine meanings from word parts) 7) เดาความหมายของคำศัพท์จากบริบท (Ability to use context to determine meanings of words) 8) บอกน้ำเสียงและอารมณ์ของผู้เขียน (Ability to determine the tone and mood of the author) 9) บอกจุดประสงค์ของผู้เขียน (Ability to determine the purpose of the author) 10) เลือกประโยคที่มีความหมายเหมือนกับประโยคในบทอ่าน (Ability to choose the best restatement) 11) อธิบายสาเหตุของบางสิ่งตามที่ได้กล่าวไว้ในบทอ่าน (Ability to explain why something in the passage occurred) 12) บอกความคิดเห็นของผู้เขียน (Ability to recognize the author's point of view) 13) บอกโครงสร้างของการนำเสนอเนื้อหาในบทอ่าน (Ability to recognize the organization of the ideas in the passage) 14) ทำนายเหตุการณ์โดยอาศัยประเด็นสำคัญหรือบริบท (Ability to make a prediction based on the main ideas or contexts) และ 15) หาข้อสรุปโดยอาศัยประเด็นสำคัญหรือบริบทของบทอ่าน (Ability to draw a conclusion by using the main points or contexts in the passage)

2) การอ่านในชีวิตประจำวันกำหนดให้ผู้เข้าสอบอ่านบทอ่านในหัวเรื่องที่คุ้นเคยทั้งในเรื่องส่วนตัวและเกี่ยวข้องกับงาน โดยมีรูปแบบของของบทอ่าน (Text Form) เป็นจดหมายหรืออีเมลส่วนตัวและที่เกี่ยวข้องกับงาน และโฆษณารับสมัครงาน และมีประเภทของเนื้อความ (Text Type) เป็นแบบโน้มน้าวใจ เชิงบรรยาย เชิงพรรณนา และเชิงโต้แย้ง

3) แหล่งที่มาของบทอ่าน ได้มาจากหนังสือพิมพ์ฉบับจริง หนังสือพิมพ์ออนไลน์ โฆษณารับสมัครงาน และตำราเรียนภาษาอังกฤษ และผ่านการพิจารณาโดยผู้วิจัยแล้วว่าไม่มีความลำเอียงในด้านหัวเรื่องต่อผู้เข้าสอบทั้ง 3 สาย ได้แก่ สายวิทยาศาสตร์ สายมนุษยศาสตร์ และสายสังคมศาสตร์

4) เทคนิคการสอบหรือรูปแบบข้อคำถามที่สามารถนำมาใช้ทดสอบทักษะการอ่านของแบบทดสอบวัดความสามารถภาษาอังกฤษควรประกอบด้วย 1) คำถามแบบให้เลือกตอบ 2) เติมคำในประโยค 3) แบบจับคู่ และ 4) แบบระบุถูกหรือผิด

5) การตรวจให้คะแนนจะเป็นรูปแบบการตอบถูกได้ 1 หรือ ตอบผิด ได้ 0

การอ่านเชิงวิชาการ (Reading in Academic Contexts)

1) การอ่านเชิงวิชาการ (Reading in Academic Contexts) มีจุดประสงค์เพื่อวัดความสามารถของผู้เข้าสอบในการอ่านสาระในบริบทวิชาการ และมีสาระภาษา (Language Knowledge/Subskills) ที่เกี่ยวข้อง คือ ความสามารถในการ 1) ระบุประเด็นหลักของบทอ่าน (Ability to identify the main idea in the passage) 2) ระบุประเด็นเฉพาะในบทอ่าน (Ability to identify a specific point in the passage) 3) เลือกคำศัพท์ที่มีความหมายเหมือนกับศัพท์ในบทอ่าน (Ability to choose a general synonym for some specific words in the reading texts) 4) อธิบายความหมายของคำศัพท์ให้สอดคล้องกับบริบท (Ability to explain a word that is specific to the reading passage) 5) ค้นหาคำหรือกลุ่มคำที่ใช้อ้างอิงถึงคำหรือกลุ่มคำอื่น ๆ ในบทอ่าน (Ability to identify

a word or phrase in the passage that refers to another word or phrase) 6) บอกความหมายของคำศัพท์โดยการพิจารณาจากส่วนประกอบของคำ (Ability to determine meanings from word parts) 7) เดาความหมายของคำศัพท์จากบริบท (Ability to use context to determine meanings of words) 8) บอกน้ำเสียงและอารมณ์ของผู้เขียน (Ability to determine the tone and mood of the author) 9) บอกจุดประสงค์ของผู้เขียน (Ability to determine the purpose of the author) 10) เลือกประโยคที่มีความหมายเหมือนกับประโยคในบทอ่าน (Ability to choose the best restatement) 11) อธิบายสาเหตุของบางสิ่งตามที่ได้กล่าวไว้ในบทอ่าน (Ability to explain why something in the passage occurred) 12) บอกความคิดเห็นของผู้เขียน (Ability to recognize the author's point of view) 13) บอกโครงสร้างของการนำเสนอเนื้อหาในบทอ่าน (Ability to recognize the organization of the ideas in the passage) 14) ทำนายเหตุการณ์โดยอาศัยประเด็นสำคัญหรือบริบท (Ability to make a prediction based on the main ideas or contexts) และ 15) หาข้อสรุปโดยอาศัยประเด็นสำคัญหรือบริบทของบทอ่าน (Ability to draw a conclusion by using the main points or contexts in the passage)

2) การอ่านเชิงวิชาการกำหนดให้ผู้เข้าสอบอ่านบทอ่านในหัวเรื่องเกี่ยวกับสิ่งแวดล้อม การศึกษา การดูแลสุขภาพ หรือบ้านและครอบครัว โดยมีรูปแบบของบทอ่าน (Text Form) เป็นบทอ่านประเภทให้ข้อมูลข่าวสารในหัวเรื่องเฉพาะต่าง ๆ บทความในหนังสือพิมพ์ วารสาร และแหล่งทรัพยากรออนไลน์ โฆษณา สื่อส่งเสริมการขายและให้ข้อมูลข่าวสาร และประเภทของเนื้อความ (Text Type) เป็นแบบโน้มน้าวใจ เชิงบรรยาย เชิงพรรณนา และเชิงโต้แย้ง

3) แหล่งที่มาของบทอ่าน ได้มาจากวารสาร หนังสือพิมพ์ และบทความออนไลน์ที่สามารถดาวน์โหลดได้ฟรี และผ่านการพิจารณาโดยผู้วิจัยแล้วว่าไม่มีความลำเอียงในด้านหัวเรื่องต่อผู้เข้าสอบทั้ง 3 สาย ได้แก่ สายวิทยาศาสตร์ สายมนุษยศาสตร์ และสายสังคมศาสตร์

4) เทคนิคการสอบหรือรูปแบบข้อคำถามที่สามารถนำมาใช้ทดสอบทักษะการอ่านของแบบทดสอบวัดความสามารถภาษาอังกฤษควรประกอบด้วย 1) คำถามแบบให้เลือกตอบ 2) เติมคำในประโยค 3) แบบจับคู่ และ 4) แบบระบุถูกหรือผิด

5) การตรวจให้คะแนนจะเป็นรูปแบบการตอบถูกได้ 1 หรือ ตอบผิด ได้ 0

การสรุปผลการศึกษานี้อาสาของแบบทดสอบด้านการเขียน

แบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test) ด้านการเขียน ควรมีจุดประสงค์เพื่อวัดความสามารถของผู้เข้าสอบในการเขียนสื่อสารในชีวิตประจำวันและเกี่ยวข้องกับงาน และประกอบด้วย 2 ตอน คือ 1) การเขียนในระบัยย่อหน้า (Paragraph Writing) และ 2) การเขียนในระดับเรียงความ (Composition Writing) โดยในแต่ละตอนมีกรอบเนื้อหาสาระที่เกี่ยวข้องดังนี้

การเขียนในระดับย่อหน้า (Paragraph Writing)

1) การเขียนในระดับย่อหน้า (Paragraph Writing) มีจุดประสงค์เพื่อวัดความสามารถของผู้เข้าสอบในการเขียนสื่อสารในระดับย่อหน้า และมีสาระภาษา (Language Knowledge/Subskills) ที่เกี่ยวข้องคือ ความสามารถในการ 1) ระบุรายละเอียดสนับสนุน (Supporting Details) ที่สัมพันธ์กับหัวเรื่อง (Topic) 2) จำแนกประโยครายละเอียดสนับสนุน (Supporting Details) ที่ไม่เกี่ยวข้องกับหัวเรื่อง (Topic) ได้ 3) เขียนเนื้อหาได้ครบถ้วนตามที่ภาระงานกำหนด และ 4) ใช้โครงสร้างภาษา ไวยากรณ์ คำศัพท์ และกลุ่มคำเฉพาะได้อย่างถูกต้อง

2) การเขียนในระดับย่อหน้า (Paragraph Writing) กำหนดให้ผู้เข้าสอบทำภาระงาน 2 ชนิด คือ 1) การเขียนโดยอ้อม (Indirect Writing) และ 2) การเขียนโดยตรง (Direct Writing) โดยการเขียนโดยอ้อมกำหนดให้ผู้เข้าสอบหาความสัมพันธ์ระหว่างหัวเรื่อง (Topic) ที่กำหนดให้กับข้อมูลที่เป็นรายละเอียดสนับสนุน (Supporting Details) และจำแนกข้อมูลที่ไม่เกี่ยวข้องกับหัวเรื่องที่กำหนดให้ หัวเรื่องที่น่ามาทดสอบเกี่ยวข้องกับชีวิตประจำวันและเกี่ยวข้องกับงาน ส่วนการเขียนโดยตรง (Direct Writing) กำหนดให้ผู้เข้าสอบเขียนระดับย่อหน้าความยาวอย่างน้อย 250 คำ ในหัวเรื่องที่เกี่ยวข้องกับชีวิตประจำวันและงาน

3) แหล่งที่มาของข้อมูลที่น่ามาใช้ในภาระงานการเขียนโดยอ้อม คือ ย่อหน้าที่ดีจากวารสาร หนังสือสารคดี หรือแหล่งทรัพยากรออนไลน์ที่อนุญาตให้ดาวน์โหลดข้อมูลได้ฟรี

4) เทคนิคการสอบหรือรูปแบบข้อคำถามที่จะนำมาใช้ในการทดสอบความสามารถในการเขียนโดยอ้อม เป็นแบบเลือกตอบ 4 ตัวเลือก และการเขียนโดยตรงเป็นการกำหนดภาระงาน

5) การตรวจให้คะแนนแบ่งเป็น 2 รูปแบบ คือ การเขียนโดยอ้อม (Indirect Writing) มีการให้คะแนนแบบตอบถูกได้ 1 หรือ ตอบผิด ได้ 0 และการเขียนโดยตรง (Direct Writing) จะให้คะแนนตามเกณฑ์การประเมินการเขียนในระดับย่อหน้าตามแนวคิดของ Honaker (2017) ซึ่งให้คะแนนระดับ 1-4 ในเกณฑ์ 5 ประการ คือ 1) ประโยคใจความสำคัญ (Topic Sentence) 2) ประโยคให้รายละเอียดสนับสนุน (Supporting Detail Sentences) 3) ประโยคขยายความรายละเอียด (Elaborating Detail Sentences) 4) ความสามารถอ่านออกได้ (Legibility) และ 5) เครื่องหมายวรรคตอนและไวยากรณ์ (Mechanics and Grammar)

การเขียนในระดับเรียงความ (Composition Writing)

1) การเขียนในระดับเรียงความ (Composition Writing) มีจุดประสงค์เพื่อวัดความสามารถของผู้เข้าสอบในการเขียนสื่อสารในระดับเรียงความ และมีสาระภาษา (Language Knowledge/Subskills) ที่เกี่ยวข้องคือ ความสามารถในการ 1) เขียนเนื้อหาได้ครอบคลุมตามที่ภาระงานกำหนด และ 2) ใช้โครงสร้างภาษา ไวยากรณ์ คำศัพท์ กลุ่มคำเฉพาะ ได้ถูกต้อง

2) การเขียนในระดับเรียงความ (Composition Writing) อาจกำหนดภาระงานให้ผู้เข้าสอบเขียนจดหมายส่วนตัว จดหมายสมัครงาน จดหมายธุรกิจ เรียงความแสดงความคิดเห็นหรือเสนอ

แนวคิดเพื่อแก้ปัญหาต่าง ๆ ตามสถานการณ์ที่กำหนด ในหัวเรื่อง (Topic) เกี่ยวข้องกับชีวิตประจำวัน และการทำงาน

3) เทคนิคการสอบหรือรูปแบบข้อคำถามที่จะนำมาใช้ในการทดสอบความสามารถในการเขียนเรียงความเป็นการกำหนดภาระงาน

4) เนื่องจากการเขียนในระดับเรียงความ กำหนดให้ผู้เข้าสอบเขียนจดหมายสมัครงาน จึงใช้เกณฑ์การให้คะแนนจดหมายธุรกิจตามแนวคิดของ Texas Education Agency (2006) โดยให้คะแนน 4 ระดับ คือ ระดับต้น (Beginning) ระดับกำลังพัฒนา (Developing) ระดับบรรลุเป้าหมาย (Accomplished) และระดับยอดเยี่ยม (Exemplary) ตามเกณฑ์การให้คะแนน 4 ประเด็น คือ เนื้อหา (Content) การเรียงเนื้อหา (Organization) รูปแบบและลักษณะ (Appearance) และการใช้ภาษา (Language Use)

การสรุปผลการศึกษานี้อาสาของแบบทดสอบด้านการพูด

แบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test) ด้านการพูด ครอบคลุมประสงค์เพื่อวัดความสามารถของผู้เข้าสอบในการพูดสื่อสารในชีวิตประจำวันและเชิงวิชาการ และประกอบด้วย 3 ตอน คือ การพูดในระดับประโยค (Sentence Level) การพูดสั้น ๆ (Short Talk) และการพูดเชิงวิชาการ (Academic Speaking) โดยในแต่ละตอนมีกรอบเนื้อหาสาระที่เกี่ยวข้องดังนี้

การพูดในระดับประโยค (Sentence Level)

1) จุดประสงค์การพูดในระดับประโยค (Sentence Level) เพื่อวัดความสามารถของผู้เข้าสอบในการพูดสื่อสารในสถานการณ์ต่าง ๆ ที่พบได้ในชีวิตประจำวันและเกี่ยวข้องกับงาน และมีสาระภาษา (Language Knowledge/Subskills) ที่เกี่ยวข้องคือ ความสามารถในการใช้ภาษาพูดสื่อสารในสถานการณ์ต่าง ๆ ได้ถูกต้องและเหมาะสม

2) สาระที่วัดการพูดในระดับประโยค (Sentence Level) เป็นการพูดโดยอ้อม (Indirect Speaking) ที่กำหนดให้ผู้เข้าสอบตอบคำถามในบทสนทนาถาม-ตอบระหว่างผู้พูด 2 คน ตามหน้าที่ภาษาและสถานการณ์ที่พบได้ในชีวิตประจำวันและเกี่ยวข้องกับงาน หน้าที่ภาษาดังกล่าว ได้แก่ 1) พูดเกี่ยวกับความชอบและความไม่ชอบ (Talking about likes and dislikes) 2) พูดเกี่ยวกับความสนใจ (Talking about interests) บรรยายบุคคล สถานที่ และสิ่งของ (Describing people, places and things) 3) ถามคำถามเพื่อให้ได้ข้อมูลและถามแบบให้ตอบใช่หรือไม่ใช่ (Asking yes / no and information questions) 4) บอกความต่าง ความเหมือนของบุคคล สถานที่ และสิ่งของ (Comparing people, places and things) 5) แนะนำตัวเองและให้ข้อมูลส่วนตัว (Introducing oneself and giving personal information) 6) สั่งอาหารในร้านอาหาร (Ordering food in a restaurant) 7) บอกความสามารถ (Expressing abilities) 8) ทำนายเหตุการณ์ (Making predictions) 9) เชื่อมโยงถึงเหตุการณ์ในอดีต (Relating past events) 10) ถามความคิดเห็น

(Asking for opinions) และแสดงความคิดเห็น (Expressing opinions) 11) แสดงความชอบมากกว่า (Showing preferences) 12) ให้ข้อเสนอแนะ (Making suggestions) 13) ตอบรับต่อข้อเสนอแนะ (Responding to suggestions) 14) ยอมรับหรือปฏิเสธ (Accepting or declining) 15) ขอหรือให้คำแนะนำ (Asking for and giving advice) 16) แสดงการเห็นด้วย (Expressing agreement) 17) แสดงการไม่เห็นด้วย (Expressing disagreement) 18) ขอความช่วยเหลือ (Asking for a favor) 19) ให้และรับของขวัญ (Giving and receiving presents) 20) ขอข้อมูล (Asking for information) 21) ขออนุญาต (Asking for permission) 22) ร้องเรียน (Making complaints) 23) คาดเดา (Guessing) 24) ตอบปฏิเสธอย่างสุภาพ (Saying “No” politely) 25) เสนอให้ความช่วยเหลือ (Offering help) 26) บอกคำเตือน (Giving warning) 27) ขอคำอธิบาย (Demanding explanations) 28) ขอความกระจ่างชัด (Asking for clarifications) 29) ให้ความกระจ่างชัด (Giving clarifications) 30) ขัดจังหวะ (Interrupting) 31) ขอรายละเอียดเพิ่มเติม (Asking for more detailed information) 32) วางแผน (Making plans) 33) แสดงระดับความแน่นอนหรือไม่แน่นอน (Expressing degrees of certainty and uncertainty) 34) ขอคำบรรยาย (Asking for descriptions) 35) ให้คำบรรยาย (Giving descriptions) 36) แสดงความผิดหวัง (Expressing disappointment) 37) กล่าวขออภัย (Making requests) 38) กล่าวขอโทษ (Apologizing) และ 39) แจ้งข่าวร้าย (Giving bad news)

3) เทคนิคการสอบหรือรูปแบบข้อคำถามที่จะนำมาใช้ในการทดสอบความสามารถในการพูด โดยอ้อม เป็นแบบเลือกตอบ 4 ตัวเลือก

4) การตรวจให้คะแนน เป็นรูปแบบการตอบถูกได้ 1 หรือ ตอบผิด ได้ 0

การพูดสั้น ๆ (Short Talk)

1) จุดประสงค์การพูดสั้น ๆ (Short Talk) เพื่อวัดความสามารถของผู้เข้าสอบในการพูดสื่อสารสั้น ๆ ในหัวข้อ (Topic) ส่วนตัวหรือเกี่ยวข้องกับงาน โดยมีความยาวประมาณ 3 นาที และมีสาระภาษา (Language Knowledge /Subskills) ที่เกี่ยวข้องคือ ความสามารถในการ 1) ใช้ภาษาพูดสื่อสารตามภาระงานที่กำหนดได้ถูกต้องและเหมาะสม และ 2) จัดระบบความคิดและสื่อสารออกมาได้อย่างคล่องแคล่ว

2) ภาระงานการพูดสั้น ๆ (Short Talk) ให้ผู้เข้าสอบบรรยายรูปภาพ โดยผู้เข้าสอบ 1) รับใบงาน 2) อ่านใบงานให้เข้าใจและเตรียมตัว 2 นาที และ 3) พูดบรรยาย แสดงความคิดเห็น หรือให้ข้อเสนอแนะที่เกี่ยวข้องกับรูปภาพ

3) รูปแบบการทดสอบหรือข้อคำถาม เป็นลักษณะของภาระงาน

4) การพูดบรรยายภาพใช้เกณฑ์การให้คะแนนตามแนวคิดของ English Language Institute (2017) โดยให้คะแนนในระดับ 0-4 ตามเกณฑ์ 3 ประเด็น คือ 1) เนื้อหาสาระ (Content) 2) ความถูกต้อง (Accuracy) และ 3) การออกเสียง (Pronunciation)

การพูดเชิงวิชาการ (Academic Speaking)

1) จุดประสงค์การพูดเชิงวิชาการ (Academic Speaking) เพื่อวัดความสามารถของผู้เข้าสอบในการพูดสื่อสารเชิงวิชาการในหัวข้อที่เป็นประเด็นร่วมสมัย มีความยาวประมาณ 5-7 นาที และผู้เข้าสอบต้อง 1) วิเคราะห์และอภิปรายความคิดของตนเองในหัวข้อที่กำหนด และ 2) แสดงความคิดเห็นและให้เหตุผล ยกตัวอย่าง สนับสนุนความคิดเห็นของตนเอง

2) สาระการพูดเชิงวิชาการ (Academic Speaking) กำหนดให้ผู้เข้าสอบอ่านใบงาน และพูดอธิบาย อภิปราย หรือนำเสนอ ตามหัวข้อที่กำหนด

3) รูปแบบการทดสอบหรือข้อคำถาม เป็นลักษณะของภาระงาน

4) การพูดเชิงวิชาการใช้แนวทางการให้คะแนนตามแนวคิดของ Study.com (2017) ให้คะแนน เป็นระดับ 1-5 ตามเกณฑ์การให้คะแนน 4 ประเด็น คือ 1) ความคล่องแคล่ว (Fluency) 2) การออกเสียงและสำเนียง (Pronunciation and Accent) 3) คำศัพท์ (Vocabulary) 4) ไวยากรณ์ (Grammar)

5) รายละเอียดตามภาระงาน (Task Details)

3.5.1.2 การกำหนดกรอบเนื้อหาสาระของแบบทดสอบ

ในการกำหนดกรอบเนื้อหาสาระของแบบทดสอบที่จะพัฒนา ผู้วิจัยได้กำหนดเนื้อหาสาระแบบทดสอบแต่ละทักษะให้คู่ขนานกับเนื้อหาสาระของแบบทดสอบมาตรฐานที่จะใช้เป็นตัวเทียบเคียงคะแนนจากการสอบ แต่ยังคงมีเนื้อหาสาระเดิมที่ได้จากการสังเคราะห์ข้อมูลอยู่ด้วย โดยในหัวข้อนี้จะนำเสนอเนื้อหาสาระของแบบทดสอบมาตรฐานแต่ละทักษะควบคู่กับเนื้อหาสาระของแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัยที่ผู้วิจัยจะพัฒนา ได้แก่ ทักษะการฟัง การใช้ภาษา การอ่าน การเขียน และการพูด โดยมีรายละเอียดดังนี้

กรอบเนื้อหาสาระของแบบทดสอบด้านการฟัง

เนื่องจากผู้วิจัยจะใช้แบบทดสอบด้านการฟังของ TOEFL (Educational Testing Service, 2014) เป็นแบบทดสอบมาตรฐานในการเทียบเคียงผลคะแนนกับแบบทดสอบด้านการฟังของแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย การกำหนดเนื้อหาสาระของแบบทดสอบด้านการฟังจึงมีลักษณะคู่ขนานกันทั้งด้านจุดประสงค์ หัวเรื่อง แหล่งที่มาของข้อมูล เทคนิคการทดสอบ และวิธีการให้คะแนน ดังแสดงในรายละเอียดต่อไปนี้

แบบทดสอบด้านการฟังของแบบทดสอบมาตรฐาน

1) จุดประสงค์การสอบ แบบทดสอบด้านการฟังของ TOEFL (Educational Testing Service, 2014) มีจุดประสงค์ เพื่อวัดความสามารถของผู้เข้าสอบในการฟังสาระที่เกี่ยวข้องกับ

ชีวิตประจำวันและเป็นที่ยึดเหนี่ยว และมีสาระภาษา (Language Knowledge/Subskills) ที่เกี่ยวข้อง คือ ความสามารถในการ Identify 1) the meaning of what the second speaker said, 2) the synonyms of the difficult words, 3) words having similar sounds, 4) the passive forms, 5) Who and What with multiple nouns, 6) negative expressions, 7) double negative expressions, 8) “almost negative” expressions, 9) negative with comparatives, expressions of agreement, 10) expressions of uncertainty and suggestion, 11) emphatic expressions of surprise, 12) untrue conditions, 13) two-and three- part words, 14) idioms, and 15) draw conclusions about who, what, where, 16) to anticipate the topics, 17) to anticipate the questions, 18) to determine the topic, และ 19) to draw conclusions about who, what, when, where.

2) หัวเรื่องที่ใช้ในการทดสอบ มักจะเกี่ยวข้องกับ เรื่องราวในชีวิตประจำวันและเป็นที่ยึดเหนี่ยว วิทยาศาสตร์ และประวัติศาสตร์

3) แหล่งที่มาของข้อมูล เป็น บทสนทนาสั้น ๆ บทสนทนายาว และการบรรยายหรืออธิบาย

4) เทคนิคการทดสอบ เป็นแบบเลือกตอบ 4 ตัวเลือก

5) การให้คะแนน เป็นรูปแบบการตอบถูกได้ 1 หรือ ตอบผิด ได้ 0

แบบทดสอบด้านการฟังของแบบทดสอบวัดความสามารถภาษาอังกฤษระดับ มหาวิทยาลัย

1) จุดประสงค์ แบบทดสอบด้านการฟังของแบบทดสอบวัดความสามารถภาษาอังกฤษระดับ มหาวิทยาลัยมีจุดประสงค์เพื่อวัดความสามารถของผู้เข้าสอบในการฟังสาระที่เกี่ยวข้องกับ ชีวิตประจำวันและเป็นที่ยึดเหนี่ยว และมีสาระภาษา (Language Knowledge/Subskills) ที่เกี่ยวข้อง คือ ความสามารถในการ Identify 1) the meaning of what the second speaker said, 2) the synonyms of the difficult words, 3) words having similar sounds, 4) the passive forms, 5) Who and What with multiple nouns, 6) negative expressions, 7) double negative expressions, 8) “almost negative” expressions, 9) negative with comparatives, expressions of agreement, 10) expressions of uncertainty and suggestion, 11) emphatic expressions of surprise, 12) untrue conditions, 13) two-and three- part words, 14) idioms, and 15) draw conclusions about who, what, where, 16) to anticipate the topics, 17) to anticipate the questions, 18) to determine the topic, และ 19) to draw conclusions about who, what, when, where.

2) หัวเรื่องใช้ในการทดสอบ มักจะเกี่ยวข้องกับ เรื่องราวในชีวิตประจำวัน เป็นที่ยึดเหนี่ยว เกี่ยวข้องกับงาน เรื่องราวส่วนตัว ธุรกิจ และวิทยาศาสตร์

3) แหล่งที่มาของบทฟัง เป็น บทสนทนาสั้น ๆ บทสนทนายาว และการบรรยายหรืออธิบาย

4) เทคนิคการทดสอบ เป็นแบบเลือกตอบ 4 ตัวเลือก

5) การให้คะแนน เป็นรูปแบบการตอบถูกได้ 1 หรือ ตอบผิด ได้ 0

กรอบเนื้อหาสาระของแบบทดสอบด้านการใช้ภาษา

เนื่องจากผู้วิจัยจะใช้แบบทดสอบด้านการใช้ภาษาของ TOEFL (Educational Testing Service, 2014) เป็นแบบทดสอบมาตรฐานในการเทียบเคียงผลคะแนนกับแบบทดสอบด้านการใช้ภาษาของแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย การกำหนดเนื้อหาสาระของแบบทดสอบด้านการใช้ภาษาจึงมีลักษณะคู่ขนานกันทั้งด้านจุดประสงค์ หัวเรื่อง แหล่งที่มาของข้อมูล เทคนิคการทดสอบ และวิธีการให้คะแนน ดังแสดงในรายละเอียดต่อไปนี้

แบบทดสอบด้านการใช้ภาษาของแบบทดสอบมาตรฐาน

1) จุดประสงค์การสอบ แบบทดสอบด้านการใช้ภาษาของ TOEFL (Educational Testing Service, 2014) มีจุดประสงค์ เพื่อวัดความสามารถของผู้เข้าสอบในความเข้าใจภาษาเขียนมาตรฐาน และความสามารถในการระบุข้อผิดพลาดในการใช้ภาษา และมีสาระภาษา (Language Knowledge/Subskills) ที่เกี่ยวข้อง คือ 1) Be sure the sentence has a subject and a verb, 2) Be careful of objects of prepositions, 3) Be careful of present participles, 4) Use reduced adjective clauses correctly, 5) Use adverb Time and Cause connectors correctly, 6) Use noun clause connector/subjects correctly, 7) Use other adverb connectors correctly, 8) Invert the subject and verb with place expressions, 9) Be careful of past participles, 10) Use noun clause connectors correctly, 11) Invert the subject and verb with conditionals, 12) Use reduced adverb clauses correctly, 13) Invert the subject and verb with negatives, 14) Use adjective clause connectors correctly, 15) Invert the subject and verb with comparisons, 16) After be, use the present participle or the past participle, 17) Form comparatives and superlatives correctly, 18) Make verbs agree after certain words, 19) After Will, Would, or other Modals, use the base form of the verb, 20) Use the correct tense with time expressions, 21) Distinguish subject and object pronouns, 22) Distinguish possessive adjectives and pronouns, 23) Use parallel structure with paired conjunctions, 24) Make verb agree after prepositional phrases, 25) Use comparatives and superlatives correctly, 26) Distinguish A and An, 27) Use the correct tense with Will and Would, 28) Distinguish Like, Alike, Unlike, and Dislike, 29) Make inverted verbs agree, 30) Distinguish countable and uncountable nouns, 31) Distinguish specific and general ideas, 32) Position adjectives and adverbs correctly, 33) Recognize active and passive meanings, 34) Use articles with singular nouns, 35) Use predicate adjectives correctly, 36) Distinguish Make and Do, 37) Use the correct form of the passive, 38) Recognize when preposition have been omitted, 39) Distinguish the person from the thing, และ 40) Use adjectives after linking verbs

- 2) หัวเรื่องที่ใช้ในการทดสอบ เป็นเนื้อหาเชิงวิชาการ
- 3) แหล่งที่มาของข้อมูล เป็น ประโยคสั้น ๆ
- 4) เทคนิคการทดสอบ เป็นแบบเลือกตอบ 4 ตัวเลือก
- 5) การให้คะแนน เป็นรูปแบบการตอบถูกได้ 1 หรือ ตอบผิด ได้ 0

แบบทดสอบด้านการใช้ภาษาของแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย

1) จุดประสงค์ แบบทดสอบด้านการใช้ภาษาของ แบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย มีจุดประสงค์เพื่อวัดความสามารถของผู้เข้าสอบในความเข้าใจภาษาเขียนมาตรฐานและความสามารถในการระบุข้อผิดพลาดในการใช้ภาษา และมีสาระภาษา (Language Knowledge/Subskills) ที่เกี่ยวข้อง คือ 1) Be sure the sentence has a subject and a verb, 2) Be careful of objects of prepositions, 3) Be careful of present participles, 4) Use reduced adjective clauses correctly, 5) Use adverb Time and Cause connectors correctly, 6) Use noun clause connector/subjects correctly, 7) Use other adverb connectors correctly, 8) Invert the subject and verb with place expressions, 9) Be careful of past participles, 10) Use noun clause connectors correctly, 11) Invert the subject and verb with conditionals, 12) Use reduced adverb clauses correctly, 13) Invert the subject and verb with negatives, 14) Use adjective clause connectors correctly, 15) Invert the subject and verb with comparisons, 16) After be, use the present participle or the past participle, 17) Form comparatives and superlatives correctly, 18) Make verbs agree after certain words, 19) After Will, Would, or other Modals, use the base form of the verb, 20) Use the correct tense with time expressions, 21) Distinguish subject and object pronouns, 22) Distinguish possessive adjectives and pronouns, 23) Use parallel structure with paired conjunctions, 24) Make verb agree after prepositional phrases, 25) Use comparatives and superlatives correctly, 26) Distinguish A and An, 27) Use the correct tense with Will and Would, 28) Distinguish Like, Alike, Unlike, and Dislike, 29) Make inverted verbs agree, 30) Distinguish countable and uncountable nouns, 31) Distinguish specific and general ideas, 32) Position adjectives and adverbs correctly, 33) Recognize active and passive meanings, 34) Use articles with singular nouns, 35) Use predicate adjectives correctly, 36) Distinguish Make and Do, 37) Use the correct form of the passive, 38) Recognize when preposition have been omitted, 39) Distinguish the person from the thing, และ 40) Use adjectives after linking verbs

- 2) หัวเรื่องที่ใช้ในการทดสอบ เป็นเนื้อหาเชิงวิชาการ

- 3) แหล่งที่มาของข้อมูลเป็นประโยคสั้น ๆ
- 4) เทคนิคการทดสอบ เป็นแบบเลือกตอบ 4 ตัวเลือก
- 5) การให้คะแนน เป็นรูปแบบการตอบถูกได้ 1 หรือ ตอบผิด ได้ 0

กรอบเนื้อหาสาระของแบบทดสอบด้านการอ่าน

เนื่องจากผู้วิจัยจะใช้แบบทดสอบด้านการอ่านของ TOEFL (Educational Testing Service, 2014) เป็นแบบทดสอบมาตรฐานในการเทียบเคียงผลคะแนนกับแบบทดสอบด้านการอ่านของแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย การกำหนดเนื้อหาสาระของแบบทดสอบด้านการอ่านจึงมีลักษณะคู่ขนานกันทั้งด้านจุดประสงค์ หัวเรื่อง แหล่งที่มาของข้อมูล เทคนิคการทดสอบ และวิธีการให้คะแนน ดังแสดงในรายละเอียดต่อไปนี้

แบบทดสอบด้านการอ่านของแบบทดสอบมาตรฐาน

- 1) จุดประสงค์การสอบ แบบทดสอบด้านการอ่านของ TOEFL (Educational Testing Service, 2014) มีจุดประสงค์วัดความสามารถของผู้เข้าสอบในการอ่านสาระเชิงวิชาการ และมีสาระภาษา (Language Knowledge/Subskills) ที่เกี่ยวข้อง คือ 1) To answer main idea question correctly, 2) To use context to determine meanings of simple words, 3) To answer transition questions correctly, 4) To answer stated detail questions correctly, 5) To find pronoun referents, 6) To determine meanings from word parts, 7) To answer implied detail question correctly และ 8) To determine where specific information is found
- 2) หัวเรื่องที่ใช้ในการทดสอบ เกี่ยวข้องกับเทคโนโลยี ศิลปะ สัตว์ ธรรมชาติ และคน
- 3) แหล่งที่มาของข้อมูล เป็น บทอ่านสั้น ๆ จากตำรา หนังสือ เชิงวิชาการ
- 4) เทคนิคการทดสอบ เป็นแบบเลือกตอบ 4 ตัวเลือก
- 5) การให้คะแนน เป็นรูปแบบการตอบถูกได้ 1 หรือ ตอบผิด ได้ 0

แบบทดสอบด้านการอ่านของแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย

จุดประสงค์ แบบทดสอบด้านการอ่านของแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย มีจุดประสงค์วัดความสามารถของผู้เข้าสอบในการอ่านสาระในชีวิตประจำวัน และเชิงวิชาการ และมีสาระภาษา (Language Knowledge/Subskills) ที่เกี่ยวข้อง คือ 1) To answer stated detail questions correctly, 2) To find “unstated” details, 3) To answer an implied detailed question correctly, 4) To determine meanings from word parts, 5) To choose a general synonym for some specific words in the reading texts, 6) To use

context to determine meanings of simple words, 7) To answer main idea question correctly, 8) To find pronoun referents, 9) To determine where specific information is found, 10) To recognize the organization of the ideas in the passage, 11) To draw a conclusion by using the main points or contexts in the passage, 12) To determine the purpose of the author, 13) To choose a general synonym for some specific words in the reading texts, 14) To determine the tone and mood of the author, และ 15) To answer transition questions correctly

2) หัวเรื่องที่ใช้ในการทดสอบ เกี่ยวข้องกับ สุขภาพ การงาน บ้าน การเรียน เทคโนโลยี เศรษฐศาสตร์ ความรัก ชีวิต และครอบครัว

3) แหล่งที่มาของบทอ่าน บทอ่านสั้น ๆ จากวารสาร สื่อสิ่งพิมพ์ที่พบได้ในชีวิตประจำวัน และตำราด้านวิชาการ

4) เทคนิคการทดสอบ เป็นแบบเลือกตอบ 4 ตัวเลือก

5) การให้คะแนน เป็นรูปแบบการตอบถูกได้ 1 หรือ ตอบผิด ได้ 0

กรอบเนื้อหาของสาระของแบบทดสอบด้านการเขียน

เนื่องจากผู้วิจัยจะใช้แบบทดสอบด้านการเขียนของ IELTS (IELTS, 2015) เป็นแบบทดสอบมาตรฐานในการเทียบเคียงผลคะแนนกับแบบทดสอบด้านการเขียนของแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย การกำหนดเนื้อหาของสาระของแบบทดสอบด้านการเขียน จึงมีลักษณะคู่ขนานกันทั้งด้านจุดประสงค์ หัวเรื่อง แหล่งที่มาของข้อมูล เทคนิคการทดสอบ และวิธีการให้คะแนน ดังแสดงในรายละเอียดต่อไปนี้

แบบทดสอบด้านการเขียนของแบบทดสอบมาตรฐาน

1) จุดประสงค์การสอบ แบบทดสอบด้านการเขียนของ IELTS (IELTS, 2015) มีจุดประสงค์เพื่อวัดความสามารถของผู้เข้าสอบในการเขียนเชิงวิชาการ และมีสาระภาษา (Language Knowledge/Subskills) ที่เกี่ยวข้อง คือ 1) the ability to describe the information objectively; 2) the ability to contrast and compare the information; 3) the ability to use specialized graphics language, not to express your own opinion and don't evaluate the data, 4) the ability to make solutions; 5) the ability to express your opinions; 6) the ability to contrast and compare the opinions and facts; 7) the ability to evaluate facts, ideas or arguments, 8) the ability to organize ideas and present them in a well-organized writing. และ 9) the ability to produce a writing task with no mechanical errors.

2) หัวเรื่องที่ใช้ในการทดสอบ มักจะเกี่ยวข้องกับหัวเรื่องที่คุ้นเคย

3) แหล่งที่มาของข้อมูล เป็น กราฟ แผนภูมิ ชาร์ต ประเด็นหัวเรื่องที่น่าสนใจ

- 4) เทคนิคการทดสอบ การเขียนตอบเป็นย่อหน้า หรือเป็นเนื้อเรื่องที่มีความยาว
- 5) การให้คะแนน ให้คะแนนตามเกณฑ์การตรวจงานเขียน ได้แก่ 1) Content 2) Organization, 3) Language และ 4) Mechanics

แบบทดสอบด้านการเขียนของแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย

1) จุดประสงค์ แบบทดสอบด้านการเขียนของแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย มีจุดประสงค์เพื่อวัดความสามารถของผู้เข้าสอบในการเขียนเชิงวิชาการ และมีสาระภาษา (Language Knowledge/Subskills) ที่เกี่ยวข้อง คือ 1) the ability to describe the information objectively; 2) the ability to contrast and compare the information; 3) the ability to use specialized graphics language, not to express your own opinion and don't evaluate the data, 4) the ability to make solutions; 5) the ability to express your opinions; 6) the ability to contrast and compare the opinions and facts; 7) the ability to evaluate facts, ideas or arguments, 8) the ability to organize ideas and present them in a well-organized writing. และ 9) the ability to produce a writing task with no mechanical errors.

- 2) หัวเรื่องที่ใช้ในการทดสอบ มักจะเกี่ยวข้องกับหัวเรื่องที่คุ้นเคย
- 3) แหล่งที่มาของข้อมูล เป็น กราฟ แผนภูมิ ชาร์ต ประเด็นหัวเรื่องที่น่าสนใจ
- 4) เทคนิคการทดสอบ การเขียนตอบเป็นย่อหน้า หรือเป็นเนื้อเรื่องที่มีความยาว
- 5) การให้คะแนน ให้คะแนนตามเกณฑ์การตรวจงานเขียน ได้แก่ 1) Content 2) Organization, 3) Language และ 4) Mechanics

กรอบเนื้อหาสาระของแบบทดสอบด้านการพูด

เนื่องจากผู้วิจัยจะใช้แบบทดสอบด้านการพูดของ IELTS (IELTS, 2015) เป็นแบบทดสอบมาตรฐานในการเทียบเคียงผลคะแนนกับแบบทดสอบด้านการพูดของแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย การกำหนดเนื้อหาสาระของแบบทดสอบด้านการพูดจึงมีลักษณะคู่ขนานกันทั้งด้านจุดประสงค์ หัวเรื่อง แหล่งที่มาของข้อมูล เทคนิคการทดสอบ และวิธีการให้คะแนน ดังแสดงในรายละเอียดต่อไปนี้

แบบทดสอบด้านการพูดของแบบทดสอบมาตรฐาน

1) จุดประสงค์การสอบ แบบทดสอบด้านการพูดของ IELTS (IELTS, 2015) มีจุดประสงค์เพื่อวัดความสามารถของผู้เข้าสอบในการพูดสื่อสารข้อมูลส่วนตัว แสดงความคิดเห็นและพูดเชิงวิชาการและมีสาระภาษา (Language Knowledge/Subskills) ที่เกี่ยวข้อง คือ 1) To answer

personal questions and give information on familiar topics, 2) To give a 2-minute-talk about how to solve a problem specified in the task card, และ 3) To express opinions and discuss some possible ways to solve the problem in Task 2 effectively by giving reasons and asking the examiner to share the ideas.

2) หัวเรื่องที่ใช้ในการทดสอบ เป็นหัวเรื่องที่คุ้นเคย

3) แหล่งที่มาของข้อมูล เป็น บทสนทนาสั้น ๆ บทสนทนายาว และการบรรยายหรืออธิบาย

4) เทคนิคการทดสอบ เป็นพูดกับผู้ทดสอบหรือให้คะแนน

5) การให้คะแนน ให้คะแนนตามเกณฑ์ 4 ประการ คือ 1) Fluency and Coherence, 2) Lexical Resource, 3) Grammatical Range and Accuracy, และ 4) Pronunciation

แบบทดสอบด้านการพูดของแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย

1) จุดประสงค์แบบทดสอบด้านการพูดของแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย มีจุดประสงค์ เพื่อวัดความสามารถของผู้เข้าสอบในการพูดสื่อสารข้อมูลส่วนตัว แสดงความคิดเห็นและพูดเชิงวิชาการ และมีสาระภาษา (Language Knowledge/Subskills) ที่เกี่ยวข้อง คือ 1) To answer personal questions and give information on familiar topics, 2) To give a 2-minute-talk about how to solve a problem specified in the task card, และ 3) To express opinions and discuss some possible ways to solve the problem in Task 2 effectively by giving reasons and asking the examiner to share the ideas.

2) หัวเรื่องที่ใช้ในการทดสอบ เป็นหัวเรื่องที่คุ้นเคย

3) แหล่งที่มาของข้อมูล เป็น บทสนทนาสั้น ๆ บทสนทนายาว และการบรรยายหรืออธิบาย

4) เทคนิคการทดสอบ เป็นพูดกับผู้ทดสอบหรือให้คะแนน

5) การให้คะแนน ให้คะแนนตามเกณฑ์ 4 ประการ คือ 1) Fluency and Coherence, 2) Lexical Resource, 3) Grammatical Range and Accuracy, และ 4) Pronunciation

3.5.2 ชั้นปฏิบัติการออกข้อสอบ (Test Operationalization)

ในชั้นปฏิบัติการของการวิจัยครั้งนี้ ผู้วิจัยได้แบ่งการดำเนินงานออกเป็น 2 ระยะ คือ ระยะที่ 1 การสร้างและพัฒนาเครื่องมือในการวิจัย และ ระยะที่ 2 การทดลองใช้เครื่องมือและหาคุณภาพของเครื่องมือ โดยมีรายละเอียดในแต่ละระยะดังนี้

3.5.2.1 การสร้างและพัฒนาเครื่องมือในการวิจัย

เครื่องมือที่เกี่ยวข้องกับการวิจัยครั้งนี้ ได้แก่ 1) แบบทดสอบวัดความสามารถภาษาอังกฤษ 2) แบบตรวจให้คะแนนแบบทดสอบอัตนัย 3) แบบประเมินคุณภาพแบบทดสอบ และ 4) แบบสอบถามความคิดเห็นผู้ใช้แบบทดสอบ โดยแต่ละอย่างมีลำดับการพัฒนา ดังนี้

1) แบบทดสอบวัดความสามารถภาษาอังกฤษ

แบบทดสอบวัดความสามารถภาษาอังกฤษที่ใช้ในการวิจัยครั้งนี้ประกอบด้วย 2 ชนิด คือ 1) แบบทดสอบวัดความสามารถภาษาอังกฤษมาตรฐาน ซึ่งผู้วิจัยนำมาใช้และให้กลุ่มตัวอย่างสอบเพื่อให้ได้คะแนนมาเทียบเคียงคุณภาพกับแบบทดสอบที่พัฒนาขึ้น และ 2) แบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test) ที่ผู้วิจัยพัฒนาขึ้น โดยสร้างให้มีคุณลักษณะของแบบทดสอบแต่ละทักษะคู่ขนานกับแบบทดสอบมาตรฐาน และจะนำเสนอตามลำดับของการจัดสอบจริง คือ ทักษะการฟัง การใช้ภาษา การอ่าน การเขียน และการพูด โดยแต่ละทักษะมีลำดับในการพัฒนาดังนี้

1.1) การจัดทำโครงสร้างกำหนดลักษณะแบบทดสอบ (Test Blueprint)

แบบทดสอบวัดความสามารถด้านการฟัง

เนื่องจากผู้วิจัยต้องการที่จะสร้างแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัยด้านการฟังให้มีลักษณะคู่ขนานกับแบบทดสอบการฟังของแบบทดสอบ TOEFL ดังรายละเอียดที่ได้กำหนดไว้แล้วในหัวข้อ 3.5.1.2 และสามารถจัดทำเป็นโครงสร้างกำหนดลักษณะแบบทดสอบ (Test Blueprint) ด้านการฟังแต่ละชนิดได้ดังนี้

ตารางที่ 8 Test Blueprint for Listening Test (TOEFL)

Total Score: 50 points

Time: 60 minutes

Part	Test Objectives	Topics	Language Content	Text Type	Text Form	Skill Level	Test Type	Weight/ Items	Scoring Criteria	Data Source
A	<p>Short Conversations (30 items)</p> <p>Amplified Objective: After listening to 30 short conversations on familiar and daily-life related topics, the test-takers are able to answer 30 M/C questions.</p> <p>Identify the meaning of what the second speaker said, the synonyms of the difficult words, words having similar sounds, the passive forms, Who and What with multiple nouns, negative expressions, double negative expressions, “almost negative” expressions, negative with comparatives, expressions of agreement, expressions of uncertainty and suggestion, emphatic expressions of surprise, untrue conditions, two-and three- part words, idioms, and draw conclusions about who, what, where.</p>	<p>Familiar and daily-life related</p>	<p>Listening for the second line, synonyms, similar sounds, questions about who – what - where, Who and What in passives, Who and What with multiple nouns, negative expressions, double negative expressions, “Almost negative” expressions, negative expressions, with comparatives, expressions of agreement, expressions of uncertainty and suggestion, emphatic expressions of</p>	<p>Asking for information Requesting Giving suggesting Expressing opinions</p>	<p>Dialogue</p>	<p>Comprehension</p>	<p>M/C</p>	<p>30/30</p>	<p>0-1</p>	<p>(Educational Testing Service, 2014)</p>

Part	Test Objectives	Topics	Language Content	Text Type	Text Form	Skill Level	Test Type	Weight/ Items	Scoring Criteria	Data Source	
B			surprise, untrue conditions, two-and three- part words, and idioms.								
	Long Conversations (8 items)										
	Amplified Objective: After listening to 2 long conversations on familiar and daily-life related topics, the test-takers are able to answer 8 M/C questions.										
	To anticipate the topics To anticipate the questions. Determine the topic Draw conclusions about who, what, when, where.	Familiar and daily-life related	Listening for main ideas and details Making inference Summarizing	Asking for information Asking for opinions	Dialogue	Comprehension	M/C	8/8	0-1	(Educational Testing Service, 2014)	
C											
	Talks and Lectures (12 items)										
	Amplified Objective: After listening to 3 talks on familiar and daily-life related topics, the test-takers are able to answer 12 M/C questions.										
	To anticipate the topics To anticipate the questions. Determine the topic Draw conclusions about who, what, when, where.	Daily-life Science History	Listening for main ideas and details Making inference Summarizing	Giving information	Monologue	Comprehension Analysis	M/C	12/12	0-1	(Educational Testing Service, 2014)	

ตารางที่ 9 Test Blueprint for Listening (University English Proficiency Test)

Total Score: 50 points

Time: 60 minutes

Part	Test Objectives	Topics	Language Content	Text Type	Text Form	Skill Level	Test Type	Weight/ Items	Scoring Criteria	Data Source
A	<p>Short Conversations (30 items)</p> <p>Amplified Objective: After listening to 30 short conversations on familiar, daily-life, and work related topics, the test-takers are able to answer 30 M/C questions.</p> <p>Identify the meaning of what the second speaker said, the synonyms of the difficult words, words having similar sounds, the passive forms, Who and What with multiple nouns, negative expressions, double negative expressions, “almost negative” expressions, negative with comparatives, expressions of agreement, expressions of uncertainty and suggestion, emphatic expressions of surprise, untrue conditions, two-and three- part words, idioms, and draw conclusions about who, what, where.</p>	<p>Familiar, daily-life, and work related</p>	<p>Listening for the second line, synonyms, similar sounds, questions about who – what – where, Who and What in passives, Who and What with multiple nouns, negative expressions, double negative expressions, “Almost negative” expressions, negative expressions with comparatives, expressions of agreement, expressions of uncertainty and suggestion, emphatic</p>	<p>Asking for information Requesting Giving suggesting Expressing opinions</p>	<p>Dialogue</p>	<p>Comprehension</p>	<p>M/C</p>	<p>30/30 (49/49)</p>	<p>0-1</p>	

Part	Test Objectives	Topics	Language Content	Situations	Text Form	Skill Level	Test Type	Weight/ Items	Scoring Criteria	Data Source	
B			expressions of surprise, untrue conditions, two-and three- part words, and idioms.								
	Long Conversations (8 items)										
	Amplified Objective: After listening to 2 long conversations on familiar and work related topics, the test-takers are able to answer 8 M/C questions.										
	To anticipate the topics To anticipate the questions. Determine the topic Draw conclusions about who, what, when, where.	Familiar and work related topics	Listening for main ideas and details Making inference Summarizing	Seeing the doctor Interviewing	Dialogue	Comprehension	M/C	8/8 (16/16)	0-1		
C											
	Talks and Lectures (12 items)										
	Amplified Objective: After listening to 3 talks on familiar and daily-life related topics, the test-takers are able to answer 12 M/C questions.										
	To anticipate the topics To anticipate the questions. Determine the topic Draw conclusions about who, what, when, where.	Personal Business Science	Listening for main ideas and details Making inference Summarizing	Giving information	Monologue	Comprehension Analysis	M/C	12/12 (19/19)	0-1		

แบบทดสอบวัดความสามารถด้านการใช้ภาษา

เนื่องจากผู้วิจัยต้องการที่จะสร้างแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัยด้านการใช้ภาษาให้มีลักษณะคู่ขนานกับแบบทดสอบการใช้ภาษาของแบบทดสอบ TOEFL ดังรายละเอียดที่ได้กำหนดไว้แล้วในหัวข้อ 3.5.1.2 และสามารถจัดทำเป็นโครงสร้างกำหนดลักษณะแบบทดสอบ (Test Blueprint) ด้านการใช้ภาษาแต่ละชนิดได้ดังนี้



ตารางที่ 10 Test Blueprint for Language Use (TOEFL)

Total Score: 40 points

Time: 25 minutes

Part	Test Objectives	Topics	Language Content	Text Type	Text Form	Skill Level	Test Type	Weight/ Items	Scoring Criteria	Data Source
A	<p>Structure</p> <p>Amplified Objective: After reading 15 incomplete sentences taken from academic reading texts, the test-takers are able to choose one word or phrase that best completes the sentence by selecting the answer from the four choices provided for each item.</p> <p>1) Be sure the sentence has a subject and a verb 2) Be careful of objects of prepositions 3) Be careful of present participles 4) Use reduced adjective clauses correctly 5) Use adverb Time and Cause connectors correctly 6) Use noun clause connector/subjects correctly 7) Use other adverb connectors correctly 8) Invert the subject and verb with place expressions 9) Be careful of past participles</p>	Academic related	<ul style="list-style-type: none"> - Agreement of subjects and verbs - Subject and verb inversion with place expressions - Subject and verb inversion with conditionals - Subject and verb inversion with negatives - Subject and verb inversion with comparisons - Objects of prepositions - Present participles - Past participles 	Informative	Incomplete sentences	Comprehension	M/C	15/15	0/1	(Educational Testing Service, 2014)

Part	Test Objectives	Topics	Language Content	Text Type	Text Form	Skill Level	Test Type	Weight/ Items	Scoring Criteria	Data Source
	10) Use noun clause connectors correctly 11) Invert the subject and verb with conditionals 12) Use reduced adverb clauses correctly 13) Invert the subject and verb with negatives 14) Use adjective clause connectors correctly 15) Invert the subject and verb with comparisons		- Noun clause connector/subjects - Noun clause connectors - Adverb. Time and Cause connectors - Adverb connectors - Reduced adverb clauses - Adjective clause connectors - Reduced adjective clauses							
B	Written Expression Amplified Objective: After reading each sentence of the 4 choices, the test-takers are able to recognize which of the 4 underlined parts contains a mistake in standard written English. 16) After be, use the present participle or the past participle 17) Form comparatives and superlatives correctly 18) Make verbs agree after certain words	Academic related	- Present/Past participle - Comparatives and superlatives - Verbs after certain words - Verbs after Will, Would, or other Modals	Informative	A sentence with the 4 underlined parts	Comprehension	M/C	25/25	25/25	(Educational Testing Service, 2014)

Part	Test Objectives	Topics	Language Content	Text Type	Text Form	Skill Level	Test Type	Weight/ Items	Scoring Criteria	Data Source
	<p>19) After Will, Would, or other Modals, use the base form of the verb</p> <p>20) Use the correct tense with time expressions</p> <p>21) Distinguish subject and object pronouns</p> <p>22) Distinguish possessive adjectives and pronouns</p> <p>23) Use parallel structure with paired conjunctions</p> <p>24) Make verb agree after prepositional phrases</p> <p>25) Use comparatives and superlatives correctly</p> <p>26) Distinguish A and An</p> <p>27) Use the correct tense with Will and Would</p> <p>28) Distinguish Like, Alike, Unlike, and Dislike</p> <p>29) Make inverted verbs agree</p> <p>30) Distinguish countable and uncountable nouns</p> <p>31) Distinguish specific and</p>		<ul style="list-style-type: none"> - Tenses and time expressions - Use of subject and object pronouns - Use of possessive adjectives and pronouns - Use of paired conjunctions - Use of verbs after prepositional phrases - Use of comparatives and superlatives - Use of A and An - Use of A and The - Use of Will and Would - Use of Like, Alike, Unlike, and Dislike - Agreement of inverted verbs - Use of countable and uncountable nouns 							

Part	Test Objectives	Topics	Language Content	Text Type	Text Form	Skill Level	Test Type	Weight/ Items	Scoring Criteria	Data Source
	<p>general ideas</p> <p>32) Position adjectives and adverbs correctly</p> <p>33) Recognize active and passive meanings</p> <p>34) Use articles with singular nouns</p> <p>35) Use predicate adjectives correctly</p> <p>36) Distinguish Make and Do</p> <p>37) Use the correct form of the passive</p> <p>38) Recognize when preposition have been omitted</p> <p>39) Distinguish the person from the thing</p> <p>40) Use adjectives after linking verbs</p>		<ul style="list-style-type: none"> - Position of adjectives and adverbs - Active and passive voice - Articles and singular nouns - Predicate adjectives - Use of Make and Do - Passive forms - Use of preposition after certain words - Use of nouns - Adjectives and linking verbs 							

ตารางที่ 11 Test Blueprint for Language Use (University English Proficiency Test)

Total Score: 40 points

Time: 25 minutes

Part	Test Objectives	Topics	Language Content	Text Type	Text Form	Skill Level	Test Type	Weight/ Items	Scoring Criteria	Data Source
A	<p>Structure</p> <p>Amplified Objective: After reading 15 incomplete sentences taken from academic reading texts, the test-takers are able to choose one word or phrase that best completes the sentence by selecting the answer from the four choices provided for each item.</p> <p>1) Be sure the sentence has a subject and a verb 2) Be careful of objects of prepositions 3) Be careful of present participles 4) Use reduced adjective clauses correctly 5) Use adverb Time and Cause connectors correctly 6) Use noun clause connector/subjects correctly 7) Use other adverb connectors correctly 8) Invert the subject and verb with place expressions 9) Be careful of past participles</p>	Academic related	<ul style="list-style-type: none"> - Agreement of subjects and verbs - Subject and verb inversion with place expressions - Subject and verb inversion with conditionals - Subject and verb inversion with negatives - Subject and verb inversion with comparisons - Objects of prepositions - Present participles - Past participles 	Informative	Incomplete sentences	Comprehension	M/C	15/15 (22/22)	0/1	(Educational Testing Service, 2014)

Part	Test Objectives	Topics	Language Content	Text Type	Text Form	Skill Level	Test Type	Weight/ Items	Scoring Criteria	Data Source
	10) Use noun clause connectors correctly 11) Invert the subject and verb with conditionals 12) Use reduced adverb clauses correctly 13) Invert the subject and verb with negatives 14) Use adjective clause connectors correctly 15) Invert the subject and verb with comparisons		<ul style="list-style-type: none"> - Noun clause connector/subjects - Noun clause connectors - Adverb. Time and Cause connectors - Adverb connectors - Reduced adverb clauses - Adjective clause connectors - Reduced adjective clauses 							
B	Written Expression Amplified Objective: After reading each sentence of the 4 choices, the test-takers are able to recognize which of the 4 underlined parts contains a mistake in standard written English. 16) After be, use the present participle or the past participle 17) Form comparatives and superlatives correctly 18) Make verbs agree after certain words	Academic related	<ul style="list-style-type: none"> - Present/Past participle - Comparatives and superlatives - Verbs after certain words - Verbs after Will, Would, or other Modals 	Informative	A sentence with the 4 underlined parts	Comprehension	M/C	25/25 (40/40)	25/25	(Educational Testing Service, 2014)

Part	Test Objectives	Topics	Language Content	Text Type	Text Form	Skill Level	Test Type	Weight/ Items	Scoring Criteria	Data Source
	19) After Will, Would, or other Modals, use the base form of the verb 20) Use the correct tense with time expressions 21) Distinguish subject and object pronouns 22) Distinguish possessive adjectives and pronouns 23) Use parallel structure with paired conjunctions 24) Make verb agree after prepositional phrases 25) Use comparatives and superlatives correctly 26) Distinguish A and An 27) Use the correct tense with Will and Would 28) Distinguish Like, Alike, Unlike, and Dislike 29) Make inverted verbs agree 30) Distinguish countable and uncountable nouns 31) Distinguish specific and general ideas		<ul style="list-style-type: none"> - Tenses and time expressions - Use of subject and object pronouns - Use of possessive adjectives and pronouns - Use of paired conjunctions - Use of verbs after prepositional phrases - Use of comparatives and superlatives - Use of A and An - Use of A and The - Use of Will and Would - Use of Like, Alike, Unlike, and Dislike - Agreement of inverted verbs - Use of countable and uncountable nouns 							

Part	Test Objectives	Topics	Language Content	Text Type	Text Form	Skill Level	Test Type	Weight/ Items	Scoring Criteria	Data Source
	32) Position adjectives and adverbs correctly 33) Recognize active and passive meanings 34) Use articles with singular nouns 35) Use predicate adjectives correctly 36) Distinguish Make and Do 37) Use the correct form of the passive 38) Recognize when preposition have been omitted 39) Distinguish the person from the thing 40) Use adjectives after linking verbs		- Position of adjectives and adverbs - Active and passive voice - Articles and singular nouns - Predicate adjectives - Use of Make and Do - Passive forms - Use of preposition after certain words - Use of nouns - Adjectives and linking verbs							

แบบทดสอบวัดความสามารถด้านการอ่าน

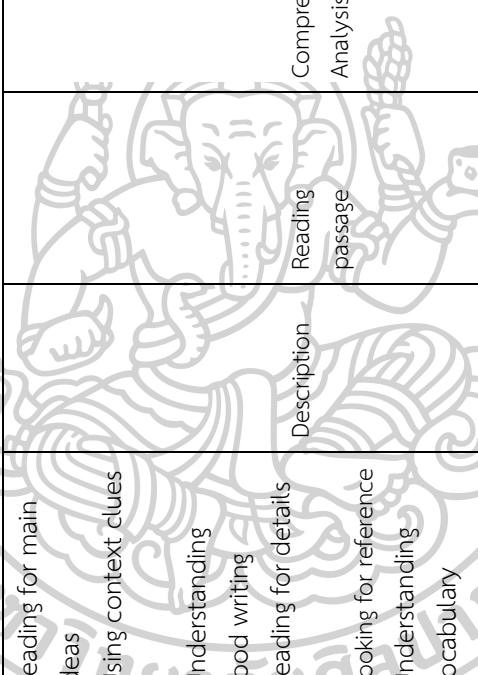
เนื่องจากผู้วิจัยต้องการที่จะสร้างแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัยด้านการอ่านให้มีลักษณะคู่ขนานกับแบบทดสอบการอ่านของแบบทดสอบ TOEFL ดังรายละเอียดที่ได้กำหนดไว้แล้วในหัวข้อ 3.5.1.2 และสามารถจัดทำเป็นโครงสร้างกำหนดลักษณะแบบทดสอบ (Test Blueprint) ด้านการอ่านแต่ละชนิดได้ดังนี้



ตารางที่ 12 Test Blueprint for Reading Test (TOEFL)

Total Score: 50 points

Time: 55 minutes

Test Objectives	Topics	Language Content	Text Type	Text Form	Skill Level	Test Type	Weight/ Items	Scoring Criteria	Data Source
Amplified Objective: After reading five passages on different topics, the test-takers are able to answer 50 M/C questions.									
<ul style="list-style-type: none"> - To answer main idea question correctly - To use context to determine meanings of simple words - To answer transition questions correctly - To answer stated detail questions correctly - To find pronoun referents - To determine meanings from word parts - To answer implied detail question correctly - To determine where specific information is found 	<ul style="list-style-type: none"> - Technology - Art - Animals - Nature - People 	Reading for main ideas Using context clues Understanding good writing Reading for details Looking for reference Understanding vocabulary Making inference Locating specific information	Description Reading passage		Comprehension Analysis	M/C	50/50	0-1	ETS (2010) Phillips (2003)

ตารางที่ 13 Test Blueprint for Reading Test (University English Proficiency Test)

Total Score: 50 points Time: 55 minutes

Part	Test Objectives	Topics	Language Content	Text Type	Text Form	Skill Level	Test Type	Weight/ Items	Scoring Criteria	Data Source	
A	English Reading for Everyday Use										
	Amplified Objective: After reading 6 types of everyday documents, the test-takers are able to answer 50 M/C questions.										
	<ul style="list-style-type: none"> - To answer stated detail questions correctly - To find “unstated” details - To answer an implied detailed question correctly - To determine meanings from word parts - To choose a general synonym for some specific words in the reading texts - To use context to determine meanings of simple words 	Health Work Home Study Familiar	Reading for details Making inference Understanding vocabulary Looking for synonym	Informative Request	- News - Advertisement - Instructions - Dictionary entry - Bill - Letter	Comprehension Analysis	M/C	50/50	0-1		
B	English Reading for Academic Purposes										
	Amplified Objective: After reading 5 academic texts, the test-takers are able to answer 50 M/C questions.										
	<ul style="list-style-type: none"> - To answer main idea question correctly - To answer stated detail questions correctly - To find “unstated” details - To answer an implied detailed question correctly. 	Technology Economy Love Family Life	Reading for main ideas Reading for details Using context clues Making inference Looking for reference Understanding	Descriptive	Passages Graph Poem	Comprehension Analysis	M/C	50/50	0/1		

Part	Test Objectives	Topics	Language Content	Text Type	Text Form	Skill Level	Test Type	Weight/ Items	Scoring Criteria	Data Source
	<ul style="list-style-type: none"> - To find pronoun referents - To determine where specific information is found - To recognize the organization of the ideas in the passage - To draw a conclusion by using the main points or contexts in the passage - To determine the purpose of the author - To choose a general synonym for some specific words in the reading texts - To determine the tone and mood of the author - To answer transition questions correctly 		vocabulary Locating specific information Recognizing text organization Restatement Passage conclusion Identifying purpose Making prediction Looking for synonym Identifying tone and mood Understanding good writing	Descriptive	Passages Graph Poem	Comprehension Analysis	M/C		0/1	

แบบทดสอบวัดความสามารถด้านการเขียน

เนื่องจากผู้วิจัยต้องการที่จะสร้างแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัยด้านการเขียนให้มีลักษณะคู่ขนานกับแบบทดสอบการเขียนแบบทดสอบ IELTS ดังรายละเอียดที่ได้กำหนดไว้แล้วในหัวข้อ 3.5.1.2 และสามารถจัดทำเป็นโครงสร้างกำหนดลักษณะแบบทดสอบ (Test Blueprint) ด้านการเขียน แต่ละชนิดได้ดังนี้



ตารางที่ 14 Test Blueprint for Writing (IELTS)

Total Score: 60 marks

Time: 60 minutes

Task	Test Objectives	Topics	Language Content	Input Text	Output Text	Skill Level	Time	Scoring Criteria	Data Source
I	<p>Task 1) Summary Writing</p> <p>Amplified Objective: After reading a writing task description, the test-takers are able to write a summary of information given in a graph, chart, table or diagram with the length of at least 150 words.</p> <ul style="list-style-type: none"> - the ability to describe the information objectively; - the ability to contrast and compare the information; - the ability to use specialized graphics language, not to express your own opinion and don't evaluate the data. 	Familiar	<ul style="list-style-type: none"> - Describing information - Contrasting and comparing information - Using specialized graphics language 	Written: Task description and a graph, chart, table, diagram, etc.	Written: A summary writing of at least 150 words in length	Application	20 mins	1) Task Achievement 2) Coherence and Cohesion 3) Lexical Resource 4) Grammatical and Accuracy	Hashemi & Thomas (2011)
II	<p>Task 2) Essay Writing</p> <p>Amplified Objective: After reading a writing task description, the test-takers are able to write a discussion essay in response to a statement of opinion or ideas with the length of at least 250 words.</p> <ul style="list-style-type: none"> - the ability to make solutions; - the ability to express your opinions; - the ability to contrast and compare the opinions and facts; - the ability to evaluate facts, ideas or arguments. 	Familiar	<ul style="list-style-type: none"> - Making solutions - Expressing opinions - Contrasting and comparing the opinions and facts - Evaluating facts, ideas or arguments 	Written: Task description	Written: An essay writing of at least 250 words in length	Application	40 mins	1) Task Achievement 2) Coherence and Cohesion 3) Lexical Resource 4) Grammatical and Accuracy	Hashemi & Thomas (2011)

ตารางที่ 15 Test Blueprint for Writing (University English Proficiency Test)

Total Score: 60 marks

Time: 60 minutes

Task	Test Objectives	Topics	Language Content	Input Text	Output Text	Skill Level	Time	Scoring Criteria	Data Source
I	<p>Task 1) Summary Writing (20 marks)</p> <p>Amplified Objective: After reading a writing task description, the test-takers are able to write a summary of information given in a graph, chart, table or diagram with the length of at least 150 words.</p> <ul style="list-style-type: none"> - the ability to describe the information objectively; - the ability to contrast and compare the information; - the ability to use specialized graphics language, not to express own opinion and not to evaluate the data. - the ability to organize ideas and present them in a well-organized writing. - the ability to produce a writing task with no mechanical errors. 	Familiar	<ul style="list-style-type: none"> - Describing information - Contrasting and comparing information - Using specialized graphics language 	Written: Task description and a graph, chart, table, diagram, etc.	Written: A summary writing of at least 150 words in length	Application	20 mins	1) Content 2) Organization 3) Language 4) Mechanics	

Task	Test Objectives	Topics	Language Content	Input Text	Output Text	Skill Level	Time	Scoring Criteria	Data Source
II	<p>Task 2) Essay Writing (40 marks)</p> <p>Amplified Objective: After reading a writing task description, the test-takers are able to write a discussion essay in response to a statement of opinion or ideas with the length of at least 250 words.</p> <ul style="list-style-type: none"> - the ability to make solutions; - the ability to express opinions; - the ability to contrast and compare the opinions and facts; - the ability to evaluate facts, ideas or arguments. - the ability to organize ideas and present them in a well-organized writing. - the ability to produce a writing task with no mechanical errors. 	Familiar	<ul style="list-style-type: none"> - Making solutions - Expressing opinions - Contrasting and comparing the opinions and facts - Evaluating facts, ideas or arguments 	Written: Task description	Written: An essay writing of at least 250 words in length	Application	40 mins	1) Content 2) Organization 3) Language 4) Mechanics	

แบบทดสอบวัดความสามารถด้านการพูด

เนื่องจากผู้วิจัยต้องการที่จะสร้างแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัยด้านการพูดให้มีลักษณะคู่ขนานกับแบบทดสอบการพูดของแบบทดสอบ IELTS ดังรายละเอียดที่ได้กำหนดไว้แล้วในหัวข้อ 3.5.1.2 และสามารถจัดทำเป็นโครงสร้างกำหนดลักษณะแบบทดสอบ (Test Blueprint) ด้านการเขียนแต่ละชนิดได้ดังนี้



ตารางที่ 16 Test Blueprint for Speaking (IELTS)

Total Score: 9 bands

Time: 10-12 minutes

Part	Test Objectives	Topics	Language Content	Input Text	Output Text	Skill Level	Time	Scoring Criteria	Data Source
I	Short Talk To answer personal questions and give information on familiar topics	Familiar	- Answering personal questions - Giving information	Spoken: examiner's questions on familiar topics	Spoken: test taker's answering questions and give information	Application	4-5 minutes	1) Fluency and Coherence 2) Lexical Resource 3) Grammatical Range and Accuracy 4) Pronunciation	Hashemi & Thomas (2011)
II	Individual long run To give a 2-minute-talk about how to solve a problem specified in the task card.	Familiar	- Giving opinions - Organizing ideas - Presenting	Written: a task card	Spoken: test taker's talking	Application	2 minutes	1) Fluency and Coherence 2) Lexical Resource 3) Grammatical Range and Accuracy 4) Pronunciation	Hashemi & Thomas (2011)

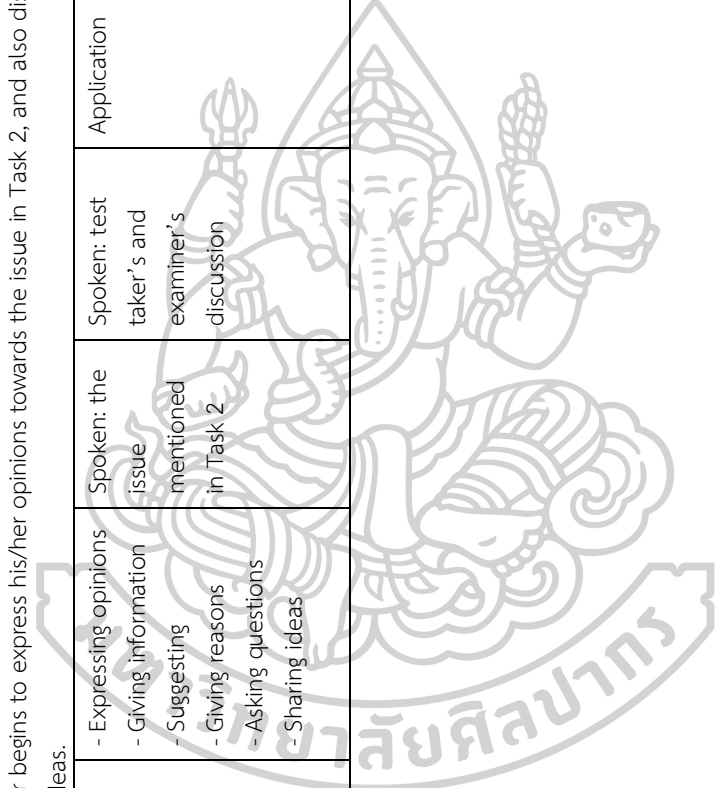
Part	Test Objectives	Topics	Language Content	Input Text	Output Text	Skill Level	Time	Scoring Criteria	Data Source
III	Two-way discussion To express opinions and discuss some possible ways to solve the problem in Task 2 effectively by giving reasons and asking the examiner to share the ideas.	Familiar	<ul style="list-style-type: none"> - Expressing opinions - Giving information - Suggesting - Giving reasons - Asking questions - Sharing ideas 	Spoken: the issue mentioned in Task 2	Spoken: test taker's and examiner's discussion	Application	4-5 minutes	1) Fluency and Coherence 2) Lexical Resource 3) Grammatical Range and Accuracy 4) Pronunciation	Hashemi & Thomas (2011)

ตารางที่ 17 Test Specifications for Speaking (University English Proficiency Test)

Total Score: 60 Time: 10-12 minutes

Part	Test Objectives	Topics	Language Content	Input Text	Output Text	Skill Level	Time	Scoring Criteria	Data Source
I	Short Talk Amplified Objective: After seated and checked the name, the test-taker is able to answer some familiar questions and interact in informal talking with the examiner. To answer personal questions and give information on familiar topics	Familiar	- Answering personal questions - Giving information	Spoken: examiner's questions on familiar topics	Spoken: test taker's answering questions and give information	Application	4-5 minutes	1) Communication and Size 2) Fluency 3) Accuracy 4) Pronunciation and Accent	
II	Individual long run Amplified Objective: After reading a task card and having 1-2 minutes for preparation, the test-taker is able to talk for about 2 minutes to solve the problem specified in the card. To give a 2-minute-talk about how to solve a problem specified in the task card.	Familiar	- Giving opinions - Organizing ideas - Presenting	Written: a task card	Spoken: test taker's talking	Application	2 minutes	1) Communication and Size 2) Fluency 3) Accuracy 4) Pronunciation and Accent	

Part	Test Objectives	Topics	Language Content	Input Text	Output Text	Skill Level	Time	Scoring Criteria	Data Source
III	Two-way discussion To express opinions and discuss some possible ways to solve the problem in Task 2 effectively by giving reasons and asking the examiner to share the ideas.	Familiar	<ul style="list-style-type: none"> - Expressing opinions - Giving information - Suggesting - Giving reasons - Asking questions - Sharing ideas 	Spoken: the issue mentioned in Task 2	Spoken: test taker's and examiner's discussion	Application	4-5 minutes	<ul style="list-style-type: none"> 1) Communication and Size 2) Fluency 3) Accuracy 4) Pronunciation and Accent 	



2) สร้างแบบทดสอบฉบับร่างครั้งที่ 1 (First draft) ตามตารางกำหนดลักษณะแบบทดสอบ (Test Blueprint) ที่ได้กำหนดไว้ข้างต้น โดยมีรายละเอียดของแบบทดสอบแต่ละทักษะดังนี้

แบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัยด้านการฟัง มีจุดประสงค์เพื่อวัดความสามารถของผู้เข้าสอบในการฟังสาระที่เกี่ยวข้องกับชีวิตประจำวันและเป็นที่ยุติเคย แบ่งออกเป็น 3 ตอน คือ ตอนที่ 1 เป็นบทสนทนาสั้น ๆ (Short conversations) จำนวน 49 ข้อ ตอนที่ 2 เป็นบทสนทนายาว (Long conversations) จำนวน 17 ข้อ และ ตอนที่ 3 เป็นบทบรรยายหรือการพูดยาว (Lectures or Talks) จำนวน 21 ข้อ รวมทั้งหมด 87 ข้อ

แบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัยด้านการใช้ภาษา มีจุดประสงค์เพื่อวัดความสามารถของผู้เข้าสอบในความเข้าใจภาษาเขียนมาตรฐานและระบุข้อผิดพลาดในการใช้ภาษา แบ่งออกเป็น 2 ตอน คือ ตอนที่ 1 เป็นโครงสร้างประโยค (Sentence structures) จำนวน 22 ข้อ และ ตอนที่ 2 เป็นสำนวนภาษาที่ใช้ในการเขียน (Written Expressions) จำนวน 38 ข้อ รวมทั้งหมด 60 ข้อ

แบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัยด้านการอ่าน มีจุดประสงค์เพื่อวัดความสามารถของผู้เข้าสอบในการอ่านสาระในชีวิตประจำวันและเชิงวิชาการ แบ่งออกเป็น 2 ตอน คือ ตอนที่ 1 การอ่านภาษาอังกฤษในชีวิตประจำวัน (English for Everyday Use) จำนวน 49 ข้อ และ ตอนที่ 2 การอ่านภาษาอังกฤษเชิงวิชาการ (English for Academic Purpose) จำนวน 48 ข้อ รวมทั้งหมด 97 ข้อ

แบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัยด้านการเขียน มีจุดประสงค์เพื่อวัดความสามารถของผู้เข้าสอบในการเขียนเชิงวิชาการ แบ่งออกเป็น 2 ตอน คือ ตอนที่ 1 การเขียนสรุปความ (Summary writing) จำนวน 1 ข้อ และ ตอนที่ 2 การเขียนแสดงความคิดเห็น (Argumentative writing) จำนวน 1 ข้อ

แบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัยด้านการพูด มีจุดประสงค์เพื่อวัดความสามารถของผู้เข้าสอบในการพูดสื่อสารข้อมูลส่วนตัว แสดงความคิดเห็นและพูดเชิงวิชาการ แบ่งออกเป็น 3 ตอน คือ ตอนที่ 1 การพูดสั้น ๆ เกี่ยวกับข้อมูลของตนเอง ตอนที่ 2 การพูดตามบัตรภาระงาน และตอนที่ 3 การพูดแสดงความคิดเห็นที่เกี่ยวข้องกับตอนที่ 2

3) นำแบบทดสอบฉบับร่างครั้งที่ 1 (First draft) ให้อาจารย์ที่ปรึกษาวิทยานิพนธ์ตรวจสอบและให้ข้อเสนอแนะเพื่อแก้ไขปรับปรุง

4) ปรับปรุงแบบทดสอบฉบับร่างครั้งที่ 1 (First draft) ตามคำแนะนำของอาจารย์ที่ปรึกษาวิทยานิพนธ์

5) นำแบบทดสอบฉบับร่างครั้งที่ 1 (First draft) ให้ผู้เชี่ยวชาญ 3 ท่าน ซึ่งเป็นผู้เชี่ยวชาญด้านการวัดและประเมินผลภาษาอังกฤษ 1 คน ผู้เชี่ยวชาญด้านการสอนภาษาอังกฤษ 1 คน และเจ้าของภาษา 1 คน (ดังรายละเอียดแสดงในภาคผนวก ก) ตรวจสอบความถูกต้อง เหมาะสมและความสอดคล้องของข้อคำถามกับจุดประสงค์การสอบรายข้อ โดยใช้แบบประเมินความสอดคล้องของข้อคำถามและจุดประสงค์การสอบ (IOC) (ดังรายละเอียดแสดงในภาคผนวก ฉ)

6) แก้ไขปรับปรุงแบบทดสอบฉบับร่างครั้งที่ 1 (First draft) ได้แบบทดสอบที่จะนำไปใช้จริงในแต่ละทักษะ ดังนี้

แบบทดสอบด้านการฟัง แก่ไขการสะกดคำผิด เรียงลำดับข้อใหม่เพื่อให้มีการกระจายของข้อคำถามที่วัดวัตถุประสงค์เดียวกัน ได้แบบทดสอบด้านการฟังที่มี 3 ตอน คือ ตอนที่ 1 เป็นบทสนทนาสั้น ๆ (Short conversations) จำนวน 49 ข้อ ตอนที่ 2 เป็นบทสนทนายาว (Long conversations) จำนวน 17 ข้อ และ ตอนที่ 3 เป็นบทบรรยายหรือการพูดยาว (Lectures or Talks) จำนวน 21 ข้อ รวมทั้งหมด 87 ข้อ

แบบทดสอบด้านการใช้ภาษา แก่ไขการสะกดคำผิด ตัดข้อคำถามที่ซ้ำ จัดเรียงลำดับข้อใหม่เพื่อให้มีการกระจายของข้อคำถามที่วัดวัตถุประสงค์เดียวกัน ได้แบบทดสอบด้านการใช้ภาษาที่มี 2 ตอน คือ ตอนที่ 1 เป็นโครงสร้างประโยค (Sentence structures) จำนวน 22 ข้อ และ ตอนที่ 2 เป็นสำนวนภาษาที่ใช้ในการเขียน (Written Expressions) จำนวน 38 ข้อ รวมทั้งหมด 60 ข้อ

แบบทดสอบด้านการอ่าน แก่ไขการสะกดคำผิด ปรับเฉลยคำตอบบางข้อ ได้แบบทดสอบด้านการอ่านที่มี 2 ตอน คือ ตอนที่ 1 การอ่านภาษาอังกฤษในชีวิตประจำวัน (English for Everyday Use) จำนวน 49 ข้อ และ ตอนที่ 2 การอ่านภาษาอังกฤษเชิงวิชาการ (English for Academic Purpose) จำนวน 48 ข้อ รวมทั้งหมด 97 ข้อ

แบบทดสอบด้านการเขียน แก่ไขการสะกดคำผิด และปรับเกณฑ์การให้คะแนนให้สอดคล้องกับวัตถุประสงค์ ได้แบบทดสอบด้านการเขียน ที่แบ่งเป็น 2 ตอน คือ ตอนที่ 1 การเขียนสรุปความ (Summary writing) จำนวน 1 ข้อ และ ตอนที่ 2 การเขียนแสดงความคิดเห็น (Argumentative writing) จำนวน 1 ข้อ

แบบทดสอบด้านการพูด แก่ไขการสะกดคำผิด และปรับเกณฑ์การให้คะแนนให้สอดคล้องกับวัตถุประสงค์ ได้แบบทดสอบด้านการพูดแบ่งออกเป็น 3 ตอน คือ ตอนที่ 1 การพูดสั้น ๆ เกี่ยวกับข้อมูลของตนเอง ตอนที่ 2 การพูดตามบัตรภาระงาน และตอนที่ 3 การพูดแสดงความคิดเห็นที่เกี่ยวข้องกับตอนที่ 2

6) นำแบบทดสอบฉบับร่างครั้งที่ 1 (First draft) ที่ได้รับการแก้ไขปรับปรุงแล้ว ไปสอบถามความคิดเห็นขั้นต้น เพื่อประเมินคุณประโยชน์ของแบบทดสอบก่อนนำไปใช้จากผู้เชี่ยวชาญ 3 ท่าน และอาจารย์ผู้สอนภาษาอังกฤษระดับมหาวิทยาลัย 7 ท่าน โดยใช้แบบประเมินคุณประโยชน์ของแบบทดสอบ (ดังรายละเอียดแสดงไว้ในภาคผนวก ฅ)

7) ปรับปรุงแบบทดสอบฉบับร่างครั้งที่ 1 (First draft) ตามผลการประเมินคุณภาพขั้นต้นของแบบทดสอบ

8) บันทึกเสียงแบบทดสอบด้านการฟัง โดยใช้เจ้าของภาษาสัญชาติอเมริกัน 2 คน เป็นผู้ชาย 1 คน และผู้หญิง 1 คน

9) ตัดต่อไฟล์เสียงของแบบทดสอบด้านการฟัง ให้มีคุณลักษณะคล้ายกับแบบทดสอบมาตรฐาน โดยให้มี 1) การพูดนำก่อนการเริ่มทำแบบทดสอบและยกตัวอย่าง 2) การพูดนำก่อนขึ้นการสอบแต่ละตอน 3) เว้นช่วงเงียบ (Silence) ระหว่างตอน 5 วินาที และระหว่างข้อ 12 วินาที

10) ตรวจสอบความสมบูรณ์ของแบบทดสอบ กระจายคำตอบ เฉลยคำตอบ และไฟล์เสียงอีกครั้ง ก่อนนำไปทดลองใช้ครั้งที่ 1

2) แบบตรวจให้คะแนนแบบทดสอบอัตโนมัติ

แบบตรวจให้คะแนนแบบทดสอบอัตนัย ประกอบด้วย 2 ชนิด คือ 1) แบบตรวจให้คะแนนแบบทดสอบการเขียน และ 2) แบบตรวจให้คะแนนแบบทดสอบการพูด โดยมีรายละเอียดในการพัฒนาดังนี้

แบบตรวจให้คะแนนแบบทดสอบการเขียน

ผู้วิจัยพัฒนาแบบตรวจให้คะแนนแบบทดสอบการเขียนตามแนวคิดของ Brown (2004) and Cambridge English (2015) ซึ่งประกอบด้วยเกณฑ์ให้คะแนน 4 ด้าน คือ 1) เนื้อหาสาระ (Content) 2) การจัดระบบความคิด (Organization) 3) ภาษา (Language) และ 4) กลไกทางภาษา (Mechanics) โดยแต่ละด้านมีระดับคะแนนตั้งแต่ 1-5 และกำหนดคะแนนเต็มของแบบทดสอบเขียนตอนที่ 1 เป็น 20 คะแนน และตอนที่ 2 เป็น 40 คะแนน รวม 60 คะแนน (ภาคผนวก ข)

แบบตรวจให้คะแนนแบบทดสอบการพูด

ผู้วิจัยพัฒนาแบบตรวจให้คะแนนแบบทดสอบการพูดโดยปรับใช้แนวคิดของ Brown (2004), Carroll (1983), and Cambridge English (2015) ซึ่งประกอบด้วยเกณฑ์ให้คะแนน 4 ด้าน คือ 1) การสื่อสารและปริมาณ (Communication and Size) 2) ความคล่องแคล่ว (Fluency) 3) ความถูกต้องของภาษา (Accuracy) และ 4) ความสามารถที่จะสื่อสารให้เข้าใจ (Comprehensibility) โดยแต่ละด้านมีระดับคะแนนตั้งแต่ 1-5 และกำหนดคะแนนเต็มของแบบทดสอบพูดแต่ละตอนเป็น 20 คะแนน รวม 60 คะแนน (ภาคผนวก ข)

3) แบบประเมินคุณภาพแบบทดสอบ

ในการสร้างแบบประเมินคุณภาพแบบทดสอบ ผู้วิจัยได้ดำเนินการพัฒนาดังนี้

- 1) ศึกษาวิธีสร้างประเมินคุณภาพแบบทดสอบตามวิธีของ Likert
- 2) กำหนดกรอบเนื้อหาสาระของแบบประเมินคุณภาพแบบทดสอบ โดยสร้างรายการประเมินคุณภาพแบบทดสอบที่มีรายละเอียดดังนี้

ตารางที่ 18 ตารางกำหนดกรอบเนื้อหาสาระของแบบประเมินคุณประโยชน์แบบทดสอบ

หัวข้อการสอบถาม	รายละเอียด
Construct Validity	1. The test objective of each section is clearly defined. 2. The test objective of each section is in accordance with the purpose of the listening test. 3. All test tasks are in accordance with the test objectives. 4. The scoring methods are appropriate, and they can reflect the test takers' English abilities. 5. The test takers' scores can predict their English abilities.
Reliability	6. The expected response of each test item is sufficiently specific for one correct answer. 7. The scores of the tests are reliable, irrespective of any raters. 8. The response to each test item is reliable, irrespective of any test-takers, any time, and any place for test-taking. 9. The test input in each part of the test does not contribute to understanding or knowledge in doing other items of other parts.
Authenticity	10. The test tasks are related to the test takers' future real-life tasks. 11. The texts are authentic, and they reflect real-life situations.
Interactiveness	12. The test tasks are familiar topics for all test takers. 13. The tasks are suitable for the English knowledge of all the test takers.
Impact	14. The test tasks provide the test takers with experience in doing future language use tasks. 15. The test tasks are similar to the learning activities that the test takers have experienced in class. 16. The purpose of the test is consistent with the goals and the values mentioned in the instructional program. 17. All test tasks can reflect test takers' English abilities currently demanded by Thai society and educational system
Practicality	18. The test can possibly be constructed with existing resources. 19. The test can be administered and scored without much difficulty.

ตารางที่ 18 ตารางกำหนดกรอบเนื้อหาสาระของแบบประเมินคุณภาพแบบทดสอบ (ต่อ)

หัวข้อการสอบถาม	รายละเอียด
	20. The test content contains suitable and bias-free topics without leading to discrimination of sex, race, and economic and social status, and it is not leading to violence.
Other comments:	Please give comments.

3) สร้างแบบประเมินคุณภาพแบบทดสอบ โดยแบ่งข้อคำถามในแบบสอบถามออกเป็น 3 ตอน และแต่ละตอนมีรายละเอียดดังนี้

ตอนที่ 1 กำหนดเป็นสถานภาพของผู้ตอบแบบสอบถาม ผู้ตอบจะต้องเลือกจาก 2 ตัวเลือก คือ 1) ผู้เชี่ยวชาญด้านการวัดและประเมินผลการสอนภาษาอังกฤษ และ 2) อาจารย์ผู้สอนภาษาอังกฤษระดับมหาวิทยาลัย

ตอนที่ 2 กำหนดเป็นความคิดเห็นของผู้ประเมินต่อแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย โดยแบ่งออกเป็น 6 หัวข้อหลัก คือ 1) ความตรงเชิงโครงสร้าง (Construct Validity) มีข้อคำถาม 5 ประเด็น 2) ความเชื่อมั่น (Reliability) มีข้อคำถาม 4 ประเด็น 3) ความสมจริง (Authenticity) มีข้อคำถาม 2 ประเด็น 4) การมีปฏิสัมพันธ์ (Interactiveness) มีข้อคำถาม 2 ประเด็น 5) ผลกระทบ (Impact) มีข้อคำถาม 4 ประเด็น และ 6) ความสะดวกใช้ (Practicality) มีข้อคำถาม 3 ประเด็น ดังแสดงไว้ในตารางที่ 18 ทั้งนี้ผู้วิจัยได้ออกแบบให้ข้อคำถาม มีลักษณะเป็นแบบมาตราส่วนประมาณค่า (Rating Scale) โดยกำหนดค่าระดับของข้อความในแบบสอบถาม ผู้วิจัยใช้เกณฑ์ของ Likert มีการกำหนดค่า ดังนี้

เห็นด้วยอย่างยิ่ง	ให้ค่าระดับเท่ากับ 5
เห็นด้วย	ให้ค่าระดับเท่ากับ 4
ไม่แน่ใจ	ให้ค่าระดับเท่ากับ 3
ไม่เห็นด้วย	ให้ค่าระดับเท่ากับ 2
ไม่เห็นด้วยอย่างยิ่ง	ให้ค่าระดับเท่ากับ 1

สำหรับการให้ความหมายของค่าที่วัดได้ ผู้วิจัยได้กำหนดเกณฑ์ที่ใช้ในการให้ความหมายโดยได้แนวคิดของ (Best 1986: 195) การให้ความหมาย โดยการให้ค่าเฉลี่ยเป็นรายชื่อ ดังนี้

4.50 – 5.00	หมายถึง	ความคิดเห็นระดับดีมาก
3.50 – 4.49	หมายถึง	ความคิดเห็นระดับดี
2.50 – 3.49	หมายถึง	ความคิดเห็นระดับดีปานกลาง
1.50 – 2.49	หมายถึง	ความคิดเห็นระดับดีน้อย
1.00 – 1.49	หมายถึง	ความคิดเห็นระดับดีน้อยที่สุด

ตอนที่ 3 กำหนดเป็นความคิดเห็นและข้อเสนอแนะอื่น ๆ โดยให้ผู้ประเมินเขียนตอบได้อย่างอิสระ

4) นำแบบประเมินที่สร้างขึ้นไปให้อาจารย์ที่ปรึกษาวิทยานิพนธ์ตรวจสอบความเหมาะสมของข้อความและรูปแบบภาษา

5) นำแบบประเมินคุณภาพแบบทดสอบที่ผ่านการตรวจแก้ไขแล้ว ให้ผู้เชี่ยวชาญตรวจสอบความสอดคล้องของข้อคำถามกับหัวเรื่อง (IOC) พร้อมทั้งให้ข้อเสนอแนะเพื่อปรับปรุง

6) หาค่า IOC ของแบบประเมิน และแก้ไขแบบประเมินตามคำแนะนำของผู้เชี่ยวชาญ ข

7) ใช้แบบประเมินคุณภาพกับผู้เชี่ยวชาญและอาจารย์มหาวิทยาลัย ในการประเมินคุณภาพแบบทดสอบก่อนการนำแบบทดสอบไปทดลองใช้

8) นำผลที่ได้จากการประเมินคุณภาพแบบทดสอบ มาวิเคราะห์หาค่าความเชื่อมั่นของแบบประเมินคุณภาพแบบทดสอบทั้งฉบับ โดยใช้สูตรสัมประสิทธิ์แอลฟา (α Coefficient) ของ Cronbach (พิมพา สุวรรณฤทธิ์ : 172-173)

4) แบบประเมินความพึงพอใจผู้เกี่ยวข้อง

แบบประเมินความพึงพอใจของผู้เกี่ยวข้องต่อแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย ดำเนินการพัฒนาดังนี้

1) ศึกษาวิธีสร้างแบบประเมินถามความพึงพอใจตามวิธีของ Likert

2) กำหนดกรอบเนื้อหาสาระของแบบประเมิน โดยในการสร้างรายการประเมินความพึงพอใจของผู้ใช้แบบทดสอบที่มีต่อแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัยผู้วิจัยสร้างขึ้น โดยมีรายละเอียดดังนี้

ตารางที่ 19 ตารางกำหนดกรอบเนื้อหาสาระของแบบประเมิน

หัวข้อการประเมิน	รายละเอียด
สถานภาพของผู้ตอบแบบประเมิน	<ul style="list-style-type: none"> - ผู้เชี่ยวชาญด้านการวัดและประเมินผลการสอนภาษาอังกฤษ - อาจารย์ผู้สอนภาษาอังกฤษระดับมหาวิทยาลัย - ผู้เข้าสอบ
คำสั่งของแบบทดสอบ	<ul style="list-style-type: none"> - มีคำสั่งเขียนไว้อย่างชัดเจนและใช้ภาษาเข้าใจง่าย - มีตัวอย่างแสดงการตอบในแต่ละตอนที่เข้าใจง่าย - ระบุน้ำหนักคะแนนหรือจำนวนข้อของแบบทดสอบแต่ละตอน - กำหนดเวลาที่ชัดเจนในการทำแบบทดสอบแต่ละตอน - กำหนดเวลาในการทำและความยาวของคำตอบสำหรับแบบทดสอบวัดการเขียนไว้อย่างชัดเจน - ผู้เข้าสอบได้รับคำแนะนำการทำสอบอย่างชัดเจน ครอบคลุมเรื่องที่ควรทราบ และไม่ทำให้เกิดความเข้าใจผิดภายหลัง
เนื้อหาสาระของแบบทดสอบ	<ul style="list-style-type: none"> - เนื้อหาของแบบทดสอบเหมาะสมกับการสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย - การจัดลำดับเนื้อหาของแบบทดสอบแต่ละตอนมีความเหมาะสม

หัวข้อการประเมิน	รายละเอียด
	<ul style="list-style-type: none"> - จำนวนข้อของแบบทดสอบเหมาะสมกับเวลาสอบ - เนื้อหาสาระของแบบทดสอบน่าสนใจ มีประโยชน์ต่อผู้เข้าสอบ - วัดและประเมินวัตถุประสงค์ของการสอบแต่ละทักษะได้อย่างเหมาะสมและเพียงพอ - วัดทักษะทางภาษาอย่างหลากหลาย (ฟัง พูด อ่าน เขียน และการใช้ภาษา) - ใช้ภาษาที่มีลีลาสุภาพ ไม่ขัดต่อจริยธรรมและศีลธรรมอันดีของแบบทดสอบสอดคล้องกับเนื้อหาที่วัด ไม่ทำให้ผู้เข้าสอบไขว้เขวหรือเข้าใจผิดได้ - ชื่อของแบบทดสอบสอดคล้องกับเนื้อหาที่วัด ไม่ทำให้ผู้เข้าสอบไขว้เขวหรือเข้าใจผิดได้ - แบบทดสอบไม่ลำเอียงต่อเชื้อชาติ พื้นภูมิลัทธิ เพศ และสภาพทางร่างกายของผู้เข้าสอบ - ปราศจากการใช้ภาษาที่ไม่คำนึงถึงความรู้สึกนึกคิด จิตใจของผู้เข้าสอบบางกลุ่ม
รูปแบบและการจัดลำดับความ	<ul style="list-style-type: none"> - แบบทดสอบมีความยาวเหมาะสมกับระดับผู้เข้าสอบ ไม่สั้นหรือยาวจนเกินไป - มีการวัดวัตถุประสงค์แต่ละข้อเหมาะสม ไม่มากเกินไปหรือน้อยเกินไป - แบบทดสอบไม่มีการชี้แนะคำตอบซึ่งกันและกัน - ทำสอบโดยใช้ความรู้ความสามารถในการเลือกคำตอบที่ถูกต้อง และไม่ใช้การเดาหรือการคัดคำตอบที่คิดว่าไม่ใช่คำตอบออกไปทีละข้อ - มีการกำหนดเวลาที่เหมาะสมโดยให้ผู้เข้าสอบที่อ่อนสามารถทำได้ด้วย - มีการใช้วิธีการวัดที่เหมาะสมกับคำตอบที่ต้องการ เช่น แบบจับคู่ มีการใช้ภาษาในตัวแบบทดสอบและคำตอบในระดับความยากง่ายเดียวกัน - มีการใช้วิธีการวัดที่หลากหลาย เพื่อวัดครอบคลุมทุกทักษะ หรือวัดทักษะย่อยที่ประสงค์ - มีการใช้ตัวกระตุ้น/ตัวแบบทดสอบ เพื่อให้ได้คำตอบที่ต้องการอย่างสมจริง เช่น มีการใช้เทปบันทึกเสียง วิดีทัศน์ ภาพประกอบ ฯลฯ - การวางรูปแบบแบบทดสอบชัดเจน สะดวกต่อการทำสอบ ไม่ทำให้ผู้เข้าสอบพลาดพลัง เช่น การเรียงข้อ และตัวเลือก a b c d ชัดเจน และถ้าเป็นแบบโคลงหรือแบบเติมคำ มีการวางตัวเลือกในหน้าเดียวกันกับตัวแบบทดสอบ เพื่อให้ผู้เข้าสอบไม่ต้องสับสน หรือเสียเวลาพลิกหน้ากระดาษ - การเว้นช่องไฟและช่องว่างระหว่างบรรทัดเหมาะสม ชัดเจน
การตรวจให้คะแนน	<ul style="list-style-type: none"> - การตรวจให้คะแนนแบบทดสอบอัตโนมัติ มีเกณฑ์ให้คะแนนที่เหมาะสมกับการะงานสอบ

หัวข้อการประเมิน	รายละเอียด
ประโยชน์ของแบบทดสอบ	- ผลการสอบมีประโยชน์ในการพัฒนาและปรับปรุงทักษะภาษาอังกฤษของผู้เข้าสอบ - ผู้เข้าสอบได้รับประสบการณ์ที่จะเป็นประโยชน์ต่อการเข้าสอบวัดความสามารถภาษาอังกฤษในอนาคต
การจัดสอบ	- การสมัครและลงทะเบียนสอบทันสมัย สะดวก และเข้าถึงได้รวดเร็ว - มีช่องทางในการติดต่อสื่อสารระหว่างผู้จัดสอบและผู้เข้าสอบ - ที่นั่งสะดวกสบายเหมาะกับการนั่งทำแบบทดสอบ - อากาศในห้องสอบเหมาะสม ไม่ร้อนหรือเย็นจนเกินไป - ปราศจากเสียงรบกวนอื่นที่เป็นการทำลายสมาธิในการทำแบบทดสอบ - แสงสว่างเพียงพอ ไม่เป็นอุปสรรคต่อการทำแบบทดสอบ - เสียงจากลำโพงในการทดสอบฟังชัดเจน ไม่เป็นอุปสรรคต่อการได้ยิน
ความคิดเห็นและข้อเสนอแนะ	โปรดแสดงความคิดเห็นหรือให้ข้อเสนอแนะ

3) สร้างแบบประเมินด้วยแอปพลิเคชัน Google Form และเก็บไฟล์ไว้ใน Google Drive โดยแบ่งข้อความในแบบประเมินออกเป็น 3 ตอน และแต่ละตอนมีรายละเอียดดังนี้

ตอนที่ 1 กำหนดเป็นสถานภาพของผู้ตอบแบบประเมิน ผู้ตอบจะต้องเลือกจาก 3 ตัวเลือก คือ 1) ผู้เชี่ยวชาญด้านการวัดและประเมินผลการสอนภาษาอังกฤษ 2) อาจารย์ผู้สอนภาษาอังกฤษระดับมหาวิทยาลัย และ 3) ผู้เข้าสอบ

ตอนที่ 2 กำหนดเป็นความพึงพอใจของผู้ตอบแบบประเมินต่อแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย โดยแบ่งออกเป็น 6 หัวข้อหลัก คือ 1) คำสั่งของแบบทดสอบ มีข้อความ 6 ประเด็น 2) เนื้อหาสาระของแบบทดสอบ มีข้อความ 10 ประเด็น 3) รูปแบบและการจัดลำดับความ มีข้อความ 10 ประเด็น 4) การตรวจให้คะแนน มีข้อความ 1 ประเด็น 5) ประโยชน์ของแบบทดสอบ มีข้อความ 2 ประเด็น และ 6) การจัดสอบ มีข้อความ 7 ประเด็น ดังแสดงไว้ในตารางที่ 8 ทั้งนี้ผู้วิจัยได้ออกแบบให้ข้อความ มีลักษณะเป็นแบบมาตราส่วนประมาณค่า (Rating Scale) โดยกำหนดค่าระดับของข้อความในแบบสอบถาม ผู้วิจัยใช้เกณฑ์ของ Likert มีการกำหนดค่า ดังนี้

พึงพอใจมากที่สุด	ให้ค่าระดับเท่ากับ 5
พึงพอใจมาก	ให้ค่าระดับเท่ากับ 4
พึงพอใจปานกลาง	ให้ค่าระดับเท่ากับ 3
พึงพอใจน้อย	ให้ค่าระดับเท่ากับ 2
พึงพอใจน้อยที่สุด	ให้ค่าระดับเท่ากับ 1

สำหรับการให้ความหมายของค่าที่วัดได้ ผู้วิจัยได้กำหนดเกณฑ์ที่ใช้ในการให้ความหมายโดยได้แนวคิดของ Best (1986: 195) การให้ความหมาย โดยการให้ค่าเฉลี่ยเป็นรายชื่อ ดังนี้

4.50 – 5.00	หมายถึง	พึงพอใจมากที่สุด
3.50 – 4.49	หมายถึง	พึงพอใจมาก
2.50 – 3.49	หมายถึง	พึงพอใจปานกลาง
1.50 – 2.49	หมายถึง	พึงพอใจน้อย
1.00 – 1.49	หมายถึง	พึงพอใจน้อยที่สุด

ตอนที่ 3 กำหนดเป็นความคิดเห็นและข้อเสนอแนะอื่น ๆ โดยให้ผู้ตอบแบบประเมินพิมพ์ตอบในแบบฟอร์มได้อย่างอิสระ

4) นำแบบประเมินที่สร้างขึ้นไปให้อาจารย์ที่ปรึกษาวิทยาลัยนพนธ์ตรวจสอบความเหมาะสมของข้อความและรูปแบบภาษา

5) นำแบบประเมินความพึงพอใจที่ผ่านการตรวจแก้ไขแล้ว ไปทดลองใช้กับผู้เข้าสอบในการทดลองครั้งที่ 1 จำนวน 28 คน และในการทดลองครั้งที่ 2 จำนวน 94 คน โดยให้ผู้เข้าสอบทำแบบประเมินความพึงพอใจหลังจากทำแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย ด้วยการใช้โทรศัพท์มือถือสแกน QR-code ที่กระดาษคำตอบก่อนส่งกระดาษคำตอบให้ผู้คุมสอบ

6) นำผลการทดลองใช้แบบประเมินความพึงพอใจในขั้นการทดลองครั้งที่ 1 และ 2 มาวิเคราะห์หาค่าความเชื่อมั่นของแบบประเมินความพึงพอใจทั้งฉบับ โดยใช้สูตรสัมประสิทธิ์แอลฟา (α Coefficient) ของ Cronbach (พินพา สุวรรณฤทธิ์ : 172-173)

7) นำแบบประเมินความพึงพอใจที่ได้หาค่าความเชื่อมั่นและปรับปรุงแก้ไขแล้ว ไปใช้กับผู้เข้าสอบที่เป็นกลุ่มตัวอย่างจริง ให้อาจารย์ผู้สอนภาษาอังกฤษ และผู้เชี่ยวชาญตอบแบบประเมินเพื่อประเมินความพึงพอใจของแบบทดสอบฉบับสุดท้ายด้วย ผู้เชี่ยวชาญ 3 คน ประกอบด้วย ผู้เชี่ยวชาญด้านการวัดและประเมินผลภาษา 1 คน ผู้เชี่ยวชาญด้านการสอนภาษาอังกฤษ 1 คน และผู้เชี่ยวชาญด้านการใช้ภาษาอังกฤษ 1 คน

3.5.2.2 การทดลองใช้เครื่องมือและวิเคราะห์ประสิทธิภาพและตรวจสอบคุณภาพของเครื่องมือ

ผู้วิจัยได้ทดลองใช้เครื่องมือวิเคราะห์ประสิทธิภาพและตรวจสอบคุณภาพของเครื่องมือ 2 ครั้ง คือ การทดลองครั้งที่ 1 ระหว่างวันที่ 16 ถึง 27 มกราคม 2562 และ ครั้งที่ 2 ระหว่างวันที่ 30 มกราคม ถึง 10 กุมภาพันธ์ 2562 โดยแบ่งการดำเนินกิจกรรมต่าง ๆ ได้เป็น 5 ขั้นตอน คือ 1) ก่อนการเก็บข้อมูล 2) ขณะเก็บข้อมูล 3) การตรวจแบบทดสอบ 4) การวิเคราะห์ประสิทธิภาพแบบทดสอบ และ 5) การตรวจสอบคุณภาพแบบทดสอบโดยรวม โดยมีรายละเอียดกิจกรรมที่ดำเนินในแต่ละครั้งดังนี้

1.1) ขั้นตอนการเก็บข้อมูล

ในขั้นก่อนการเก็บข้อมูล ผู้วิจัยได้ดำเนินการตามลำดับดังนี้

1.1.1) การจัดการกรรมการคุมสอบและผู้ตรวจให้คะแนนแบบทดสอบอัตโนมัติ ในการจัดคนคุมสอบ ผู้วิจัยได้ขอความร่วมมืออาจารย์ในสาขาวิชาภาษาตะวันตกเป็นผู้ร่วมคุมสอบกับผู้วิจัย 1 คน

และเป็นผู้ตรวจแบบทดสอบการเขียน 3 คน และดำเนินทดสอบการพูด 3 คน ดังรายชื่อแนบในภาคผนวก 5

1.1.2) การประชุมกรรมการคุมสอบและผู้ตรวจให้คะแนน ผู้วิจัยได้เชิญกรรมการคุมสอบและผู้ตรวจให้คะแนนทั้งการเขียนและการพูดมาประชุม 1 สัปดาห์ก่อนวันสอบ โดยอธิบายจุดประสงค์ของการวิจัย ขั้นตอนในการทำวิจัยเรื่องนี้ ตลอดจนขั้นตอนต่างๆ ที่เกี่ยวข้องกับการเก็บข้อมูล และอธิบายเกณฑ์การให้คะแนน พร้อมทั้งแบบฟอร์มให้คะแนน เพื่อให้กรรมการทุกคนได้รับทราบข้อมูลและเข้าใจทุกอย่างตรงกัน

1.1.3) จอห้องสอบผ่านระบบการจองห้องในเว็บไซต์มหาวิทยาลัย ได้กำหนดเวลาสอบ คือ การทดลองครั้งที่ 1 สอบวันพุธที่ 16 และ 23 มกราคม 2562 และการทดลองครั้งที่ 2 สอบวันพุธที่ 30 มกราคม และ 6 กุมภาพันธ์ 2562 เวลา 13.00 -17.00 น. ณ ห้องประชุม 15403

1.1.4) จัดพิมพ์แบบทดสอบและกระดาษคำตอบ และทำสำเนาให้เพียงพอตามจำนวนผู้เข้าสอบที่ได้ลงทะเบียนไว้แล้ว พร้อมทั้งให้มีแบบทดสอบสำรองอย่างน้อย 5 ชุด

1.1.5) ตรวจสอบความพร้อมของห้องสอบ โดยเฉพาะระบบเสียง และไฟล์เสียงที่จะใช้เปิดสำหรับการทดสอบทักษะการฟังให้อยู่สภาพที่พร้อมใช้งานก่อนวันสอบ 1 วัน

1.2) ขั้นตอนเก็บข้อมูล

ในขั้นการเก็บข้อมูลของการวิจัยครั้งนี้ ได้แบ่งออกเป็น 2 รอบ คือ 1) การเก็บข้อมูลด้วยการสอบโดยใช้แบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test) ที่ผู้วิจัยสร้างขึ้น และ 2) การเก็บข้อมูลด้วยการสอบโดยใช้แบบทดสอบมาตรฐาน (Standard English Proficiency Test) โดยมีกิจกรรมที่เกี่ยวข้องในการจัดสอบแต่ละชนิดดังนี้

1.2.1) การสอบโดยใช้แบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย

การเก็บข้อมูลด้วยแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test) สำหรับการทดลองครั้งที่ 1 ดำเนินการในวันพุธที่ 16 มกราคม 2562 และการทดลองครั้งที่ 2 ดำเนินการในวันพุธที่ 30 มกราคม 2562 เวลา 13.00 – 17.00 ณ ห้อง 15403 อาคาร 15 มหาวิทยาลัยทักษิณ วิทยาเขตสงขลา ผู้วิจัยได้ดำเนินการตามลำดับดังนี้

1.2.1.1) ให้ผู้เข้าสอบเข้าประจำที่นั่งในห้องสอบก่อนเวลา 15 นาที

1.2.1.2) ผู้วิจัยชี้แจงวัตถุประสงค์ของการสอบ และอธิบายถึงขั้นตอนและเวลาในการทำแบบทดสอบ และเน้นย้ำว่าผลการสอบไม่มีผลใด ๆ เกี่ยวกับคะแนนสอบในรายวิชาภาษาอังกฤษที่ผู้เข้าสอบกำลังเรียนอยู่

1.2.1.3) กรรมการคุมสอบ แจกแบบทดสอบและกระดาษคำตอบของแบบทดสอบวัดความสามารถ ภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test) โดยใช้เวลาในการทำแบบทดสอบแต่ละตอน ดังนี้

การฟัง	ใช้เวลา 50 นาที (13.00 – 13.50)
การใช้ภาษา	ใช้เวลา 40 นาที (13.50 – 14.30)
การอ่าน	ใช้เวลา 90 นาที (14.30 – 16.00)
การเขียน	ใช้เวลา 60 นาที (16.00 – 17.00)

1.2.1.4) ผู้เข้าสอบเริ่มทำแบบทดสอบ เวลา 13.00 น.

1.2.1.5) ผู้เข้าสอบส่งแบบทดสอบและกระดาษคำตอบคืนกรรมการคุมสอบ และสแกน QR-Code เพื่อทำแบบสอบถามความคิดเห็นต่อแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย

1.2.1.6) แจ้งผู้เข้าสอบให้ทราบ ผู้เข้าสอบบางคนจะได้รับการติดต่อทางโทรศัพท์ให้มาทดสอบการพูด โดยการทดลองครั้งที่ 1 สอบการพูดในวันจันทร์ที่ 21 มกราคม 2562 การทดลองครั้งที่ 2 สอบพูดในวันจันทร์ที่ 28 มกราคม 62 เวลา 16.00 – 18.00 น. ณ ห้องประชุมสาขาวิชาภาษาตะวันตก คณะมนุษยศาสตร์และสังคมศาสตร์ มหาวิทยาลัยทักษิณ วิทยาเขตสงขลา

1.2.1.7) จัดสอบทักษะการพูดให้กับผู้เข้าสอบที่เป็นกลุ่มเป้าหมาย โดยในการทดลองครั้งที่ 1 มีจำนวน 18 คน และการทดลองครั้งที่ 2 มีจำนวน 15 คน โดยเป็นผู้ที่ได้คะแนนรวมจากการสอบรอบแรก อยู่ในกลุ่มเก่ง ปานกลาง และอ่อน คัดเลือกโดยใช้เกณฑ์ 20 % ของจำนวนผู้เข้าสอบ

1.2.2) การสอบโดยใช้แบบทดสอบวัดความสามารถภาษาอังกฤษมาตรฐาน

การเก็บข้อมูลด้วยแบบทดสอบวัดความสามารถภาษาอังกฤษมาตรฐาน (Standard English Proficiency Test) สำหรับการทดลองครั้งที่ 1 ดำเนินการในวันพุธที่ 23 มกราคม 2562 และการทดลองครั้งที่ 2 ดำเนินการในวันพุธที่ 6 กุมภาพันธ์ 2562 เวลา 13.00 – 16.05 ณ ห้อง 15403 อาคาร 15 มหาวิทยาลัยทักษิณ วิทยาเขตสงขลา ผู้วิจัยได้ดำเนินการตามลำดับดังนี้

1.2.2.1) ให้ผู้เข้าสอบเข้าประจำที่นั่งในห้องสอบก่อนเวลา 15 นาที

1.2.2.2) ผู้วิจัยชี้แจงวัตถุประสงค์ของการสอบ และอธิบายถึงขั้นตอนและเวลาในการทำแบบทดสอบ และเน้นย้ำว่าผลการสอบไม่มีผลใด ๆ เกี่ยวกับคะแนนสอบในรายวิชาภาษาอังกฤษที่ผู้เข้าสอบกำลังเรียนอยู่

1.2.2.3) กรรมการคุมสอบ แจกแบบทดสอบและกระดาษคำตอบของแบบทดสอบวัดความสามารถ ภาษาอังกฤษมาตรฐาน (Standard English Proficiency Test) โดยใช้เวลาในการทำแบบทดสอบแต่ละตอน ดังนี้

การฟัง	ใช้เวลา 50 นาที (13.00 – 13.50)
การใช้ภาษา	ใช้เวลา 25 นาที (13.50 – 14.15)
การอ่าน	ใช้เวลา 50 นาที (14.15 – 15.05)
การเขียน	ใช้เวลา 60 นาที (15.05 – 16.05)

- 1.2.2.4) ผู้เข้าสอบเริ่มทำแบบทดสอบ เวลา 13.00 น.
- 1.2.2.5) ผู้เข้าสอบส่งแบบทดสอบและกระดาษคำตอบคืนกรรมการคุมสอบ
- 1.2.2.6) แจ้งผู้เข้าสอบให้ทราบ ผู้เข้าสอบบางคนจะได้รับการติดต่อทางโทรศัพท์ให้มาทดสอบการพูด สำหรับการทดลองครั้งที่ 1 ทดสอบพูดในวันจันทร์ที่ 28 มกราคม 2562 และการทดลองครั้งที่ 1 ทดสอบพูดในวันจันทร์ที่ 11 กุมภาพันธ์ 2562 เวลา 16.00 – 18.00 น. ณ ห้องประชุมสาขาวิชา ภาษาตะวันตก คณะมนุษยศาสตร์และสังคมศาสตร์ มหาวิทยาลัยทักษิณ วิทยาเขตสงขลา
- 1.2.2.7) จัดสอบทักษะการพูดให้กับผู้เข้าสอบที่เป็นกลุ่มเป้าหมาย โดยในการทดลองครั้งที่ 1 มีจำนวน 18 คน และการทดลองครั้งที่ 2 มีจำนวน 15 คน โดยเป็นผู้ที่ได้คะแนนอยู่ในกลุ่มเก่ง ปานกลาง และอ่อน

1.3) ขั้นตอนการตรวจแบบทดสอบ

ในขั้นตอนการตรวจแบบทดสอบ ผู้วิจัยได้แบ่งกระดาษคำตอบออกเป็น 2 ชนิด คือ 1) แบบทดสอบปรนัย และ 2) แบบทดสอบอัตนัย โดยแบบทดสอบแต่ละชนิดมีขั้นตอนในการตรวจให้คะแนนดังนี้

1.3.1) การตรวจแบบทดสอบปรนัย

ในการตรวจแบบทดสอบปรนัย ผู้วิจัยได้ดำเนินการตามขั้นตอนดังนี้

- 1.3.1.1) ตรวจสอบกระดาษคำตอบว่า ผู้เข้าสอบได้เขียนชื่อ-สกุล รหัสประจำตัว และให้คำตอบครบถ้วนทุกข้อหรือไม่
- 1.3.1.2) เปิดโปรแกรมวิเคราะห์แบบทดสอบ B-Index Version 7.0 (สาคร แสงผิ้ง, 2550) และป้อนคำตอบของแบบทดสอบการฟัง การใช้ภาษา และการอ่าน ลงในโปรแกรม
- 1.3.1.3) คีย์คำตอบของผู้เข้าสอบเข้าโปรแกรมจนครบทุกคน
- 1.3.1.4) วิเคราะห์ข้อสอบรายข้อ และสิ่งพิมพ์รายงานผลการวิเคราะห์นำมาใช้ในการวิเคราะห์ประสิทธิภาพแบบทดสอบต่อไป

1.3.2) การตรวจแบบทดสอบแบบทดสอบอัตนัย

แบบทดสอบอัตนัยของการทดสอบครั้งนี้มี 2 ชนิด คือ 1) แบบทดสอบการเขียน และ 2) แบบทดสอบการพูด โดยมีขั้นตอนในการตรวจให้คะแนนดังนี้

13.2.1) การตรวจแบบทดสอบการเขียน

ในการตรวจแบบทดสอบการเขียน ผู้วิจัยได้ดำเนินการดังนี้

คัดแยกกระดาษคำตอบการเขียน และทำสำเนา เป็น 3 ชุด

ส่งสำเนากระดาษคำตอบและแบบประเมินการเขียนให้คณะกรรมการตรวจแบบทดสอบการเขียนทั้ง 3 คน (ดังรายชื่อในภาคผนวก จ)

คณะกรรมการตรวจแบบทดสอบการเขียนให้คะแนนงานเขียนตามเกณฑ์การประเมิน และกรอกผลการคะแนนลงในแบบฟอร์มที่กำหนด และผู้วิจัยตรวจสอบความสอดคล้องในการตรวจให้คะแนนของกรรมการผู้ตรวจทั้ง 3 คน ว่าสอดคล้องหรือไม่ โดยการหาความเชื่อมั่นของการให้คะแนนแบบทดสอบอัตนัย ใช้วิธี Inter-Rater Reliability สถิติที่ใช้คือ The Spearman-Brown Prophecy Formula (Henning, 1987: 83) โดยมีสูตรในการคำนวณ ดังนี้

$$r_{tt} = \frac{nr_{A,B}}{1 + (n-1)r_{A,B}}$$

โดยที่ r_{tt} = ค่าสัมประสิทธิ์ความเชื่อมั่นระหว่างผู้ให้คะแนน
 n = จำนวนผู้ให้คะแนนที่ถูกนำมาใช้ในการคำนวณคะแนนสุดท้ายของ ผู้เข้าสอบ
 $r_{A,B}$ = ความสัมพันธ์ระหว่างผู้ให้คะแนน หรือ ค่าเฉลี่ยความสัมพันธ์ในกลุ่มผู้ให้คะแนนทั้งหมดถ้ามีผู้ให้คะแนนมากกว่า 2 คน

ได้ค่าความเชื่อมั่นของการให้คะแนนแบบทดสอบอัตนัย (การเขียน) $r_{tt} = .99$ โดยทั้ง 2 ภาระงานเขียนย่อย Task A และ Task B มีค่า $r_{tt} = .98$ เมื่อผู้ตรวจให้คะแนนตรวจการเขียนมีความสอดคล้องกัน จึงดำเนินการตรวจข้อสอบเขียนต่อไปจนจบ

คณะกรรมการส่งสำเนาแบบทดสอบและผลคะแนนคืนผู้วิจัย
 ผู้วิจัยตรวจสอบการให้คะแนนของคณะกรรมการตรวจแบบทดสอบการเขียน และบันทึกผลคะแนน

1.3.2.2) การตรวจแบบทดสอบการพูด

ในการตรวจให้คะแนนการสอบพูด ดำเนินการในวันที่มีการสอบพูด ผู้วิจัยได้ดำเนินการดังนี้

จัดผู้เข้าสอบที่คัดเลือกมาแล้ว เข้าห้องสอบเพื่อพบกับคณะกรรมการ 3 คน โดยใช้เวลาในการทดสอบพูดคนละ ประมาณ 10 นาที

คณะกรรมการให้ผู้เข้าสอบแนะนำตัว และถามคำถามที่เกี่ยวกับข้อมูลส่วนบุคคลประมาณ 3 นาที

คณะกรรมการให้ผู้เข้าสอบอ่านบัตรภาระงาน และพูดตามภาระงานที่กำหนด ใช้เวลาประมาณ 3 นาที

คณะกรรมการพูดคุยสอบถามความคิดเห็นผู้เข้าสอบเกี่ยวกับประเด็นในบัตรภาระงาน โดยให้ผู้เข้าสอบพูดแสดงความคิดเห็น ใช้เวลาประมาณ 4 นาที

คณะกรรมการให้คะแนนตามแบบประเมินการพูดที่กำหนดและผู้วิจัยตรวจสอบความสอดคล้องในการตรวจให้คะแนนของกรรมการผู้สอบทั้ง 3 คน ว่าสอดคล้องหรือไม่ โดยการหาความเชื่อมั่นของการให้คะแนนแบบทดสอบอัตนัย ใช้วิธี Inter-Rater Reliability สถิติที่ใช้คือ The Spearman-Brown Prophecy Formula (Henning, 1987: 83) โดยมีสูตรในการคำนวณ ดังนี้

$$r_{tt} = \frac{nr_{A,B}}{1 + (n-1)r_{A,B}}$$

โดยที่ r_{tt} = ค่าสัมประสิทธิ์ความเชื่อมั่นระหว่างผู้ให้คะแนน
 n = จำนวนผู้ให้คะแนนที่ถูกนำมาใช้ในการคำนวณคะแนนสุดท้ายของ ผู้เข้าสอบ
 $r_{A,B}$ = ความสัมพันธ์ระหว่างผู้ให้คะแนน หรือ ค่าเฉลี่ยความสัมพันธ์ในกลุ่มผู้ให้คะแนนทั้งหมดถ้ามีผู้ให้คะแนนมากกว่า 2 คน

ได้ค่าความเชื่อมั่นของการให้คะแนนแบบทดสอบอัตนัย (การพูด) $r_{tt} = .98$ โดยภาระงานพูดย่อย Task A มีค่า $r_{tt} = .97$ และทั้งภาระงานพูด Task B และ C มีค่า $r_{tt} = .98$ เมื่อผู้สอบให้คะแนนตรวจการพูดมีความสอดคล้องกัน จึงดำเนินการสอบพูดต่อไปจนจบ

คณะกรรมการประเมินส่งผลการสอบให้ผู้วิจัย
 ผู้วิจัยตรวจสอบการให้คะแนนของคณะกรรมการทดสอบการพูดและบันทึกคะแนนสอบ

1.4) ขั้นตอนการวิเคราะห์ประสิทธิภาพและตรวจสอบคุณภาพแบบทดสอบโดยรวม

1.4.1) การวิเคราะห์ประสิทธิภาพและตรวจสอบคุณภาพแบบทดสอบโดยรวมหลังการทดลองครั้งที่ 1

1.4.1.1) การวิเคราะห์ประสิทธิภาพแบบทดสอบ

การวิจัยครั้งนี้ แบ่งการวิเคราะห์ประสิทธิภาพของแบบทดสอบออกเป็น 4 ชนิด คือ 1) การวิเคราะห์ค่าสถิติพื้นฐานของคะแนนสอบ 2) การวิเคราะห์แบบทดสอบรายข้อ 3) การวิเคราะห์ความเชื่อมั่นของแบบทดสอบ ดังรายละเอียดดังต่อไปนี้

1.4.1.1.1) การวิเคราะห์ค่าสถิติพื้นฐานของคะแนนสอบ

ค่าสถิติพื้นฐานของคะแนนสอบที่ผู้วิจัยต้องการวิเคราะห์ คือ คะแนนสูงสุด คะแนนต่ำสุด คะแนนเฉลี่ย และส่วนเบี่ยงเบนมาตรฐาน โดยดำเนินการตามขั้นตอนดังนี้

ก) สั่งพิมพ์ผลคะแนนรายบุคคลของผู้เข้าสอบในแต่ละทักษะที่เข้าสอบจากโปรแกรมวิเคราะห์แบบทดสอบ B-Index Version 7.0 โดยใช้เกณฑ์ 27% ในการคัดเลือกผู้เข้าสอบที่ได้คะแนนในกลุ่มสูงและกลุ่มต่ำ (สาคร แสงผึ้ง, 2550)

ข) พิมพ์คะแนนรายบุคคลของผู้เข้าสอบ ทั้งแบบตอนย่อยและรวมแต่ละทักษะ ลงในตารางรายงานค่าสถิติพื้นฐานของคะแนนสอบจากการทดลองครั้งที่ 1

ค) ได้ค่าสถิติพื้นฐานของคะแนนสอบจากการทดลองครั้งที่ 1 ดังแสดงในตารางที่ 20 - 23



ตารางที่ 20 รายงานผลคะแนนการสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย – การทดลองครั้งที่ 1

ที่	รหัสประจำตัว	ฟัง			รวม	%	การใช้ภาษา		รวม	%	อ่าน		รวม	%	เขียน		รวม	%	รวมทุกทักษะ	%	กลุ่ม	
		A	B	C			A	B			A	B			A	B						
1	601011569	24	14	11	49	56.32	16	22	38	60	100	49	48	97	100	20	40	60	100	304	100	
2	601011595	31	9	10	50	57.47	19	30	49	60	81.67	37	27	64	65.98	5	14	19	31.67	182	59.87	
3	601011536	24	13	10	47	54.02	9	16	25	26	41.67	35	28	63	64.95		31.33	31.33	52.22	166.33	54.71	
4	581031308	13	6	10	29	33.33	10	16	26	26	43.33	33	22	55	56.7	15	26	41	68.33	151	49.67	
5	601011562	18	6	9	33	37.93	8	18	26	26	43.33	28	25	53	54.64	4.33	16	20.33	33.88	132.33	43.53	
6	591011521	13	8	10	31	35.63	15	20	35	35	58.33	32	34	66	68.04					132	43.42	
7	601011600	17	7	7	31	35.63	13	21	34	34	56.67	32	30	62	63.92					127	41.78	
8	591011545	18	7	9	34	39.08	13	22	35	35	58.33	32	24	56	57.73					125	41.12	
9	581031305	7	6	11	24	27.59	4	19	23	23	38.33	23	21	44	45.36	11.67	20	31.67	52.78	122.67	40.35	
10	581031299	13	6	7	26	29.89	11	11	22	22	36.67	28	27	55	56.7	5	10.67	15.67	26.12	118.67	39.04	
11	581031093	14	7	9	30	34.48	10	13	23	23	38.33	32	16	48	49.48	9.33	5.33	14.66	24.43	115.66	38.05	
12	581031295	16	10	10	36	41.38	13	18	31	31	51.67	32	14	46	47.42					113	37.17	
13	591011517	12	5	8	25	28.74	8	20	28	28	46.67	23	26	49	50.52	11		11	18.33	113	37.17	
14	581031315	11	9	10	30	34.48	9	18	27	27	45	32	21	53	54.64					110	36.18	
15	601011572	19	7	9	35	40.23	8	17	25	25	41.67	25	24	49	50.52					109	35.86	
16	581031302	18	6	7	31	35.63	7	19	26	26	43.33	27	23	50	51.55	1.33		1.33	2.217	108.33	35.63	
17	581031311	10	6	8	24	27.59	9	11	20	20	33.33	28	27	55	56.7	2	4	6	10	105	34.54	

ตาราง 20 รายงานผลคะแนนการสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย – การทดลองครั้งที่ 1 (ต่อ)

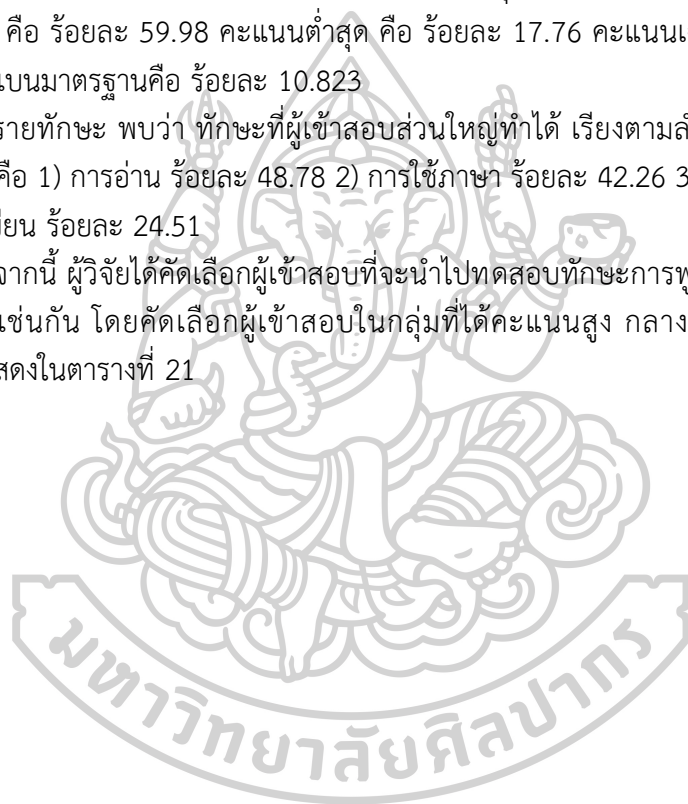
ที่	รหัสประจำตัว	ฟัง			รวม	%	การใช้ภาษา		รวม	%	อ่าน		รวม	%	เขียน		รวม	%	รวมทุกทักษะ	%	กลุ่ม
		A	B	C			A	B			A	B			A	B					
18	591071009	15	2	6	23	26.44	11	21	32	53.33	24	17	41	42.27					96	31.58	
19	581031309	11	3	10	24	27.59	6	18	24	40	25	18	43	44.33	1		1	1.667	92	30.26	
20	581031398	12	7	8	27	31.03	8	11	19	31.67	21	18	39	40.21	2	4	6	10	91	29.93	
21	591071030	8	7	8	23	26.44	6	15	21	35	26	13	39	40.21					83	27.30	
22	591071098	10	6	9	25	28.74	6	9	15	25	27	9	36	37.11	3		3	5	79	25.99	
23	591071106	14	4	7	25	28.74	3	7	10	16.67	21	20	41	42.27		2	2	3.333	78	25.66	
24	591071065	10	6	9	25	28.74	7	17	24	40	16	11	27	27.84					76	25.00	
25	591071006	10	4	6	20	22.99	5	11	16	26.67	21	14	35	36.08	2		2	3.333	73	24.01	
26	591071084	12	2	9	23	26.44	8	7	15	25	19	16	35	36.08					73	24.01	
27	591071052	6	5	5	16	18.39	6	13	19	31.67	19	17	36	37.11					71	23.36	
28	581031400	7	0	6	13	14.94	9	13	22	36.67	18	1	19	19.59					54	17.76	
	Max	31.00	14.00	11.00	50.00	57.47	19.00	30.00	49.00	81.67	37.00	34.00	66.00	68.04	15.00	31.33	41.00	68.33	182.34	59.98	
	Min	6.00	0.00	5.00	13.00	14.94	3.00	7.00	10.00	16.67	16.00	1.00	19.00	19.59	1.00	2.00	1.00	1.67	54.00	17.76	
	\bar{X}	14.04	6.36	8.50	28.89	33.21	9.18	16.18	25.36	42.26	26.86	20.46	47.32	48.78	5.81	14.00	14.71	24.51	109.98	36.18	
	S.D.	5.680	2.984	1.622	8.757	10.066	3.692	5.179	8.098	13.496	5.842	7.331	11.851	12.218	4.510	9.793	12.943	21.571	32.902	10.823	

ตารางที่ 20 แสดงผลคะแนนการสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัยหลังการทดลองครั้งที่ 1 ซึ่งประกอบด้วย 4 ทักษะ คือ การฟัง การใช้ภาษา การอ่านและการเขียน และแสดงคะแนนโดยจัดเรียงจากคะแนนสูงสุดไปหาต่ำสุด พบว่า จากจำนวนผู้เข้าสอบ 28 คน มีผู้เข้าสอบที่ทำแบบทดสอบครบทั้ง 4 ทักษะ จำนวน 16 คน และมีผู้เข้าสอบที่ไม่ทำแบบทดสอบการเขียน 12 คน ซึ่งส่วนใหญ่เป็นผู้ที่อยู่ในกลุ่มที่ได้คะแนนต่ำ นอกจากนี้ ในจำนวนผู้ที่ทำแบบทดสอบการเขียน 16 คน สามารถแบ่งได้เป็น 2 กลุ่ม คือ 1) ผู้ที่ทำครบทั้ง 2 ภาระงาน จำนวน 9 คน และ 2) ผู้ที่ทำเพียงภาระงานเดียว จำนวน 7 คน ซึ่งส่วนใหญ่เป็นผู้ที่อยู่ในกลุ่มที่ได้คะแนนต่ำและเลือกทำในภาระงานที่ 1 ซึ่งเป็นการเขียนบรรยายกราฟ

เมื่อพิจารณาคะแนนสอบ พบว่า จากคะแนนรวมทุกทักษะ ค่าร้อยละของคะแนนสูงสุดที่ผู้เข้าสอบทำได้ คือ ร้อยละ 59.98 คะแนนต่ำสุด คือ ร้อยละ 17.76 คะแนนเฉลี่ย คือ ร้อยละ 36.18 และส่วนเบี่ยงเบนมาตรฐานคือ ร้อยละ 10.823

เมื่อพิจารณารายทักษะ พบว่า ทักษะที่ผู้เข้าสอบส่วนใหญ่ทำได้ เรียงตามลำดับตามค่าร้อยละของคะแนนเฉลี่ย คือ 1) การอ่าน ร้อยละ 48.78 2) การใช้ภาษา ร้อยละ 42.26 3) การฟัง ร้อยละ 33.21 และ 4) การเขียน ร้อยละ 24.51

นอกจากนี้ ผู้วิจัยได้คัดเลือกผู้เข้าสอบที่จะนำไปทดสอบทักษะการพูดจากผลคะแนนรวมในตารางที่ 20 เช่นกัน โดยคัดเลือกผู้เข้าสอบในกลุ่มที่ได้คะแนนสูง กลาง ต่ำ กลุ่มละ 6 คน ดังรายละเอียดแสดงในตารางที่ 21

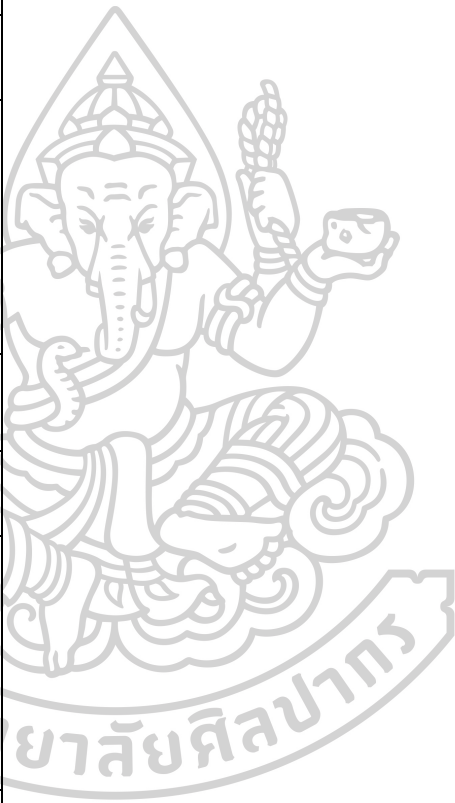


ตารางที่ 21 รายงานผลคะแนนการสอบการพูดโดยแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย – การทดลองครั้งที่ 1

ที่	รหัสประจำตัว	Task A			รวม	%	Task B			รวม	%	Task C			รวม	%	Speaking			รวม	%	กลุ่ม		
		1	2	3			1	2	3			1	2	3			1	2	3					
1	601011595	20.00	15.00	14.00	20.00	100.00	20.00	17.00	16.00	20.00	100.00	20.00	15.00	16.00	20.00	20.00	100.00	60.00	60.00	60.00	60.00	100.00		
2	601011569	13.00	15.00	16.00	14.67	73.33	14.00	14.00	17.00	15.00	75.00	13.00	15.00	14.00	14.00	14.00	14.00	70.00	40.00	44.00	47.00	43.67	72.78	
3	601011536	14.00	15.00	16.00	15.00	75.00	16.00	16.00	15.00	15.00	75.00	14.00	15.00	16.00	15.00	15.00	75.00	44.00	44.00	47.00	45.00	75.00	สูง	
4	581031295	14.00	13.00	12.00	13.00	65.00	14.00	14.00	14.00	15.00	71.67	14.00	13.00	14.00	13.67	68.33	42.00	40.00	41.00	41.00	41.00	68.33	สูง	
5	601011572	15.00	14.00	15.00	14.67	73.33	15.00	15.00	12.00	14.00	70.00	15.00	14.00	15.00	14.67	73.33	45.00	43.00	42.00	42.00	43.33	72.22		
6	591011545	13.00	13.00	14.00	13.33	66.67	14.00	14.00	13.00	13.00	66.67	13.00	13.00	14.00	13.33	66.67	40.00	39.00	41.00	41.00	40.00	66.67		
7	581031315	10.00	12.00	11.00	11.00	55.00	10.00	10.00	11.00	11.00	51.67	10.00	11.00	11.00	10.67	53.33	30.00	33.00	33.00	33.00	32.00	53.33		
8	581031308	12.00	12.00	10.00	11.33	56.67	12.00	11.00	10.00	10.00	55.00	12.00	10.00	10.00	10.67	53.33	36.00	33.00	30.00	33.00	33.00	55.00		
9	581031398	11.00	11.00	12.00	11.33	56.67	10.00	10.00	10.00	10.33	51.67	11.00	10.00	12.00	11.00	55.00	32.00	32.00	34.00	32.67	32.67	54.44		
10	581031299	11.00	10.00	13.00	11.33	56.67	11.00	11.00	11.00	10.67	53.33	11.00	11.00	13.00	11.67	58.33	33.00	31.00	37.00	33.67	33.67	56.11		
11	591011517	11.00	12.00	12.00	11.67	58.33	11.00	11.00	11.00	10.67	53.33	11.00	11.00	12.00	11.33	56.67	33.00	33.00	35.00	33.67	33.67	56.11		
12	591071065	10.00	12.00	11.00	11.00	55.00	10.00	10.00	11.00	10.33	51.67	10.00	11.00	11.00	10.67	53.33	30.00	33.00	33.00	32.00	32.00	53.33		
13	591071009	5.00	6.00	5.00	5.33	26.67	7.00	8.00	6.00	7.00	35.00	8.00	6.00	5.00	6.33	31.67	20.00	20.00	16.00	18.67	18.67	31.11		
14	591071030	6.00	3.00	5.00	4.67	23.33	5.00	6.00	5.00	5.33	26.67	6.00	3.00	5.00	4.67	23.33	17.00	12.00	15.00	14.67	14.67	24.44		
15	591071084	7.00	6.00	4.00	5.67	28.33	5.00	6.00	5.00	5.33	26.67	6.00	6.00	4.00	5.33	26.67	18.00	18.00	13.00	16.33	16.33	27.22		
16	591071006	5.00	4.00	5.00	4.67	23.33	5.00	6.00	4.00	5.00	25.00	6.00	4.00	5.00	5.00	25.00	16.00	14.00	14.00	14.67	14.67	24.44		
17	591071052	6.00	4.00	4.00	4.67	23.33	5.00	5.00	4.00	4.67	23.33	5.00	4.00	4.00	4.33	21.67	16.00	13.00	12.00	13.67	13.67	22.78		
18	581031400	4.00	5.00	6.00	5.00	25.00	5.00	5.00	5.00	5.33	26.67	5.00	6.00	6.00	5.67	28.33	14.00	16.00	18.00	16.00	16.00	26.67		

ตารางที่ 21 รายงานผลคะแนนการสอบการพูดโดยแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย – การทดลองครั้งที่ 1 (ต่อ)

ที่	Task A			รวม	%	Task B			รวม	%	Task C			รวม	%	Speaking			รวม	%	กลุ่ม
	1	2	3			1	2	3			1	2	3			1	2	3			
รหัสประจำตัว	20.00	20.00	20.00	20.00	100.00	20.00	20.00	20.00	20.00	100.00	20.00	20.00	20.00	20.00	100.00	60.00	60.00	60.00	60.00	100.00	
Max	15.00	16.00	16.00	15.00	75.00	17.00	16.00	17.00	16.00	80.00	15.00	16.00	16.00	15.67	78.33	47.00	48.00	47.00	46.67	77.78	
Min	4.00	3.00	4.00	4.67	23.33	5.00	5.00	4.00	4.67	23.33	5.00	3.00	4.00	4.33	21.67	14.00	12.00	12.00	13.67	22.78	
\bar{X}	10.11	10.17	10.28	10.19	50.93	10.33	10.22	10.06	10.20	51.02	10.28	9.94	10.39	10.20	51.02	30.72	30.33	30.72	30.59	50.99	
S.D.	3.71	4.30	4.30	4.02	20.08	4.17	3.61	4.17	3.92	19.60	3.49	4.15	4.37	3.95	19.75	11.28	11.92	12.65	11.84	19.73	
r_{tt}	.97					.98					.98					.98					



ตารางที่ 21 แสดงผลคะแนนการสอบการพูดโดยแบบทดสอบวัดความสามารถภาษาอังกฤษ ระดับมหาวิทยาลัยหลังการทดลองใช้แบบทดสอบครั้งที่ 1 และมีจำนวนผู้เข้าสอบที่ได้รับการคัดเลือก มาตามคะแนนที่ได้ในกลุ่มสูง กลาง และต่ำ กลุ่มละ 6 คน รวม 18 คน พบว่า ผู้เข้าสอบทุกคนได้เข้า ทดสอบการพูดครบทั้ง 3 ภาระงาน โดยในภาพรวมของทุกภาระงาน ค่าร้อยละของคะแนนสูงสุดที่ผู้ เข้าสอบทำได้ คือ ร้อยละ 77.78 คะแนนต่ำสุด คือ ร้อยละ 22.78 คะแนนเฉลี่ย คือ ร้อยละ 50.99 และส่วนเบี่ยงเบนมาตรฐานคือ ร้อยละ 19.73 และค่าความเชื่อมั่นของผู้ให้คะแนน (r_{tt}) เท่ากับ .98

เมื่อพิจารณารายภาระงาน พบว่า ภาระงานที่ผู้เข้าสอบส่วนใหญ่ทำได้ เรียงตามลำดับตามค่า ร้อยละของคะแนนเฉลี่ย คือ 1) ภาระงาน B และ C ร้อยละ 51.02 โดยทั้งภาระงาน B และ C มีค่า ความเชื่อมั่นของผู้ให้คะแนน (r_{tt}) เท่ากับ .98 และ 2) ภาระงาน A ร้อยละ 50.93 และค่าความ เชื่อมั่นของผู้ให้คะแนน (r_{tt}) เท่ากับ .97



ตารางที่ 22 รายงานผลคะแนนการสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัยโดยใช้แบบทดสอบมาตรฐาน- การทดลองครั้งที่ 1

ที่	รหัสประจำตัว	ฟัง		รวม	%	การใช้ภาษา		รวม	%	อ่าน		%	เขียน		รวม	%	รวมทุกทักษะ	%	กลุ่ม
		A	B-C			A	B			A	B		A	B					
		30	20	50	100	15	25	40	100	50	50	100	20	40	60	100	200	100	
1	601011569	13	15	28	56	9	16	25	62.5	26	26	52	8	20	28	46.67	107.00	53.50	
2	581031295	15	7	22	44	7	10	17	42.5	29	29	58	12	23.33	35.33	58.88	103.33	51.67	
3	601011595	15	13	28	56	11	18	29	72.5	28	28	56	5	13	18	30	103.00	51.50	สูง
4	601011536	7	12	19	38	7	13	20	50	27	27	54		28	28	46.67	94.00	47.00	
5	581031308	10	9	19	38	6	10	16	40	21	21	42	11	22.67	33.67	56.12	89.67	44.84	
6	581031305	7	4	11	22	7	12	19	47.5	18	18	36	9	27.33	36.33	60.55	84.33	42.17	
7	601011562	12	10	22	44	8	12	20	50	23	23	46	4	15	19	31.67	84.00	42.00	
8	591011517	10	9	19	38	7	10	17	42.5	29	29	58	12		12	20	77.00	38.50	
9	591011521	12	10	22	44	9	12	21	52.5	28	28	56			0	0	71.00	35.50	
10	581031299	6	9	15	30	5	8	13	32.5	19	19	38	9	12.33	21.33	35.55	68.33	34.17	
11	581031093	7	7	14	28	5	5	10	25	22	22	44		15.33	15.33	25.55	61.33	30.67	
12	581031302	12	7	19	38	7	13	20	50	19	19	38		2	2	3.33	60.00	30.00	
13	601011600	5	9	14	28	5	19	24	60	21	21	42			0	0	59.00	29.50	
14	591011545	8	5	13	26	7	18	25	62.5	21	21	42			0	0	59.00	29.50	กลาง
15	591071098	8	9	17	34	4	11	15	37.5	15	15	30	9		9	15	56.00	28.00	
16	581031309	7	5	12	24	7	11	18	45	16	16	32	4	4	8	13.33	54.00	27.00	
17	581031311	7	9	16	32	5	11	16	40	15	15	30	2	4	6	10	53.00	26.50	

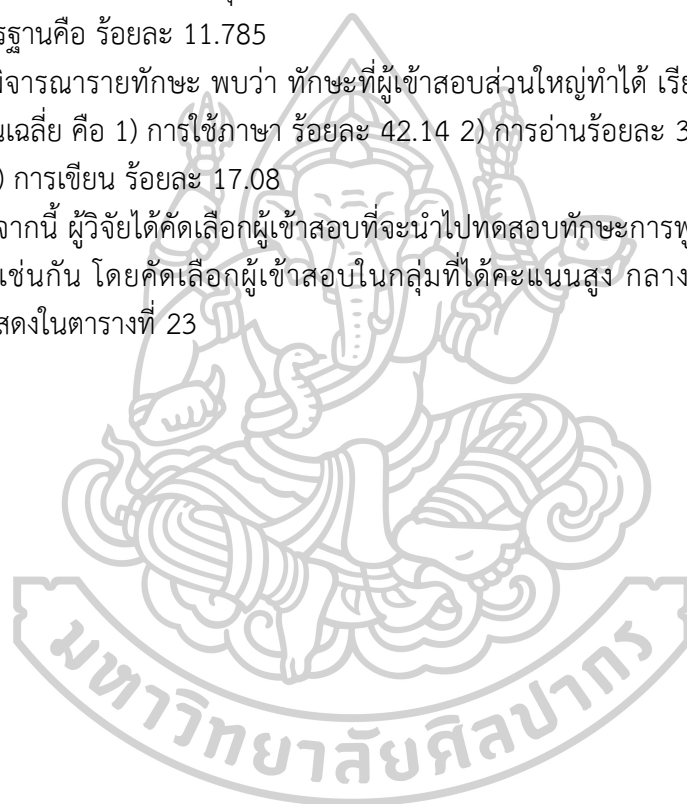
ตารางที่ 22 รายงานผลคะแนนการสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัยโดยใช้แบบทดสอบมาตรฐาน- การทดลองครั้งที่ 1 (ต่อ)

ที่	รหัสประจำตัว	ฟัง		รวม	%	การใช้ภาษา		รวม	%	อ่าน	%	เขียน		รวม	%	รวมทุกทักษะ	%	กลุ่ม
		A	B-C			A	B					A	B					
18	601011572	30	20	50	100	15	25	40	100	50	100	20	40	60	100	200	100	
19	591071006	7	8	15	42	7	9	16	40	15	30			0	0	52.00	26.00	
20	581031315	5	3	8	16	8	11	19	47.5	14	28			0	0	41.00	20.50	
21	581031398	5	4	9	18	3	5	8	20	13	26	6	4	10	16.67	40.00	20.00	
22	581031400	3	6	9	18	6	7	13	32.5	15	30	1		1	1.667	38.00	19.00	
23	591071009	6	2	8	16	5	8	13	32.5	15	30			0	0	36.00	18.00	
24	591071106	6	2	8	16	5	7	12	30	16	32			0	0	36.00	18.00	
25	591071084	6	7	13	26	4	7	11	27.5	10	20			0	0	34.00	17.00	
26	591071030	5	7	12	24	7	5	12	30	10	20			0	0	34.00	17.00	
27	591071065	5	5	10	20	6	8	14	35	10	20			0	0	34.00	17.00	
28	591071052	6	3	9	18	5	12	17	42.5	8	16			0	0	34.00	17.00	
	Max	15.00	15.00	28.00	56.00	11.00	19.00	29.00	72.50	29.00	58.00	12.00	28.00	36.33	60.55	107.00	53.50	
	Min	3.00	2.00	8.00	16.00	3.00	5.00	8.00	20.00	8.00	16.00	1.00	2.00	0.00	0.00	34.00	17.00	
	X̄	8.04	7.39	15.43	30.86	6.32	10.54	16.86	42.14	18.57	37.14	6.71	13.79	10.25	17.08	61.11	30.55	
	S.D.	3.214	3.292	5.802	11.604	1.744	3.844	5.009	12.522	6.191	12.382	3.791	9.470	12.397	20.661	23.569	11.785	

ตารางที่ 22 แสดงผลคะแนนการสอบวัดความสามารถภาษาอังกฤษโดยใช้แบบทดสอบมาตรฐานหลังการทดลองครั้งที่ 1 ซึ่งประกอบด้วย 4 ทักษะ คือ การฟัง การใช้ภาษา การอ่านและการเขียน และแสดงคะแนนโดยจัดเรียงจากคะแนนสูงสุดไปหาต่ำสุด พบว่า จากจำนวนผู้เข้าสอบ 28 คน มีผู้เข้าสอบที่ทำแบบทดสอบครบทั้ง 4 ทักษะ จำนวน 17 คน และมีผู้เข้าสอบที่ไม่ทำแบบทดสอบการเขียน 11 คน ซึ่งส่วนใหญ่เป็นผู้ที่อยู่ในกลุ่มที่ได้คะแนนต่ำ นอกจากนี้ ในจำนวนผู้ที่ทำแบบทดสอบการเขียน 17 คน สามารถแบ่งได้เป็น 2 กลุ่ม คือ 1) ผู้ที่ทำครบทั้ง 2 ภาระงาน จำนวน 11 คน และ 2) ผู้ที่ทำเพียงภาระงานเดียว จำนวน 6 คน ซึ่งส่วนใหญ่เป็นผู้ที่อยู่ในกลุ่มที่ได้คะแนนต่ำ เมื่อพิจารณาคะแนนสอบ พบว่า จากคะแนนรวมทุกทักษะ ค่าร้อยละของคะแนนสูงสุดที่ผู้เข้าสอบทำได้ คือ ร้อยละ 53.50 คะแนนต่ำสุด คือ ร้อยละ 17.00 คะแนนเฉลี่ย คือ ร้อยละ 30.55 และส่วนเบี่ยงเบนมาตรฐานคือ ร้อยละ 11.785

เมื่อพิจารณารายทักษะ พบว่า ทักษะที่ผู้เข้าสอบส่วนใหญ่ทำได้ เรียงตามลำดับตามค่าร้อยละของคะแนนเฉลี่ย คือ 1) การใช้ภาษา ร้อยละ 42.14 2) การอ่านร้อยละ 37.14 3) การฟัง ร้อยละ 30.86 และ 4) การเขียน ร้อยละ 17.08

นอกจากนี้ ผู้วิจัยได้คัดเลือกผู้เข้าสอบที่จะนำไปทดสอบทักษะการพูดจากผลคะแนนรวมในตารางที่ 22 เช่นกัน โดยคัดเลือกผู้เข้าสอบในกลุ่มที่ได้คะแนนสูง กลาง ต่ำ กลุ่มละ 6 คน ดังรายละเอียดแสดงในตารางที่ 23



ตารางที่ 23 รายงานผลคะแนนการสอบการพูดโดยแบบทดสอบวัดความสามารภภาษาอังกฤษมาตรฐาน- การทดลองครั้งที่ 1

ที่	รหัสประจำตัว	Task A			รวม	%	Task B			รวม	%	Task C			รวม	%	Speaking			รวม	%	กลุ่ม
		1	2	3			1	2	3			1	2	3			1	2	3			
1	601011569	14.00	15.00	15.00	14.67	73.33	16.00	15.00	15.00	15.33	76.67	16.00	16.00	15.00	15.67	78.33	46.00	46.00	45.00	45.67	76.11	สูง
2	581031295	16.00	16.00	14.00	15.33	76.67	14.00	15.00	15.00	14.67	73.33	15.00	16.00	16.00	15.67	78.33	45.00	47.00	45.00	45.67	76.11	สูง
3	601011595	16.00	15.00	13.00	14.67	73.33	14.00	17.00	17.00	16.00	80.00	15.00	14.00	14.00	14.33	71.67	45.00	46.00	44.00	45.00	75.00	สูง
4	601011562	15.00	13.00	15.00	14.33	71.67	15.00	12.00	14.00	13.67	68.33	14.00	15.00	14.00	14.33	71.67	44.00	40.00	43.00	42.33	70.56	สูง
5	591011521	12.00	13.00	14.00	13.00	65.00	14.00	15.00	15.00	14.67	73.33	13.00	14.00	14.00	13.67	68.33	39.00	42.00	43.00	41.33	68.89	สูง
6	601011572	14.00	13.00	13.00	13.33	66.67	13.00	13.00	13.00	13.00	65.00	13.00	14.00	14.00	13.67	68.33	40.00	40.00	40.00	40.00	66.67	สูง
12	581031093	11.00	13.00	11.00	11.67	58.33	10.00	11.00	11.00	10.67	53.33	11.00	13.00	12.00	12.00	60.00	32.00	37.00	34.00	34.33	57.22	กลาง
13	591071006	12.00	12.00	11.00	11.67	58.33	11.00	10.00	11.00	10.67	53.33	10.00	12.00	12.00	11.33	56.67	33.00	34.00	34.00	33.67	56.11	กลาง
14	601011600	9.00	12.00	11.00	10.67	53.33	10.00	11.00	11.00	10.67	53.33	11.00	12.00	12.00	11.67	58.33	30.00	35.00	34.00	33.00	55.00	กลาง
15	581031311	10.00	11.00	10.00	10.33	51.67	10.00	11.00	11.00	10.67	53.33	11.00	11.00	11.00	11.00	55.00	31.00	33.00	32.00	32.00	53.33	กลาง
16	591011545	10.00	11.00	10.00	10.33	51.67	10.00	11.00	11.00	10.67	53.33	11.00	11.00	11.00	11.00	55.00	31.00	33.00	32.00	32.00	53.33	กลาง
17	581031299	12.00	10.00	12.00	11.33	56.67	11.00	10.00	10.00	10.33	51.67	10.00	10.00	10.00	10.00	50.00	33.00	30.00	32.00	31.67	52.78	กลาง
23	581031398	6.00	5.00	5.00	5.33	26.67	8.00	6.00	6.00	6.67	33.33	6.00	5.00	5.00	5.33	26.67	20.00	16.00	16.00	17.33	28.89	ต่ำ
24	591071052	6.00	7.00	5.00	6.00	30.00	6.00	5.00	5.00	5.33	26.67	6.00	4.00	4.00	4.67	23.33	18.00	16.00	14.00	16.00	26.67	ต่ำ
25	581031315	4.00	5.00	5.00	4.67	23.33	6.00	4.00	5.00	5.00	25.00	4.00	5.00	5.00	4.67	23.33	14.00	14.00	15.00	14.33	23.89	ต่ำ
26	591071106	5.00	4.00	4.00	4.33	21.67	5.00	6.00	3.00	4.67	23.33	6.00	6.00	4.00	5.33	26.67	16.00	16.00	11.00	14.33	23.89	ต่ำ

ตารางที่ 23 รายงานผลคะแนนการสอบการพูดโดยแบบทดสอบวัดความเข้าใจภาษาอังกฤษมาตรฐาน- การทดลองครั้งที่ 1 (ต่อ)

ที่	รหัสประจำตัว	Task A			รวม	%	Task B			รวม	%	Task C			รวม	%	Speaking			รวม	%	กลุ่ม
		1	2	3			1	2	3			1	2	3								
27	581031400	3.00	6.00	3.00	4.00	20.00	6.00	5.00	5.00	5.33	26.67	3.00	5.00	5.00	4.33	21.67	12.00	16.00	13.00	13.67	22.78	
28	591071009	4.00	6.00	4.00	4.67	23.33	5.00	4.00	4.00	4.33	21.67	4.00	4.00	5.00	4.33	21.67	13.00	14.00	13.00	13.33	22.22	
	Max	16.00	16.00	15.00	15.33	76.67	16.00	17.00	17.00	16.00	80.00	16.00	16.00	16.00	15.67	78.33	46.00	47.00	45.00	45.67	76.11	
	Min	3.00	4.00	3.00	4.00	20.00	5.00	4.00	3.00	4.33	21.67	3.00	4.00	4.00	4.33	21.67	12.00	14.00	11.00	13.33	22.22	
	\bar{X}	9.94	10.39	9.72	10.02	50.09	10.22	10.06	10.11	10.13	50.65	9.94	10.39	10.17	10.17	50.83	30.11	30.83	30.00	30.31	50.52	
	S.D.	4.35	3.88	4.21	4.07	20.35	3.61	4.17	4.40	4.01	20.03	4.15	4.37	4.27	4.22	21.12	11.92	12.21	12.74	12.22	20.37	

ตารางที่ 23 แสดงผลคะแนนการสอบการพูดโดยแบบทดสอบมาตรฐานหลังการทดลองครั้งที่ 1 และมีจำนวนผู้เข้าสอบที่ได้รับการคัดเลือกมาตามคะแนนที่ได้ในกลุ่มสูง กลาง และต่ำ กลุ่มละ 6 คน รวม 18 คน พบว่า ผู้เข้าสอบทุกคนได้เข้าทดสอบการพูดครบทั้ง 3 ภาระงาน โดยในภาพรวมของทุกภาระงาน ค่าร้อยละของคะแนนสูงสุดที่ผู้เข้าสอบทำได้ คือ ร้อยละ 76.11 คะแนนต่ำสุด คือ ร้อยละ 22.22 คะแนนเฉลี่ย คือ ร้อยละ 50.52 และส่วนเบี่ยงเบนมาตรฐานคือ ร้อยละ 20.37

เมื่อพิจารณารายภาระงาน พบว่า ภาระงานที่ผู้เข้าสอบส่วนใหญ่ทำได้ เรียงตามลำดับตามค่าร้อยละของคะแนนเฉลี่ย คือ 1) ภาระงาน C ร้อยละ 50.83 2) ภาระงาน B ร้อยละ 50.65 และ 3) ภาระงาน A ร้อยละ 50.09

1.4.1.1.2) การวิเคราะห์แบบทดสอบรายข้อ

หลังการทดลองครั้งที่ 1 ผู้วิจัยนำกระดาษคำตอบจำนวน 28 ชุด มาวิเคราะห์แบบทดสอบรายข้อ ซึ่งประกอบด้วยการวิเคราะห์ 3 ชนิด คือ 1) การวิเคราะห์ค่าความยากง่าย 2) การวิเคราะห์ค่าอำนาจจำแนก และ 3) การวิเคราะห์ดัชนีความลำเอียง โดยการวิเคราะห์แต่ละชนิดมีรายละเอียดและสูตรการคำนวณ ดังนี้

ก) การวิเคราะห์ค่าความยากง่าย

การหาค่าความยากง่ายแบ่งได้เป็น 2 ชนิด คือ ความยากง่ายของแบบทดสอบปรนัย และความยากง่ายของแบบทดสอบอัตนัย โดยมีรายละเอียดและสูตรการคำนวณ ดังนี้

การวิเคราะห์หาค่าความยากง่าย (Item facility / Item difficulty / Item easiness) ของแบบทดสอบปรนัย รายข้อ ตอน ทักษะ และแบบทดสอบทั้งฉบับ ใช้สูตรในการคำนวณ (Brown, 2005: 66) ดังนี้

$$P = \frac{N_{\text{correct}}}{N_{\text{total}}}$$

โดยที่

P = ค่าความยากง่ายของข้อสอบแต่ละข้อ

N_{correct} = จำนวนผู้สอบที่ตอบถูก

N_{total} = จำนวนผู้สอบที่เข้าสอบทั้งหมด

การวิเคราะห์ค่าความยากง่ายของแบบทดสอบอัตนัย (Item Difficulty) มีสูตรในการคำนวณ (Scannell and Tracy, 1975) ดังนี้

$$P = \frac{SU + SL + (NT)X_{\text{min}}}{NT (X_{\text{max}} + X_{\text{min}})}$$

เมื่อ S_U = คะแนนรวมของ 27 % ของจำนวนผู้เข้าสอบที่ได้คะแนนสูง

- S_L = คะแนนรวมของ 27 % ของจำนวนผู้เข้าสอบที่ได้คะแนนต่ำ
 N_T = จำนวนผู้เข้าสอบทั้งหมดที่ใช้ในการวิเคราะห์
 X_{max} = คะแนนสูงสุด
 X_{min} = คะแนนต่ำสุด

การตีความค่าความยากง่าย (Ebel, 1978) ใช้เกณฑ์ดังนี้

	0.40 – 0.60	ดีมาก (ยากง่ายปานกลาง)
	0.31 – 0.39, 0.61 – 0.69	ดี (ค่อนข้างยาก, ค่อนข้างง่าย)
	0.20 – 0.29, 0.70 – 0.80	พอใช้ (ค่อนข้างยาก, ค่อนข้างง่าย)
	≥ 0.81	ควรปรับปรุง (ง่ายเกินไป)
≤ 0.19		ควรปรับปรุง (ยากเกินไป)

ข) การวิเคราะห์ค่าอำนาจจำแนก

การหาค่าอำนาจจำแนก แบ่งได้เป็น 2 ชนิด คือ ค่าอำนาจจำแนกของแบบทดสอบปรนัย และค่าอำนาจจำแนกของแบบทดสอบอัตนัย โดยมีรายละเอียดและสูตรการคำนวณ ดังนี้ การวิเคราะห์หาค่าอำนาจจำแนก (Item Discrimination) ของแบบทดสอบปรนัยทั้งรายข้อ ตอน ทักษะ และทั้งฉบับ มีสูตรในการคำนวณ (Brown, 2005: 69) ดังนี้

$$R = IF_{upper} - IF_{lower}$$

โดยที่

R = ค่าอำนาจจำแนกของแบบทดสอบแต่ละข้อ

IF_{upper} = ค่าความยากสำหรับกลุ่มสูงในการทำแบบทดสอบทั้งหมด

IF_{lower} = ค่าความยากสำหรับกลุ่มต่ำในการทำแบบทดสอบทั้งหมด

การวิเคราะห์ค่าอำนาจจำแนกของแบบทดสอบอัตนัย มีสูตรที่ใช้ในการคำนวณ (Scannell and Tracy, 1975) ดังนี้

$$R = \frac{S_U + S_L}{N_U (X_{max} - X_{min})}$$

เมื่อ R = ค่าอำนาจจำแนก

S_U = คะแนนรวมของกลุ่มที่ได้คะแนนสูง

S_L = คะแนนรวมของกลุ่มที่ได้คะแนนต่ำ

N_U = จำนวนผู้เข้าสอบในกลุ่มที่ได้คะแนนสูง

X_{max} = คะแนนสูงสุด

X_{min} = คะแนนต่ำสุด

การตีความค่าอำนาจจำแนก (Ebel, 1978) ใช้เกณฑ์ดังนี้

0.60 ขึ้นไป	ดีมาก (จำแนกได้ดีมาก)
0.35 – 0.59	ดี (จำแนกได้ดี)
0.20 – 0.34	พอใช้ (จำแนกพอใช้ได้)
ต่ำกว่า 0.20	ควรปรับปรุง (จำแนกได้ไม่ดี)

ค) การวิเคราะห์ดัชนีความลำเอียง

การวิเคราะห์ดัชนีความลำเอียง (Angoff and Ford, 1973) มีสูตรในการคำนวณ คือ

$$d_g = \frac{(a\Delta g_1 - \Delta g_2 + b)}{\sqrt{a^2 + 1}}$$

$$\text{เมื่อ } a = \frac{(s_2^2 - s_1^2) + \sqrt{(s_2^2 - s_1^2)^2 + 4r_{12}^2 s_1^2 s_2^2}}{2r_{12} s_1 s_2}$$

$$b = \bar{x}_2 - a\bar{x}_1$$

d_g = ระยะห่างระหว่างข้อสอบข้อ g ไปยังเส้นแกนหลัก

Δg_1 = ค่าความยากมาตรฐานของข้อสอบข้อ g ในกลุ่ม 1

Δg_2 = ค่าความยากมาตรฐานของข้อสอบข้อ g ในกลุ่ม 2

ค่าความลำเอียงของข้อสอบ

S_1 = ส่วนเบี่ยงเบนมาตรฐานของค่าความยากมาตรฐานกลุ่ม 1

S_2 = ส่วนเบี่ยงเบนมาตรฐานของค่าความยากมาตรฐานกลุ่ม 2

\bar{x}_1 = ค่าเฉลี่ยของค่าความยากมาตรฐานในกลุ่ม 1

\bar{x}_2 = ค่าเฉลี่ยของค่าความยากมาตรฐานในกลุ่ม 2

r_{12} = สหสัมพันธ์ระหว่างค่าความยากมาตรฐานของกลุ่ม 1 และ 2

โดยใช้เกณฑ์ 0.75 หรือก็คือ $d_g < -0.75$ และ $d_g > +0.75$ จึงจะถือว่าข้อสอบนั้นมีความลำเอียง

เนื่องจากมีจำนวนผู้เข้าสอบค่อนข้างจำกัด ไม่สามารถแบ่งเป็น 2 กลุ่ม ตามเพศ เช่น ชาย-หญิง หรือตามสายวิชา เช่น สายวิทย์-สายศิลป์ ได้ในปริมาณที่พอจะคำนวณค่าได้ ผู้วิจัยจึงไม่ได้วิเคราะห์หาค่าดัชนีความลำเอียงในการทดลองครั้งที่ 1 แต่ได้วิเคราะห์หาค่าความลำเอียงหลังการทดลองใช้แบบทดสอบครั้งที่ 2 โดยการคัดเลือกผู้เข้าสอบตามสายวิชา คือ สายวิทยาศาสตร์ และสายศิลปะศาสตร์ โดยผู้เข้าสอบสายวิทย์ คือ นิสิตที่เรียนในคณะศึกษาศาสตร์ วิชาเอกคณิตศาสตร์ เคมี ชีววิทยา และฟิสิกส์ จำนวน 20 คน และผู้เข้าสอบสายศิลป์ คือ นิสิตที่เรียนในคณะมนุษยศาสตร์และ

สังคมศาสตร์ เศรษฐศาสตร์และบริหารธุรกิจ ในวิชาเอก การบัญชี ภาษาอังกฤษ ภาษาญี่ปุ่น ภาษาจีน เป็นต้น จำนวน 20 คน

1.4.1.1.3) การวิเคราะห์ความเชื่อมั่น

การวิเคราะห์ค่าความเชื่อมั่น แบ่งออกเป็น 2 ชนิด คือ 1) ความเชื่อมั่นของแบบทดสอบปรนัย (KR-20) และ 2) ความเชื่อมั่นของแบบทดสอบอัตนัย (Cronbach Alpha Coefficient) โดยมีสูตรในการคำนวณ ดังนี้

ก) ความเชื่อมั่นแบบทดสอบปรนัย ((KR-20)

ความเชื่อมั่นของแบบทดสอบปรนัย คำนวณโดยใช้วิธี Kuder-Richardson Formula 20 (Henning, 1987: 84) โดยมีสูตรในการคำนวณ ดังนี้

$$r_{tt} = \frac{n}{(n-1)} \frac{(S_t^2 - \sum S_i^2)}{S_t^2}$$

โดยที่ r_{tt} = ค่าสัมประสิทธิ์ความเชื่อมั่นที่คำนวณได้โดย วิธี KR 20
 N = จำนวนข้อของแบบทดสอบ
 S_t^2 = ความแปรปรวนของคะแนนสอบทั้งฉบับ
 S_i^2 = ความแปรปรวนของคะแนนข้อที่ i ซึ่งอาจจะแสดงได้ในรูปของ $\sum p_i q_i$
 เมื่อ p_i = สัดส่วนผู้ตอบถูกข้อที่ i
 q_i = สัดส่วนผู้ตอบผิดข้อที่ i

ข) ความเชื่อมั่นของแบบทดสอบอัตนัย (Cronbach Alpha Coefficient)

ความเชื่อมั่นของแบบทดสอบอัตนัย คำนวณโดยใช้วิธี Cronbach Alpha Coefficient ในรูปสัมประสิทธิ์สหสัมพันธ์แอลฟา (Alpha Coefficient) พัฒนามาจากสูตร KR-20 ซึ่งไม่จำเป็นต้องตรวจให้คะแนน แบบถูก 1 ผิด 0 ก็ได้ (Cronbach, 1951) โดยมีสูตรในการคำนวณ ดังนี้

$$\alpha = \frac{n}{n-1} \left\{ 1 - \frac{\sum S_i^2}{S_t^2} \right\}$$

โดยที่

α แทน ค่าความเชื่อมั่น
 n แทน จำนวนข้อของแบบทดสอบ
 S_i^2 แทน ความแปรปรวนรายข้อของแบบทดสอบ

S^2 แทน ความแปรปรวนทั้งฉบับ

ความเชื่อมั่นแบบทดสอบปรนัย ((KR-20) และ ความเชื่อมั่นของแบบทดสอบอัตนัย (Cronbach Alpha Coefficient) คำนวณได้จากโปรแกรมวิเคราะห์แบบทดสอบ B-Index Version 7.0 (สาคร แสงผิ่ง, 2550)

1.4.1.2) การตรวจสอบคุณภาพแบบทดสอบโดยรวม

การตรวจสอบคุณภาพแบบทดสอบโดยรวม แบ่งออกเป็น 2 ชนิด คือ 1) การตรวจสอบข้อมูลเชิงคุณภาพ และ 2) การตรวจสอบข้อมูลเชิงปริมาณ โดยมีรายละเอียดดังนี้

1.4.1.2.1) การตรวจสอบข้อมูลเชิงคุณภาพ เป็นการสอบถามความพึงพอใจของผู้ใช้แบบทดสอบเกี่ยวกับคุณภาพแบบทดสอบ 6 ด้าน ได้แก่ คำสั่งของแบบทดสอบ เนื้อหาสาระของแบบทดสอบ รูปแบบและการจัดลำดับความ การตรวจให้คะแนน ประโยชน์ของแบบทดสอบ และการจัดสอบ โดยให้ผู้ใช้แบบทดสอบตอบแบบสอบถามหลังจากทำแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test) และนำแบบสอบถามความพึงพอใจมาคำนวณหาคะแนนเฉลี่ย (\bar{X}) และส่วนเบี่ยงเบนมาตรฐาน (S.D.)

1.4.1.2.2) การตรวจสอบข้อมูลเชิงปริมาณ เป็นการหาความสัมพันธ์ระหว่างแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test) และแบบทดสอบมาตรฐาน หรือที่เรียกว่าเป็นการวิเคราะห์ความตรงร่วมสมัย (Concurrent Validity) โดยหาค่าความสัมพันธ์หรือความสอดคล้องของคะแนนสอบที่ได้จากแบบทดสอบทั้งสองฉบับของผู้เข้าสอบทุกคน รายชื่อ ตอน ทักษะ และแบบทดสอบทั้งฉบับ โดยสถิติในการคำนวณ คือ The Pearson Product-Moment Correlation Coefficient มีสูตรดังนี้

$$r_{xy} = \frac{\sum(X - M_X)(Y - M_Y)}{NS_X S_Y}$$

โดยที่

r_{xy} = ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (x) และแบบทดสอบวัดความสามารถภาษาอังกฤษมาตรฐาน (y)

X = คะแนนของผู้เข้าสอบแต่ละคนจากแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย

M_X = คะแนนเฉลี่ยจากแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย

S_X = ส่วนเบี่ยงเบนมาตรฐานของคะแนนสอบจากแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย

Y = คะแนนของนักเรียนแต่ละคนจากแบบทดสอบวัดความสามารถภาษาอังกฤษมาตรฐาน

M_Y = คะแนนเฉลี่ยจากแบบทดสอบวัดความสามารถภาษาอังกฤษมาตรฐาน

S_Y = ส่วนเบี่ยงเบนมาตรฐานของคะแนนสอบจากแบบทดสอบวัดความสามารถภาษาอังกฤษ

มาตรฐาน
 N = จำนวนผู้เข้าสอบทั้งหมด

1.4.1.3) สรุปผลคุณภาพแบบทดสอบและประสิทธิภาพของเครื่องมือที่เกี่ยวข้องหลังการทดลองครั้งที่ 1

ผู้วิจัยดำเนินการเก็บข้อมูลกับกลุ่มทดลองครั้งที่ 1 จำนวน 28 คน โดยมีการใช้เครื่องมือที่เกี่ยวข้อง คือ 1) แบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test) 2) แบบประเมินความพึงพอใจของผู้ใช้แบบทดสอบต่อแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัยที่ผู้วิจัยสร้างขึ้น และ 3) แบบทดสอบวัดความสามารถภาษาอังกฤษมาตรฐานสากล และพอที่จะสรุปผลการคัดเลือกหรือปรับปรุงแบบทดสอบเพื่อที่จะนำไปทดลองในครั้งที่ 2 ดังนี้

1.4.1.3.1) แบบทดสอบทักษะการฟัง มีข้อคำถามทั้งหมด 87 ข้อ จาก 3 ตอน คือ ตอน A มีข้อคำถาม 49 ข้อ ตอน B มีข้อคำถาม 49 ข้อ และตอน C มีข้อคำถาม 21 ข้อ โดยมีค่าความเชื่อมั่น $r = .71$ และมีการดำเนินการในแต่ละตอนหลังจากการวิเคราะห์แบบทดสอบรายข้อ ดังนี้

ตอน A มีข้อคำถามทั้งหมด 49 ข้อ มีค่าความเชื่อมั่น $r = .73$ และมีข้อคำถามที่ไม่ได้คุณภาพ ได้แก่ ยากเกินไป ง่ายเกินไป อำนาจจำแนกไม่ดี ไม่มีอำนาจจำแนก และอำนาจจำแนกติดลบ รวมทั้งหมด 14 ข้อ ผู้วิจัยได้ตัดข้อคำถามออก 14 ข้อ คงไว้ใช้ในการทดลองครั้งที่ 2 จำนวน 35 ข้อ

ตอน B มีข้อคำถามทั้งหมด 17 ข้อ มีค่าความเชื่อมั่น $r = .66$ และมีข้อคำถามที่ไม่ได้คุณภาพ ได้แก่ ยากเกินไป ไม่มีอำนาจจำแนก หรืออำนาจจำแนกเป็นลบ จำนวน 3 ข้อ ผู้วิจัยได้ตัดข้อคำถามออก 3 ข้อ คงไว้ใช้ในการทดลองครั้งที่ 2 จำนวน 14 ข้อ

ตอน C มีข้อคำถามทั้งหมด 21 ข้อ มีค่าความเชื่อมั่น $r = .75$ และมีข้อคำถามที่ไม่ได้คุณภาพ ได้แก่ ยากเกินไป อำนาจจำแนกไม่ดี ไม่มีอำนาจจำแนก และอำนาจจำแนกติดลบ จำนวน 5 ข้อ ผู้วิจัยได้ตัดข้อคำถามออก 5 ข้อ คงไว้ใช้ในการทดลองครั้งที่ 2 จำนวน 16 ข้อ

1.4.1.3.2) แบบทดสอบทักษะการใช้ภาษา มีข้อคำถามทั้งหมด 60 ข้อ จาก 2 ตอน คือ ตอน A มีข้อคำถาม 22 ข้อ และตอน B มีข้อคำถาม 38 ข้อ โดยมีค่าความเชื่อมั่น $r = .69$ และมีการดำเนินการในแต่ละตอนหลังจากการวิเคราะห์แบบทดสอบรายข้อ ดังนี้

ตอน A มีข้อคำถามทั้งหมด 22 ข้อ มีค่าความเชื่อมั่น $r = .68$ และมีข้อคำถามที่ไม่ได้คุณภาพ ได้แก่ ยากเกินไป และอำนาจจำแนกไม่ดี จำนวน 2 ข้อ ผู้วิจัยได้ตัดข้อคำถามออก 2 ข้อ คงไว้ใช้ในการทดลองครั้งที่ 2 จำนวน 20 ข้อ

ตอน B มีข้อคำถามทั้งหมด 38 ข้อ มีค่าความเชื่อมั่น $r = .70$ และมีข้อคำถามที่ไม่ได้คุณภาพ ได้แก่ ยากเกินไป ง่ายเกินไป อำนาจจำแนกไม่ดี ไม่มีอำนาจจำแนก และอำนาจจำแนกติดลบ จำนวน 8 ข้อ ผู้วิจัยได้ตัดข้อคำถามออก 8 ข้อ คงไว้ใช้ในการทดลองครั้งที่ 2 จำนวน 30 ข้อ

1.4.1.3.3) แบบทดสอบทักษะการอ่าน มีข้อคำถามทั้งหมด 97 ข้อ จาก 2 ตอน คือ ตอน A มีข้อคำถาม 49 ข้อ และตอน B มีข้อคำถาม 48 ข้อ โดยมีค่าความเชื่อมั่น $r = .76$ และมีการดำเนินการในแต่ละตอนหลังจากการวิเคราะห์แบบทดสอบรายข้อ ดังนี้

ตอน A มีข้อคำถามทั้งหมด 49 ข้อ มีค่าความเชื่อมั่น $r = .72$ และมีข้อคำถามที่ไม่ได้คุณภาพ ได้แก่ ยากเกินไป อำนาจจำแนกไม่ดี ไม่มีอำนาจจำแนก และอำนาจจำแนกติดลบ จำนวน 9 ข้อ ผู้วิจัยได้ตัดข้อคำถามออก 9 ข้อ คงไว้ใช้ในการลองครั้งที่ 2 จำนวน 40 ข้อ

ตอน B มีข้อคำถามทั้งหมด 48 ข้อ มีค่าความเชื่อมั่น $r = .81$ และมีข้อคำถามที่ไม่ได้คุณภาพ ได้แก่ ยากเกินไป อำนาจจำแนกไม่ดี ไม่มีอำนาจจำแนก และอำนาจจำแนกติดลบ จำนวน 6 ข้อ ผู้วิจัยได้ตัดข้อคำถามออก 6 ข้อ คงไว้ใช้ในการลองครั้งที่ 2 จำนวน 42 ข้อ

1.4.1.3.4) แบบทดสอบทักษะการเขียน โดยรวมมีค่าความเชื่อมั่น (Cronbach Alpha Coefficient) เท่ากับ .44 ประกอบด้วย 2 ภาระงาน คือ ภาระงาน A ซึ่งมีค่าความเชื่อมั่น (Cronbach Alpha Coefficient) เท่ากับ .45 และภาระงาน B ซึ่งมีค่าความเชื่อมั่น (Cronbach Alpha Coefficient) เท่ากับ .43 และสามารถนำไปใช้ทดลองต่อในครั้งที่ 2 ได้ทั้ง 2 ภาระงาน

1.4.1.3.5) แบบทดสอบทักษะการพูด โดยรวมมีค่าความเชื่อมั่น (Cronbach Alpha Coefficient) เท่ากับ .99 ประกอบด้วย 3 ภาระงาน คือ ภาระงาน A, B และ C ซึ่งมีค่าความเชื่อมั่น (Cronbach Alpha Coefficient) เท่ากับ .99 ทั้ง 3 ภาระงาน และสามารถนำไปใช้ทดลองต่อในครั้งที่ 2 ได้ทั้ง 3 ภาระงาน

1.4.1.3.6) แบบสอบถามความพึงพอใจผู้ใช้แบบทดสอบต่อแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test) ได้รับความพึงพอใจจากผู้ใช้แบบทดสอบอยู่ในระดับพึงพอใจมาก ($\bar{X} = 4.08$) และมีค่าอัลฟา เท่ากับ .95

1.4.1.3.7) การหาความตรงร่วมสมัย (Concurrent Validity) ของแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test) และแบบทดสอบมาตรฐาน (Standard English Proficiency Test) โดยภาพรวมแบบทดสอบมีค่าความตรงร่วมสมัย (Concurrent Validity) $r_{xy} = .87$ และในรายทักษะย่อย มีค่าความตรงร่วมสมัย (Concurrent Validity) เรียงจากสูงไปหาค่า คือ 1) ทักษะการพูด $r_{xy} = .99$ 2) ทักษะการเขียน $r_{xy} = .96$ 3) การใช้ภาษา $r_{xy} = .80$ 4) ทักษะการฟัง $r_{xy} = .75$ และ 5) ทักษะการอ่าน $r_{xy} = .70$

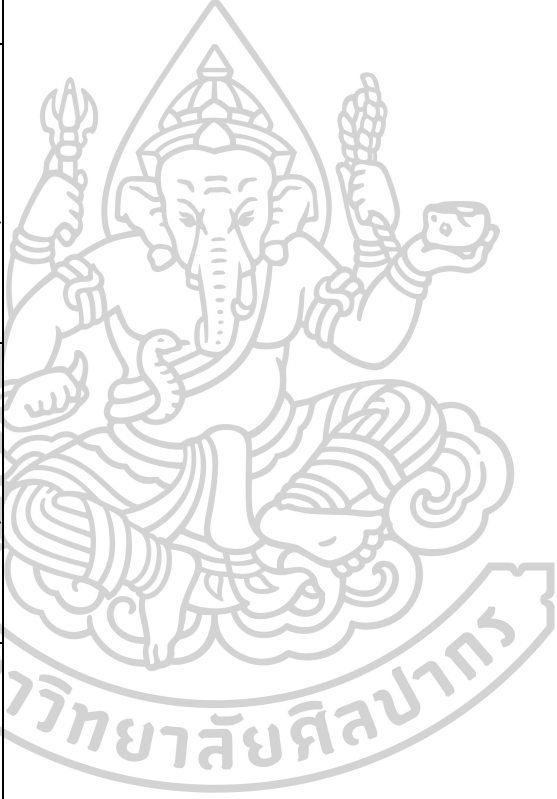
โดยมีสรุปผลคุณภาพแบบทดสอบและประสิทธิภาพของเครื่องมือที่เกี่ยวข้องหลังจากการทดลองใช้ครั้งที่ 1 แสดงไว้ในตารางที่ 24

ตารางที่ 24 สรุปคุณภาพแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัยหลังจากการทดลองใช้ครั้งที่ 1

แบบทดสอบ	ครั้งที่ 1	Max	Min	\bar{X}	S.D.	P	R	Reliability	ตัดออก	เก็บไว้
การฟัง	87	50	13	28.89	8.757	.36	.41	.71	22	65
A	49	31	6	14.03	5.577	.30	.39	.73	14	35
B	17	14	0	6.35	2.930	.38	.51	.66	3	14
C	21	11	5	8.50	1.592	.41	.35	.75	5	16
การใช้ภาษา	60	49	10	25.36	8.098	.42	.49	.69	10	50
A	22	19	3	9.17	3.625	.43	.53	.68	2	20
B	38	30	7	16.17	5.085	.42	.46	.70	8	30
การอ่าน	97	66	19	47.32	11.851	.50	.44	.76	15	82
A	49	37	16	26.85	5.73	.55	.45	.72	9	40
B	48	34	1	20.46	7.198	.45	.43	.81	6	42
การเขียน	2	41	1.33	15.18	12.136	.38	.71	.44	-	2
A	1	15	1.33	5.44	4.420	.26	.38	.45	-	1
B	1	30	2.67	9.75	9.770	.31	.59	.43	-	1
การพูด	3	46.67	13.67	31.61	11.319	.47	.88	.99	-	3
A	1	15	4.67	10.50	3.880	.50	.50	.99	-	1
B	1	16	4.67	10.56	3.730	.51	.48	.99	-	1
C	1	15.67	4.33	10.56	3.770	.51	.49	.99	-	1
แบบสอบถาม ความคิดเห็น ผู้ใช้แบบทดสอบ	36	4.53	3.67	4.08	0.867	-	-	Alpha = .95	-	-

ตารางที่ 25 สรุปความตรงร่วมสมัย (Concurrent Validity) ของแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (X) และแบบทดสอบมาตรฐาน (y) หลังจากการทดลองใช้ครั้งที่ 1

ทักษะการฟัง		ทักษะการใช้ภาษา		ทักษะการอ่าน		ทักษะการเขียน		ทักษะการพูด		ทั้งหมด	
UEPT	TOEFL	UEPT	TOEFL	UEPT	TOEFL	UEPT	IELTS	UEPT	IELTS	UEPT	IELTS
$r_{xy} = .75$		$r_{xy} = .80$		$r_{xy} = .70$		$r_{xy} = .96$		$r_{xy} = .99$		$r_{xy} = .87$	



1.4.2) การวิเคราะห์ประสิทธิภาพและตรวจสอบคุณภาพแบบทดสอบโดยรวมหลังการทดลองครั้งที่ 2

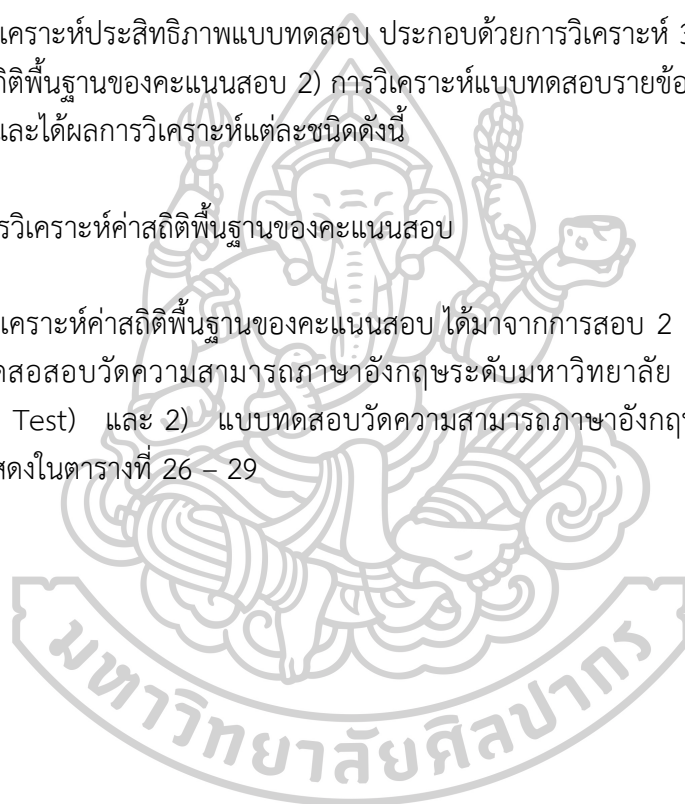
ผู้วิจัยได้เก็บข้อมูลครั้งที่ 2 ในวันพุธ ที่ 30 มกราคม 2562 เวลา 13.00 – 16.30 น. ณ ห้อง 15403 อาคารเรียน 15 มหาวิทยาลัยทักษิณ วิทยาเขตสงขลา โดยมีผู้เข้าสอบ จำนวน 94 คน และได้นำกระดาษคำตอบจำนวน 94 ชุด มาวิเคราะห์และตรวจสอบคุณภาพ โดยใช้วิธีการและสูตรต่างๆ เหมือนกับการทดลองครั้งที่ 1 โดยมีรายละเอียดต่อไปนี้

1.4.2.1) การวิเคราะห์ประสิทธิภาพแบบทดสอบ

การวิเคราะห์ประสิทธิภาพแบบทดสอบ ประกอบด้วยการวิเคราะห์ 3 ชนิด คือ 1) การวิเคราะห์ค่าสถิติพื้นฐานของคะแนนสอบ 2) การวิเคราะห์แบบทดสอบรายข้อ และ 3) การวิเคราะห์ความเชื่อมั่น และได้ผลการวิเคราะห์แต่ละชนิดดังนี้

1.4.2.1.1) การวิเคราะห์ค่าสถิติพื้นฐานของคะแนนสอบ

การวิเคราะห์ค่าสถิติพื้นฐานของคะแนนสอบ ได้มาจากการสอบ 2 ครั้ง คือ การสอบโดยใช้ 1) แบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test) และ 2) แบบทดสอบวัดความสามารถภาษาอังกฤษมาตรฐาน และผลการวิเคราะห์ดังแสดงในตารางที่ 26 – 29



ตารางที่ 26 รายงานผลคะแนนการสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย - การทดลองครั้งที่ 2

ที่	รหัสประจำตัว	ฟัง		รวม	%	การใช้ภาษา		รวม	%	อ่าน		รวม	%	เขียน		รวม	%	รวมทุกทักษะ	%
		A	BC			A	B			A	B			A	B				
1	591031245	26	20	46	70.77	19	24	43	86.00	33	36	69	84.14	16.83	33.67	50.5	84.1667	208.5	81.13
2	591031251	24	20	44	67.69	18	22	40	80.00	32	33	65	79.26	15	32	47	78.3333	196	76.26
3	591031248	24	23	47	72.31	19	21	40	80.00	28	26	54	65.85	16.67	33.33	50	83.3333	191	74.32
4	591031253	25	17	42	64.62	17	17	34	68.00	30	29	59	71.95	14.17	32.33	46.5	77.5	181.5	70.62
5	591031235	28	24	52	80.00	15	19	34	68.00	33	23	56	68.29	9.5	18.33	27.83	46.3833	169.83	66.08
6	591031250	21	18	39	60.00	17	21	38	76.00	31	12	43	52.43	15.83	33.33	49.16	81.9333	169.16	65.82
7	591031255	16	15	31	47.69	11	16	27	54.00	31	29	60	73.17	16.5	32.33	48.83	81.3833	166.83	64.91
8	591031263	28	17	45	69.23	11	19	30	60.00	29	28	57	69.51	13.33	19	32.33	53.8833	164.33	63.94
9	591031236	19	17	36	55.38	12	19	31	62.00	28	28	56	68.29	12.67	27.67	40.34	67.2333	163.34	63.56
10	591031246	22	17	39	60.00	14	17	31	62.00	26	29	55	67.07	9.67	27.33	37	61.6667	162	63.04
11	591011517	14	22	36	55.38	11	22	33	66.00	28	31	59	71.95	13.5	18.33	31.83	53.05	159.83	62.19
12	601011577	25	19	44	67.69	13	24	37	74.00	26	31	57	69.51	1	18.33	19.33	32.2167	157.33	61.22
13	591031243	14	18	32	49.23	17	20	37	74.00	27	20	47	57.31	9.83	21.33	31.16	51.9333	147.16	57.26
14	591031254	12	17	29	44.62	10	22	32	64.00	27	28	55	67.07	13	17.67	30.67	51.1167	146.67	57.07
15	611011632	24	24	48	73.85	16	23	39	78.00	24	33	57	69.51			0	0	144	56.03
16	591031262	14	17	31	47.69	12	15	27	54.00	29	24	53	64.63	15.5	15.87	31.37	52.2833	142.37	55.40

ตารางที่ 26 รายงานผลคะแนนการสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย - การทดลองครั้งที่ 2 (ต่อ)

ที่	รหัสประจำตัว	ฟัง		รวม	%	การใช้ภาษา		รวม	%	อ่าน		รวม	%	เขียน		รวม	%	รวมทุกทักษะ	%
		A	BC			A	B			A	B			A	B				
17	591031234	35	30	65	100	20	30	50	100	40	42	82	100	20	40	60	100	257	100
18	591011521	11	18	29	44.62	15	22	37	74.00	26	34	60	73.17			0	0	126	49.03
19	601011602	18	13	31	47.69	12	20	32	64.00	21	18	39	47.56	6.33	17	23.33	38.8833	125.33	48.77
20	601011627	7	10	17	26.15	12	18	30	60.00	26	22	48	58.53	3.33	19.67	23	38.3333	118	45.91
21	591031249	18	17	35	53.85	14	17	31	62.00	29	21	50	60.97			0	0	116	45.14
22	611011649	16	24	40	61.54	10	14	24	48.00	13	10	23	28.04	8.33	20	28.33	47.2167	115.33	44.88
23	591011534	17	12	29	44.62	13	20	33	66.00	23	22	45	54.87	5.33	0	5.33	8.88333	112.33	43.71
24	601011624	17	18	35	53.85	10	15	25	50.00	21	31	52	63.41			0	0	112	43.58
25	601011576	15	13	28	43.08	10	24	34	68.00	21	28	49	59.75			0	0	111	43.19
26	601011611	14	14	28	43.08	13	17	30	60.00	23	27	50	60.97			0	0	108	42.02
27	601011607	12	16	28	43.08	12	16	28	56.00	30	19	49	59.75			0	0	105	40.86
28	611081030	9	12	21	32.31	11	14	25	50.00	18	20	38	46.34	6.67	14	20.67	34.45	104.67	40.73
29	601011588	18	14	32	49.23	12	18	30	60.00	18	24	42	51.21			0	0	104	40.47
30	591011524	13	14	27	41.54	11	17	28	56.00	25	23	48	58.53			0	0	103	40.08
31	601011597	20	17	37	56.92	11	12	23	46.00	25	11	36	43.90			0	0	96	37.35
32	601011616	13	13	26	40.00	11	12	23	46.00	20	27	47	57.31			0	0	96	37.35

ตารางที่ 26 รายงานผลคะแนนการสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย - การทดลองครั้งที่ 2 (ต่อ)

ที่	รหัสประจำตัว	ฟัง		รวม	%	การใช้ภาษา		รวม	%	อ่าน		รวม	%	เขียน		รวม	%	รวมทุกทักษะ	%
		A	BC			A	B			A	B			A	B				
33	601011610	35	30	65	100	20	30	50	100	40	42	82	100	20	40	60	100	257	100
34	611011589	14	12	26	40.00	13	19	32	64.00	21	15	36	43.90			0	0	94	36.58
35	601011567	12	11	23	35.38	9	12	21	42.00	16	10	26	31.70	9.33	14	23.33	38.8833	93.33	36.32
36	581031341	17	21	38	58.46	9	14	23	46.00	12	18	30	36.58			0	0	91	35.41
37	591011544	13	18	31	47.69	12	11	23	46.00	17	19	36	43.90			0	0	90	35.02
38	591011363	9	12	21	32.31	16	15	31	62.00	17	19	36	43.90			0	0	88	34.24
39	611011638	21	18	39	60.00	9	7	16	32.00	15	17	32	39.02			0	0	87	33.85
40	581031401	15	11	26	40.00	12	10	22	44.00	13	26	39	47.56			0	0	87	33.85
41	601011551	9	13	22	33.85	9	7	16	32.00	7	12	19	23.17	10	18	28	46.6667	85	33.07
42	601011581	10	10	20	30.77	9	12	21	42.00	21	15	36	43.90		8	8	13.3333	85	33.07
43	601011580	11	9	20	30.77	13	18	31	62.00	18	16	34	41.46			0	0	85	33.07
44	581031321	5	13	18	27.69	8	11	19	38.00	15	16	31	37.80	2	14.67	16.67	27.7833	84.67	32.95
45	601011631	13	16	29	44.62	11	17	28	56.00	11	16	27	32.92			0	0	84	32.68
46	601011622	9	8	17	26.15	9	18	27	54.00	18	22	40	48.78			0	0	84	32.68
47	601011585	9	15	24	36.92	8	14	22	44.00	15	21	36	43.90			0	0	82	31.91
48	611011569	9	20	29	44.62	9	11	20	40.00	15	16	31	37.80			0	0	80	31.13
48	611011569	16	8	24	36.92	11	12	23	46.00	17	15	32	39.02			0	0	79	30.74

ตารางที่ 26 รายงานผลคะแนนการสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย - การทดลองครั้งที่ 2 (ต่อ)

ที่	รหัสประจำตัว	ฟัง		รวม	%	การใช้ภาษา		รวม	%	อ่าน		รวม	%	เขียน		รวม	%	รวมทุกทักษะ	%
		A	BC			A	B			A	B			A	B				
49	591071014	35	30	65	100	20	30	50	100	40	42	82	100	20	40	60	100	257	100
50	591071057	10	13	23	35.38	9	11	20	40.00	18	16	34	41.46			0	0	77	29.96
51	601011605	5	13	18	27.69	9	9	18	36.00	28	13	41	50.00			0	0	77	29.96
52	601011565	10	12	22	33.85	13	19	32	64.00	9	14	23	28.04			0	0	77	29.96
53	611011588	11	12	23	35.38	8	13	21	42.00	13	19	32	39.02			0	0	76	29.57
54	611011572	12	13	25	38.46	8	14	22	44.00	13	15	28	34.14			0	0	75	29.18
55	601011596	10	15	25	38.46	6	8	14	28.00	13	12	25	30.48	2.67	8	10.67	17.7833	74.67	29.05
56	581031330	16	8	24	36.92	14	16	30	60.00	12	8	20	24.39			0	0	74	28.79
57	581031320	10	10	20	30.77	8	13	21	42.00	8	11	19	23.17	1	12	13	21.6667	73	28.40
58	601011594	6	9	15	23.08	10	14	24	48.00	11	7	18	21.95	1	14.67	15.67	26.1167	72.67	28.28
59	601011579	11	16	27	41.54	11	10	21	42.00	18	6	24	29.26			0	0	72	28.02
60	601011544	15	13	28	43.08	8	11	19	38.00	12	11	23	28.04			0	0	70	27.24
61	611031545	12	13	25	38.46	7	13	20	40.00	15	9	24	29.26			0	0	69	26.85
62	601011564	11	13	24	36.92	8	12	20	40.00	14	11	25	30.48			0	0	69	26.85
63	591011390	5	14	19	29.23	9	14	23	46.00	16	9	25	30.48			0	0	67	26.07
64	611011554	5	14	19	29.23	6	14	20	40.00	13	14	27	32.92			0	0	66	25.68
64	611011554	7	12	19	29.23	11	9	20	40.00	10	16	26	31.70			0	0	65	25.29

ตารางที่ 26 รายงานผลคะแนนการสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย - การทดลองครั้งที่ 2 (ต่อ)

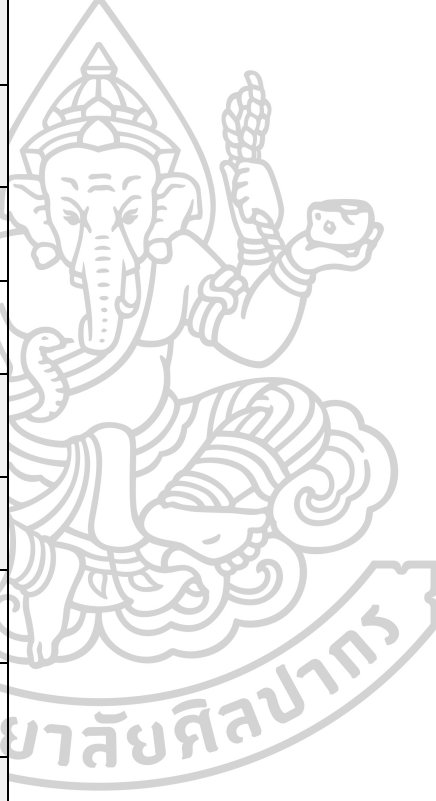
ที่	รหัสประจำตัว	ฟัง		รวม	%	การใช้ภาษา		รวม	%	อ่าน		รวม	%	เขียน		รวม	%	รวมทุกทักษะ	%
		A	BC			A	B			A	B			A	B				
		35	30	65	100	20	30	50	100	40	42	82	100	20	40	60	100	257	100
65	591011364	10	12	22	33.85	5	8	13	26.00	12	16	28	34.14			0	0	63	24.51
66	591011386	9	13	22	33.85	8	14	22	44.00	10	8	18	21.95			0	0	62	24.12
67	611011583	5	12	17	26.15	8	10	18	36.00	17	10	27	32.92			0	0	62	24.12
68	591011373	11	11	22	33.85	6	12	18	36.00	13	8	21	25.60			0	0	61	23.74
69	591071077	5	13	18	27.69	5	10	15	30.00	14	14	28	34.14			0	0	61	23.74
70	581031318	6	15	21	32.31	10	7	17	34.00	10	11	21	25.60	1	0	1	1.66667	60	23.35
71	581031327	8	6	14	21.54	11	15	26	52.00	10	10	20	24.39			0	0	60	23.35
72	581031332	10	10	20	30.77	9	10	19	38.00	11	10	21	25.60			0	0	60	23.35
73	591011378	11	13	24	36.92	7	12	19	38.00	9	5	14	17.07	3	0	3	5	60	23.35
74	591011360	7	13	20	30.77	7	8	15	30.00	10	14	24	29.26			0	0	59	22.96
75	591011371	11	9	20	30.77	6	13	19	38.00	11	8	19	23.17			0	0	58	22.57
76	581031334	10	7	17	26.15	10	14	24	48.00	6	9	15	18.29			0	0	56	21.79
77	591011357	8	9	17	26.15	11	10	21	42.00	10	8	18	21.95			0	0	56	21.79
78	591071028	6	12	18	27.69	8	11	19	38.00	7	12	19	23.17			0	0	56	21.79
79	581031328	8	13	21	32.31	6	7	13	26.00	10	10	20	24.39			0	0	54	21.01
80	581031343	4	15	19	29.23	6	9	15	30.00	10	10	20	24.39			0	0	54	21.01

ตารางที่ 26 รายงานผลคะแนนการสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย - การทดลองครั้งที่ 2 (ต่อ)

ที่	รหัสประจำตัว	ฟัง		รวม	%	การใช้ภาษา		รวม	%	อ่าน		รวม	%	เขียน		รวม	%	รวมทุกทักษะ	%
		A	BC			A	B			A	B			A	B				
81	591011355	35	30	65	100	20	30	50	100	40	42	82	100	20	40	60	100	257	100
82	581031331	9	11	20	30.77	6	7	13	26.00	12	9	21	25.6098			0	0	54	21.01
83	581031336	6	12	18	27.69	7	10	17	34.00	5	13	18	21.95			0	0	53	20.62
84	591011352	8	9	17	26.15	7	7	14	28.00	9	12	21	25.60			0	0	52	20.23
85	591011367	6	11	17	26.15	4	12	16	32.00	11	8	19	23.17			0	0	52	20.23
86	591011374	9	5	14	21.54	5	10	15	30.00	13	10	23	28.04			0	0	52	20.23
87	591071055	15	7	22	33.85	4	9	13	26.00	11	6	17	20.73			0	0	52	20.23
88	591011347	7	10	17	26.15	6	12	18	36.00	8	9	17	20.73			0	0	52	20.23
89	591071027	10	9	19	29.23	7	7	14	28.00	9	9	18	21.95			0	0	51	19.84
90	581031338	8	6	14	21.54	4	5	9	18.00	6	13	19	23.17		8	8	13.3333	50	19.46
91	591071086	6	6	12	18.46	7	10	17	34.00	11	9	20	24.39			0	0	49	19.07
92	591071113	5	10	15	23.08	3	4	7	14.00	13	14	27	32.92			0	0	49	19.07
93	591011362	7	7	14	21.54	4	10	14	28.00	9	12	21	25.60			0	0	49	19.07
94	581031344	12	5	17	26.15	7	5	12	24.00	13	6	19	23.17			0	0	48	18.68
		7	8	15	23.08	7	11	18	36.00	8	5	13	15.85			0	0	46	17.90

ตารางที่ 26 รายงานผลคะแนนการสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย – การทดลองครั้งที่ 2 (ต่อ)

ที่	รหัสประจำตัว	ฟัง		รวม	%	การใช้ภาษา		รวม	%	อ่าน		รวม	%	เขียน		รวม	%	รวมทุกทักษะ	%
		A	BC			A	B			A	B			A	B				
		35	30	65	100	20	30	50	100	40	42	82	100	20	40	60	100	257	100
	Max	28.00	24.00	52.00	80.00	19.00	24.00	43.00	86.00	33.00	36.00	69.00	84.15	16.83	33.67	50.50	84.17	208.50	81.13
	Min	4.00	5.00	12.00	18.46	3.00	4.00	7.00	14.00	5.00	5.00	13.00	15.85	1.00	0.00	0.00	0.00	46.00	17.90
	\bar{X}	12.33	13.37	25.70	39.54	9.91	13.76	23.67	47.34	17.06	16.60	33.66	41.05	9.04	18.29	8.82	14.70	91.85	35.74
	S.D.	5.800	4.365	9.129	14.044	3.546	4.865	7.861	15.722	7.632	7.941	14.550	17.743	5.473	9.810	15.152	25.254	40.373	15.709

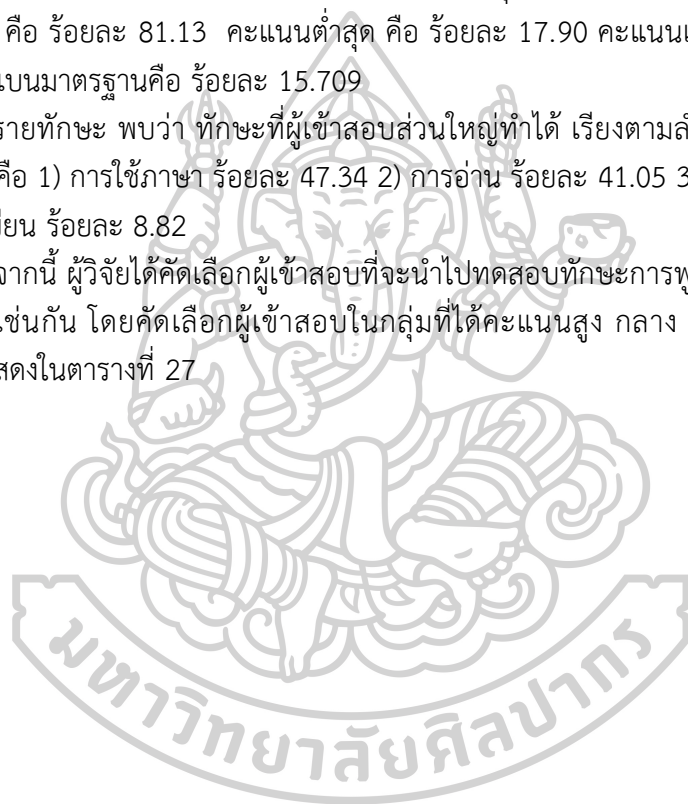


ตารางที่ 26 แสดงผลคะแนนการสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัยหลังการทดลองครั้งที่ 2 ซึ่งประกอบด้วย 4 ทักษะ คือ การฟัง การใช้ภาษา การอ่านและการเขียน และแสดงคะแนนโดยจัดเรียงจากคะแนนสูงสุดไปหาต่ำสุด พบว่า จากจำนวนผู้เข้าสอบ 94 คน มีผู้เข้าสอบที่ทำแบบทดสอบครบทั้ง 4 ทักษะ จำนวน 31 คน และมีผู้เข้าสอบที่ไม่ทำแบบทดสอบการเขียน 63 คน ซึ่งส่วนใหญ่เป็นผู้ที่อยู่ในกลุ่มที่ได้คะแนนปานกลางและต่ำ นอกจากนี้ ในจำนวนผู้ที่ทำแบบทดสอบการเขียน 31 คน สามารถแบ่งได้เป็น 2 กลุ่ม คือ 1) ผู้ที่ทำครบทั้ง 2 ภาระงาน จำนวน 9 คน และ 2) ผู้ที่ทำเพียงภาระงานเดียว จำนวน 2 คน ซึ่งเป็นผู้ที่อยู่ในกลุ่มที่ได้คะแนนปานกลางและต่ำและเลือกทำเฉพาะภาระงานที่ 2 ซึ่งเป็นการเขียนเรียงความ

เมื่อพิจารณาคะแนนสอบ พบว่า จากคะแนนรวมทุกทักษะ ค่าร้อยละของคะแนนสูงสุดที่ผู้เข้าสอบทำได้ คือ ร้อยละ 81.13 คะแนนต่ำสุด คือ ร้อยละ 17.90 คะแนนเฉลี่ย คือ ร้อยละ 35.74 และส่วนเบี่ยงเบนมาตรฐานคือ ร้อยละ 15.709

เมื่อพิจารณารายทักษะ พบว่า ทักษะที่ผู้เข้าสอบส่วนใหญ่ทำได้ เรียงตามลำดับตามค่าร้อยละของคะแนนเฉลี่ย คือ 1) การใช้ภาษา ร้อยละ 47.34 2) การอ่าน ร้อยละ 41.05 3) การฟัง ร้อยละ 39.54 และ 4) การเขียน ร้อยละ 8.82

นอกจากนี้ ผู้วิจัยได้คัดเลือกผู้เข้าสอบที่จะนำไปทดสอบทักษะการพูดจากผลคะแนนรวมในตารางที่ 25 เช่นกัน โดยคัดเลือกผู้เข้าสอบในกลุ่มที่ได้คะแนนสูง กลาง ต่ำ กลุ่มละ 13 คน ดังรายละเอียดแสดงในตารางที่ 27



ตารางที่ 26 รายงานผลคะแนนการสอบพูดโดยแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย – การทดลองครั้งที่ 2

ที่	รหัส ประจำตัว	Task A			รวม	%	Task B			รวม	%	Task C			รวม	%	Speaking			รวม	%	กลุ่ม	
		1	2	3			1	2	3			1	2	3			1	2	3				
1	591031253	15.00	16.00	14.00	15.00	75.00	20.00	16.00	15.00	15.33	76.67	20.00	14.00	15.00	14.67	73.33	60.00	46.00	44.00	45.00	60.00	100.00	75.00
2	591031246	15.00	16.00	14.00	15.00	75.00	20.00	16.00	15.00	15.33	76.67	20.00	14.00	15.00	14.67	73.33	60.00	46.00	44.00	45.00	60.00	100.00	75.00
3	591031250	15.00	16.00	14.00	15.00	75.00	20.00	16.00	15.00	15.33	76.67	20.00	13.00	15.00	14.33	71.67	60.00	46.00	44.00	44.67	60.00	100.00	74.44
4	591031245	14.00	14.00	14.00	14.00	70.00	20.00	14.00	15.00	14.33	71.67	20.00	15.00	14.00	14.67	73.33	60.00	43.00	43.00	43.00	60.00	100.00	71.67
5	601011577	15.00	15.00	15.00	15.00	75.00	20.00	15.00	12.00	14.00	70.00	20.00	15.00	15.00	14.00	70.00	60.00	42.00	42.00	43.00	60.00	100.00	71.67
6	591031251	15.00	15.00	15.00	15.00	75.00	20.00	15.00	12.00	14.00	70.00	20.00	13.00	15.00	13.33	66.67	60.00	42.00	42.00	42.33	60.00	100.00	70.56
7	591031235	13.00	14.00	14.00	13.67	68.33	20.00	14.00	13.00	14.00	70.00	20.00	15.00	13.00	14.33	71.67	60.00	42.00	42.00	42.00	60.00	100.00	70.00
8	591031255	13.00	14.00	14.00	13.67	68.33	20.00	14.00	13.00	14.00	70.00	20.00	15.00	13.00	14.33	71.67	60.00	42.00	42.00	42.00	60.00	100.00	70.00
9	591031263	14.00	15.00	15.00	14.67	73.33	20.00	15.00	14.00	13.67	68.33	20.00	15.00	12.00	13.67	68.33	60.00	41.00	41.00	42.00	60.00	100.00	70.00
10	591011517	13.00	14.00	14.00	13.67	68.33	20.00	14.00	13.00	14.00	70.00	20.00	14.00	15.00	14.00	70.00	60.00	42.00	42.00	41.67	60.00	100.00	69.44
11	591031236	14.00	14.00	13.00	13.67	68.33	20.00	14.00	14.00	13.67	68.33	20.00	13.00	13.00	13.33	66.67	60.00	41.00	40.00	40.67	60.00	100.00	67.78
12	591031248	13.00	14.00	13.00	13.33	66.67	20.00	14.00	13.00	13.33	66.67	20.00	14.00	13.00	13.33	66.67	60.00	40.00	39.00	40.00	60.00	100.00	66.67
13	591031243	13.00	14.00	13.00	13.33	66.67	20.00	14.00	13.00	13.33	66.67	20.00	13.00	13.00	13.00	65.00	60.00	40.00	39.00	39.67	60.00	100.00	66.11
14	601011580	11.00	11.00	10.00	10.67	53.33	20.00	11.00	11.00	11.00	55.00	20.00	11.00	11.00	11.00	55.00	60.00	33.00	32.00	32.67	60.00	100.00	54.44
15	601011585	11.00	11.00	10.00	10.67	53.33	20.00	11.00	11.00	11.00	55.00	20.00	11.00	11.00	11.00	55.00	60.00	33.00	32.00	32.67	60.00	100.00	54.44
16	601011565	11.00	11.00	10.00	10.67	53.33	20.00	11.00	11.00	11.00	55.00	20.00	10.00	11.00	10.67	53.33	60.00	33.00	32.00	32.33	60.00	100.00	53.89

ตารางที่ 27 รายงานผลคะแนนการสอบพหุรูปแบบทดสอบวัดความเข้าใจภาษาอังฤษระดับมหาวิทยาลัย – การทดลองครั้งที่ 2 (ต่อ)

ที่	รหัสประจำตัว	Task A			รวม	%	Task B			รวม	%	Task C			รวม	%	Speaking			รวม	%	กลุ่ม	
		1	2	3			1	2	3			1	2	3			1	2	3				
17	601011631	10.00	12.00	11.00	11.00	55.00	20.00	10.00	10.00	10.67	53.33	20.00	10.00	10.00	10.33	51.67	60.00	32.00	31.00	32.00	60.00	100.00	53.33
18	601011581	10.00	11.00	10.00	10.33	51.67	11.00	10.00	11.00	10.67	53.33	11.00	11.00	10.00	10.67	53.33	32.00	32.00	31.00	32.00	31.67	52.78	52.78
19	581031321	11.00	10.00	10.00	10.33	51.67	10.00	11.00	11.00	10.67	53.33	10.00	11.00	11.00	10.67	53.33	31.00	32.00	32.00	31.67	52.78	52.78	
20	591071014	11.00	10.00	11.00	10.67	53.33	10.00	11.00	10.00	10.33	51.67	11.00	10.00	11.00	10.67	53.33	32.00	31.00	32.00	31.67	52.78	52.78	
21	591071057	10.00	11.00	10.00	10.33	51.67	11.00	10.00	11.00	10.67	53.33	11.00	10.00	10.00	10.67	53.33	32.00	32.00	31.00	31.67	52.78	52.78	
22	601011605	10.00	11.00	10.00	10.33	51.67	11.00	10.00	11.00	10.67	53.33	11.00	11.00	10.00	10.67	53.33	32.00	32.00	31.00	31.67	52.78	52.78	
23	611011588	10.00	11.00	10.00	10.33	51.67	11.00	10.00	11.00	10.67	53.33	11.00	11.00	10.00	10.67	53.33	32.00	32.00	31.00	31.67	52.78	52.78	
24	611011569	10.00	11.00	10.00	10.33	51.67	11.00	10.00	11.00	10.67	53.33	10.00	11.00	10.00	10.33	51.67	31.00	32.00	31.00	31.33	52.22	52.22	
25	601011551	10.00	10.00	11.00	10.33	51.67	10.00	10.00	10.00	10.00	50.00	11.00	10.00	10.00	10.33	51.67	31.00	30.00	31.00	30.67	51.11	51.11	
26	601011622	10.00	10.00	11.00	10.33	51.67	10.00	10.00	10.00	10.00	50.00	11.00	10.00	10.00	10.33	51.67	31.00	30.00	31.00	30.67	51.11	51.11	
27	581031336	7.00	5.00	6.00	6.00	30.00	5.00	7.00	5.00	5.67	28.33	6.00	5.00	7.00	6.00	30.00	18.00	17.00	18.00	17.67	29.44	29.44	
28	581031344	7.00	5.00	6.00	6.00	30.00	5.00	7.00	5.00	5.67	28.33	6.00	5.00	7.00	6.00	30.00	18.00	17.00	18.00	17.67	29.44	29.44	
29	581031331	6.00	5.00	6.00	5.67	28.33	5.00	6.00	5.00	5.33	26.67	8.00	5.00	6.00	6.33	31.67	19.00	16.00	17.00	17.33	28.89	28.89	
30	591011374	7.00	5.00	6.00	6.00	30.00	5.00	7.00	5.00	5.67	28.33	5.00	5.00	7.00	5.67	28.33	17.00	17.00	18.00	17.33	28.89	28.89	
31	591011347	6.00	5.00	6.00	5.67	28.33	5.00	6.00	5.00	5.33	26.67	8.00	5.00	6.00	6.33	31.67	19.00	16.00	17.00	17.33	28.89	28.89	
32	591071113	7.00	5.00	6.00	6.00	30.00	5.00	7.00	5.00	5.67	28.33	5.00	5.00	7.00	5.67	28.33	17.00	17.00	18.00	17.33	28.89	28.89	

ตารางที่ 27 รายงานผลคะแนนการสอบพูดโดยแบบทดสอบวัดความสามารภภาษาอังกฤษระดับมหาวิทยาลัย – การทดลองครั้งที่ 2 (ต่อ)

ที่	รหัสประจำตัว	Task A			รวม	%	Task B			รวม	%	Task C			รวม	%	Speaking			รวม	%	กลุ่ม	
		1	2	3			1	2	3			1	2	3			1	2	3				
33	591011367	6.00	5.00	6.00	5.67	28.33	5.00	6.00	5.00	5.33	26.67	6.00	5.00	6.00	5.67	28.33	17.00	16.00	17.00	16.67	27.78		
34	591071086	6.00	5.00	6.00	5.67	28.33	5.00	6.00	5.00	5.33	26.67	6.00	5.00	6.00	5.67	28.33	17.00	16.00	17.00	16.67	27.78		
35	591011362	6.00	5.00	6.00	5.67	28.33	5.00	6.00	4.00	5.00	25.00	5.00	4.00	6.00	5.00	25.00	16.00	15.00	16.00	15.67	26.11		
36	591011352	5.00	5.00	6.00	5.33	26.67	5.00	5.00	4.00	4.67	23.33	6.00	4.00	5.00	5.00	25.00	16.00	14.00	15.00	15.00	25.00		
37	581031338	5.00	5.00	6.00	5.33	26.67	5.00	5.00	4.00	4.67	23.33	6.00	4.00	5.00	5.00	25.00	16.00	14.00	15.00	15.00	25.00		
38	591071055	5.00	5.00	6.00	5.33	26.67	5.00	5.00	4.00	4.67	23.33	5.00	4.00	5.00	4.67	23.33	15.00	14.00	15.00	14.67	24.44		
39	591071027	4.00	5.00	6.00	5.00	25.00	5.00	4.00	5.00	4.67	23.33	6.00	5.00	4.00	5.00	25.00	15.00	14.00	15.00	14.67	24.44		
	Max	15.00	16.00	15.00	15.00	75.00	16.00	15.00	15.00	15.33	76.67	15.00	15.00	15.00	14.67	73.33	45.00	46.00	44.00	45.00	75.00		
	Min	4.00	5.00	6.00	5.00	25.00	5.00	4.00	4.00	4.67	23.33	5.00	4.00	4.00	4.67	23.33	15.00	14.00	15.00	14.67	24.44		
	\bar{X}	10.10	10.15	10.10	10.12	50.60	10.15	10.10	9.74	10.00	50.00	10.28	9.74	10.10	10.04	50.21	30.54	30.00	29.95	30.16	50.27		
	S.D.	3.44	4.08	3.35	3.59	17.96	4.08	3.44	3.94	3.77	18.86	3.45	3.94	3.44	3.54	17.72	10.81	11.31	10.56	10.88	18.14		
	(r_{tt})	0.98				0.98				0.97				0.99									

ตารางที่ 27 แสดงผลคะแนนการสอบการพูดโดยแบบทดสอบมาตรฐานหลังการทดลองครั้งที่ 2 และมีจำนวนผู้เข้าสอบที่ได้รับการคัดเลือกมาตามคะแนนที่ได้ในกลุ่มสูง กลาง และต่ำ กลุ่มละ 13 คน รวม 39 คน พบว่า ผู้เข้าสอบทุกคนได้เข้าทดสอบการพูดครบทั้ง 3 ภาระงาน โดยในภาพรวมของทุกภาระงาน ค่าร้อยละของคะแนนสูงสุดที่ผู้เข้าสอบทำได้ คือ ร้อยละ 75 คะแนนต่ำสุด คือ ร้อยละ 24.44 คะแนนเฉลี่ย คือ ร้อยละ 50.27 และส่วนเบี่ยงเบนมาตรฐานคือ ร้อยละ 18.14 และค่าความเชื่อมั่นของผู้ให้คะแนน (r_{tt}) เท่ากับ .99

เมื่อพิจารณารายภาระงาน พบว่า ภาระงานที่ผู้เข้าสอบส่วนใหญ่ทำได้ เรียงตามลำดับตามค่าร้อยละของคะแนนเฉลี่ย คือ 1) ภาระงาน A ร้อยละ 50.60 และค่าความเชื่อมั่นของผู้ให้คะแนน (r_{tt}) เท่ากับ .98 2) ภาระงาน C ร้อยละ 50.21 และค่าความเชื่อมั่นของผู้ให้คะแนน (r_{tt}) เท่ากับ .97 และ 3) ภาระงาน B ร้อยละ 50.00 และค่าความเชื่อมั่นของผู้ให้คะแนน (r_{tt}) เท่ากับ .98



ตารางที่ 28 รายงานผลคะแนนการสอบวัดความสามารถภาษาอังกฤษโดยใช้แบบทดสอบมาตรฐาน- การทดลองครั้งที่ 2

ที่	รหัสประจำตัว	ฟัง		รวม	%	การใช้ภาษา		รวม	%	อ่าน	%	เขียน		รวม	%	รวมทักษะ	%
		A	B-C			A	B					A	B				
		30	20	50	100	15	25	40	100	50	100	20	40	60	100	200	100
1	591031248	15.00	9.00	24.00	48.00	7.00	17.00	24.00	60.00	29.00	58.00	15.50	32.67	48.17	80.28	125.17	62.59
2	591031250	12.00	10.00	22.00	44.00	10.00	11.00	21.00	52.50	18.00	36.00	10.67	23.67	34.34	57.23	95.34	47.67
3	591031235	20.00	9.00	29.00	58.00	8.00	13.00	21.00	52.50	14.00	28.00	8.67	19.00	27.67	46.12	91.67	45.84
4	601011577	17.00	8.00	25.00	50.00	5.00	16.00	21.00	52.50	29.00	58.00			0.00	0.00	75.00	37.50
5	591031253	18.00	7.00	25.00	50.00	8.00	15.00	23.00	57.50	26.00	52.00			0.00	0.00	74.00	37.00
6	611011632	15.00	5.00	20.00	40.00	6.00	12.00	18.00	45.00	24.00	48.00			0.00	0.00	62.00	31.00
7	591011368	14.00	6.00	20.00	40.00	5.00	8.00	13.00	32.50	22.00	44.00			0.00	0.00	55.00	27.50
8	591011363	8.00	8.00	16.00	32.00	4.00	12.00	16.00	40.00	16.00	32.00			0.00	0.00	48.00	24.00
9	591071057	7.00	4.00	11.00	22.00	7.00	13.00	20.00	50.00	16.00	32.00			0.00	0.00	47.00	23.50
10	591011364	7.00	6.00	13.00	26.00	3.00	8.00	11.00	27.50	18.00	36.00			0.00	0.00	42.00	21.00
11	611011638	8.00	3.00	11.00	22.00	7.00	8.00	15.00	37.50	14.00	28.00			0.00	0.00	40.00	20.00
12	591011386	8.00	2.00	10.00	20.00	8.00	7.00	15.00	37.50	14.00	28.00			0.00	0.00	39.00	19.50
13	611011583	7.00	6.00	13.00	26.00	5.00	5.00	10.00	25.00	16.00	32.00			0.00	0.00	39.00	19.50
14	591011367	9.00	7.00	16.00	32.00	3.00	5.00	8.00	20.00	14.00	28.00			0.00	0.00	38.00	19.00
15	591071077	4.00	1.00	5.00	10.00	6.00	9.00	15.00	37.50	18.00	36.00			0.00	0.00	38.00	19.00
16	601011551	5.00	6.00	11.00	22.00	7.00	8.00	15.00	37.50	12.00	24.00			0.00	0.00	38.00	19.00

ตารางที่ 28 รายงานผลคะแนนการสอบวัดความสามารถภาษาอังกฤษโดยใช้แบบทดสอบมาตรฐาน- การทดลองครั้งที่ 2 (ต่อ)

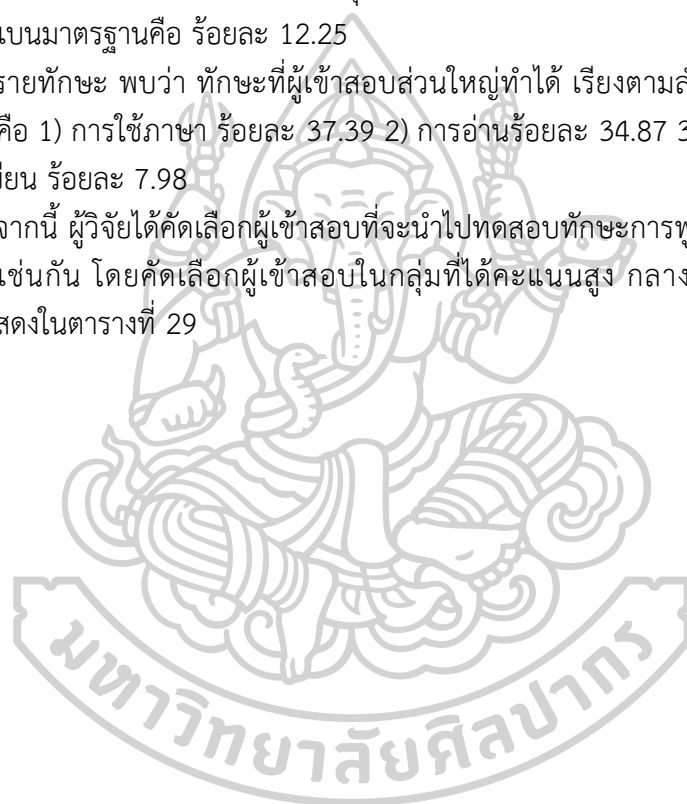
ที่	รหัสประจำตัว	ฟัง		รวม	%	การใช้ภาษา		รวม	%	อ่าน	%	เขียน		รวม	%	รวมทักษะ	%
		A	B-C			A	B					A	B				
17	591011355	30	20	50	100	15	25	40	100	50	100	20	40	60	100	200	100
18	591011360	7.00	8.00	15.00	30.00	2.00	3.00	5.00	12.50	17.00	34.00			0.00	0.00	37.00	18.50
19	591011390	4.00	4.00	8.00	16.00	6.00	9.00	15.00	37.50	14.00	28.00			0.00	0.00	37.00	18.50
20	591011357	5.00	5.00	10.00	20.00	6.00	5.00	11.00	27.50	14.00	28.00			0.00	0.00	35.00	17.50
21	591071086	4.00	3.00	7.00	14.00	2.00	9.00	11.00	27.50	17.00	34.00			0.00	0.00	35.00	17.50
22	591011347	4.00	3.00	7.00	14.00	4.00	7.00	11.00	27.50	15.00	30.00			0.00	0.00	33.00	16.50
23	591071055	5.00	4.00	9.00	18.00	3.00	7.00	10.00	25.00	13.00	26.00			0.00	0.00	32.00	16.00
	Max	20.00	10.00	29.00	58.00	10.00	17.00	24.00	60.00	29.00	58.00	15.50	32.67	48.17	80.28	125.17	62.59
	Min	4.00	1.00	5.00	10.00	2.00	3.00	5.00	12.50	11.00	22.00	8.67	19.00	0.00	0.00	32.00	16.00
	\bar{X}	9.09	5.61	14.70	29.39	5.61	9.35	14.96	37.39	17.43	34.87	11.61	25.11	4.79	7.98	51.88	25.94
	S.D.	4.990	2.407	6.779	13.557	2.126	3.700	5.022	12.556	5.133	10.266	3.511	6.948	13.034	21.723	24.451	12.225

ตารางที่ 28 แสดงผลคะแนนการสอบวัดความสามารถภาษาอังกฤษโดยใช้แบบทดสอบมาตรฐานหลังการทดลองครั้งที่ 2 ซึ่งประกอบด้วย 4 ทักษะ คือ การฟัง การใช้ภาษา การอ่านและการเขียน และแสดงคะแนนโดยจัดเรียงจากคะแนนสูงสุดไปหาต่ำสุด พบว่า จากจำนวนผู้เข้าสอบ 23 คน มีผู้เข้าสอบที่ทำแบบทดสอบครบทั้ง 4 ทักษะ จำนวน 3 คน ซึ่งเป็นผู้ที่ได้คะแนนสอบอยู่ในกลุ่มสูง และมีผู้เข้าสอบที่ไม่ทำแบบทดสอบการเขียน 20 คน ซึ่งส่วนใหญ่เป็นผู้ที่อยู่ในกลุ่มที่ได้คะแนนปานกลางและต่ำ

เมื่อพิจารณาคะแนนสอบ พบว่า จากคะแนนรวมทุกทักษะ ค่าร้อยละของคะแนนสูงสุดที่ผู้เข้าสอบทำได้ คือ ร้อยละ 62.59 คะแนนต่ำสุด คือ ร้อยละ 16.00 คะแนนเฉลี่ย คือ ร้อยละ 25.94 และส่วนเบี่ยงเบนมาตรฐานคือ ร้อยละ 12.25

เมื่อพิจารณารายทักษะ พบว่า ทักษะที่ผู้เข้าสอบส่วนใหญ่ทำได้ เรียงตามลำดับตามค่าร้อยละของคะแนนเฉลี่ย คือ 1) การใช้ภาษา ร้อยละ 37.39 2) การอ่านร้อยละ 34.87 3) การฟัง ร้อยละ 29.39 และ 4) การเขียน ร้อยละ 7.98

นอกจากนี้ ผู้วิจัยได้คัดเลือกผู้เข้าสอบที่จะนำไปทดสอบทักษะการพูดจากผลคะแนนรวมในตารางที่ 27 เช่นกัน โดยคัดเลือกผู้เข้าสอบในกลุ่มที่ได้คะแนนสูง กลาง ต่ำ กลุ่มละ 5 คน ดังรายละเอียดแสดงในตารางที่ 29



ตารางที่ 27 รายงานผลคะแนนการสอบพูดโดยแบบทดสอบวัดความสามารถภาษาอังกฤษมาตรฐาน- การทดลองครั้งที่ 2

ที่	รหัสประจำตัว	Task A			รวม	%	Task B			รวม	%	Task C			รวม	%	Speaking			รวม	%	กลุ่ม
		1.00	2.00	3.00			1.00	2.00	3.00			1.00	2.00	3.00			1.00	2.00	3.00			
1	591031248	20.00	14.00	14.00	20.00	100.00	20.00	20.00	20.00	20.00	100.00	20.00	20.00	20.00	20.00	100.00	60.00	60.00	60.00	60.00	100.00	กลุ่ม
2	591031250	15.00	15.00	15.00	15.00	75.00	15.00	15.00	15.00	14.00	70.00	14.00	15.00	15.00	14.00	73.33	43.00	43.00	43.00	43.00	71.67	สูง
3	591031235	13.00	14.00	13.00	13.33	66.67	14.00	13.00	13.00	13.33	66.67	14.00	13.00	13.00	13.33	66.67	41.00	40.00	39.00	40.00	66.67	สูง
4	601011577	15.00	16.00	14.00	15.00	75.00	16.00	15.00	15.00	15.33	76.67	14.00	15.00	15.00	14.67	73.33	45.00	46.00	44.00	45.00	75.00	สูง
5	591031253	13.00	14.00	14.00	13.67	68.33	14.00	13.00	13.00	14.00	70.00	15.00	15.00	13.00	14.33	71.67	42.00	42.00	42.00	42.00	70.00	สูง
6	591011364	10.00	10.00	11.00	10.33	51.67	10.00	10.00	10.00	10.00	50.00	11.00	10.00	10.00	10.33	51.67	31.00	30.00	31.00	30.67	51.11	กลาง
7	611011638	11.00	11.00	10.00	10.67	53.33	11.00	11.00	11.00	11.00	55.00	11.00	11.00	11.00	11.00	55.00	33.00	33.00	32.00	32.67	54.44	กลาง
8	591011386	10.00	10.00	10.00	10.33	51.67	11.00	10.00	11.00	10.67	53.33	11.00	11.00	10.00	10.67	53.33	32.00	32.00	31.00	31.67	52.78	กลาง
9	611011583	11.00	10.00	10.00	10.33	51.67	10.00	11.00	11.00	10.67	53.33	10.00	11.00	11.00	10.67	53.33	31.00	32.00	32.00	31.67	52.78	กลาง
10	591011367	10.00	12.00	11.00	11.00	55.00	12.00	10.00	10.00	10.67	53.33	11.00	10.00	10.00	10.33	51.67	33.00	32.00	31.00	32.00	53.33	กลาง
11	591011390	6.00	5.00	6.00	5.67	28.33	5.00	6.00	6.00	5.33	26.67	8.00	5.00	6.00	6.33	31.67	19.00	16.00	17.00	17.33	28.89	ต่ำ
12	591011357	7.00	5.00	6.00	6.00	30.00	5.00	7.00	5.00	5.67	28.33	6.00	5.00	7.00	6.00	30.00	18.00	17.00	18.00	17.67	29.44	ต่ำ
13	591071086	5.00	5.00	6.00	5.33	26.67	5.00	5.00	4.00	4.67	23.33	6.00	4.00	5.00	5.00	25.00	16.00	14.00	15.00	15.00	25.00	ต่ำ
14	591011347	6.00	5.00	6.00	5.67	28.33	5.00	6.00	5.00	5.33	26.67	6.00	5.00	6.00	5.67	28.33	17.00	16.00	17.00	16.67	27.78	ต่ำ

ตารางที่ 29 รายงานผลคะแนนการสอบพูดโดยแบบทดสอบวัดความสามารถภาษาอังกฤษมาตรฐาน- การทดลองครั้งที่ 2 (ต่อ)

ที่	รหัสประจำตัว	Task A			รวม	%	Task B			รวม	%	Task C			รวม	%	Speaking			รวม	%	กลุ่ม
		1.00	2.00	3.00			1.00	2.00	3.00			1.00	2.00	3.00			1.00	2.00	3.00			
15	591071055	7.00	5.00	6.00	6.00	30.00	5.00	7.00	5.00	5.67	28.33	5.00	5.00	7.00	5.67	28.33	17.00	17.00	18.00	17.33	28.89	
	Max	15.00	16.00	15.00	15.00	75.00	16.00	15.00	15.00	15.33	76.67	15.00	15.00	15.00	14.67	73.33	45.00	46.00	44.00	45.00	75.00	
	Min	5.00	5.00	6.00	5.33	26.67	5.00	5.00	4.00	4.67	23.33	5.00	4.00	5.00	5.00	25.00	16.00	14.00	15.00	15.00	25.00	
	\bar{X}	10.20	10.13	10.13	10.16	50.78	10.13	10.20	9.80	10.04	50.22	10.40	9.80	10.20	10.13	50.67	30.73	30.13	30.13	30.33	50.56	
	S.D.	3.38	4.14	3.42	3.62	18.09	4.14	3.38	4.02	3.80	19.00	3.48	4.02	3.38	3.57	17.85	10.83	11.40	10.68	10.96	18.26	
	(r_{tt})	.98					.98					.98					.99					

ตารางที่ 29 แสดงผลคะแนนการสอบการพูดโดยแบบทดสอบมาตรฐานหลังการทดลองครั้งที่ 2 มีจำนวนผู้เข้าสอบที่ได้รับการคัดเลือกมาตามคะแนนที่ได้ในกลุ่มสูง กลาง และต่ำ กลุ่มละ 5 คน รวม 15 คน พบว่า ผู้เข้าสอบทุกคนได้เข้าทดสอบการพูดครบทั้ง 3 ภาระงาน โดยในภาพรวมของทุก ภาระงาน ค่าร้อยละของคะแนนสูงสุดที่ผู้เข้าสอบทำได้ คือ ร้อยละ 75.00 คะแนนต่ำสุด คือ ร้อยละ 25 คะแนนเฉลี่ย คือ ร้อยละ 50.56 และส่วนเบี่ยงเบนมาตรฐานคือ ร้อยละ 18.26 และค่าความเชื่อมั่นของผู้ให้คะแนน (r_{tt}) เท่ากับ .99

เมื่อพิจารณารายภาระงาน พบว่า ภาระงานที่ผู้เข้าสอบส่วนใหญ่ทำได้ เรียงตามลำดับตามค่า ร้อยละของคะแนนเฉลี่ย คือ 1) ภาระงาน A ร้อยละ 50.78 และค่าความเชื่อมั่นของผู้ให้คะแนน (r_{tt}) เท่ากับ .98 2) ภาระงาน C ร้อยละ 50.67 และค่าความเชื่อมั่นของผู้ให้คะแนน (r_{tt}) เท่ากับ .98 และ 3) ภาระงาน B ร้อยละ 50.22 และค่าความเชื่อมั่นของผู้ให้คะแนน (r_{tt}) เท่ากับ .98

1.4.2.1.2) การวิเคราะห์แบบทดสอบรายข้อ

การวิเคราะห์แบบทดสอบรายข้อ ประกอบด้วยการวิเคราะห์ 3 ชนิด คือ การวิเคราะห์ค่า ความยากง่าย และการวิเคราะห์ค่าอำนาจจำแนก (ดังแสดงในภาคผนวก รฐ) และ การวิเคราะห์ดัชนี ความลำเอียง (ดังแสดงในภาคผนวก ฎ)

1.4.2.1.3) การวิเคราะห์ความเชื่อมั่น

การวิเคราะห์ความเชื่อมั่นของแบบทดสอบ ประกอบด้วยการวิเคราะห์ 2 ชนิด คือ 1) ความ เชื่อมั่นแบบทดสอบปรนัย ((KR-20) (ภาคผนวก รฐ) และ 2) ความเชื่อมั่นของการให้คะแนน แบบทดสอบอัตนัย (ภาคผนวก ท)

1.4.2.2) การตรวจสอบคุณภาพแบบทดสอบโดยรวม

การตรวจสอบคุณภาพแบบทดสอบโดยรวม ประกอบด้วย 1) การตรวจสอบข้อมูลเชิง คุณภาพ โดยการสอบถามความพึงพอใจของผู้เกี่ยวข้องต่อแบบทดสอบวัดความสามารถภาษาอังกฤษ (ดังรายละเอียดแสดงในภาคผนวก ซ และ 2) การตรวจสอบข้อมูลเชิงปริมาณ ซึ่งเป็นการเทียบ คะแนนที่ได้จากสอบด้วยแบบทดสอบที่พัฒนาขึ้น และแบบทดสอบมาตรฐาน (ดังรายละเอียดแสดง ในภาคผนวก ฉ)

1.4.2.3) สรุปผลคุณภาพแบบทดสอบและประสิทธิภาพของเครื่องมือที่เกี่ยวข้องหลังการ ทดลองครั้งที่ 2

ผู้วิจัยดำเนินการเก็บข้อมูลกับกลุ่มทดลองครั้งที่ 2 จำนวน 94 คน โดยมีการใช้เครื่องมือที่เกี่ยวข้อง คือ 1) แบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test) 2) แบบประเมินความพึงพอใจของผู้ใช้แบบทดสอบต่อแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัยที่ผู้วิจัยสร้างขึ้น และ 3) แบบทดสอบวัดความสามารถภาษาอังกฤษมาตรฐานสากล โดยมีรายละเอียดผลคุณภาพแบบทดสอบในภาคผนวก 15 และพอที่จะสรุปผลการคัดเลือกหรือปรับปรุงแบบทดสอบเพื่อที่จะนำไปใช้ในการเก็บข้อมูลจริง ดังนี้

1.4.1.3.1) แบบทดสอบทักษะการฟัง มีข้อคำถามทั้งหมด 65 ข้อ จาก 3 ตอน คือ ตอน A มีข้อคำถาม 35 ข้อ ตอน B มีข้อคำถาม 14 ข้อ และตอน C มีข้อคำถาม 16 ข้อ โดยมีค่าความเชื่อมั่น $r = .74$ และมีการดำเนินการในแต่ละตอนหลังจากการวิเคราะห์แบบทดสอบรายข้อ ดังนี้

ตอน A มีข้อคำถามทั้งหมด 35 ข้อ มีค่าความเชื่อมั่น $r = .80$ และมีข้อคำถามที่ไม่ได้คุณภาพ คือ มีความลำเอียงเข้าผู้เข้าสอบสายศิลป์ จำนวน 5 ข้อ ผู้วิจัยได้ตัดข้อคำถามออก 5 ข้อ คงไว้ใช้ในการเก็บข้อมูลจริง จำนวน 30 ข้อ

ตอน B-C มีข้อคำถามทั้งหมด 30 ข้อ มีค่าความเชื่อมั่น $r = .68$ และมีข้อคำถามที่ไม่ได้คุณภาพ คือ อำนาจจำแนกติดลบ และมีความลำเอียงเข้าหาผู้เข้าสอบสายศิลป์ รวม 3 ข้อ ผู้วิจัยได้ตัดข้อคำถามออก 3 ข้อ คงไว้ใช้ในการเก็บข้อมูลจริง จำนวน 27 ข้อ

1.4.1.3.2) แบบทดสอบทักษะการใช้ภาษา มีข้อคำถามทั้งหมด 50 ข้อ จาก 2 ตอน คือ ตอน A มีข้อคำถาม 20 ข้อ และตอน B มีข้อคำถาม 30 ข้อ โดยมีค่าความเชื่อมั่น $r = .71$ และมีการดำเนินการในแต่ละตอนหลังจากการวิเคราะห์แบบทดสอบรายข้อ ดังนี้

ตอน A มีข้อคำถามทั้งหมด 20 ข้อ มีค่าความเชื่อมั่น $r = .70$ และมีข้อคำถามที่ไม่ได้คุณภาพ คือ มีลำเอียงถึงผู้เข้าสอบสายวิทย์ จำนวน 1 ข้อ ผู้วิจัยได้ตัดข้อคำถามออก 1 ข้อ คงไว้ใช้ในการเก็บข้อมูลจริง จำนวน 19 ข้อ

ตอน B มีข้อคำถามทั้งหมด 30 ข้อ มีค่าความเชื่อมั่น $r = .73$ และมีข้อคำถามที่ไม่ได้คุณภาพ คือ อำนาจจำแนกติดลบ และ มีความลำเอียงเข้าหาสายวิทย์ รวม 2 ข้อ ผู้วิจัยได้ตัดข้อคำถามออก 2 ข้อ คงไว้ใช้ในการเก็บข้อมูลจริง จำนวน 28 ข้อ

1.4.1.3.3) แบบทดสอบทักษะการอ่าน มีข้อคำถามทั้งหมด 82 ข้อ จาก 2 ตอน คือ ตอน A มีข้อคำถาม 40 ข้อ และตอน B มีข้อคำถาม 42 ข้อ โดยมีค่าความเชื่อมั่น $r = .86$ และมีการดำเนินการในแต่ละตอนหลังจากการวิเคราะห์แบบทดสอบรายข้อ ดังนี้

ตอน A มีข้อคำถามทั้งหมด 40 ข้อ มีค่าความเชื่อมั่น $r = .86$ และมีข้อคำถามที่ไม่ได้คุณภาพ คือ อำนาจจำแนกติดลบ ลำเอียงเข้าหาสายวิทย์ และลำเอียงเข้าหาสายศิลป์ 9 ข้อ เนื่องจากบางข้อเป็นข้อสอบต่อเนื่อง ผู้วิจัยจึงตัดข้อคำถามออก 3 ข้อ คงไว้ใช้ในการเก็บข้อมูลจริง จำนวน 37 ข้อ

ตอน B มีข้อคำถามทั้งหมด 42 ข้อ มีค่าความเชื่อมั่น $r = .86$ และมีข้อคำถามที่ไม่ได้คุณภาพ คือ อำนาจจำแนกติดลบ ลำเอียงเข้าหาสายวิทย์ และลำเอียงเข้าหาสายศิลป์ จำนวน 7 ข้อ ผู้วิจัยจึงตัดข้อคำถามออก 7 ข้อ คงไว้ใช้ในการเก็บข้อมูลจริง จำนวน 35 ข้อ

1.4.1.3.4) แบบทดสอบทักษะการเขียน โดยรวมมีค่าความเชื่อมั่น (Cronbach Alpha Coefficient) เท่ากับ .78 ประกอบด้วย 2 ภาระงาน คือ ภาระงาน A ซึ่งมีค่าความเชื่อมั่น

(Cronbach Alpha Coefficient) เท่ากับ .79 และภาระงาน B ซึ่งมีค่าความเชื่อมั่น (Cronbach Alpha Coefficient) เท่ากับ .77 และสามารถนำไปใช้เก็บข้อมูลกับกลุ่มตัวอย่าง ได้ทั้ง 2 ภาระงาน 1.4.1.3.5) แบบทดสอบทักษะการพูด โดยรวมมีค่าความเชื่อมั่น เชื่อมั่น (Cronbach Alpha Coefficient) เท่ากับ .99 ประกอบด้วย 3 ภาระงาน คือ ภาระงาน A, B และ C ซึ่งมีค่าความเชื่อมั่น (Cronbach Alpha Coefficient) เท่ากับ .99 ทั้ง 3 ภาระงาน และสามารถนำไปใช้เก็บข้อมูลกับกลุ่มตัวอย่าง ได้ทั้ง 3 ภาระงาน

1.4.1.3.6) แบบสอบถามความพึงพอใจผู้เกี่ยวข้องต่อแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test) ได้ความพึงพอใจอยู่ในระดับพึงพอใจมาก ($\bar{X} = 4.08$) และมีค่าอัลฟา เท่ากับ .98

1.4.1.3.7) การหาความตรงร่วมสมัย (Concurrent Validity) ของแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test) และแบบทดสอบมาตรฐาน (Standard English Proficiency Test) โดยภาพรวมมีค่าความสัมพันธ์ $r = .95$ และในรายทักษะย่อย มีค่าความเชื่อมั่นเรียงจากสูงไปหาต่ำ คือ 1) ทักษะการพูด $r = .92$ 2) ทักษะการฟัง $r = .89$ 3) การใช้ภาษา $r = .78$ 4) ทักษะการเขียน $r = .77$ 5) ทักษะการอ่าน $r = .73$

โดยมีสรุปผลคุณภาพแบบทดสอบและประสิทธิภาพของเครื่องมือที่เกี่ยวข้องหลังจากการทดลองใช้ครั้งที่ 2 แสดงไว้ในตารางที่ 30-31



ตารางที่ 28 สรุปคุณภาพแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัยหลังจากการทดลองใช้ครั้งที่ 2

แบบทดสอบ	จำนวนข้อ	Max	Min	\bar{X}	S.D.	P	R	Reliability	คุณภาพรายชื่อ	ลำเอียงเข้าหา (ข้อที่)		รวมตัดออก	เก็บไว้
										สายวิทย์	สายศิลป์		
การฟัง	65	52	12	25.70	9.129	.41	.44	.74				8	57
A	35	28	4	12.31	5.754	.37	.47	.80			12 16 25 27 29	5	30
B-C	30	24	5	13.38	4.346	.46	.42	.68	อำนาจจำแนกติดลบ ข้อที่ 43		43 55 59	3	27
การใช้ภาษา	50	43	7	23.67	7.861	.50	.48	.71				3	47
A	20	19	3	9.90	3.525	.52	.50	.70			20	1	19
B	30	24	4	13.75	4.839	.48	.47	.73	อำนาจจำแนกติดลบ ข้อที่ 26		38	2	28

ตารางที่ 30 สรุปคุณภาพแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัยหลังจาการทดลองใช้ครั้งที่ 2 (ต่อ)

แบบทดสอบ	จำนวนข้อ	Max	Min	\bar{X}	S.D.	P	R	Reliability	คุณภาพรายชื่อ	ลำเอียงเข้าหา (ข้อที่)		รวมตัดออก	เก็บไว้
										สายวิทย์	สายศิลป์		
การอ่าน	82	69	13	33.66	14.550	.45	.54	.86				10	72
A	40	33	5	17.06	7.591	.47	.56	.86	อำนาจจำแนกดีดลบ ข้อที่ 28	35	11	3	37
											24		
											25		
											29		
B	42	36	5	16.59	7.898	.43	.52	.86	อำนาจจำแนกดีดลบ ข้อที่ 19 อำนาจจำแนกดีดลบ ข้อที่ 34	11	4	7	35
											37		
											38		
											34		

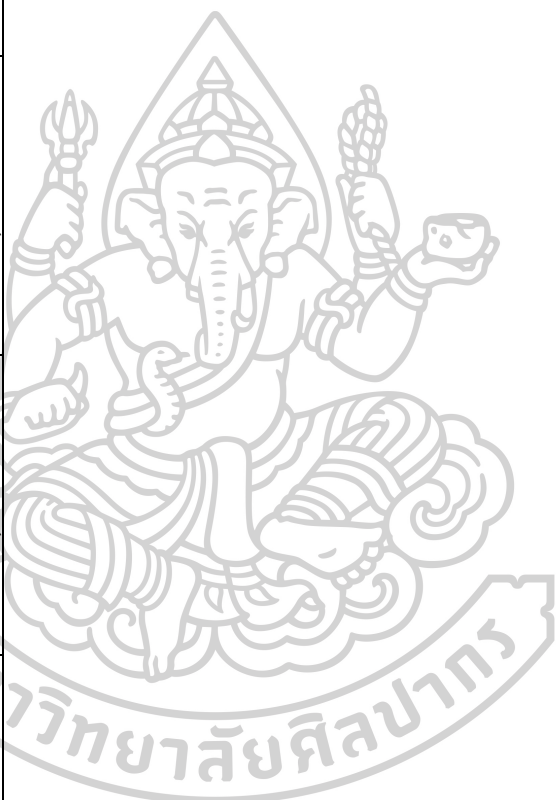
ตารางที่ 30 สรุปคุณภาพแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัยหลังจากการทดลองใช้ครั้งที่ 2 (ต่อ)

แบบทดสอบ	ครั้งที่ 2	Max	Min	\bar{X}	S.D.	P	R	Reliability	ลำเอียงเข้าหา		รวม ตัดออก	เก็บไว้
									(ข้อที่)	สายวิทย์		
การเขียน	2	50	0	27.80	14.463	.46	.63	.78			-	2
A	1	16.83	1	8.97	5.580	.43	.63	.79			-	1
B	1	33.67	0	18.84	9.770	.48	.63	.77			-	1
การพูด	3	45	14.67	31.20	10.802	.50	.45	.99			-	3
A	1	15	5	10.46	3.520	.51	.44	.99			-	1
B	1	15.33	4.67	10.33	3.710	.51	.46	.99			-	1
C	1	14.67	4.67	10.41	3.610	.50	.44	.99			-	1
แบบสอบถาม ความคิดเห็น ผู้ใช้ แบบทดสอบ	36	4.56	3.70	4.17	.856			Alpha = .98				

หมายเหตุ การอ่าน Part A ไม่ตัดข้อ 24-25 และ ข้อ 31, 35, 39 และ 40 ออก เพราะเป็นข้อสอบต่อเนื่อง

ตารางที่ 29 สรุปความตรงร่วมสมัย (Concurrent Validity) ของแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (X) และแบบทดสอบมาตรฐาน (Y) ที่ใช้ในการทดลองใช้ครั้งที่ 2

ทักษะการฟัง		ทักษะการใช้ภาษา		ทักษะการอ่าน		ทักษะการเขียน		ทักษะการพูด		ทั้งหมด	
UEPT	TOEFL	UEPT	TOEFL	UEPT	TOEFL	UEPT	IELTS	UEPT	IELTS	UEPT	IELTS
$r_{xy} = .89$		$r_{xy} = .78$		$r_{xy} = .73$		$r_{xy} = .77$		$r_{xy} = .92$		$r_{xy} = .95$	



ตารางที่ 30 สรุปผลการวิเคราะห์ข้อมูลตามจุดประสงค์ของการวิจัยและสถิติที่ใช้

จุดประสงค์การวิจัย	สถิติที่ใช้	การทดลองครั้งที่ 1	การทดลองครั้งที่ 2
1) เพื่อประเมินคุณภาพข้อสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test)	การหาค่าคุณประโยชน์ของแบบทดสอบค่าเฉลี่ย (\bar{X}) ส่วนเบี่ยงเบนมาตรฐาน (S.D.)	จะนำเสนอผลในทที่ 4	
2) เพื่อประเมินคุณภาพรายข้อและความเชื่อมั่นของแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test)	การวิเคราะห์แบบทดสอบรายข้อ ค่าความยากง่าย ค่าอำนาจจำแนก	<p>ฟัง P = .36</p> <p>ใช้ภาษา P = .42</p> <p>อ่าน P = .50</p> <p>เขียน P = .38</p> <p>พูด P = .47</p> <p>ฟัง R = .41</p> <p>ใช้ภาษา R = .49</p> <p>อ่าน R = .44</p> <p>เขียน R = .71</p> <p>พูด R = .88</p> <p>ไม่ได้ทำเนื่องจากจำนวนผู้เข้าสอบมีน้อยเกินไป</p>	<p>ฟัง P = .41</p> <p>ใช้ภาษา P = .50</p> <p>อ่าน P = .45</p> <p>เขียน P = .46</p> <p>พูด P = .50</p> <p>ฟัง R = .44</p> <p>ใช้ภาษา R = .48</p> <p>อ่าน R = .54</p> <p>เขียน R = .63</p> <p>พูด R = .45</p> <p>ฟัง $d_g = -3.10 - 2.68$</p> <p>ใช้ภาษา $d_g = -1.22 - 3.90$</p> <p>อ่าน $d_g = -3.54 - 3.18$</p>
	ค่าดัชนีความลำเอียง		

ตารางที่ 32 สรุปผลการวิเคราะห์ข้อมูลตามจุดประสงค์ของการวิจัยและสถิติที่ใช้ (ต่อ)

จุดประสงค์การวิจัย	สถิติที่ใช้	การทดลองครั้งที่ 1	การทดลองครั้งที่ 2
	ค่าความเชื่อมั่น	ฟัง $r = .71$ ใช้ภาษา $r = .69$ อ่าน $r = .76$ เขียน $r = .44$ พูด $r = .99$	ฟัง $r = .74$ ใช้ภาษา $r = .71$ อ่าน $r = .86$ เขียน $r = .78$ พูด $r = .99$
3) เพื่อหาความสัมพันธ์ระหว่างคะแนนแต่ละด้านของแบบทดสอบที่สร้างขึ้นกับแบบทดสอบมาตรฐาน	การตรวจสอบข้อมูลเชิงปริมาณ ค่าความสัมพันธ์ (The Pearson Product-Moment Correlation Coefficient)	ฟัง $r_{xy} = .75$ ใช้ภาษา $r_{xy} = .80$ อ่าน $r_{xy} = .70$ เขียน $r_{xy} = .96$ พูด $r_{xy} = .99$ โดยรวม $r_{xy} = .87$	ฟัง $r_{xy} = .89$ ใช้ภาษา $r_{xy} = .78$ อ่าน $r_{xy} = .73$ เขียน $r_{xy} = .77$ พูด $r_{xy} = .92$ โดยรวม $r_{xy} = .95$
4) เพื่อศึกษาความพึงพอใจของผู้เกี่ยวข้องต่อแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test)	การตรวจสอบข้อมูลเชิงคุณภาพ ค่าเฉลี่ย (\bar{X}) ส่วนเบี่ยงเบนมาตรฐาน (S.D.) Cronbach's Alpha Coefficient	$\bar{X} = 4.08$ S.D. = 0.867 Alpha = .95	$\bar{X} = 4.17$ S.D. = 0.856 Alpha = .98

จากข้อมูลในตารางที่ 32 แสดงให้เห็นว่า ผู้วิจัยได้นำเครื่องมือการวิจัย ได้แก่ แบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย แบบทดสอบวัดความสามารถภาษาอังกฤษมาตรฐาน และแบบประเมินความประเมินความพึงพอใจผู้เกี่ยวข้องต่อแบบทดสอบที่สร้างขึ้น ไปทดลองใช้ 2 ครั้ง กับกลุ่มทดลองจำนวน 28 และ 94 คน ตามลำดับ และได้ตัดข้อคำถามที่ไม่มีคุณภาพตามเกณฑ์ออกไป เพื่อให้เครื่องมือมีประสิทธิภาพในการตอบสนองจุดประสงค์ต่าง ๆ ของการวิจัยในการเก็บข้อมูลจริงกับกลุ่มตัวอย่าง จำนวน 45 คน โดยสามารถสรุปประสิทธิภาพของเครื่องมือจากผลที่ได้ดังนี้

1) แบบประเมินคุณภาพเบื้องต้น ตรวจสอบคุณภาพโดยการหาค่า IOC ดัชนีความสอดคล้องระหว่างข้อคำถามกับจุดประสงค์ เท่ากับ 1.00

2) แบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย มีคุณภาพรายข้อและความเชื่อมั่น ดังนี้

ด้านการฟัง มีค่าความยากง่ายระดับปานกลาง ($P = .41$) อำนาจจำแนกจำแนกได้ดี ($R = .44$) ความเชื่อมั่นค่อนข้างสูง ($r = .74$)

ด้านการใช้ภาษา มีค่าความยากง่ายระดับปานกลาง ($P = .50$) อำนาจจำแนกจำแนกได้ดี ($R = .48$) ความเชื่อมั่นค่อนข้างสูง ($r = .71$)

ด้านการอ่าน มีค่าความยากง่ายระดับปานกลาง ($P = .45$) อำนาจจำแนกจำแนกได้ดี ($R = .54$) ความเชื่อมั่นค่อนข้างสูง ($r = .86$)

ด้านการเขียน มีค่าความยากง่ายระดับปานกลาง ($P = .46$) อำนาจจำแนกจำแนกได้ดีมาก ($R = .63$) ความเชื่อมั่นค่อนข้างสูง ($\text{Alpha} = .78$)

ด้านการพูด มีค่าความยากง่ายระดับปานกลาง ($P = .50$) อำนาจจำแนกจำแนกได้ดี ($R = .45$) ความเชื่อมั่นสูงมาก ($\text{Alpha} = .99$)

3) การหาความตรงร่วมสมัย (Concurrent Validity) โดยการเทียบคะแนนกับแบบทดสอบมาตรฐาน พบว่า ความตรงร่วมสมัย (Concurrent Validity) โดยรวมทั้งฉบับเท่ากับ .95 และความตรงร่วมสมัย (Concurrent Validity) ของแต่ละทักษะ คือ ด้านการฟัง สูงกว่าเกณฑ์ ($r_{xy} = .89$) ด้านการใช้ภาษา สูงกว่าเกณฑ์ ($r_{xy} = .78$) ด้านการอ่าน สูงกว่าเกณฑ์ ($r_{xy} = .73$) ด้านการเขียน สูงกว่าเกณฑ์ ($r_{xy} = .77$) และด้านการพูด สูงกว่าเกณฑ์ ($r_{xy} = .89$)

4) ความพึงพอใจของผู้เกี่ยวข้องต่อแบบทดสอบที่สร้างขึ้นอยู่ในระดับมาก ($\bar{X} = 4.17$, S.D. = 0.856) และค่า $\text{Alpha} = .98$

บทที่ 4

ผลการวิเคราะห์ข้อมูล

การวิจัยเรื่องการพัฒนาแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (Development of a University English Proficiency Test) มีวัตถุประสงค์ 2 ข้อ คือ เพื่อสร้างและพัฒนาแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test) ให้เป็นแบบทดสอบมาตรฐานที่มีลักษณะเฉพาะ คือ

- คุณประโยชน์ของแบบทดสอบเป็นไปตามเกณฑ์ที่กำหนด
- คุณภาพรายข้อและความเชื่อมั่นเป็นไปตามเกณฑ์ที่กำหนด
- ความตรงร่วมสมัย (Concurrent Validity) สูง เมื่อเทียบกับแบบทดสอบมาตรฐาน

เพื่อศึกษาความพึงพอใจของผู้เกี่ยวข้องต่อแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test)

โดยมีรายละเอียดต่อไปนี้

4.1 คุณประโยชน์ของแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test)

ในการประเมินคุณประโยชน์ของแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test) ผู้วิจัยได้สำรวจความคิดเห็นของผู้เชี่ยวชาญ 3 ท่าน และอาจารย์ผู้สอนภาษาอังกฤษระดับมหาวิทยาลัย 7 ท่าน (ดังรายชื่อในภาคผนวก ง) โดยใช้แบบประเมินคุณประโยชน์ของแบบทดสอบ (ดังภาคผนวก ฉ) และกำหนดเกณฑ์ไว้ว่าแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test) มีความมียุทธประโยชน์อยู่ในระดับมาก ($\bar{X} \geq 3.5$) โดยมีผลการประเมินและข้อเสนอแนะสำหรับการพัฒนาแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัยแต่ละทักษะ ดังนี้

4.1.1 ทักษะการฟัง

แบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test) ด้านการฟัง มีผลการประเมินคุณประโยชน์และข้อเสนอแนะที่เป็นประโยชน์ดังนี้

4.1.1.1 ผลการประเมินคุณประโยชน์ของแบบทดสอบด้านการฟัง

ในการประเมินคุณประโยชน์ของแบบทดสอบด้านการฟังโดยใช้แบบประเมินคุณประโยชน์แบบทดสอบก่อนนำไปใช้ ที่มีข้อคำถามแบบมาตราส่วนประมาณค่า จำนวน 20 ข้อ และสอบถามเกี่ยวกับคุณประโยชน์แบบทดสอบ 7 ด้าน คือ 1) ความตรงเชิงโครงสร้าง (Construct Validity) 2) ความเชื่อมั่น (Reliability) 3) ความสมจริง (Authenticity) 4) การมีปฏิสัมพันธ์ (Interactiveness) 5) ผลกระทบ (Impact) 6) ความสะดวกใช้ (Practicality) และ 7) ความเป็นธรรม (Fairness) ได้ผลการประเมินคุณประโยชน์ในแต่ละด้าน ดังรายละเอียดในตารางที่ 33

ตารางที่ 31 ผลการประเมินคุณประโยชน์ของแบบทดสอบด้านการฟัง

คุณประโยชน์ของแบบทดสอบ (Test Usefulness)	\bar{X}	S.D.	แปลผล
ความตรงเชิงโครงสร้าง (Construct Validity)	4.52	0.110	มากที่สุด
1. มีการระบุวัตถุประสงค์การสอบของแบบทดสอบการฟังแต่ละตอนไว้อย่างชัดเจน	4.60	0.699	มากที่สุด
2. วัตถุประสงค์ย่อยของแบบทดสอบแต่ละตอนสอดคล้องกับจุดประสงค์หลัก	4.40	0.516	มาก
3. ภาระงานสอบด้านการฟังทุกตอนสอดคล้องกับวัตถุประสงค์การสอบ	4.60	0.516	มากที่สุด
4. วิธีการให้คะแนนมีความเหมาะสมและสามารถสะท้อนความสามารถด้านการฟังของผู้เข้าสอบ	4.40	0.516	มาก
5. คะแนนสอบสามารถทำนายความสามารถด้านการฟังของผู้เข้าสอบ	4.60	0.516	มากที่สุด
ความเชื่อมั่น (Reliability)	4.25	0.191	มาก
6. ข้อสอบแต่ละข้อมีตัวเลือกที่เป็นคำตอบถูกเพียงตัวเลือกเดียวเท่านั้น	4.40	0.516	มาก
7. คะแนนสอบเชื่อถือได้ไม่ว่าใครเป็นคนตรวจ	4.20	0.789	มาก
8. คำตอบของข้อสอบแต่ละข้อเชื่อถือได้ไม่ว่าใครจะตอบ เวลาใดและสถานที่ใด	4.00	0.000	มาก
9. เนื้อหาในแบบทดสอบแต่ละตอนไม่เอื้อประโยชน์หรือชี้แนะในการทำข้ออื่นๆ ในตอนอื่น ๆ	4.40	0.516	มาก
ความสมจริง (Authenticity)	4.40	0.000	มาก
10. ภาระงานด้านการฟังสัมพันธ์กับการฟังในชีวิตจริงในขนาดของผู้เข้าสอบ	4.40	0.516	มาก
11. บทฟังมีความสมจริงและสะท้อนสถานการณ์ในชีวิตจริง	4.40	0.516	มาก
การมีปฏิสัมพันธ์ (Interactiveness)	4.30	0.224	มาก
12. ภาระงานด้านการฟังมีหัวข้อที่คุ้นเคยสำหรับผู้เข้าสอบทุกคน	4.00	0.000	มาก
13. ภาระงานด้านการฟังเหมาะกับระดับความรู้ของผู้เข้าสอบ	4.30	0.483	มาก
ผลกระทบ (Impact)	4.40	0.432	มาก
14. ภาระงานฟังช่วยให้ผู้เข้าสอบมีประสบการณ์ในการทำแบบทดสอบการฟังในอนาคต	4.00	0.000	มาก
15. ภาระงานการฟังคล้ายกับกิจกรรมการฟังที่ผู้เข้าสอบได้ประสบในห้องเรียน	5.00	0.000	มากที่สุด

ตารางที่ 33 ผลการประเมินคุณประโยชน์ของแบบทดสอบด้านการฟัง (ต่อ)

คุณประโยชน์ของแบบทดสอบ (Test Usefulness)	\bar{X}	S.D.	แปลผล
16. จุดประสงค์ของแบบทดสอบด้านการฟังสอดคล้องกับเป้าหมายและความต้องการที่ระบุไว้ในหลักสูตร	4.40	0.516	มาก
17. ภาระงานการฟังสามารถสะท้อนความสามารถการฟังของผู้เข้าสอบตามที่สังคมและระบบการศึกษาไทยต้องการ	4.20	0.422	มาก
ความสะดวกใช้ (Practicality)	4.30	0.141	มาก
18. แบบทดสอบด้านการฟังนี้สามารถสร้างได้จริงด้วยทรัพยากรที่มีอยู่	4.40	0.516	มาก
19. แบบทดสอบด้านการฟังสามารถนำไปจัดสอบและตรวจให้คะแนนได้โดยไม่มียาก	4.20	0.422	มาก
ความเป็นธรรม (Fairness)	4.50	0.527	มากที่สุด
20. เนื้อหาสาระแบบทดสอบเหมาะสมและเป็นหัวเรื่องที่ปราศจากความลำเอียง โดยไม่มีการกล่าวถึงการแบ่งแยกเพศ เชื้อชาติ สถานทางเศรษฐกิจและสังคม และไม่นำไปสู่ความรุนแรงใด ๆ	4.50	0.527	มากที่สุด
โดยรวม	4.38	0.232	มาก

จากตารางที่ 33 แสดงว่า แบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัยด้านการฟังมีคุณประโยชน์ด้านต่าง ๆ แบ่งได้เป็น 2 ระดับ คือ ด้านที่มีคุณประโยชน์มากที่สุด และด้านที่มีคุณประโยชน์มาก โดยมีรายละเอียดดังนี้

ด้านที่มีคุณประโยชน์มากที่สุด มี 2 ด้าน คือ ความตรงเชิงโครงสร้าง (Construct Validity) ($\bar{X} = 4.52$, $S = 0.110$) และ ด้านความเป็นธรรม (Fairness) ($\bar{X} = 4.50$, $S.D. = 0.527$) โดยด้านความตรงเชิงโครงสร้างมีข้อคำถามย่อย 5 ข้อ แบ่งเป็นข้อคำถามที่มีคุณประโยชน์มากที่สุด 3 ข้อ ได้แก่ 1) ภาระงานด้านการฟังทุกตอนสอดคล้องกับวัตถุประสงค์การสอบ ($\bar{X} = 4.60$, $S.D. = 0.516$) 2) คะแนนสอบสามารถทำนายความสามารถด้านการฟังของผู้เข้าสอบ ($\bar{X} = 4.60$, $S.D. = 0.516$) และ 3) มีการระบุวัตถุประสงค์การสอบของแบบทดสอบการฟังแต่ละตอนไว้อย่างชัดเจน ($\bar{X} = 4.60$, $S.D. = 0.699$) และข้อคำถามที่มีคุณประโยชน์มาก 2 ข้อ ได้แก่ 1) วัตถุประสงค์ย่อยของแบบทดสอบแต่ละตอนสอดคล้องกับจุดประสงค์หลัก ($\bar{X} = 4.40$, $S.D. = 0.516$) และ 2) วิธีการให้คะแนนมีความเหมาะสมและสามารถสะท้อนความสามารถของผู้เข้าสอบได้ ($\bar{X} = 4.40$, $S.D. = 0.516$) และด้านความเป็นธรรม มีข้อคำถามย่อย 1 ข้อ และเป็นข้อคำถามที่มีคุณประโยชน์มาก คือ เนื้อหาสาระแบบทดสอบเหมาะสมและเป็นหัวเรื่องที่ปราศจากความลำเอียง โดยไม่มีการกล่าวถึงการแบ่งแยกเพศ เชื้อชาติ สถานทางเศรษฐกิจและสังคม และไม่นำไปสู่ความรุนแรงใด ๆ ($\bar{X} = 4.50$, $S.D. = 0.527$)

ด้านที่มีคุณประโยชน์มาก มี 5 ด้าน คือ 1) ด้านความสมจริง (Authenticity) ($\bar{X} = 4.40$, $S.D. = 0.000$) 2) ด้านผลกระทบ (Impact) ($\bar{X} = 4.40$, $S.D. = 0.432$) 3) ด้านความสะดวกใช้ (Practicality) ($\bar{X} = 4.30$, $S.D. = 0.141$) 4) ด้าน การมีปฏิสัมพันธ์ (Interactiveness) ($\bar{X} = 4.30$, $S.D. = 0.224$) และ 5) ด้านความเชื่อมั่น (Reliability) ($\bar{X} = 4.25$, $S.D. = 0.191$) โดยแต่ละด้านมีคุณประโยชน์ของข้อคำถามย่อยที่เกี่ยวข้องดังรายละเอียดต่อไปนี้

ด้านความสมจริง (Authenticity) ($\bar{X} = 4.40$, S.D. = 0.000) มีข้อคำถามย่อย 2 ข้อ และมีคุณประโยชน์มากที่สุด 2 ข้อ คือ 1) ภาระงานด้านการฟังสัมพันธ์กับการฟังในชีวิตจริงในอนาคตของผู้เข้าสอบ ($\bar{X} = 4.40$, S.D. = 0.516) และ 2) บทฟังมีความสมจริงและสะท้อนสถานการณ์ในชีวิตจริง ($\bar{X} = 4.40$, S.D. = 0.516)

ด้านผลกระทบ (Impact) ($\bar{X} = 4.40$, S.D. = 0.432) มีข้อคำถามย่อย 4 ข้อ แบ่งเป็นข้อคำถามที่มีคุณประโยชน์มากที่สุด 1 ข้อ คือ 1) ภาระงานการฟังคล้ายกับกิจกรรมการฟังที่ผู้เข้าสอบได้ประสบในห้องเรียน ($\bar{X} = 5.00$, S.D. = 0.000) และข้อคำถามที่มีคุณประโยชน์มากที่สุด 3 ข้อ คือ 1) จุดประสงค์ของแบบทดสอบด้านการฟังสอดคล้องกับเป้าหมายและความต้องการที่ระบุไว้ในหลักสูตร ($\bar{X} = 4.40$, S.D. = 0.516) 2) ภาระงานการฟังสามารถสะท้อนความสามารถการฟังของผู้เข้าสอบตามที่สังคมและระบบการศึกษาไทยต้องการ ($\bar{X} = 4.20$, S.D. = 0.422) และ 3) ภาระงานฟังช่วยให้ผู้เข้าสอบมีประสบการณ์ในการทำแบบทดสอบการฟังในอนาคต ($\bar{X} = 4.00$, S.D. = 0.000)

ด้านความสะดวกใช้ (Practicality) ($\bar{X} = 4.30$, S.D. = 0.141) มีข้อคำถามย่อย 2 ข้อ และมีคุณประโยชน์มากที่สุด 2 ข้อ คือ 1) แบบทดสอบด้านการฟังนี้สามารถสร้างได้จริงด้วยทรัพยากรที่มีอยู่ ($\bar{X} = 4.40$, S.D. = 0.516) และ 2) แบบทดสอบด้านการฟังสามารถนำไปจัดสอบและตรวจให้คะแนนได้โดยไม่ยาก ($\bar{X} = 4.20$, S.D. = 0.422)

ด้าน การมีปฏิสัมพันธ์ (Interactiveness) ($\bar{X} = 4.30$, S.D. = 0.224) มีข้อคำถามย่อย 2 ข้อ และมีคุณประโยชน์มากที่สุด 2 ข้อ คือ 1) ภาระงานด้านการฟังเหมาะกับระดับความรู้ของผู้เข้าสอบ ($\bar{X} = 4.30$, S.D. = 0.483) และ 2) ภาระงานด้านการฟังเหมาะกับระดับความรู้ของผู้เข้าสอบ ($\bar{X} = 4.00$, S.D. = 0.000)

ด้านความเชื่อมั่น (Reliability) ($\bar{X} = 4.25$, S.D. = 0.191) มีข้อคำถามย่อย 4 ข้อ และมีคุณประโยชน์มากที่สุด 4 ข้อ คือ 1) ข้อสอบแต่ละข้อมีตัวเลือกที่เป็นคำตอบถูกเพียงตัวเลือกเดียวเท่านั้น ($\bar{X} = 4.40$, S.D. = 0.516) 2) เนื้อหาในแบบทดสอบแต่ละตอนไม่เอื้อประโยชน์หรือชี้แนะในการทำข้ออื่น ๆ ในตอนอื่น ๆ ($\bar{X} = 4.40$, S.D. = 0.516) 3) คะแนนสอบเชื่อถือได้ไม่ว่าใครเป็นคนตรวจ ($\bar{X} = 4.20$, S.D. = 0.789) และ 4) คำตอบของข้อสอบแต่ละข้อเชื่อถือได้ไม่ว่าใครจะตอบเวลาใดและสถานที่ใด ($\bar{X} = 4.00$, S.D. = 0.000)

4.1.1.2 ข้อเสนอแนะสำหรับแบบทดสอบด้านการฟัง

ไม่มีข้อเสนอแนะ

สรุปได้ว่า ในภาพรวม แบบทดสอบด้านการฟัง มีคุณประโยชน์อยู่ในระดับ มาก ($\bar{X} = 4.38$, S.D. = 0.232)

4.1.2 ทักษะการใช้ภาษา

แบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test) ด้านการใช้ภาษา มีผลการประเมินคุณประโยชน์และข้อเสนอแนะที่เป็นประโยชน์ ดังนี้

4.1.2.1 ผลการประเมินคุณประโยชน์ของแบบทดสอบด้านการใช้ภาษา

ในการประเมินคุณประโยชน์ของแบบทดสอบด้านการใช้ภาษาโดยใช้แบบประเมินคุณประโยชน์แบบทดสอบก่อนนำไปใช้ ที่มีข้อคำถามแบบมาตราส่วนประมาณค่า จำนวน 20 ข้อ และสอบถามเกี่ยวกับคุณประโยชน์แบบทดสอบ 7 ด้าน คือ 1) ความตรงเชิงโครงสร้าง (Construct Validity) 2) ความเชื่อมั่น (Reliability) 3) ความสมจริง (Authenticity) 4) การมีปฏิสัมพันธ์ (Interactiveness) 5) ผลกระทบ (Impact) 6) ความสะดวกใช้ (Practicality) และ 7) ความเป็นธรรม (Fairness) ได้ผลการประเมินความตรงในแต่ละด้าน ดังรายละเอียดในตารางที่ 34

ตารางที่ 32 ผลการประเมินคุณประโยชน์ของแบบทดสอบด้านการใช้ภาษา

คุณประโยชน์ของแบบทดสอบ (Test Usefulness)	\bar{X}	S.D.	แปลผล
ความตรงเชิงโครงสร้าง (Construct Validity)	4.20	0.100	มาก
1. มีการระบุวัตถุประสงค์การสอบของแบบทดสอบการใช้ภาษาแต่ละตอนไว้อย่างชัดเจน	4.30	0.675	มาก
2. วัตถุประสงค์ย่อยของแบบทดสอบแต่ละตอนสอดคล้องกับจุดประสงค์หลัก	4.20	0.789	มาก
3. ภาระงานสอบด้านการใช้ภาษาทุกตอนสอดคล้องกับวัตถุประสงค์การสอบ	4.10	0.738	มาก
4. วิธีการให้คะแนนมีความเหมาะสมและสามารถสะท้อนความสามารถด้านการใช้ภาษาของผู้เข้าสอบ	4.10	0.738	มาก
5. คะแนนสอบสามารถทำนายความสามารถด้านการใช้ภาษาของผู้เข้าสอบ	4.30	0.675	มาก
ความเชื่อมั่น (Reliability)	4.18	0.096	มาก
6. ข้อสอบแต่ละข้อมีตัวเลือกที่เป็นคำตอบถูกเพียงตัวเลือกเดียวเท่านั้น	4.20	0.789	มาก
7. คะแนนสอบเชื่อถือได้ไม่ว่าใครเป็นคนตรวจ	4.10	0.738	มาก
8. คำตอบของข้อสอบแต่ละข้อเชื่อถือได้ไม่ว่าใครจะตอบ เวลาใดและสถานที่ใด	4.10	0.738	มาก
9. เนื้อหาในแบบทดสอบแต่ละตอนไม่เอื้อประโยชน์หรือชี้แนะในการทำข้ออื่นๆ ในตอนอื่น ๆ	4.30	0.675	มาก
ความสมจริง (Authenticity)	4.20	0.141	มาก
10. ภาระงานด้านการใช้ภาษาสัมพันธ์กับการใช้ภาษาในชีวิตจริงในอนาคตของผู้เข้าสอบ	4.10	0.738	มาก
11. ตัวข้อความในประโยคมีความสมจริงและสะท้อนสถานการณ์ในชีวิตจริง	4.30	0.675	มาก

ตารางที่ 34 ผลการประเมินคุณประโยชน์ของแบบทดสอบด้านการใช้ภาษา (ต่อ)

คุณประโยชน์ของแบบทดสอบ (Test Usefulness)	\bar{X}	S.D.	แปลผล
การมีปฏิสัมพันธ์ (Interactiveness)	4.16	0.089	มาก
12. ภาระงานด้านการใช้ภาษามีหัวเรื่องที่คุ้นเคยสำหรับผู้เข้าสอบทุกคน	4.20	0.789	มาก
13. ภาระงานด้านการใช้ภาษาเหมาะกับระดับความรู้ของผู้เข้าสอบ	4.10	0.738	มาก
ผลกระทบ (Impact)	4.25	0.100	มาก
14. ภาระงานการใช้ภาษาช่วยให้ผู้เข้าสอบมีประสบการณ์ในการทำแบบทดสอบการใช้ภาษาในอนาคต	4.30	0.675	มาก
15. ภาระงานการใช้ภาษาคลายกับกิจกรรมการใช้ภาษาที่ผู้เข้าสอบได้ประสบในห้องเรียน	4.30	0.675	มาก
16. จุดประสงค์ของแบบทดสอบด้านการใช้ภาษาสอดคล้องกับเป้าหมายและความต้องการที่ระบุไว้ในหลักสูตร	4.10	0.738	มาก
17. ภาระงานการใช้ภาษาสามารถสะท้อนความสามารถการใช้ภาษาของผู้เข้าสอบตามที่สังคมและระบบการศึกษาไทยต้องการ	4.30	0.675	มาก
ความสะดวกใช้ (Practicality)	4.15	0.071	มาก
18. แบบทดสอบด้านการฟังนี้สามารถสร้างได้จริงด้วยทรัพยากรที่มีอยู่	4.20	0.789	มาก
19. แบบทดสอบด้านการฟังสามารถนำไปจัดสอบและตรวจให้คะแนนได้โดยไม่มียาก	4.10	0.738	มาก
ความเป็นธรรม (Fairness)	4.70	0.483	มากที่สุด
20. เนื้อหาสาระแบบทดสอบเหมาะสมและเป็นหัวเรื่องที่ปราศจากความลำเอียง โดยไม่มีการกล่าวถึงการแบ่งแยกเพศ เชื้อชาติ สถานทางเศรษฐกิจและสังคม และไม่นำไปสู่ความรุนแรงใด ๆ	4.70	0.483	มากที่สุด
โดยรวม	4.26	0.154	มาก

จากตารางที่ 34 แสดงว่า แบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัยด้านการใช้ภาษามีคุณประโยชน์ด้านต่าง ๆ ที่เหมาะสม แบ่งได้เป็น 2 ระดับ คือ ด้านที่มีคุณประโยชน์มากที่สุด และด้านที่มีคุณประโยชน์มาก โดยมีรายละเอียดดังนี้

ด้านที่มีคุณประโยชน์มากที่สุด มี 1 ด้าน คือ ความเป็นธรรม (Fairness) ($\bar{X} = 4.70$, S.D. = 0.483) โดยมีข้อคำถามย่อย 1 ข้อ เนื้อหาสาระแบบทดสอบเหมาะสมและเป็นหัวเรื่องที่ปราศจากความลำเอียง โดยไม่มีการกล่าวถึงการแบ่งแยกเพศ เชื้อชาติ สถานทางเศรษฐกิจและสังคม และไม่นำไปสู่ความรุนแรงใด ๆ ($\bar{X} = 4.70$, S.D. = 0.483)

ด้านที่มีคุณประโยชน์มาก มี 6 ด้าน คือ 1) ด้านผลกระทบ (Impact) ($\bar{X} = 4.25$, S.D. = 0.100) 2) ด้านความตรงเชิงโครงสร้าง (Construct Validity) ($\bar{X} = 4.20$, S.D. = 0.100) 3) ด้านความสมจริง (Authenticity) ($\bar{X} = 4.20$, S.D. = 0.141) 4) ด้านความเชื่อมั่น (Reliability) ($\bar{X} = 4.18$, S.D. = 0.096) 5) ด้านการมีปฏิสัมพันธ์ (Interactiveness) ($\bar{X} = 4.16$, S.D. = 0.089) และ 6) ด้านความสะดวกใช้ (Practicality) ($\bar{X} = 4.15$, S.D. = 0.071) โดยแต่ละด้านมีคุณประโยชน์ของข้อคำถามย่อยที่เกี่ยวข้องดังรายละเอียดต่อไปนี้

ด้านผลกระทบ (Impact) ($\bar{X} = 4.25$, S.D. = 0.100) มีข้อคำถามย่อย 4 ข้อ เป็นข้อคำถามที่มีคุณประโยชน์มากที่สุดทั้ง 4 ข้อ คือ 1) ภาระงานการใช้ภาษาช่วยให้ผู้เข้าสอบมีประสบการณ์ในการทำแบบทดสอบการใช้ภาษาในอนาคต ($\bar{X} = 4.30$, S.D. = 0.675) 2) ภาระงานการใช้ภาษาคล้ายกับกิจกรรมการใช้ภาษาที่ผู้เข้าสอบได้ประสบในห้องเรียน ($\bar{X} = 4.30$, S.D. = 0.675) 3) ภาระงานการใช้ภาษาสามารถสะท้อนความสามารถการใช้ภาษาของผู้เข้าสอบตามที่สังคมและระบบการศึกษาไทยต้องการ ($\bar{X} = 4.30$, S.D. = 0.675) และ 4) จุดประสงค์ของแบบทดสอบด้านการใช้ภาษาสอดคล้องกับเป้าหมายและความต้องการที่ระบุไว้ในหลักสูตร ($\bar{X} = 4.10$, S.D. = 0.738)

ด้านความตรงเชิงโครงสร้าง (Construct Validity) ($\bar{X} = 4.20$, S.D. = 0.100) มีข้อคำถามย่อย 5 ข้อ และเป็นข้อคำถามที่มีคุณประโยชน์มากที่สุดทั้ง 5 ข้อ คือ 1) มีการระบุวัตถุประสงค์การสอบของแบบทดสอบการใช้ภาษาแต่ละตอนไว้อย่างชัดเจน ($\bar{X} = 4.30$, S.D. = 0.675) 2) คะแนนสอบสามารถทำนายความสามารถด้านการใช้ภาษาของผู้เข้าสอบ ($\bar{X} = 4.30$, S.D. = 0.675) 3) วัตถุประสงค์ย่อยของแบบทดสอบแต่ละตอนสอดคล้องกับจุดประสงค์หลัก ($\bar{X} = 4.20$, S.D. = 0.789) 4) ภาระงานสอบด้านการใช้ภาษาทุกตอนสอดคล้องกับวัตถุประสงค์การสอบ ($\bar{X} = 4.10$, S.D. = 0.738) และ 5) วิธีการให้คะแนนมีความเหมาะสมและสามารถสะท้อนความสามารถด้านการใช้ภาษาของผู้เข้าสอบ ($\bar{X} = 4.10$, S.D. = 0.738)

ด้านความสมจริง (Authenticity) ($\bar{X} = 4.20$, S.D. = 0.141) มีข้อคำถามย่อย 2 ข้อ และเป็นข้อคำถามที่มีคุณประโยชน์มากที่สุดทั้ง 2 ข้อ คือ 1) ตัวข้อความในประโยคมีความสมจริงและสะท้อนสถานการณ์ในชีวิตจริง ($\bar{X} = 4.30$, S.D. = 0.675) และ 2) ภาระงานด้านการใช้ภาษาสัมพันธ์กับการใช้ภาษาในชีวิตจริงในอนาคตของผู้เข้าสอบ ($\bar{X} = 4.10$, S.D. = 0.738)

ด้านความเชื่อมั่น (Reliability) ($\bar{X} = 4.18$, S.D. = 0.096) มีข้อคำถามย่อย 4 ข้อ และเป็นข้อคำถามที่มีคุณประโยชน์มากที่สุดทั้ง 4 ข้อ คือ 1) เนื้อหาในแบบทดสอบแต่ละตอนไม่เอื้อประโยชน์หรือชี้แนะในการทำข้ออื่นๆ ในตอนอื่นๆ ($\bar{X} = 4.30$, S.D. = 0.675) 2) ข้อสอบแต่ละข้อมีตัวเลือกที่เป็นคำตอบถูกเพียงตัวเลือกเดียวเท่านั้น ($\bar{X} = 4.20$, S.D. = 0.789) 3) คะแนนสอบเชื่อถือได้ไม่ว่าใครเป็นคนตรวจ ($\bar{X} = 4.10$, S.D. = 0.738) และ 4) คำตอบของข้อสอบแต่ละข้อเชื่อถือได้ไม่ว่าใครจะตอบ เวลาใดและสถานที่ใด ($\bar{X} = 4.10$, S.D. = 0.738)

ด้านการมีปฏิสัมพันธ์ (Interactiveness) ($\bar{X} = 4.16$, S.D. = 0.089) มีข้อคำถามย่อย 2 ข้อ และเป็นข้อคำถามที่มีคุณประโยชน์มากที่สุดทั้ง 2 ข้อ คือ 1) ภาระงานด้านการใช้ภาษามีหัวเรื่องที่คุ้นเคยสำหรับผู้เข้าสอบทุกคน ($\bar{X} = 4.20$, S.D. = 0.789) และ 2) ภาระงานด้านการใช้ภาษาเหมาะสมกับระดับความรู้ของผู้เข้าสอบ ($\bar{X} = 4.10$, S.D. = 0.738)

ด้านความสะดวกใช้ (Practicality) ($\bar{X} = 4.15$, S.D. = 0.071) มีข้อคำถามย่อย 2 ข้อ และเป็นข้อคำถามที่มีคุณประโยชน์มากที่สุดทั้ง 2 ข้อ คือ 1) แบบทดสอบด้านการฟังนี้สามารถสร้างได้จริงด้วยทรัพยากรที่มีอยู่ ($\bar{X} = 4.20$, S.D. = 0.789) และ 2) แบบทดสอบด้านการฟังสามารถนำไปจัดสอบและตรวจให้คะแนนได้โดยง่าย ($\bar{X} = 4.10$, S.D. = 0.738)

4.1.2.2 ข้อเสนอแนะสำหรับแบบทดสอบด้านการใช้ภาษา

ไม่มีข้อเสนอแนะ

สรุปได้ว่า ในภาพรวม แบบทดสอบด้านการใช้ภาษามีคุณประโยชน์อยู่ในระดับ มาก ($\bar{X} = 4.26$, S.D. = 0.154)

4.1.3 ทักษะการอ่าน

แบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test) ด้านการอ่าน มีผลการประเมินคุณประโยชน์และข้อเสนอแนะที่เป็นประโยชน์ดังนี้

4.1.3.1 ผลการประเมินคุณประโยชน์ของแบบทดสอบด้านการอ่าน

ในการประเมินความตรงของแบบทดสอบด้านการอ่านโดยใช้แบบประเมินคุณประโยชน์แบบทดสอบก่อนนำไปใช้ ที่มีข้อคำถามแบบมาตราส่วนประมาณค่า จำนวน 20 ข้อ และสอบถามเกี่ยวกับคุณประโยชน์แบบทดสอบ 7 ด้าน คือ 1) ความตรงเชิงโครงสร้าง (Construct Validity) 2) ความเชื่อมั่น (Reliability) 3) ความสมจริง (Authenticity) 4) การมีปฏิสัมพันธ์ (Interactiveness) 5) ผลกระทบ (Impact) 6) ความสะดวกใช้ (Practicality) และ 7) ความเป็นธรรม (Fairness) ได้ผลการประเมินความตรงในแต่ละด้าน ดังรายละเอียดในตารางที่ 35

ตารางที่ 33 ผลการประเมินคุณประโยชน์ของแบบทดสอบด้านการอ่าน

คุณประโยชน์ของแบบทดสอบ (Test Usefulness)	\bar{X}	S.D.	แปลผล
ความตรงเชิงโครงสร้าง (Construct Validity)	4.14	0.207	มาก
1. มีภาระบัญญัติประสงค์การสอบของแบบทดสอบการอ่านแต่ละตอนไว้อย่างชัดเจน	4.30	0.675	มาก
2. วัตถุประสงค์ย่อยของแบบทดสอบแต่ละตอนสอดคล้องกับจุดประสงค์หลัก	4.10	0.568	มาก
3. ภาระงานสอบด้านการอ่านทุกตอนสอดคล้องกับวัตถุประสงค์การสอบ	4.20	0.789	มาก
4. วิธีการให้คะแนนมีความเหมาะสมและสามารถสะท้อนความสามารถด้านการอ่านของผู้เข้าสอบ	4.30	0.675	มาก
5. คะแนนสอบสามารถทำนายความสามารถด้านการอ่านของผู้เข้าสอบ	3.80	0.789	มาก
ความเชื่อมั่น (Reliability)	4.10	0.216	มาก
6. ข้อสอบแต่ละข้อมีตัวเลือกที่เป็นคำตอบถูกเพียงตัวเลือกเดียวเท่านั้น	4.00	0.667	มาก
7. คะแนนสอบเชื่อถือได้ไม่ว่าใครเป็นคนตรวจ	4.40	0.516	มาก
8. คำตอบของข้อสอบแต่ละข้อเชื่อถือได้ไม่ว่าใครจะตอบ เวลาใดและสถานที่ใด	4.10	0.876	มาก

ตารางที่ 35 ผลการประเมินคุณประโยชน์ของแบบทดสอบด้านการอ่าน (ต่อ)

คุณประโยชน์ของแบบทดสอบ (Test Usefulness)	\bar{X}	S.D.	แปลผล
9. เนื้อหาในแบบทดสอบแต่ละตอนไม่เอื้อประโยชน์หรือซ้ำในการทำข้ออื่น ๆ ในตอนอื่น ๆ	3.90	0.876	มาก
ความสมจริง (Authenticity)	4.10	0.141	มาก
10. ภาระงานด้านการอ่านสัมพันธ์กับการอ่านในชีวิตจริงในอนาคตของผู้เข้าสอบ	4.00	0.471	มาก
11. ตัวบทอ่านมีความสมจริงและสะท้อนสถานการณ์ในชีวิตจริง	4.20	0.789	มาก
การมีปฏิสัมพันธ์ (Interactiveness)	4.14	0.055	มาก
12. ภาระงานด้านการอ่านมีหัวข้อที่คุ้นเคยสำหรับผู้เข้าสอบทุกคน	4.20	0.632	มาก
13. ภาระงานด้านการอ่านเหมาะกับระดับความรู้ของผู้เข้าสอบ	4.20	0.789	มาก
ผลกระทบ (Impact)	4.23	0.096	มาก
14. ภาระงานการอ่านช่วยให้ผู้เข้าสอบมีประสบการณ์ในการทำแบบทดสอบการอ่านในอนาคต	4.30	0.675	มาก
15. ภาระงานการอ่านคล้ายกับกิจกรรมการอ่านที่ผู้เข้าสอบได้ประสบในห้องเรียน	4.30	0.675	มาก
16. จุดประสงค์ของแบบทดสอบด้านการอ่านสอดคล้องกับเป้าหมายและความต้องการที่ระบุไว้ในหลักสูตร	4.20	0.789	มาก
17. ภาระงานการอ่านสามารถสะท้อนความสามารถการอ่านของผู้เข้าสอบตามที่สังคมและระบบการศึกษาไทยต้องการ	4.10	0.568	มาก
ความสะดวกใช้ (Practicality)	3.95	0.212	มาก
18. แบบทดสอบด้านการอ่านนี้สามารถสร้างได้จริงด้วยทรัพยากรที่มีอยู่	3.80	0.789	มาก
19. แบบทดสอบด้านการอ่านสามารถนำไปจัดสอบและตรวจให้คะแนนได้โดยไม่ยาก	4.10	0.738	มาก
ความเป็นธรรม (Fairness)	4.70	0.483	มากที่สุด
20. เนื้อหาสาระแบบทดสอบเหมาะสมและเป็นหัวเรื่องที่ปราศจากความลำเอียง โดยไม่มีการกล่าวถึงการแบ่งแยกเพศ เชื้อชาติ สถานทางเศรษฐกิจและสังคม และไม่นำไปสู่ความรุนแรงใด ๆ	4.70	0.483	มากที่สุด
โดยรวม	4.19	0.200	มาก

จากตารางที่ 35 แสดงว่า แบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัยด้านการอ่านมีคุณประโยชน์ด้านต่าง ๆ ที่เหมาะสม แบ่งได้เป็น 2 ระดับ คือ ด้านที่มีคุณประโยชน์มากที่สุด และด้านที่มีคุณประโยชน์มาก โดยมีรายละเอียดดังนี้

ด้านที่มีคุณประโยชน์มากที่สุด มี 1 ด้าน คือ ความเป็นธรรม (Fairness) ($\bar{X} = 4.70$, S.D. = 0.483) โดยมีข้อคำถามย่อย 1 ข้อ เนื้อหาสาระแบบทดสอบเหมาะสมและเป็นหัวเรื่องที่ปราศจากความลำเอียง โดยไม่มีการกล่าวถึงการแบ่งแยกเพศ เชื้อชาติ สถานทางเศรษฐกิจและสังคม และไม่นำไปสู่ความรุนแรงใด ๆ ($\bar{X} = 4.70$, S.D. = 0.483)

ด้านที่มีคุณประโยชน์มาก มี 6 ด้าน คือ 1) ด้านผลกระทบ (Impact) ($\bar{X} = 4.23$, S.D. = 0.096) 2) ด้านการมีปฏิสัมพันธ์ (Interactiveness) ($\bar{X} = 4.14$, S.D. = 0.055) 3) ด้านความตรงเชิงโครงสร้าง (Construct Validity) ($\bar{X} = 4.14$, S.D. = 0.207) 4) ด้านความสมจริง (Authenticity) ($\bar{X} = 4.10$, S.D. = 0.141) 5) ด้านความเชื่อมั่น (Reliability) ($\bar{X} = 4.10$, S.D. = 0.216) และ 6) ด้านความสะดวกใช้ (Practicality) ($\bar{X} = 3.95$, S.D. = 0.212) โดยแต่ละด้านมีคุณประโยชน์ของข้อคำถามย่อยที่เกี่ยวข้องดังรายละเอียดต่อไปนี้

ด้านผลกระทบ (Impact) ($\bar{X} = 4.23$, S.D. = 0.096) มีข้อคำถามย่อย 4 ข้อ เป็นข้อคำถามที่มีคุณประโยชน์มากที่สุดทั้ง 4 ข้อ คือ 1) ภาระงานการอ่านช่วยให้ผู้เข้าสอบมีประสบการณ์ในการทำแบบทดสอบการอ่านในอนาคต ($\bar{X} = 4.30$, S.D. = 0.675) 2) ภาระงานการอ่านคล้ายกับกิจกรรมการอ่านที่ผู้เข้าสอบได้ประสบในห้องเรียน ($\bar{X} = 4.30$, S.D. = 0.675) 3) จุดประสงค์ของแบบทดสอบด้านการอ่านสอดคล้องกับเป้าหมายและความต้องการที่ระบุไว้ในหลักสูตร ($\bar{X} = 4.20$, S.D. = 0.789) และ 4) ภาระงานการอ่านสามารถสะท้อนความสามารถการอ่านของผู้เข้าสอบตามที่สังคมและระบบการศึกษาไทยต้องการ ($\bar{X} = 4.10$, S.D. = 0.568)

ด้านการมีปฏิสัมพันธ์ (Interactiveness) ($\bar{X} = 4.14$, S.D. = 0.055) มีข้อคำถามย่อย 2 ข้อ และเป็นข้อคำถามที่มีคุณประโยชน์มากที่สุดทั้ง 2 ข้อ คือ 1) ภาระงานด้านการอ่านมีหัวเรื่องที่คุ้นเคยสำหรับผู้เข้าสอบทุกคน ($\bar{X} = 4.20$, S.D. = 0.632) และ 2) ภาระงานด้านการอ่านเหมาะสมกับระดับความรู้ของผู้เข้าสอบ ($\bar{X} = 4.20$, S.D. = 0.789)

ด้านความตรงเชิงโครงสร้าง (Construct Validity) ($\bar{X} = 4.14$, S.D. = 0.207) มีข้อคำถามย่อย 5 ข้อ และเป็นข้อคำถามที่มีคุณประโยชน์มากที่สุดทั้ง 5 ข้อ คือ 1) มีการระบุวัตถุประสงค์การสอบของแบบทดสอบการอ่านแต่ละตอนไว้อย่างชัดเจน ($\bar{X} = 4.30$, S.D. = 0.675) 2) วิธีการให้คะแนนมีความเหมาะสมและสามารถสะท้อนความสามารถด้านการอ่านของผู้เข้าสอบ ($\bar{X} = 4.30$, S.D. = 0.675) 3) ภาระงานสอบด้านการอ่านทุกตอนสอดคล้องกับวัตถุประสงค์การสอบ ($\bar{X} = 4.20$, S.D. = 0.789) 4) วัตถุประสงค์ย่อยของแบบทดสอบแต่ละตอนสอดคล้องกับจุดประสงค์หลัก ($\bar{X} = 4.10$, S.D. = 0.568) และ 5) คะแนนสอบสามารถทำนายความสามารถด้านการอ่านของผู้เข้าสอบ ($\bar{X} = 3.80$, S.D. = 0.789)

ด้านความสมจริง (Authenticity) ($\bar{X} = 4.10$, S.D. = 0.141) มีข้อคำถามย่อย 2 ข้อ และเป็นข้อคำถามที่มีคุณประโยชน์มากที่สุดทั้ง 2 ข้อ คือ 1) ตัวบทอ่านมีความสมจริงและสะท้อนสถานการณ์ในชีวิตจริง ($\bar{X} = 4.20$, S.D. = 0.789) และ 2) ภาระงานด้านการอ่านสัมพันธ์กับการอ่านในชีวิตจริงในอนาคตของผู้เข้าสอบ ($\bar{X} = 4.00$, S.D. = 0.471)

ด้านความเชื่อมั่น (Reliability) ($\bar{X} = 4.10$, S.D. = 0.216) มีข้อคำถามย่อย 4 ข้อ และเป็นข้อคำถามที่มีคุณประโยชน์มากที่สุดทั้ง 4 ข้อ คือ 1) คะแนนสอบเชื่อถือได้ไม่ว่าใครเป็นคนตรวจ ($\bar{X} = 4.40$, S.D. = 0.516) 2) คำตอบของข้อสอบแต่ละข้อเชื่อถือได้ไม่ว่าใครจะตอบ เวลาใดและสถานที่ใด ($\bar{X} = 4.10$, S.D. = 0.876) 3) ข้อสอบแต่ละข้อมีตัวเลือกที่เป็นคำตอบถูกเพียงตัวเลือกเดียวเท่านั้น ($\bar{X} = 4.00$, S.D. = 0.667) และ 4) เนื้อหาในแบบทดสอบแต่ละตอนไม่เอื้อประโยชน์หรือชี้แนะในการทำข้ออื่นๆ ในตอนอื่น ๆ ($\bar{X} = 3.90$, S.D. = 0.876)

ด้านความสะดวกใช้ (Practicality) ($\bar{X} = 3.95$, S.D. = 0.212) มีข้อคำถามย่อย 2 ข้อ และเป็นข้อคำถามที่มีคุณประโยชน์มากทั้ง 2 ข้อ คือ 1) แบบทดสอบด้านการอ่านสามารถนำไปจัดสอบและตรวจให้คะแนนได้โดยไม่ยาก ($\bar{X} = 4.10$, S.D. = 0.738) และ 2) แบบทดสอบด้านการอ่านนี้สามารถสร้างได้จริงด้วยทรัพยากรที่มีอยู่ ($\bar{X} = 3.80$, S.D. = 0.789)

4.1.3.2 ข้อเสนอแนะสำหรับแบบทดสอบด้านการอ่าน

ไม่มีข้อเสนอแนะ

สรุปได้ว่า ในภาพรวม แบบทดสอบด้านการอ่านมีคุณประโยชน์อยู่ในระดับมาก ($\bar{X} = 4.19$, S.D. = 0.200)

4.1.4 ทักษะการเขียน

แบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test) ด้านการเขียน มีผลการประเมินคุณประโยชน์และข้อเสนอแนะที่เป็นประโยชน์ ดังนี้

4.1.4.1 ผลการประเมินคุณประโยชน์ของแบบทดสอบด้านการเขียน

ในการประเมินคุณประโยชน์ของแบบทดสอบด้านการเขียนโดยใช้แบบประเมินคุณประโยชน์แบบทดสอบก่อนนำไปใช้ ที่มีข้อคำถามแบบมาตราส่วนประมาณค่า จำนวน 20 ข้อ และสอบถามเกี่ยวกับคุณประโยชน์แบบทดสอบ 7 ด้าน คือ 1) ความตรงเชิงโครงสร้าง (Construct Validity) 2) ความเชื่อมั่น (Reliability) 3) ความสมจริง (Authenticity) 4) การมีปฏิสัมพันธ์ (Interactiveness) 5) ผลกระทบ (Impact) 6) ความสะดวกใช้ (Practicality) และ 7) ความเป็นธรรม (Fairness) ได้ผลการประเมินความตรงในแต่ละด้าน ดังรายละเอียดในตารางที่ 36

ตารางที่ 34 ผลการประเมินคุณประโยชน์ของแบบทดสอบด้านการเขียน

คุณประโยชน์ของแบบทดสอบ (Test Usefulness)	\bar{X}	S.D.	แปลผล
ความตรงเชิงโครงสร้าง (Construct Validity)	4.20	0.100	มาก
1. มีการระบุวัตถุประสงค์การสอบของแบบทดสอบการเขียนแต่ละตอนไว้อย่างชัดเจน	4.10	0.738	มาก
2. วัตถุประสงค์ย่อยของแบบทดสอบแต่ละตอนสอดคล้องกับจุดประสงค์หลัก	4.30	0.675	มาก
3. ภาระงานสอบด้านการเขียนทุกตอนสอดคล้องกับวัตถุประสงค์การสอบ	4.10	0.738	มาก
4. วิธีการให้คะแนนมีความเหมาะสมและสามารถสะท้อนความสามารถด้านการเขียนของผู้เข้าสอบ	4.30	0.675	มาก
5. คะแนนสอบสามารถทำนายความสามารถด้านการเขียนของผู้เข้าสอบ	4.20	0.789	มาก

ตารางที่ 36 ผลการประเมินคุณประโยชน์ของแบบทดสอบด้านการเขียน (ต่อ)

คุณประโยชน์ของแบบทดสอบ (Test Usefulness)	\bar{X}	S.D.	แปลผล
ความเชื่อมั่น (Reliability)	4.20	0.115	มาก
6. เกณฑ์การให้คะแนนมีความชัดเจนเพียงพอที่จะทำให้เกิดความเป็นธรรมในการให้คะแนนงานเขียนของผู้เข้าสอบทุกคน	4.10	0.738	มาก
7. คะแนนสอบเชื่อถือได้ไม่ว่าใครเป็นคนตรวจ	4.30	0.675	มาก
8. คำตอบของข้อสอบแต่ละข้อเชื่อถือได้ไม่ว่าใครจะตอบ เวลาใดและสถานที่ใด	4.30	0.675	มาก
9. เนื้อหาในแบบทดสอบแต่ละตอนไม่เอื้อประโยชน์หรือชี้แนะในการทำข้ออื่นๆ ในตอนอื่น ๆ	4.10	0.738	มาก
ความสมจริง (Authenticity)	4.15	0.071	มาก
10. ภาระงานด้านการเขียนสัมพันธ์กับการเขียนในชีวิตจริงในขนาดของผู้เข้าสอบ	4.10	0.738	มาก
11. ภาระงานด้านการเขียนมีความสมจริงและสะท้อนสถานการณ์ในชีวิตจริง	4.20	0.789	มาก
การมีปฏิสัมพันธ์ (Interactiveness)	4.16	0.089	มาก
12. ภาระงานด้านการเขียนมีหัวข้อที่คุ้นเคยสำหรับผู้เข้าสอบทุกคน	4.10	0.738	มาก
13. ภาระงานด้านการเขียนเหมาะสมกับระดับความรู้ของผู้เข้าสอบ	4.10	0.738	มาก
ผลกระทบ (Impact)	4.15	0.100	มาก
14. ภาระงานการเขียนช่วยให้ผู้เข้าสอบมีประสบการณ์ในการทำแบบทดสอบการเขียนในอนาคต	4.10	0.738	มาก
15. ภาระงานการเขียนคล้ายกับกิจกรรมการเขียนที่ผู้เข้าสอบได้ประสบในห้องเรียน	4.10	0.738	มาก
16. จุดประสงค์ของแบบทดสอบด้านการเขียนสอดคล้องกับเป้าหมายและความต้องการที่ระบุไว้ในหลักสูตร	4.10	0.738	มาก
17. ภาระงานการเขียนสามารถสะท้อนความสามารถการเขียนของผู้เข้าสอบตามที่สังคมและระบบการศึกษาไทยต้องการ	4.30	0.675	มาก
ความสะดวกใช้ (Practicality)	4.15	0.071	มาก
18. แบบทดสอบด้านการเขียนนี้สามารถสร้างได้จริงด้วยทรัพยากรที่มีอยู่	4.20	0.789	มาก
19. แบบทดสอบด้านการเขียนสามารถนำไปจัดสอบและตรวจให้คะแนนได้โดยไม่ยาก	4.10	0.738	มาก
ความเป็นธรรม (Fairness)	4.70	0.483	มากที่สุด
20. เนื้อหาสาระแบบทดสอบเหมาะสมและเป็นหัวเรื่องที่ปราศจากความลำเอียง โดยไม่มีการกล่าวถึงการแบ่งแยกเพศ เชื้อชาติ สถานทางเศรษฐกิจและสังคม และไม่นำไปสู่ความรุนแรงใด ๆ	4.70	0.483	มากที่สุด
โดยภาพรวม	4.24	0.147	มาก

จากตารางที่ 36 แสดงว่า แบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัยด้านการเขียนมีคุณประโยชน์ด้านต่าง ๆ ที่เหมาะสม แบ่งได้เป็น 2 ระดับ คือ ด้านที่มีคุณประโยชน์มากที่สุด และด้านที่มีคุณประโยชน์มาก โดยมีรายละเอียดดังนี้

ด้านที่มีคุณประโยชน์มากที่สุด มี 1 ด้าน คือ ความเป็นธรรม (Fairness) ($\bar{X} = 4.70$, S.D. = 0.483) โดยมีข้อคำถามย่อย 1 ข้อ เนื้อหาสาระแบบทดสอบเหมาะสมและเป็นหัวเรื่องที่ปราศจากความลำเอียง โดยไม่มีการกล่าวถึงการแบ่งแยกเพศ เชื้อชาติ สถานทางเศรษฐกิจและสังคม และไม่นำไปสู่ความรุนแรงใด ๆ ($\bar{X} = 4.70$, S.D. = 0.483)

ด้านที่มีคุณประโยชน์มาก มี 6 ด้าน คือ 1) ด้านความตรงเชิงโครงสร้าง (Construct Validity) ($\bar{X} = 4.20$, S.D. = 0.100) 2) ด้านความเชื่อมั่น (Reliability) ($\bar{X} = 4.20$, S.D. = 0.115) 3) ด้านการมีปฏิสัมพันธ์ (Interactiveness) ($\bar{X} = 4.16$, S.D. = 0.089) 4) ด้านความสมจริง (Authenticity) ($\bar{X} = 4.15$, S.D. = 0.071) 5) ด้านความสะดวกใช้ (Practicality) ($\bar{X} = 4.15$, S.D. = 0.071) และ 6) ด้านผลกระทบ (Impact) ($\bar{X} = 4.15$, S.D. = 0.100) โดยแต่ละด้านมีคุณประโยชน์ของข้อคำถามย่อยที่เกี่ยวข้องดังรายละเอียดต่อไปนี้

ด้านความตรงเชิงโครงสร้าง (Construct Validity) ($\bar{X} = 4.20$, S.D. = 0.100) มีข้อคำถามย่อย 5 ข้อ และเป็นข้อคำถามที่มีคุณประโยชน์มากที่สุดทั้ง 5 ข้อ คือ 1) วัตถุประสงค์ย่อยของแบบทดสอบแต่ละตอนสอดคล้องกับจุดประสงค์หลัก ($\bar{X} = 4.30$, S.D. = 0.675) 2) วิธีการให้คะแนนมีความเหมาะสมและสามารถสะท้อนความสามารถด้านการเขียนของผู้เข้าสอบ ($\bar{X} = 4.30$, S.D. = 0.675) 3) คะแนนสอบสามารถทำนายความสามารถด้านการเขียนของผู้เข้าสอบ ($\bar{X} = 4.20$, S.D. = 0.789) 4) มีการระบุวัตถุประสงค์การสอบของแบบทดสอบการเขียนแต่ละตอนไว้อย่างชัดเจน ($\bar{X} = 4.10$, S.D. = 0.738) และ 5) ภาระงานสอบด้านการเขียนทุกตอนสอดคล้องกับวัตถุประสงค์การสอบ ($\bar{X} = 4.10$, S.D. = 0.738)

ด้านความเชื่อมั่น (Reliability) ($\bar{X} = 4.20$, S.D. = 0.115) มีข้อคำถามย่อย 4 ข้อ และเป็นข้อคำถามที่มีคุณประโยชน์มากที่สุดทั้ง 4 ข้อ คือ 1) คะแนนสอบเชื่อถือได้ไม่ว่าใครเป็นคนตรวจ ($\bar{X} = 4.30$, S.D. = 0.675) 2) คำตอบของข้อสอบแต่ละข้อเชื่อถือได้ไม่ว่าใครจะตอบ เวลาใดและสถานที่ใด ($\bar{X} = 4.30$, S.D. = 0.675) 3) เกณฑ์การให้คะแนนมีความชัดเจนเพียงพอที่จะทำให้เกิดความเป็นธรรมในการให้คะแนนงานเขียนของผู้เข้าสอบทุกคน ($\bar{X} = 4.10$, S.D. = 0.738) และ 4) เนื้อหาในแบบทดสอบแต่ละตอนไม่เอื้อประโยชน์หรือชี้แนะในการทำข้ออื่น ๆ ในตอนอื่น ๆ ($\bar{X} = 4.10$, S.D. = 0.738)

ด้านการมีปฏิสัมพันธ์ (Interactiveness) ($\bar{X} = 4.16$, S.D. = 0.089) มีข้อคำถามย่อย 2 ข้อ และเป็นข้อคำถามที่มีคุณประโยชน์มากที่สุดทั้ง 2 ข้อ คือ 1) ภาระงานด้านการเขียนมีหัวเรื่องที่คุ้นเคยสำหรับผู้เข้าสอบทุกคน ($\bar{X} = 4.10$, S.D. = 0.738) และ 2) ภาระงานด้านการเขียนเหมาะสมกับระดับความรู้ของผู้เข้าสอบ ($\bar{X} = 4.10$, S.D. = 0.738)

ด้านความสมจริง (Authenticity) ($\bar{X} = 4.15$, S.D. = 0.071) มีข้อคำถามย่อย 2 ข้อ และเป็นข้อคำถามที่มีคุณประโยชน์มากที่สุดทั้ง 2 ข้อ คือ 1) ภาระงานด้านการเขียนมีความสมจริงและสะท้อนสถานการณ์ในชีวิตจริง ($\bar{X} = 4.20$, S.D. = 0.789) และ 2) ภาระงานด้านการเขียนสัมพันธ์กับการเขียนในชีวิตจริงในโอกาสของผู้เข้าสอบ ($\bar{X} = 4.10$, S.D. = 0.738)

ด้านความสะดวกใช้ (Practicality) ($\bar{X} = 4.15$, S.D. = 0.071) มีข้อคำถามย่อย 2 ข้อ และเป็นข้อคำถามที่มีคุณประโยชน์มากทั้ง 2 ข้อ คือ 1) แบบทดสอบด้านการเขียนนี้สามารถสร้างได้จริงด้วยทรัพยากรที่มีอยู่ ($\bar{X} = 4.20$, S.D. = 0.789) และ 2) แบบทดสอบด้านการเขียนสามารถนำไปจัดสอบและตรวจให้คะแนนได้โดยไมยาก ($\bar{X} = 4.10$, S.D. = 0.738)

ด้านผลกระทบ (Impact) ($\bar{X} = 4.15$, S.D. = 0.100) มีข้อคำถามย่อย 4 ข้อ และเป็นข้อคำถามที่มีคุณประโยชน์มากทั้ง 4 ข้อ คือ 1) ภาระงานการเขียนสามารถสะท้อนความสามารถการเขียนของผู้เข้าสอบตามที่สังคมและระบบการศึกษาไทยต้องการ ($\bar{X} = 4.30$, S.D. = 0.675) 2) ภาระงานการเขียนช่วยให้ผู้เข้าสอบมีประสบการณ์ในการทำแบบทดสอบการเขียนในอนาคต ($\bar{X} = 4.10$, S.D. = 0.738) 3) ภาระงานการเขียนคล้ายกับกิจกรรมการเขียนที่ผู้เข้าสอบได้ประสบในห้องเรียนอนาคต ($\bar{X} = 4.10$, S.D. = 0.738) และ 4) จุดประสงค์ของแบบทดสอบด้านการเขียนสอดคล้องกับเป้าหมายและความต้องการที่ระบุไว้ในหลักสูตร อนาคต ($\bar{X} = 4.10$, S.D. = 0.738)

4.1.4.2 ข้อเสนอแนะสำหรับแบบทดสอบด้านการเขียน
ไม่มีข้อเสนอแนะ

สรุปได้ว่า ในภาพรวม แบบทดสอบด้านการเขียนมีคุณประโยชน์อยู่ในระดับ มาก ($\bar{X} = 4.24$, S.D. = 0.147)

4.1.5 ทักษะการพูด

แบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test) ด้านการพูด มีผลการประเมินคุณประโยชน์และข้อเสนอแนะที่เป็นประโยชน์ดังนี้

4.1.5.1 ผลการประเมินคุณประโยชน์ของแบบทดสอบด้านการพูด

ในการประเมินคุณประโยชน์ของแบบทดสอบด้านการพูดโดยใช้แบบประเมินคุณประโยชน์แบบทดสอบก่อนนำไปใช้ ที่มีข้อคำถามแบบมาตราส่วนประมาณค่า จำนวน 20 ข้อ และสอบถามเกี่ยวกับคุณประโยชน์แบบทดสอบ 7 ด้าน คือ 1) ความตรงเชิงโครงสร้าง (Construct Validity) 2) ความเชื่อมั่น (Reliability) 3) ความสมจริง (Authenticity) 4) การมีปฏิสัมพันธ์ (Interactiveness) 5) ผลกระทบ (Impact) 6) ความสะดวกใช้ (Practicality) และ 7) ความเป็นธรรม (Fairness) ได้ผลการประเมินความตรงในแต่ละด้าน ดังรายละเอียดในตาราง 37

ตารางที่ 35 ผลการประเมินคุณประโยชน์ของแบบทดสอบด้านการพูด

คุณประโยชน์ของแบบทดสอบ (Test Usefulness)	\bar{X}	S.D.	แปลผล
ความตรงเชิงโครงสร้าง (Construct Validity)	4.16	0.089	มาก
1. มีการระบุวัตถุประสงค์การสอบของแบบทดสอบการพูดแต่ละตอนไว้อย่างชัดเจน	4.20	0.789	มาก
2. วัตถุประสงค์ย่อยของแบบทดสอบแต่ละตอนสอดคล้องกับจุดประสงค์หลัก	4.10	0.738	มาก
3. ภาระงานสอบด้านการพูดทุกตอนสอดคล้องกับวัตถุประสงค์การสอบ	4.10	0.738	มาก
4. วิธีการให้คะแนนมีความเหมาะสมและสามารถสะท้อนความสามารถด้านการพูดของผู้เข้าสอบ	4.30	0.675	มาก
5. คะแนนสอบสามารถทำนายความสามารถด้านการพูดของผู้เข้าสอบ	4.10	0.738	มาก
ความเชื่อมั่น (Reliability)	4.25	0.100	มาก
6. เกณฑ์การให้คะแนนมีความชัดเจนเพียงพอที่จะทำให้เกิดความเป็นธรรมในการให้คะแนนการพูดของผู้เข้าสอบทุกคน	4.30	0.675	มาก
7. คะแนนสอบเชื่อถือได้ไม่ว่าใครเป็นคนตรวจ	4.30	0.675	มาก
8. คำตอบของข้อสอบแต่ละข้อเชื่อถือได้ไม่ว่าใครจะตอบ เวลาใดและสถานที่ใด	4.10	0.738	มาก
9. เนื้อหาในแบบทดสอบแต่ละตอนไม่เอื้อประโยชน์หรือชี้แนะในการทำข้ออื่นๆ ในตอนอื่น ๆ	4.30	0.675	มาก
ความสมจริง (Authenticity)	4.15	0.071	มาก
10. ภาระงานด้านการพูดสัมพันธ์กับการพูดในชีวิตจริงในขนาดของผู้เข้าสอบ	4.20	0.789	มาก
11. ภาระงานด้านการพูดมีความสมจริงและสะท้อนสถานการณ์ในชีวิตจริง	4.10	0.738	มาก
การมีปฏิสัมพันธ์ (Interactiveness)	4.16	0.089	มาก
12. ภาระงานด้านการพูดมีหัวข้อที่คุ้นเคยสำหรับผู้เข้าสอบทุกคน	4.20	0.789	มาก
13. ภาระงานด้านการพูดเหมาะกับระดับความรู้ของผู้เข้าสอบ	4.10	0.738	มาก
ผลกระทบ (Impact)	4.18	0.096	มาก
14. ภาระงานการพูดช่วยให้ผู้เข้าสอบมีประสบการณ์ในการทำแบบทดสอบการพูดในอนาคต	4.10	0.738	มาก
15. ภาระงานการพูดคล้ายกับกิจกรรมการพูดที่ผู้เข้าสอบได้ประสบในห้องเรียน	4.30	0.675	มาก
16. จุดประสงค์ของแบบทดสอบด้านการพูดสอดคล้องกับเป้าหมายและความต้องการที่ระบุไว้ในหลักสูตร	4.20	0.789	มาก
17. ภาระงานการพูดสามารถสะท้อนความสามารถการพูดของผู้เข้าสอบตามที่สังคมและระบบการศึกษาไทยต้องการ	4.10	0.738	มาก
ความสะดวกใช้ (Practicality)	4.15	0.071	มาก
18. แบบทดสอบด้านการพูดนี้สามารถสร้างได้จริงด้วยทรัพยากรที่มีอยู่	4.20	0.789	มาก
19. แบบทดสอบด้านการพูดสามารถนำไปจัดสอบและตรวจให้คะแนนได้โดยไม่มียาก	4.10	0.738	มาก

ตารางที่ 37 ผลการประเมินคุณประโยชน์ของแบบทดสอบด้านการพูด (ต่อ)

คุณประโยชน์ของแบบทดสอบ (Test Usefulness)	\bar{X}	S.D.	แปลผล
ความเป็นธรรม (Fairness)	4.70	0.483	มากที่สุด
20. เนื้อหาสาระแบบทดสอบเหมาะสมและเป็นหัวเรื่องที่ปราศจากความลำเอียง โดยไม่มีการกล่าวถึงการแบ่งแยกเพศ เชื้อชาติ สถานทางเศรษฐกิจ และสังคม และไม่นำไปสู่ความรุนแรงใด ๆ	4.70	0.483	มากที่สุด
โดยภาพรวม	4.25	0.142	มาก

จากตารางที่ 37 แสดงว่า แบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัยด้านการพูดมีคุณประโยชน์ด้านต่าง ๆ ที่เหมาะสม แบ่งได้เป็น 2 ระดับ คือ ด้านที่มีคุณประโยชน์มากที่สุด และด้านที่มีคุณประโยชน์มาก โดยมีรายละเอียดดังนี้

ด้านที่มีคุณประโยชน์มากที่สุด มี 1 ด้าน คือ ความเป็นธรรม (Fairness) ($\bar{X} = 4.70$, S.D. = 0.483) โดยมีข้อคำถามย่อย 1 ข้อ เนื้อหาสาระแบบทดสอบเหมาะสมและเป็นหัวเรื่องที่ปราศจากความลำเอียง โดยไม่มีการกล่าวถึงการแบ่งแยกเพศ เชื้อชาติ สถานทางเศรษฐกิจและสังคม และไม่นำไปสู่ความรุนแรงใด ๆ ($\bar{X} = 4.70$, S.D. = 0.483)

ด้านที่มีคุณประโยชน์มาก มี 6 ด้าน คือ 1) ด้านความเชื่อมั่น (Reliability) ($\bar{X} = 4.25$, S.D. = 0.100) 2) ด้านผลกระทบ (Impact) ($\bar{X} = 4.18$, S.D. = 0.096) 3) ด้านความตรงเชิงโครงสร้าง (Construct Validity) ($\bar{X} = 4.16$, S.D. = 0.089) 4) ด้านการมีปฏิสัมพันธ์ (Interactiveness) ($\bar{X} = 4.16$, S.D. = 0.089) 5) ด้านความสมจริง (Authenticity) ($\bar{X} = 4.15$, S.D. = 0.071) และ 6) ด้านความสะดวกใช้ (Practicality) ($\bar{X} = 4.15$, S.D. = 0.071) โดยแต่ละด้านมีคุณประโยชน์ของข้อคำถามย่อยที่เกี่ยวข้องดังรายละเอียดต่อไปนี้

ด้านความเชื่อมั่น (Reliability) ($\bar{X} = 4.25$, S.D. = 0.100) มีข้อคำถามย่อย 4 ข้อ และเป็นข้อคำถามที่มีคุณประโยชน์มากที่สุดทั้ง 4 ข้อ คือ 1) เกณฑ์การให้คะแนนมีความชัดเจนเพียงพอที่จะทำให้เกิดความเป็นธรรมในการให้คะแนนการพูดของผู้เข้าสอบทุกคน ($\bar{X} = 4.30$, S.D. = 0.675) 2) คะแนนสอบเชื่อถือได้ไม่ว่าใครเป็นคนตรวจ ($\bar{X} = 4.30$, S.D. = 0.675) 3) เนื้อหาในแบบทดสอบแต่ละตอนไม่เอื้อประโยชน์หรือชี้แนะในการทำข้ออื่น ๆ ในตอนอื่น ๆ ($\bar{X} = 4.30$, S.D. = 0.675) และ 4) คำตอบของข้อสอบแต่ละข้อเชื่อถือได้ไม่ว่าใครจะตอบ เวลาใดและสถานที่ใด ($\bar{X} = 4.10$, S.D. = 0.738)

ด้านผลกระทบ (Impact) ($\bar{X} = 4.18$, S.D. = 0.096) มีข้อคำถามย่อย 4 ข้อ และเป็นข้อคำถามที่มีคุณประโยชน์มากที่สุดทั้ง 4 ข้อ คือ 1) ภาระงานการพูดคล้ายกับกิจกรรมการพูดที่ผู้เข้าสอบได้ประสบในห้องเรียน ($\bar{X} = 4.30$, S.D. = 0.675) 2) จุดประสงค์ของแบบทดสอบด้านการพูดสอดคล้องกับเป้าหมายและความต้องการที่ระบุไว้ในหลักสูตร ($\bar{X} = 4.20$, S.D. = 0.789) 3) ภาระงานการพูดช่วยให้ผู้เข้าสอบมีประสบการณ์ในการทำแบบทดสอบการพูดในอนาคต ($\bar{X} = 4.10$, S.D. = 0.738) และ 4) ภาระงานการพูดสามารถสะท้อนความสามารถการพูดของผู้เข้าสอบตามที่สังคมและระบบการศึกษาไทยต้องการ ($\bar{X} = 4.10$, S.D. = 0.738)

ด้านความตรงเชิงโครงสร้าง (Construct Validity) ($\bar{X} = 4.16$, S.D. = 0.089) มีข้อคำถามย่อย 5 ข้อ และเป็นข้อคำถามที่มีคุณประโยชน์มากที่สุดทั้ง 5 ข้อ คือ 1) วิธีการให้คะแนนมีความเหมาะสม และสามารถสะท้อนความสามารถด้านการพูดของผู้เข้าสอบ ($\bar{X} = 4.30$, S.D. = 0.675) 2) มีการระบุวัตถุประสงค์การสอบของแบบทดสอบการพูดแต่ละตอนไว้อย่างชัดเจน ($\bar{X} = 4.20$, S.D. = 0.789) 3) วัตถุประสงค์ย่อยของแบบทดสอบแต่ละตอนสอดคล้องกับจุดประสงค์หลัก ($\bar{X} = 4.10$, S.D. = 0.738) 4) ภาระงานสอบด้านการพูดทุกตอนสอดคล้องกับวัตถุประสงค์การสอบ ($\bar{X} = 4.10$, S.D. = 0.738) และ 5) คะแนนสอบสามารถทำนายความสามารถด้านการพูดของผู้เข้าสอบ ($\bar{X} = 4.10$, S.D. = 0.738)

ด้านการมีปฏิสัมพันธ์ (Interactiveness) ($\bar{X} = 4.16$, S.D. = 0.089) มีข้อคำถามย่อย 2 ข้อ และเป็นข้อคำถามที่มีคุณประโยชน์มากที่สุดทั้ง 2 ข้อ คือ 1) ภาระงานด้านการพูดมีหัวเรื่องที่คุ้นเคยสำหรับผู้เข้าสอบทุกคน ($\bar{X} = 4.20$, S.D. = 0.789) และ 2) ภาระงานด้านการพูดเหมาะกับระดับความรู้ของผู้เข้าสอบ ($\bar{X} = 4.10$, S.D. = 0.738)

ด้านความสมจริง (Authenticity) ($\bar{X} = 4.15$, S.D. = 0.071) มีข้อคำถามย่อย 2 ข้อ และเป็นข้อคำถามที่มีคุณประโยชน์มากที่สุดทั้ง 2 ข้อ คือ 1) ภาระงานด้านการพูดสัมพันธ์กับการพูดในชีวิตจริงในขนาดของผู้เข้าสอบ ($\bar{X} = 4.20$, S.D. = 0.789) และ 2) ภาระงานด้านการพูดมีความสมจริง และสะท้อนสถานการณ์ในชีวิตจริง ($\bar{X} = 4.10$, S.D. = 0.738)

ด้านความสะดวกใช้ (Practicality) ($\bar{X} = 4.15$, S.D. = 0.071) มีข้อคำถามย่อย 2 ข้อ และเป็นข้อคำถามที่มีคุณประโยชน์มากที่สุดทั้ง 2 ข้อ คือ 1) แบบทดสอบด้านการพูดนี้สามารถสร้างได้จริงด้วยทรัพยากรที่มีอยู่ ($\bar{X} = 4.20$, S.D. = 0.789) และ 2) แบบทดสอบด้านการพูดสามารถนำไปจัดสอบและตรวจให้คะแนนได้โดยไม่ยาก ($\bar{X} = 4.10$, S.D. = 0.738)

4.1.5.2 ข้อเสนอแนะสำหรับแบบทดสอบด้านการพูด

ไม่มีข้อเสนอแนะ

สรุปได้ว่า ในภาพรวม แบบทดสอบด้านการพูดมีคุณประโยชน์อยู่ในระดับ มาก ($\bar{X} = 4.25$, S.D. = 0.142)

ตารางที่ 36 สรุประดับคุณประโยชน์ของแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัยในแต่ละด้าน

คุณประโยชน์/ทักษะ	การฟัง (\bar{X}) (S.D.)	การใช้ ภาษา (\bar{X}) (S.D.)	การอ่าน (\bar{X}) (S.D.)	การ เขียน (\bar{X}) (S.D.)	การพูด (\bar{X}) (S.D.)	รวม (\bar{X}) (S.D.)	อันดับ
ความตรงเชิงโครงสร้าง (Construct Validity)	4.52 .110	4.20 .100	4.14 .207	4.20 .100	4.16 .089	4.24 .121	2
ความเชื่อมั่น (Reliability)	4.25 .191	4.18 .096	4.10 .216	4.20 .115	4.25 .100	4.19 .143	5
ความสมจริง (Authenticity)	4.40 .000	4.20 .141	4.10 .141	4.15 .071	4.15 .071	4.20 .084	4
การมีปฏิสัมพันธ์ (Interactiveness)	4.30 .224	4.16 .089	4.14 .055	4.16 .089	4.16 .089	4.18 .109	6
ผลกระทบ (Impact)	4.40 .432	4.25 .100	4.23 .096	4.15 .100	4.18 .096	4.24 .164	3
ความสะดวกใช้ (Practicality)	4.30 .141	4.15 .071	3.95 .212	4.15 .071	4.15 .071	4.14 .113	7
ความเป็นธรรม (Fairness)	4.50 .527	4.70 .483	4.70 .483	4.70 .483	4.70 .483	4.66 .491	1
รวม	4.38 .232	4.26 .154	4.19 .200	4.24 .147	4.25 .142	4.26 .175	
อันดับ	1	2	5	4	3		

จากตารางที่ 38 สรุปได้ว่า ในภาพรวมแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัยมีคุณประโยชน์ตามเกณฑ์ที่กำหนด คือ มีคุณประโยชน์อยู่ในระดับดี ($\bar{X} = 4.26$, S.D. = .175) และเมื่อแยกรายด้านสามารถเรียงตามอันดับคุณประโยชน์ได้ คือ 1) ความเป็นธรรม (Fairness) ($\bar{X} = 4.66$, S.D. = .491) 2) ความตรงเชิงโครงสร้าง (Construct Validity) ($\bar{X} = 4.24$, S.D. = .121) 3) ผลกระทบ (Impact) ($\bar{X} = 4.24$, S.D. = .164) 4) ความสมจริง (Authenticity) ($\bar{X} = 4.20$, S.D. = .084) 5) ความเชื่อมั่น (Reliability) ($\bar{X} = 4.19$, S.D. = .143) 6) การมีปฏิสัมพันธ์ (Interactiveness) ($\bar{X} = 4.18$, S.D. = .109) และ 7) ความสะดวกใช้ (Practicality) ($\bar{X} = 4.14$, S.D. = .113)

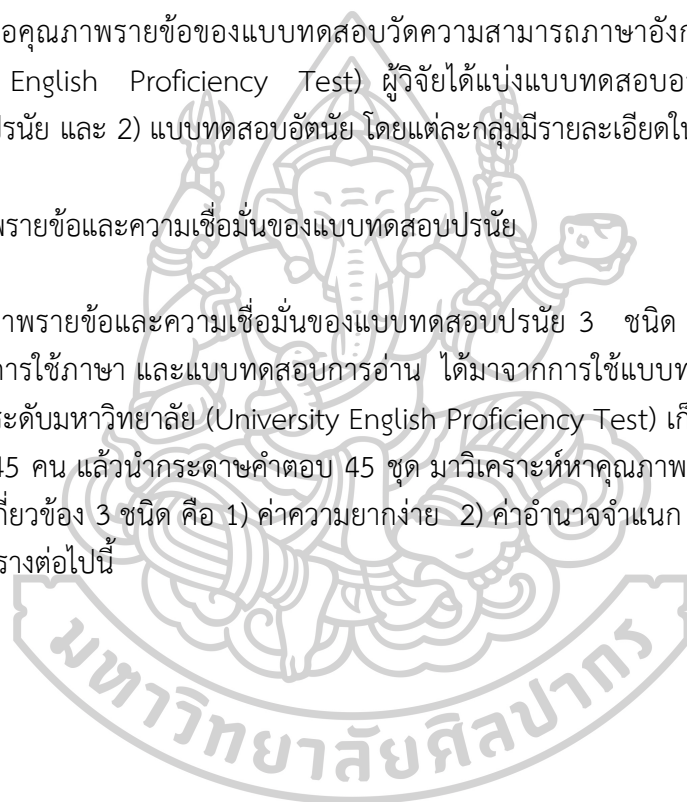
4.2 คุณภาพรายข้อและความเชื่อมั่นของแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test)

คุณภาพรายข้อของแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test) ได้มาจากการนำกระดาษคำตอบ และคะแนนของกลุ่มตัวอย่างจริง จำนวน 45 คน มาหาค่า ความยากง่าย (P) ค่าอำนาจจำแนก (R) และค่าความเชื่อมั่น (Reliability) ของแบบทดสอบรายข้อตามเกณฑ์ที่กำหนดไว้ในบทที่ 3

ในการนำเสนอคุณภาพรายข้อของแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test) ผู้วิจัยได้แบ่งแบบทดสอบออกเป็น 2 กลุ่ม คือ 1) แบบทดสอบปรนัย และ 2) แบบทดสอบอัตนัย โดยแต่ละกลุ่มมีรายละเอียดในการวิเคราะห์ผลดังนี้

4.2.1 คุณภาพรายข้อและความเชื่อมั่นของแบบทดสอบปรนัย

คุณภาพรายข้อและความเชื่อมั่นของแบบทดสอบปรนัย 3 ชนิด คือ แบบทดสอบการฟัง แบบทดสอบการใช้ภาษา และแบบทดสอบการอ่าน ได้มาจากการใช้แบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test) เก็บข้อมูลกับกลุ่มตัวอย่างจริง จำนวน 45 คน แล้วนำกระดาษคำตอบ 45 ชุด มาวิเคราะห์หาคุณภาพรายข้อและความเชื่อมั่น โดยใช้สถิติที่เกี่ยวข้อง 3 ชนิด คือ 1) ค่าความยากง่าย 2) ค่าอำนาจจำแนก และ 3) ค่าความเชื่อมั่น ดังแสดงในตารางต่อไปนี้



ตารางที่ 37 คุณภาพรายข้อและค่าความเชื่อมั่นของแบบทดสอบด้านการฟัง

ข้อที่	P	R	แปลความ	จำนวนข้อ	%
คุณประโยชน์ดี				38	76.00
3	.67	.37	ค่อนข้างง่าย อำนาจจำแนกดี	2	4.00
38	.71	.30			
6	.61	.68	ค่อนข้างง่าย อำนาจจำแนกดีมาก	12	24.00
1	.35	.63			
22	.78	.76			
23	.61	.68			
26	.65	.63			
27	.65	.63			
29	.34	.85			
30	.66	.85			
37	.61	.68			
42	.65	.63			
43	.64	.44			
31	.27	.51			
32	.30	.82	ยากง่ายปานกลาง อำนาจจำแนกพอใช้ได้	3	6.00
17	.45	.26			
36	.45	.26			
40	.55	.26	ยากง่ายปานกลาง อำนาจจำแนกดี	4	8.00
11	.50	.34			



ตารางที่ 39 คุณภาพรายข้อและค่าความเชื่อมั่นของแบบทดสอบด้านการฟัง (ต่อ)

ข้อที่	P	R	แปลความ	จำนวนข้อ	%		
33	.59	.34					
45	.59	.34					
46	.41	.34					
4	.55	.58	ยากง่ายปานกลาง อำนาจจำแนกดีมาก	12	24.00		
8	.50	.50					
9	.55	.58					
12	.54	.43					
18	.54	.43					
19	.55	.58					
20	.41	.89					
25	.55	.58					
47	.45	.58					
48	.50	.50					
49	.50	.65					
50	.57	.73					
15	.39	.68				ค่อนข้างยาก อำนาจจำแนกดีมาก	3
16	.39	.68					
34	.27	.51					
39	.19	.72	ยากมาก อำนาจจำแนกดีมาก	1	2.00		
ควรปรับปรุง				12	24.00		
2	.25	-	ค่อนข้างยาก ไม่มีอำนาจจำแนก	12	24.00		
5	.63	-.27	ค่อนข้างง่าย ไม่ดี คนเก่งหลงทำผิด				
14	.67	.19	ค่อนข้างง่าย อำนาจจำแนกไม่ดี				
7	.58	-	ยากง่ายปานกลาง ไม่มีอำนาจจำแนก				
44	.50	-					
10	.41	.18					
28	.41	.18					
35	.46	.08	ยากง่ายปานกลาง อำนาจจำแนกไม่ดี				
41	.50	.16					
13	.41	.18					
24	.33	-					
21	.21	.11	ค่อนข้างยาก ไม่มีอำนาจจำแนก				
			ค่อนข้างยาก อำนาจจำแนกไม่ดี				
เฉลี่ย	.50	.49	จำนวนข้อ	50	100.00		
แปลผล	ดีมาก	จำแนกได้ดี	ความเชื่อมั่น KR-20	.83			

จากตารางที่ 39 แสดงคุณภาพรายข้อและค่าความเชื่อมั่นของแบบทดสอบด้านการฟัง พบว่า ในภาพรวมแบบทดสอบด้านการฟังทั้งหมด 50 ข้อ มีค่าความยากง่ายอยู่ในระดับดีมาก ($P = .50$) และสามารถจำแนกได้ดี ($R = .49$) โดยมีค่าความเชื่อมั่น ($KR-20$) เท่ากับ .83 เมื่อพิจารณาคุณภาพรายข้อ สามารถแบ่งออกได้เป็น 2 กลุ่ม คือ 1) ข้อสอบที่ได้คุณประโยชน์ จำนวน 38 ข้อ คิดเป็น 76.00 % และ 2) ข้อที่ควรปรับปรุง จำนวน 12 ข้อ คิดเป็น 24 %

ตารางที่ 38 คุณภาพรายข้อและค่าความเชื่อมั่นของแบบทดสอบด้านการใช้ภาษา

ข้อที่	P	R	แปลความ	จำนวนข้อ	%
คุณประโยชน์ดี				36	90.00
11	.81	.37	ง่ายมาก อำนาจจำแนกดี	2	5.00
22	.81	.37			
35	.84	.68	ง่ายมาก อำนาจจำแนกดีมาก	1	2.50
33	.75	.21	ค่อนข้างง่าย อำนาจจำแนกพอใช้ได้	1	2.50
25	.67	.37	ค่อนข้างง่าย อำนาจจำแนกดี	1	2.50
15	.61	.68	ค่อนข้างง่าย อำนาจจำแนกดีมาก	7	17.50
18	.64	.44			
21	.70	.57			
28	.67	.37			
34	.64	.44			
38	.74	.79			
40	.74	.79			
8	.45	.26	ยากง่ายปานกลาง อำนาจจำแนกพอใช้ได้	1	2.50
19	.55	.26			
31	.45	.26			
9	.59	.34	ยากง่ายปานกลาง อำนาจจำแนกดี	4	10.00
24	.59	.34			
27	.41	.34			
37	.50	.34			
3	.50	.50	ยากง่ายปานกลาง อำนาจจำแนกดีมาก	11	27.50
12	.59	.51			
14	.50	.80			
16	.43	.73			
17	.50	.50			
20	.54	.43			
23	.59	.51			
30	.43	.73			
32	.50	.50			
36	.50	.80			
39	.59	.51			

ตารางที่ 40 คุณภาพรายข้อและค่าความเชื่อมั่นของแบบทดสอบด้านการใช้ภาษา (ต่อ)

ข้อที่	P	R	แปลความ	จำนวนข้อ	%
1	.26	.79	ค่อนข้างยาก อำนาจจำแนกดีมาก	4	10.00
5	.30	.57			
10	.30	.82			
13	.39	.68			
4	.13	.63	ยากมาก อำนาจจำแนกดีมาก	2	5.00
7	.19	.72			
ควรปรับปรุง				4	10.00
6	.71	.10	ค่อนข้างง่าย อำนาจจำแนกไม่ดี	4	10.00
2	.58	-			
26	.59	.18			
29	.37	.10			
เฉลี่ย	.54	.52	จำนวนข้อ	40	100.00
แปลผล	ดีมาก	จำแนกได้ดี	ความเชื่อมั่น KR-20	.81	

จากตารางที่ 40 แสดงคุณภาพรายข้อและค่าความเชื่อมั่นของแบบทดสอบด้านการใช้ภาษา พบว่า ในภาพรวมแบบทดสอบด้านการใช้ภาษาทั้งหมด 40 ข้อ มีค่าความยากง่ายอยู่ในระดับดีมาก ($P = .54$) และสามารถจำแนกได้ดี ($R = .52$) โดยมีค่าความเชื่อมั่น (KR-20) เท่ากับ .81 เมื่อพิจารณาคุณภาพรายข้อ สามารถแบ่งออกได้เป็น 2 กลุ่ม คือ 1) ข้อสอบที่ได้คุณประโยชน์ จำนวน 36 ข้อ คิดเป็น 90 % และ 2) ข้อที่ควรปรับปรุง จำนวน 4 ข้อ คิดเป็น 10 %

ตารางที่ 39 คุณภาพรายข้อและค่าความเชื่อมั่นของแบบทดสอบด้านการอ่าน

ข้อที่	P	R	แปลความ	จำนวนข้อ	%
คุณประโยชน์ดี				48	80.00
60	.81	.37	ง่ายมาก อำนาจจำแนกดี	1	1.66
24	.87	.63			
50	.87	.63	ง่ายมาก อำนาจจำแนกดีมาก	4	6.66
55	.81	.72			
59	.87	.63			
3	.75	.21	ค่อนข้างง่าย อำนาจจำแนกพอใช้ได้	2	3.33
43	.63	.27			
4	.67	.37	ค่อนข้างง่าย อำนาจจำแนกดี	4	6.66
12	.71	.30			
30	.71	.30			
42	.71	.30			
2	.65	.63	ค่อนข้างง่าย อำนาจจำแนกดีมาก	11	18.33
7	.64	.44			
8	.73	.51			

ตารางที่ 41 คุณภาพรายข้อและค่าความเชื่อมั่นของแบบทดสอบด้านการอ่าน (ต่อ)

ข้อที่	P	R	แปลความ	จำนวนข้อ	%
10	.77	.45			
18	.65	.63			
22	.61	.68			
25	.64	.44			
28	.77	.45			
29	.65	.63			
44	.64	.44			
53	.65	.63			
20	.45	.26	ยากง่ายปานกลาง อำนาจจำแนกพอใช้ได้	2	3.33
35	.45	.26			
6	.59	.34			
11	.41	.34			
16	.41	.34	ยากง่ายปานกลาง อำนาจจำแนกดี	6	10.00
31	.50	.34			
33	.59	.34			
36	.50	.34			
1	.50	.80			
9	.46	.43			
13	.45	.58	ยากง่ายปานกลาง อำนาจจำแนกดีมาก	13	21.66
23	.54	.43			
26	.59	.51			
27	.59	.89			
32	.50	.65			
37	.46	.43			
39	.55	.58			
48	.50	.80			
51	.55	.58			
56	.41	.89			
58	.59	.89			
47	.33	.37	ค่อนข้างยาก อำนาจจำแนกดี	1	1.66
5	.36	.44			
45	.36	.44	ค่อนข้างยาก อำนาจจำแนกดีมาก	4	6.66
46	.35	.63			
52	.35	.63			
ควรปรับปรุง				12	20.00
14	.54	.08	ยากง่ายปานกลาง อำนาจจำแนกไม่ดี		
15	.88	.18	ง่ายมาก อำนาจจำแนกไม่ดี	12	20.00
17	.54	.08	ยากง่ายปานกลาง อำนาจจำแนกไม่ดี		

ตารางที่ 41 คุณภาพรายข้อและค่าความเชื่อมั่นของแบบทดสอบด้านการอ่าน (ต่อ)

ข้อที่	P	R	แปลความ	จำนวนข้อ	%
19	.71	-.10	ค่อนข้างง่าย ไม่ดี คนเก่งหลงทำผิด		
21	.21	-.11	ค่อนข้างยาก ไม่ดี คนเก่งหลงทำผิด		
34	.67	-	ค่อนข้างง่าย ไม่มีอำนาจจำแนก		
38	.63	.10	ค่อนข้างง่าย อำนาจจำแนกไม่ดี		
40	.50	-	ยากง่ายปานกลาง ไม่มีอำนาจจำแนก		
41	.54	.08	ยากง่ายปานกลาง อำนาจจำแนกไม่ดี		
49	.83	-	ง่ายมาก ไม่มีอำนาจจำแนก		
54	.67	.19	ค่อนข้างง่าย อำนาจจำแนกไม่ดี		
57	.88	.18	ง่ายมาก อำนาจจำแนกไม่ดี		
เฉลี่ย	.60	.46	จำนวนข้อ		
แปลผล	ดีมาก	จำแนกได้ดี	ความเชื่อมั่น KR-20	.83	

จากตารางที่ 41 แสดงคุณภาพรายข้อและค่าความเชื่อมั่นของแบบทดสอบด้านการอ่าน พบว่า ในภาพรวมแบบทดสอบด้านการอ่านทั้งหมด 60 ข้อ มีค่าความยากง่ายอยู่ในระดับดีมาก ($P = .60$) และสามารถจำแนกได้ดี ($R = .46$) โดยมีค่าความเชื่อมั่น (KR-20) เท่ากับ .83 เมื่อพิจารณาคุณภาพรายข้อ สามารถแบ่งออกได้เป็น 2 กลุ่ม คือ 1) ข้อสอบที่ได้คุณประโยชน์ จำนวน 48 ข้อ คิดเป็น 80 % และ 2) ข้อสอบที่ควรปรับปรุง จำนวน 12 ข้อ คิดเป็น 20 %

4.2.2 คุณภาพรายข้อและความเชื่อมั่นของแบบทดสอบอัตนัย

คุณภาพรายข้อและความเชื่อมั่นของแบบทดสอบอัตนัย 2 ชนิด คือ แบบทดสอบการเขียน และแบบทดสอบการพูด ได้มาจากการใช้แบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test) ในทักษะการเขียนกับกลุ่มตัวอย่างจริง จำนวน 45 คน และทักษะการพูดกับกลุ่มตัวอย่าง ที่มีความสามารถอยู่ในกลุ่มสูง กลาง และต่ำ กลุ่มละ 7 คน รวมจำนวน 21 คน สำหรับการตรวจให้คะแนนแบบทดสอบการเขียน ดำเนินการตรวจโดยกรรมการ 3 คน และแบบทดสอบการพูด ดำเนินการสอบและให้คะแนนโดยกรรมการ 3 คน แล้วนำผลคะแนนที่ได้มาค่าความยากง่าย ค่าอำนาจจำแนก และค่าความเชื่อมั่น (Alpha Coefficient) ดังผลการวิเคราะห์ที่แสดงในตารางต่อไปนี้

ตารางที่ 40 คุณภาพรายข้อและค่าความเชื่อมั่นของแบบทดสอบด้านการเขียน

การเขียน	จำนวนผู้ตอบ	คะแนนเต็ม	\bar{X}	S.D.	P	R	ความเชื่อมั่น (Alpha Coefficient)
TASK A	14	20	13.43	2.380	.68	.21	.61
TASK B	14	40	24.21	5.090	.60	.31	
ทั้งหมด	14	60	37.64	6.736	.60	.26	

จากตารางที่ 42 แสดงคุณภาพรายข้อและค่าความเชื่อมั่นของแบบทดสอบด้านการเขียน พบว่า ในภาพรวมแบบทดสอบด้านการเขียนทั้งหมด 2 ภาระงาน มีความยากง่ายอยู่ในระดับดีมาก ($P = .60$) และอำนาจจำแนกอยู่ในระดับจำแนกได้พอใช้ ($R = .26$) โดยมีความเชื่อมั่น (Alpha Coefficient) เท่ากับ .61 และเมื่อพิจารณารายการภาระงาน พบว่า Task A มีความยากง่ายอยู่ในระดับดี ($P = .68$) และมีอำนาจจำแนกอยู่ในระดับจำแนกได้พอใช้ ($R = .21$) ส่วน Task B มีความยากง่ายอยู่ในระดับดีมาก ($P = .60$) และอำนาจจำแนกอยู่ในระดับจำแนกได้ดี ($R = .31$)

ตารางที่ 41 คุณภาพรายข้อและค่าความเชื่อมั่นของแบบทดสอบด้านการพูด

การเขียน	จำนวนผู้ตอบ	คะแนนเต็ม	\bar{X}	S.D.	P	R	ความเชื่อมั่น (Alpha Coefficient)
TASK A	18	20	10.50	3.880	.50	.50	.99
TASK B	18	20	10.56	3.730	.51	.48	
TASK C	18	20	10.56	3.770	.51	.49	
ทั้งฉบับ	18	60	31.61	11.319	.50	.49	

จากตารางที่ 43 แสดงคุณภาพรายข้อและค่าความเชื่อมั่นของแบบทดสอบด้านการพูด พบว่า ในภาพรวมแบบทดสอบด้านการพูดทั้งหมด 3 ภาระงาน มีความยากง่ายอยู่ในระดับดีมาก ($P = .50$) และอำนาจจำแนกอยู่ในระดับจำแนกได้ดีมาก ($R = .49$) โดยมีความเชื่อมั่น (Alpha Coefficient) เท่ากับ .99 และเมื่อพิจารณารายการภาระงาน พบว่า ทั้ง 3 ภาระงาน มีความยากง่ายอยู่ในระดับดีมาก และมีอำนาจจำแนกอยู่ในระดับจำแนกได้ดีมากเช่นเดียวกัน โดย Task A มีค่าความยากง่ายเท่ากับ .50 ค่าอำนาจจำแนกเท่ากับ .50 สำหรับ Task B มีค่าความยากง่ายเท่ากับ .51 และค่าอำนาจจำแนกเท่ากับ .48 และ Task C มีค่าความยากง่ายเท่ากับ .51 และค่าอำนาจจำแนกเท่ากับ .49

ตารางที่ 42 สรุปคุณภาพรายข้อและความเชื่อมั่นของแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย

ทักษะ	จำนวนข้อ	ความยาก ง่าย (P)	อำนาจ จำแนก (R)	ความเชื่อมั่น	คุณประโยชน์	
					ดี (ข้อ/%)	ควรปรับปรุง (ข้อ/%)
การฟัง	50	.50	.49	.83	38/76%	12/24%
A	30	.50	.44	.82	21/70%	9/30%
B	8	.45	.43	.83	7/87.50%	1/12.50%
C	12	.50	.45	.84	10/83.33%	2/16.67%
การใช้ภาษา	40	.54	.52	.82	36/90%	4/10%
A	15	.46	.51	.81	13/86.67%	2/13.33%
B	25	.58	.46	.83	23/92%	2/8%
การอ่าน	60	.60	.46	.83	48/80%	12/20%
A	30	.60	.40	.84	25/83.33%	5/16.67%
B	30	.59	.42	.83	23/76.67%	7/23.33%
ปรนัย	150	.55	.49	.83	122/81.33%	28/18.67%
การเขียน	2	.60	.26	.61	2/100%	-
A	1	.68	.21	.61	1/100%	-
B	1	.60	.31	.61	1/100%	-
การพูด	3	.50	.49	.99	3/100%	-
A	1	.50	.50	.99	1/100%	-
B	1	.51	.48	.99	1/100%	-
C	1	.51	.49	.99	1/100%	-
อัตนัย	5	.55	.38	.80	5/100%	

ตารางที่ 44 สรุปคุณภาพรายข้อและความเชื่อมั่นของแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัยทั้ง 2 ชนิด คือ แบบทดสอบปรนัย และแบบทดสอบอัตนัยโดยมีรายละเอียด ดังนี้

แบบทดสอบปรนัยซึ่งประกอบด้วยการฟัง การใช้ภาษา และการอ่าน มีข้อคำถามทั้งหมด 150 ข้อ มีค่าความยากง่ายอยู่ในระดับดีมาก (P = .55) ค่าอำนาจจำแนกอยู่ในระดับจำแนกได้ดี (R = .49) และค่าความเชื่อมั่น (KR-20) เท่ากับ .83 มีข้อคำถามคุณประโยชน์ดี 122 ข้อ คิดเป็น 81.33 % และข้อคำถามที่ควรปรับปรุง 28 ข้อ คิดเป็น 18.67 % และสามารถแยกคุณประโยชน์รายทักษะได้ คือ แบบทดสอบการฟัง จำนวน 50 ข้อ มีค่าความยากง่ายอยู่ในระดับดีมาก (P = .50) ค่าอำนาจจำแนกอยู่ในระดับจำแนกได้ดี (R = .49) และค่าความเชื่อมั่น (KR-20) เท่ากับ .83 มีข้อคำถามคุณประโยชน์ดี 38 ข้อ คิดเป็น 76 % และข้อคำถามที่ควรปรับปรุง 12 ข้อ คิดเป็น 24 % แบบทดสอบการใช้ภาษา จำนวน 40 ข้อ มีค่าความยากง่ายอยู่ในระดับดีมาก (P = .54) ค่าอำนาจจำแนกอยู่ในระดับจำแนกได้ดี (R = .52) และค่าความเชื่อมั่น (KR-20) เท่ากับ .82 มีข้อคำถาม

คุณประโยชน์ดี 36 ข้อ คิดเป็น 90 % และข้อคำถามที่ควรปรับปรุง 4 ข้อ คิดเป็น 10 % และแบบทดสอบการอ่าน จำนวน 60 ข้อ มีค่าความยากง่ายอยู่ในระดับดีมาก ($P = .60$) ค่าอำนาจจำแนกอยู่ในระดับจำแนกได้ดี ($R = .46$) และค่าความเชื่อมั่น (KR-20) เท่ากับ .83 มีข้อคำถามคุณประโยชน์ดี 48 ข้อ คิดเป็น 80 % และข้อคำถามที่ควรปรับปรุง 12 ข้อ คิดเป็น 20 % แบบทดสอบอัตนัยซึ่งประกอบด้วย การเขียนและการพูด มีข้อคำถามทั้งหมด 5 ข้อ มีค่าความยากง่ายอยู่ในระดับดีมาก ($P = .55$) ค่าอำนาจจำแนกอยู่ในระดับจำแนกได้ดี ($R = .38$) และค่าความเชื่อมั่น (Alpha Coefficient) เท่ากับ .80 และเป็นข้อคำถามคุณประโยชน์ดีทั้ง 5 ข้อ คิดเป็น 100 % และสามารถแยกคุณประโยชน์รายทักษะได้ คือ แบบทดสอบการเขียน จำนวน 2 ข้อ มีค่าความยากง่ายอยู่ในระดับดีมาก ($P = .60$) ค่าอำนาจจำแนกอยู่ในระดับจำแนกได้พอใช้ ($R = .26$) และค่าความเชื่อมั่น (Alpha Coefficient) เท่ากับ .61 และเป็นข้อคำถามคุณประโยชน์ดีทั้ง 2 ข้อ คิดเป็น 100 % และแบบทดสอบการพูด จำนวน 3 ข้อ มีค่าความยากง่ายอยู่ในระดับดีมาก ($P = .50$) ค่าอำนาจจำแนกอยู่ในระดับจำแนกได้ดี ($R = .49$) และค่าความเชื่อมั่น (Alpha Coefficient) เท่ากับ .99 และเป็นข้อคำถามคุณประโยชน์ดีทั้ง 3 ข้อ คิดเป็น 100 %

4.3 ความตรงร่วมสมัยของแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (Concurrent Validity)

ความตรงร่วมสมัย (Concurrent Validity) ของแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย คือ ความสัมพันธ์ระหว่างคะแนนของแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test) กับคะแนนของแบบทดสอบมาตรฐาน ได้มาจากการใช้แบบทดสอบ 2 ชนิด คือ 1) แบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test) และ 2) แบบทดสอบวัดความสามารถภาษาอังกฤษมาตรฐาน (Standard English Proficiency Test) โดยผู้วิจัยจะนำเสนอความสัมพันธ์ของคะแนนจากแบบทดสอบ 2 กลุ่ม คือ แบบทดสอบปรนัย และแบบทดสอบอัตนัย โดยมีรายละเอียดดังนี้

4.3.1 ความตรงร่วมสมัยของแบบทดสอบปรนัย

ในการหาความตรงร่วมสมัยของแบบทดสอบปรนัย ผู้วิจัยได้เก็บข้อมูลคะแนนสอบของแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัยกับคะแนนสอบของแบบทดสอบ TOEFL (Test of English as a Foreign Language) ซึ่งได้เก็บข้อมูลกับกลุ่มตัวอย่าง 2 สัปดาห์ หลังจากการสอบโดยใช้แบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย ประกอบด้วย 3 ทักษะ คือ ทักษะการฟัง ทักษะการใช้ภาษา และทักษะการอ่าน โดยผู้วิจัยได้ดำเนินการ 2 ขั้นตอน คือ 1) การจัดทำรายงานคะแนนการสอบ และ 2) การหาความสัมพันธ์ของคะแนนสอบ ซึ่งมีรายละเอียดดังนี้

4.3.1.1 การจัดทำรายงานคะแนนการสอบ

ในการจัดทำรายงานคะแนนสอบ ผู้วิจัยได้นำคะแนนของผู้เข้าสอบทุกคน ที่ได้จากการตรวจให้คะแนนแบบปรนัยในแต่ละทักษะมาคำนวณหาค่าสูงสุด (Max) ค่าต่ำสุด (Min) ค่าเฉลี่ย (\bar{X}) ส่วนเบี่ยงเบนมาตรฐาน (S.D.) และร้อยละ (Percentage) และจัดเรียงลำดับคะแนนที่ผู้เข้าสอบทำได้ ดังแสดงในตารางต่อไปนี้

ตารางที่ 43 รายงานผลคะแนนการสอบวัดความสามารถภาษาอังกฤษโดยใช้แบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัยของกลุ่มตัวอย่าง

ที่	รหัสประจำตัว	ฟัง	%	การใช้ภาษา	%	อ่าน	%	รวมทุกทักษะ	%
		50	100	40	100	60	100	150	100
1	601011529	35	70.00	32	80.00	50	83.33	117	78.00
2	611011556	35	70.00	32	80.00	49	81.67	116	77.33
3	611011551	38	76.00	30	75.00	47	78.33	115	76.67
4	601031260	36	72.00	32	80.00	46	76.67	114	76.00
5	581011528	34	68.00	33	82.50	46	76.67	113	75.33
6	591011369	37	74.00	31	77.50	45	75.00	113	75.33
7	591081265	36	72.00	31	77.50	46	76.67	113	75.33
8	611011580	33	66.00	29	72.50	50	83.33	112	74.67
9	601031239	36	72.00	28	70.00	47	78.33	111	74.00
10	591011521	37	74.00	27	67.50	46	76.67	110	73.33
11	601011583	36	72.00	31	77.50	42	70.00	109	72.67
12	601011575	27	54.00	29	72.50	45	75.00	101	67.33
13	601011530	29	58.00	21	52.50	39	65.00	89	59.33
14	581011503	27	54.00	21	52.50	39	65.00	87	58.00
15	601031481	25	50.00	23	57.50	37	61.67	85	56.67
16	611011559	28	56.00	23	57.50	34	56.67	85	56.67
17	611011581	29	58.00	16	40.00	39	65.00	84	56.00
18	591081251	25	50.00	23	57.50	35	58.33	83	55.33
19	611011575	29	58.00	20	50.00	34	56.67	83	55.33
20	611011576	28	56.00	20	50.00	35	58.33	83	55.33
21	591081170	26	52.00	22	55.00	34	56.67	82	54.67

ตารางที่ 45 รายงานผลคะแนนการสอบวัดความสามารถภาษาอังกฤษโดยใช้แบบทดสอบ
วัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (ต่อ)

ที่	รหัสประจำตัว	ฟัง	%	การใช้ ภาษา	%	อ่าน	%	รวมทุก ทักษะ	%
		50	100	40	100	60	100	150	100
22	601011553	28	56.00	18	45.00	36	60.00	82	54.67
23	611011553	31	62.00	20	50.00	31	51.67	82	54.67
24	611011665	27	54.00	22	55.00	33	55.00	82	54.67
25	581011538	25	50.00	21	52.50	35	58.33	81	54.00
26	601011552	26	52.00	18	45.00	37	61.67	81	54.00
27	591081066	27	54.00	20	50.00	33	55.00	80	53.33
28	591081111	28	56.00	18	45.00	34	56.67	80	53.33
29	581011540	21	42.00	24	60.00	34	56.67	79	52.67
30	611011560	26	52.00	20	50.00	32	53.33	78	52.00
31	611011563	12	24.00	21	52.50	39	65.00	72	48.00
32	601011571	18	36.00	14	35.00	25	41.67	57	38.00
33	601011623	16	32.00	13	32.50	27	45.00	56	37.33
34	581011521	18	36.00	12	30.00	25	41.67	55	36.67
35	591081122	14	28.00	14	35.00	27	45.00	55	36.67
36	591011377	14	28.00	15	37.50	25	41.67	54	36.00
37	591011350	17	34.00	14	35.00	22	36.67	53	35.33
38	601011533	15	30.00	12	30.00	26	43.33	53	35.33
39	601011619	16	32.00	14	35.00	23	38.33	53	35.33
40	611011600	17	34.00	13	32.50	23	38.33	53	35.33
41	591081262	13	26.00	13	32.50	26	43.33	52	34.67
42	581011539	16	32.00	12	30.00	23	38.33	51	34.00
43	591011381	14	28.00	15	37.50	22	36.67	51	34.00
44	591081055	11	22.00	14	35.00	26	43.33	51	34.00
45	581011523	15	30.00	10	25.00	23	38.33	48	32.00
	Max	38.00	76.00	33.00	82.50	50.00	83.33	117.00	78.00
	Min	11.00	22.00	10.00	25.00	22.00	36.67	48.00	32.00
	\bar{X}	25.13	50.27	20.91	52.28	34.93	58.22	80.98	
	S.D.	8.231	16.463	6.832	17.080	8.742	14.570	22.876	

ตารางที่ 45 แสดงรายงานผลคะแนนการสอบวัดความสามารถภาษาอังกฤษโดยใช้แบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัยเฉพาะแบบทดสอบปรนัย ได้แก่ ทักษะการฟัง การใช้ภาษา และการอ่าน โดยมีกลุ่มตัวอย่างที่เข้าสอบ 45 คน พบว่า จากคะแนน 150 คะแนน คะแนนสูงสุดที่ผู้เข้าสอบทำได้คือ 117 คะแนน หรือ 78% คะแนนต่ำสุด คือ 48 คะแนน หรือ 32 % ค่าเฉลี่ยเท่ากับ 80.98 และส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 22.876 เมื่อพิจารณาค่าร้อยละของคะแนนเฉลี่ยรายทักษะ พบว่า ลำดับของทักษะที่ผู้เข้าสอบสามารถทำแบบทดสอบได้ เรียงจากมากไปหาน้อย คือ ทักษะการอ่าน การใช้ภาษา และการฟัง ดังรายละเอียดต่อไปนี้

ในทักษะการอ่าน จากคะแนนเต็ม 60 คะแนน คะแนนสูงสุดที่ผู้เข้าสอบทำได้คือ 50 คะแนน หรือ 83.33 % คะแนนต่ำสุด คือ 22 คะแนน หรือ 36.67 % ค่าเฉลี่ยเท่ากับ 34.93 และส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 8.742

ในทักษะการใช้ภาษา จากคะแนนเต็ม 40 คะแนน คะแนนสูงสุดที่ผู้เข้าสอบทำได้คือ 33 คะแนน หรือ 82.50 % คะแนนต่ำสุด คือ 10 คะแนน หรือ 25.00 % ค่าเฉลี่ยเท่ากับ 20.91 และส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 6.832

ในทักษะการฟัง จากคะแนนเต็ม 50 คะแนน คะแนนสูงสุดที่ผู้เข้าสอบทำได้คือ 38 คะแนน หรือ 76.00 % คะแนนต่ำสุด คือ 11 คะแนน หรือ 22.00 % ค่าเฉลี่ยเท่ากับ 25.13 และส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 8.231

ตารางที่ 44 รายงานผลคะแนนการสอบวัดความสามารถภาษาอังกฤษโดยใช้แบบทดสอบ TOEFL (Test of English as a Foreign Language) ของกลุ่มตัวอย่าง

ที่	รหัสประจำตัว	ฟัง	%	การใช้ภาษา	%	อ่าน	%	รวมทุกทักษะ	%
		50	100	40	100	50	100	140	100
1	601011529	28	56.00	28	70.00	38	63.33	94	67.14
2	601031260	25	50.00	14	35.00	35	58.33	74	52.86
3	581011528	27	54.00	27	67.50	34	56.67	88	62.86
4	591081265	30	60.00	16	40.00	33	55.00	79	56.43
5	611011580	23	46.00	29	72.50	32	53.33	84	60.00
6	601031239	32	64.00	27	67.50	33	55.00	92	65.71
7	591011521	33	66.00	24	60.00	37	61.67	94	67.14
8	601011575	24	48.00	19	47.50	33	55.00	76	54.29
9	601011530	22	44.00	16	40.00	32	53.33	70	50.00
10	581011503	25	50.00	17	42.50	35	58.33	77	55.00
11	601031481	24	48.00	20	50.00	37	61.67	81	57.86
12	611011559	21	42.00	19	47.50	33	55.00	73	52.14

ตารางที่ 46 รายงานผลคะแนนการสอบวัดความสามารถภาษาอังกฤษโดยใช้แบบทดสอบ TOEFL (Test of English as a Foreign Language) ของกลุ่มตัวอย่าง (ต่อ)

ที่	รหัสประจำตัว	ฟัง	%	การใช้ภาษา	%	อ่าน	%	รวมทุกทักษะ	%
		50	100	40	100	50	100	140	100
13	611011581	18	36.00	15	37.50	30	50.00	63	45.00
14	611011575	14	28.00	18	45.00	28	46.67	60	42.86
15	611011576	24	48.00	18	45.00	22	36.67	64	45.71
16	591081170	18	36.00	21	52.50	26	43.33	65	46.43
17	601011553	22	44.00	19	47.50	24	40.00	65	46.43
18	611011553	20	40.00	21	52.50	28	46.67	69	49.29
19	611011665	20	40.00	18	45.00	29	48.33	67	47.86
20	581011538	19	38.00	22	55.00	27	45.00	68	48.57
21	601011552	19	38.00	16	40.00	31	51.67	66	47.14
22	611011560	18	36.00	18	45.00	28	46.67	64	45.71
23	611011563	8	16.00	19	47.50	29	48.33	56	40.00
24	601011571	12	24.00	17	42.50	24	40.00	53	37.86
25	601011623	8	16.00	12	30.00	21	35.00	41	29.29
26	591011350	10	20.00	10	25.00	19	31.67	39	27.86
27	601011533	9	18.00	11	27.50	18	30.00	38	27.14
28	591081262	8	16.00	12	30.00	17	28.33	37	26.43
29	581011539	9	18.00	11	27.50	16	26.67	36	25.71
30	591011381	7	14.00	9	22.50	18	30.00	34	24.29
31	591081055	9	18.00	8	20.00	19	31.67	36	25.71
32	581011523	9	18.00	8	20.00	12	20.00	29	20.71
Max		33.00	66.00	29.00	72.50	38.00	63.33	94.00	67.14
Min		7.00	14.00	8.00	20.00	12.00	20.00	29.00	20.71
\bar{X}		18.59	37.19	17.47	43.67	27.44	45.73	63.50	45.36
S.D.		7.783	15.566	5.697	14.241	7.043	11.739	18.928	13.520

ตารางที่ 46 แสดงรายงานผลคะแนนการสอบวัดความสามารถภาษาอังกฤษโดยใช้แบบทดสอบ แบบทดสอบ TOEFL (Test of English as a Foreign Language) เฉพาะแบบทดสอบปรนัย ได้แก่ ทักษะการฟัง การใช้ภาษา และการอ่าน โดยมีกลุ่มตัวอย่างที่เข้าสอบ 32 คน พบว่า

จากคะแนน 140 คะแนน คะแนนสูงสุดที่ผู้เข้าสอบทำได้คือ 94 คะแนน หรือ 67.14 % คะแนนต่ำสุด คือ 29 คะแนน หรือ 20.71 % ค่าเฉลี่ยเท่ากับ 63.50 และส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 18.928

เมื่อพิจารณาคำร้อยละของคะแนนเฉลี่ยรายทักษะ พบว่า ลำดับของทักษะที่ผู้เข้าสอบสามารถทำแบบทดสอบได้ เรียงจากมากไปหาน้อย คือ ทักษะการอ่าน การใช้ภาษา และการฟัง ดังรายละเอียดต่อไปนี้

ในทักษะการอ่าน จากคะแนนเต็ม 50 คะแนน คะแนนสูงสุดที่ผู้เข้าสอบทำได้คือ 38 คะแนน หรือ 63.33 % คะแนนต่ำสุด คือ 12 คะแนน หรือ 20.00 % ค่าเฉลี่ยเท่ากับ 27.44 และส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 7.043

ในทักษะการใช้ภาษา จากคะแนนเต็ม 40 คะแนน คะแนนสูงสุดที่ผู้เข้าสอบทำได้คือ 29 คะแนน หรือ 72.50 % คะแนนต่ำสุด คือ 8 คะแนน หรือ 20.00 % ค่าเฉลี่ยเท่ากับ 17.47 และส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 5.697

ในทักษะการฟัง จากคะแนนเต็ม 50 คะแนน คะแนนสูงสุดที่ผู้เข้าสอบทำได้คือ 33 คะแนน หรือ 66.00 % คะแนนต่ำสุด คือ 7 คะแนน หรือ 14.00 % ค่าเฉลี่ยเท่ากับ 18.59 และส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 7.783

4.3.1.2 ความสัมพันธ์ระหว่างคะแนนสอบแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัยกับคะแนนสอบแบบทดสอบ TOEFL

ในการหาความสัมพันธ์ระหว่างคะแนนสอบแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัยกับคะแนนสอบแบบทดสอบ TOEFL ผู้วิจัยได้นำคะแนนรายบุคคลของผู้เข้าสอบที่ได้จากการทำแบบทดสอบวัดความสามารถภาษาอังกฤษ (University English Proficiency Test) และแบบทดสอบ TOEFL (Test of English as a Foreign Language) มาหาความสัมพันธ์ เนื่องจากมีผู้เข้าสอบบางคน que เข้าสอบโดยใช้แบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test) แต่ไม่ได้เข้าสอบโดยใช้แบบทดสอบ TOEFL (Test of English as a Foreign Language) ดังนั้นจึงมีผลคะแนนที่สามารถนำมาเปรียบเทียบได้ในภาพรวม และรายทักษะของผู้เข้าสอบ 32 คน ดังรายละเอียดในตารางที่ 47

ตารางที่ 45 ความสัมพันธ์ระหว่างคะแนนแบบทดสอบปรนัยของแบบทดสอบวัดความสามารถ
ภาษาอังกฤษระดับมหาวิทยาลัยและแบบทดสอบ TOEFL (Test of English as
a Foreign Language)

ที่	รหัสประจำตัว	ฟัง (%)		ใช้ภาษา (%)		อ่าน (%)		รวม (%)	
		UEPT	TOEFL	UEPT	TOEFL	UEPT	TOEFL	UEPT	TOEFL
1	601011529	70.00	56.00	80.00	70.00	83.33	76.00	78.00	67.14
2	601031260	72.00	50.00	80.00	35.00	76.67	70.00	76.00	52.86
3	581011528	68.00	54.00	82.50	67.50	76.67	68.00	75.33	62.86
4	591081265	72.00	60.00	77.50	40.00	76.67	66.00	75.33	56.43
5	611011580	66.00	46.00	72.50	72.50	83.33	64.00	74.67	60.00
6	601031239	72.00	64.00	70.00	67.50	78.33	66.00	74.00	65.71
7	591011521	74.00	66.00	67.50	60.00	76.67	74.00	73.33	67.14
8	601011575	54.00	48.00	72.50	47.50	75.00	66.00	67.33	54.29
9	601011530	58.00	44.00	52.50	40.00	65.00	64.00	59.33	50.00
10	581011503	54.00	50.00	52.50	42.50	65.00	70.00	58.00	55.00
11	601031481	50.00	48.00	57.50	50.00	61.67	74.00	56.67	57.86
12	611011559	56.00	42.00	57.50	47.50	56.67	66.00	56.67	52.14
13	611011581	58.00	36.00	40.00	37.50	65.00	60.00	56.00	45.00
14	611011575	58.00	28.00	50.00	45.00	56.67	56.00	55.33	42.86
15	611011576	56.00	48.00	50.00	45.00	58.33	44.00	55.33	45.71
16	591081170	52.00	36.00	55.00	52.50	56.67	52.00	54.67	46.43
17	601011553	56.00	44.00	45.00	47.50	60.00	48.00	54.67	46.43
18	611011553	62.00	40.00	50.00	52.50	51.67	56.00	54.67	49.29
19	611011665	54.00	40.00	55.00	45.00	55.00	58.00	54.67	47.86
20	581011538	50.00	38.00	52.50	55.00	58.33	54.00	54.00	48.57
21	601011552	52.00	38.00	45.00	40.00	61.67	62.00	54.00	47.14
22	611011560	52.00	36.00	50.00	45.00	53.33	56.00	52.00	45.71
23	611011563	24.00	16.00	52.50	47.50	65.00	58.00	48.00	40.00
24	601011571	36.00	24.00	35.00	42.50	41.67	48.00	38.00	37.86
25	601011623	32.00	16.00	32.50	30.00	45.00	42.00	37.33	29.29
26	591011350	34.00	20.00	35.00	25.00	36.67	38.00	35.33	27.86
27	601011533	30.00	18.00	30.00	27.50	43.33	36.00	35.33	27.14
28	591081262	26.00	16.00	32.50	30.00	43.33	34.00	34.67	26.43
29	581011539	32.00	18.00	30.00	27.50	38.33	32.00	34.00	25.71

ตารางที่ 47 ความสัมพันธ์ระหว่างคะแนนแบบทดสอบปรนัยของแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัยและแบบทดสอบ TOEFL (Test of English as a Foreign Language) (ต่อ)

ที่	รหัสประจำตัว	ฟัง (%)		ใช้ภาษา (%)		อ่าน (%)		รวม (%)	
		UEPT	TOEFL	UEPT	TOEFL	UEPT	TOEFL	UEPT	TOEFL
30	591011381	28.00	14.00	37.50	22.50	36.67	36.00	34.00	24.29
31	591081055	22.00	18.00	35.00	20.00	43.33	38.00	34.00	25.71
32	581011523	30.00	18.00	25.00	20.00	38.33	24.00	32.00	20.71
Correlation (r)		0.93		0.76		0.87		0.95	

ตารางที่ 47 แสดงความสัมพันธ์ของคะแนนแบบทดสอบปรนัยระหว่างคะแนนจากแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัยและแบบทดสอบ TOEFL (Test of English as a Foreign Language) ของผู้เข้าสอบ จำนวน 32 คน โดยภาพรวมหรือแบบทดสอบปรนัย พบว่าคะแนนสอบมีความสัมพันธ์กันในระดับสูง โดยมีค่าความสัมพันธ์เท่ากับ .95

เมื่อพิจารณารายทักษะ พบว่าค่าความสัมพันธ์ของคะแนนในแต่ละทักษะยังคงอยู่ในระดับสูงทุกทักษะ โดยทักษะการฟังมีความสัมพันธ์ของคะแนนเท่ากับ .93 ทักษะการอ่านมีความสัมพันธ์ของคะแนนเท่ากับ .87 และทักษะการฟังมีความสัมพันธ์ของคะแนนเท่ากับ .76

สรุปได้ว่า คะแนนแบบทดสอบปรนัยของแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test) และแบบทดสอบ TOEFL (Test of English as a Foreign Language) มีความสัมพันธ์กันสูงกว่าเกณฑ์ที่กำหนด ทั้งภาพรวมและรายทักษะ ($r \geq .75$)

4.3.2 ความตรงร่วมสมัยของแบบทดสอบอัตนัย

ในการหาความสัมพันธ์ระหว่างคะแนนของแบบทดสอบอัตนัย ผู้วิจัยได้เก็บข้อมูลคะแนนสอบของแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัยกับคะแนนสอบของแบบทดสอบ IELTS (International English Language Testing System) ใน 2 ทักษะ คือ ทักษะการเขียนและทักษะการพูด สำหรับการเขียนจะเก็บข้อมูลพร้อมกับการสอบแบบทดสอบปรนัย คือ 2 สัปดาห์หลังการสอบด้วยแบบทดสอบที่พัฒนาขึ้น และทักษะการพูด จะดำเนินการสอบ 3 วัน หลังจากการสอบปรนัย และดำเนินการตรวจโดยการให้ผู้ตรวจทักษะละ 3 คนมาให้คะแนนตามเกณฑ์ที่กำหนด และหาค่าความเชื่อมั่นของผู้ให้คะแนนหลายคน (Inter-rater Reliability) และเนื่องจากผู้เข้าสอบที่ทำแบบทดสอบการเขียนและการพูดมีจำนวนไม่เท่ากัน ผู้วิจัยจึงนำเสนอรายงานผลคะแนนแยกทักษะ ดังรายละเอียดดังนี้

4.3.2.1 ความสัมพันธ์ระหว่างคะแนนของแบบทดสอบด้านการเขียน

ในการหาความสัมพันธ์ระหว่างคะแนนของแบบทดสอบด้านการเขียน ผู้วิจัยได้ดำเนินการ 2 ขั้นตอน คือ 1) การจัดทำรายงานคะแนนการสอบ และ 2) การหาความสัมพันธ์ของคะแนนสอบ ดังรายละเอียดต่อไปนี้

การจัดทำรายงานคะแนนการสอบ

ในการจัดทำรายงานคะแนนสอบ ผู้วิจัยได้นำคะแนนของผู้เข้าสอบทุกคน ที่ได้จากการตรวจแบบทดสอบรายข้อในแต่ละทักษะมาคำนวณหาค่าสูงสุด (Max) ค่าต่ำสุด (Min) ค่าเฉลี่ย (\bar{X}) ส่วนเบี่ยงเบนมาตรฐาน (S.D.) และร้อยละ (Percentage) และจัดเรียงลำดับคะแนนที่ผู้เข้าสอบทำได้ ดังแสดงในตารางต่อไปนี้

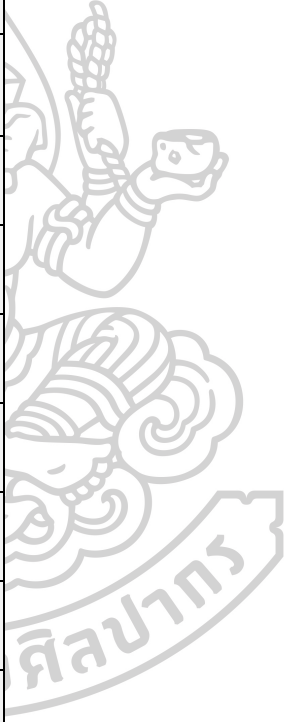


ตารางที่ 46 รายงานผลคะแนนการสอบวัดความสามารถภาษาอังกฤษโดยใช้แบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย-การเขียน

ที่	รหัสประจำตัว	Task A			รวม	เฉลี่ย	%	Task B			รวม	เฉลี่ย	%	Writing			รวม	เฉลี่ย	%
		1	2	3				1	2	3				1	2	3			
1	601011529	15	16.5	17	48.5	16.17	80.83	32	28	34	94	31.33	78.33	47	44.5	51	142.5	47.50	79.17
2	611011556	14	16	18	48	16.00	80.00	28	32	30	90	30.00	75.00	42	48	48	138	46.00	76.67
3	611011580	13.5	13	14	40.5	13.50	67.50	28	30	34	92	30.67	76.67	41.5	43	48	132.5	44.17	73.61
4	611011551	14	16	18	48	16.00	80.00	24	28	32	84	28.00	70.00	38	44	50	132	44.00	73.33
5	591011369	12.5	13	12.5	38	12.67	63.33	24	28	30	82	27.33	68.33	36.5	41	42.5	120	40.00	66.67
6	601031260	17	16	16	49	16.33	81.67	18	24	26	68	22.67	56.67	35	40	42	117	39.00	65.00
7	601011575	16	12	14	42	14.00	70.00	20	24	28	72	24.00	60.00	36	36	42	114	38.00	63.33
8	601011583	14	10	12	36	12.00	60.00	30	22	26	78	26.00	65.00	44	32	38	114	38.00	63.33
9	591011521	12	8	10	30	10.00	50.00	18	26	28	72	24.00	60.00	30	34	38	102	34.00	56.67
10	611011581	15	12	10	37	12.33	61.67	24	18	22	64	21.33	53.33	39	30	32	101	33.67	56.11
11	581011528	14	13	15	42	14.00	70.00	16	20	18	54	18.00	45.00	30	33	33	96	32.00	53.33
12	601031239	10	9.5	10	29.5	9.83	49.17	16	22	24	62	20.67	51.67	26	31.5	34	91.5	30.50	50.83

ตารางที่ 48 รายงานผลคะแนนการสอบวัดความสามารถภาษาอังกฤษโดยใช้แบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย-การเขียน (ต่อ)

ที่	รหัสประจำตัว	Task A			รวม	เฉลี่ย	%	Task B			รวม	เฉลี่ย	%	Writing			รวม	เฉลี่ย	%
		1	2	3				1	2	3				1	2	3			
13	581011503	10	12	14	36	12.00	60.00	18	16	14	48	16.00	40.00	28	28	28	84	28.00	46.67
14	591081265	11	9.5	9.5	30	10.00	50.00	14	18	16	48	16.00	40.00	25	27.5	25.5	78	26.00	43.33
	Max	17.00	16.50	18.00	49.00	16.33	81.67	32.00	32.00	34.00	94.00	31.33	78.33	47.00	48.00	51.00	142.50	47.50	79.17
	Min	10.00	8.00	9.50	29.50	9.83	49.17	14.00	16.00	14.00	48.00	16.00	40.00	25.00	27.50	25.50	78.00	26.00	43.33
	\bar{X}	13.43	12.61	13.57	39.61	13.20	66.01	22.14	24.00	25.86	72.00	24.00	60.00	35.57	36.61	39.43	111.61	37.20	62.00
	S.D.	2.111	2.754	3.012	7.031	2.344	11.72	5.789	4.899	6.395	15.551	5.184	12.96	6.906	6.717	8.180	20.303	6.768	11.28



ตารางที่ 48 แสดงรายงานผลคะแนนการสอบวัดความสามารถภาษาอังกฤษโดยใช้แบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัยเฉพาะทักษะการเขียน พบว่า จากจำนวนผู้เข้าสอบ 45 คน มีผู้เข้าสอบที่ทำแบบทดสอบการเขียน 14 คน จากคะแนน 60 คะแนน คะแนนสูงสุดที่ผู้เข้าสอบทำได้คือ 47.50 คะแนน หรือ 79.17 % คะแนนต่ำสุด คือ 26.00 คะแนน หรือ 43.33 % ค่าเฉลี่ยเท่ากับ 37.20 และส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 6.768

เมื่อพิจารณาคำร้อยละของคะแนนเฉลี่ยรายการงาน พบว่า ผู้เข้าสอบสามารถทำ Task A ได้ดีกว่า Task B โดย Task A มีคะแนนเต็ม 20 คะแนน คะแนนสูงสุดที่ผู้เข้าสอบทำได้คือ 16.33 คะแนน หรือ 81.67 % คะแนนต่ำสุด คือ 9.83 คะแนน หรือ 49.17 % ค่าเฉลี่ยเท่ากับ 13.20 และส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 2.344 ส่วน Task B มีคะแนนเต็ม 40 คะแนน คะแนนสูงสุดที่ผู้เข้าสอบทำได้คือ 31.33 คะแนน หรือ 78.33 % คะแนนต่ำสุด คือ 16.00 คะแนน หรือ 40.00 % ค่าเฉลี่ยเท่ากับ 24.00 และส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 5.184



ตารางที่ 49 แสดงรายงานผลคะแนนการสอบวัดความสามารถภาษาอังกฤษโดยใช้แบบทดสอบ IELTS (International English Language Testing System) เฉพาะทักษะการเขียน พบว่า จากจำนวนผู้เข้าสอบ 32 คน มีผู้เข้าสอบที่ทำแบบทดสอบการเขียน 10 คน จากคะแนน 60 คะแนน คะแนนสูงสุดที่ผู้เข้าสอบทำได้คือ 45.33 คะแนน หรือ 75.56 % คะแนนต่ำสุด คือ 26.00 คะแนน หรือ 43.33 % ค่าเฉลี่ยเท่ากับ 35.18 และส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 5.805

เมื่อพิจารณาค่าร้อยละของคะแนนเฉลี่ยรายภาระงาน พบว่า ผู้เข้าสอบสามารถทำ Task A ได้ดีกว่า Task B โดย Task A มีคะแนนเต็ม 20 คะแนน คะแนนสูงสุดที่ผู้เข้าสอบทำได้คือ 15.33 คะแนน หรือ 76.67 % คะแนนต่ำสุด คือ 11.33 คะแนน หรือ 56.67 % ค่าเฉลี่ยเท่ากับ 13.32 และส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 1.282

ส่วน Task B มีคะแนนเต็ม 40 คะแนน คะแนนสูงสุดที่ผู้เข้าสอบทำได้คือ 30.67 คะแนน หรือ 76.67 % คะแนนต่ำสุด คือ 14.67 คะแนน หรือ 36.67 % ค่าเฉลี่ยเท่ากับ 21.87 และส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 4.856

ความสัมพันธ์ระหว่างคะแนนของแบบทดสอบด้านการเขียน

ในการหาความสัมพันธ์ของคะแนนสอบการเขียน ผู้วิจัยได้นำคะแนนรายบุคคลของผู้เข้าสอบที่ได้จากการทำแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test) และแบบทดสอบ IELTS (International English Language Testing System) มาหาความสัมพันธ์ เนื่องจากมีผู้เข้าสอบบางคน que เข้าสอบโดยใช้แบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test) แต่ไม่ได้เข้าสอบโดยใช้แบบทดสอบ IELTS (International English Language Testing System) ดังนั้นจึงมีผลคะแนนการเขียนที่สามารถนำมาเปรียบเทียบได้จากผู้เข้าสอบ 10 คน ดังรายละเอียดในตารางที่ 50

ตารางที่ 48 ความสัมพันธ์ของคะแนนแบบทดสอบอัตนัย (การเขียน) ระหว่างคะแนน จากแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัยและแบบทดสอบ IELTS (International English Language Testing System)

ที่	รหัสประจำตัว	Task A		Task B		รวม	
		UEPT	IELTS	UEPT	IELTS	UEPT	IELTS
		20	20	40	40	60	60
1	581011503	12.00	12.67	16.00	19.33	28.00	32.00
2	581011528	14.00	14.67	18.00	16.67	32.00	31.33
3	591011521	10.00	13.33	24.00	20.00	34.00	33.33
4	591081265	10.00	11.33	16.00	14.67	26.00	26.00
5	601011529	16.17	16.17	31.33	30.67	47.50	46.83
6	601011575	14.00	13.67	24.00	24.00	38.00	37.67
7	601031239	9.83	11.83	20.67	20.67	30.50	32.50
8	601031260	16.33	17.00	22.67	21.33	39.00	38.33
9	611011580	13.50	13.00	30.67	26.67	44.17	39.67
10	611011581	12.33	12.33	21.33	23.33	33.67	35.67
Correlation (r)		0.88		0.90		0.95	

หมายเหตุ: UEPT = University English Proficiency Test, IELTS = International English Language Testing System

ตารางที่ 50 แสดงความสัมพันธ์ของคะแนนแบบทดสอบอัตนัย (การเขียน) ระหว่างคะแนน ของแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัยและแบบทดสอบ IELTS (International English Language Testing System) ของผู้เข้าสอบ จำนวน 10 คน โดยภาพรวม พบว่า คะแนนสอบในทักษะการเขียนมีความสัมพันธ์กันในระดับสูง โดยมีค่าความสัมพันธ์เท่ากับ .95

เมื่อพิจารณารายภาระงาน พบว่าค่าความสัมพันธ์ของคะแนนการเขียนทั้ง 2 ภาระงานยังคง อยู่ในระดับสูง โดย Task A มีค่าความสัมพันธ์ของคะแนนเท่ากับ .88 และ Task B มีค่าความสัมพันธ์ ของคะแนนเท่ากับ .90

สรุปได้ว่า คะแนนแบบทดสอบการเขียนของแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test) และแบบทดสอบ IELTS (International English Language Testing System) มีความสัมพันธ์กันสูงกว่าเกณฑ์ที่กำหนด ทั้ง ภาพรวมและรายภาระงาน ($r > .75$)

4.3.2.2 ความสัมพันธ์ระหว่างคะแนนของแบบทดสอบด้านการพูด

ในการเปรียบเทียบความตรงของแบบทดสอบอัตนัย-การพูด ผู้วิจัยได้ดำเนินการ 2 ขั้นตอนคือ 1) การจัดทำรายงานคะแนนการสอบ และ 2) การหาความสัมพันธ์ของคะแนนสอบ ดังรายละเอียดต่อไปนี้

การจัดทำรายงานคะแนนการสอบ

ในการจัดทำรายงานคะแนนการสอบพูด ผู้วิจัยได้นำคะแนนของผู้เข้าสอบจำนวน 21 คน ที่เข้าสอบโดยใช้แบบทดสอบทักษะการพูดของแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test) และแบบทดสอบ แบบทดสอบ IELTS (International English Language Testing System) โดยดำเนินการสอบโดยกรรมการทดสอบการพูด 3 คนและให้คะแนนตามเกณฑ์การให้คะแนนการสอบพูด (ภาคผนวก 5) และผู้วิจัยนำคะแนนทั้ง 3 ภาระงาน ของผู้เข้าสอบทุกคนที่ได้จากคณะกรรมการทดสอบการพูด คำนวณหาค่าสูงสุด (Max) ค่าต่ำสุด (Min) ค่าเฉลี่ย (\bar{X}) ส่วนเบี่ยงเบนมาตรฐาน (S.D.) และร้อยละ (Percentage) และจัดเรียงลำดับคะแนนที่ผู้เข้าสอบทำได้ ดังแสดงในตารางที่ 51



ตารางที่ 49 รายงานผลคะแนนการสอบวัดความสามารถภาษาอังกฤษโดยใช้แบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย-การพูด

ที่	รหัสประจำตัว	Task A			รวม	%	Task B			รวม	%	Task C			รวม	%	Speaking			รวม	%
		1	2	3			1	2	3			1	2	3			1	2	3		
1	581011523	7	7	6	6.67	33.33	8	6	5	6.33	31.67	8	5	6	6.33	31.67	23	18	17	19.33	32.22
2	581011528	13	12	12	12.33	61.67	12	13	15	13.33	66.67	12	15	13	13.33	66.67	37	40	40	39.00	65.00
3	581011538	8	11	9	9.33	46.67	8	10	9	9.00	45.00	12	11	10	11.00	55.00	28	32	28	29.33	48.89
4	581011539	8	8	6	7.33	36.67	7	7	4	6.00	30.00	7	4	6	5.67	28.33	22	19	16	19.00	31.67
5	591011381	9	5	7	7.00	35.00	8	6	7	7.00	35.00	7	5	6	6.00	30.00	24	16	20	20.00	33.33
6	591011521	15	13	14	14.00	70.00	12	15	13	13.33	66.67	13	15	13	13.67	68.33	40	43	40	41.00	68.33
7	591081055	6	6	6	6.00	30.00	5	8	5	6.00	30.00	6	7	6	6.33	31.67	17	21	17	18.33	30.56
8	591081262	7	5	7	6.33	31.67	8	6	5	6.33	31.67	8	8	6	7.33	36.67	23	19	18	20.00	33.33
9	601011529	12	14	15	13.67	68.33	14	15	12	13.67	68.33	14	12	16	14.00	70.00	40	41	43	41.33	68.89
10	601011533	7	8	6	7.00	35.00	6	6	5	5.67	28.33	9	5	9	7.67	38.33	22	19	20	20.33	33.89
11	601011553	12	9	10	10.33	51.67	9	10	11	10.00	50.00	10	11	10	10.33	51.67	31	30	31	30.67	51.11
12	601011575	13	11	13	12.33	61.67	12	13	13	12.67	63.33	11	13	13	12.33	61.67	36	37	39	37.33	62.22
13	601031239	12	14	13	13.00	65.00	13	13	13	13.00	65.00	12	13	13	12.67	63.33	37	40	39	38.67	64.44
14	601031260	12	16	14	14.00	70.00	16	15	12	14.33	71.67	13	15	13	13.67	68.33	41	46	39	42.00	70.00
15	601031481	8	11	11	10.00	50.00	12	10	9	10.33	51.67	10	9	10	9.67	48.33	30	30	30	30.00	50.00
16	611011553	10	11	10	10.33	51.67	11	10	11	10.67	53.33	11	11	10	10.67	53.33	32	32	31	31.67	52.78

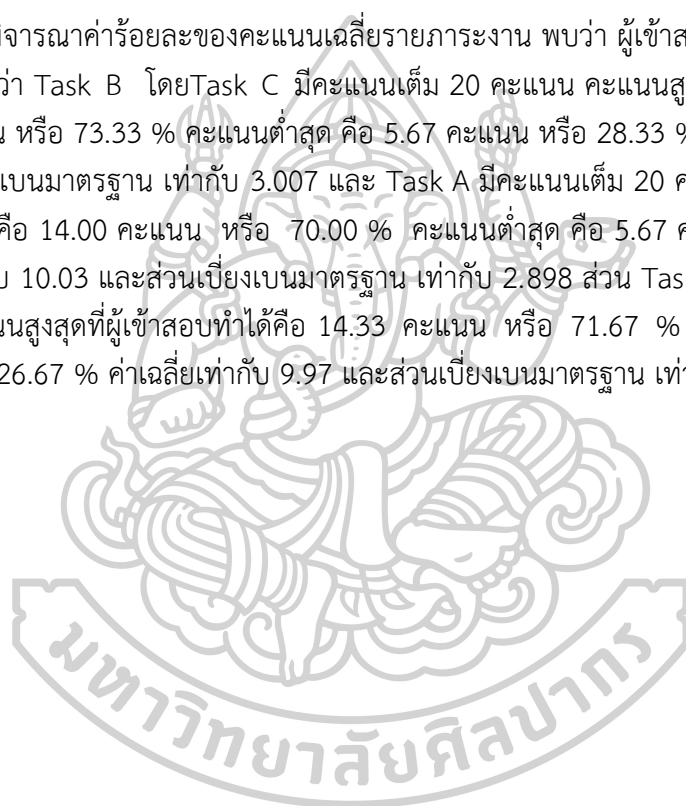
ตารางที่ 51 รายงานผลคะแนนการสอบวัดความสามารถภาษาอังกฤษโดยใช้แบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย-การพูด (ต่อ)

ที่	รหัสประจำตัว	Task A			รวม	%	Task B			รวม	%	Task C			รวม	%	Speaking			รวม	%	
		1	2	3			1	2	3			1	2	3								
17	611011559	10	12	10	10.67	53.33	11	12	11	11.33	56.67	12	11	10	11.00	55.00	33	35	31	33.00	55.00	
18	611011576	12	11	10	11.00	55.00	10	10	11	10.33	51.67	13	11	10	11.33	56.67	35	32	31	32.67	54.44	
19	611011580	13	14	15	14.00	70.00	14	14	15	14.33	71.67	16	15	13	14.67	73.33	43	43	43	43.00	71.67	
20	611011600	5	5	7	5.67	28.33	5	6	5	5.33	26.67	8	5	6	6.33	31.67	18	16	18	17.33	28.89	
21	611011665	8	11	10	9.67	48.33	11	9	11	10.33	51.67	11	11	12	11.33	56.67	30	31	33	31.33	52.22	
	Max	15.00	16.00	15.00	14.00	70.00	16.00	15.00	15.00	14.33	71.67	16.00	15.00	16.00	14.67	73.33	43.00	46.00	43.00	43.00	43.00	71.67
	Min	5.00	5.00	6.00	5.67	28.33	5.00	6.00	4.00	5.33	26.67	6.00	4.00	6.00	5.67	28.33	17.00	16.00	16.00	17.33	28.89	
	\bar{X}	9.86	10.19	10.05	10.03	50.16	10.10	10.19	9.62	9.97	49.84	10.62	10.10	10.05	10.25	51.27	30.57	30.48	29.71	30.25	50.42	
	S.D.	2.798	3.281	3.138	2.898	14.490	3.064	3.281	3.598	3.173	15.863	2.617	3.727	3.090	3.007	15.036	7.859	9.943	9.535	8.993	14.988	

ตารางที่ 51 แสดงรายงานผลคะแนนการสอบวัดความสามารถภาษาอังกฤษโดยใช้แบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัยเฉพาะทักษะการพูด จากจำนวนผู้เข้าสอบ 45 คน มีการคัดเลือกผู้เข้าสอบที่ได้คะแนนการสอบครั้งที่ 1 อยู่ในกลุ่มสูง กลาง และต่ำ กลุ่มละ 7 คน รวม 21 คน มาทดสอบการพูด พบว่า

จากคะแนน 60 คะแนน คะแนนสูงสุดที่ผู้เข้าสอบทำได้คือ 43.00 คะแนน หรือ 71.67 % คะแนนต่ำสุด คือ 17.33 คะแนน หรือ 28.89 % ค่าเฉลี่ยเท่ากับ 30.25 และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 8.995

เมื่อพิจารณาค่าร้อยละของคะแนนเฉลี่ยรายการงาน พบว่า ผู้เข้าสอบสามารถทำ Task C และ A ได้ดีกว่า Task B โดย Task C มีคะแนนเต็ม 20 คะแนน คะแนนสูงสุดที่ผู้เข้าสอบทำได้คือ 14.67 คะแนน หรือ 73.33 % คะแนนต่ำสุด คือ 5.67 คะแนน หรือ 28.33 % ค่าเฉลี่ยเท่ากับ 10.25 และส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 3.007 และ Task A มีคะแนนเต็ม 20 คะแนน คะแนนสูงสุดที่ผู้เข้าสอบทำได้คือ 14.00 คะแนน หรือ 70.00 % คะแนนต่ำสุด คือ 5.67 คะแนน หรือ 28.33 % ค่าเฉลี่ยเท่ากับ 10.03 และส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 2.898 ส่วน Task B มีคะแนนเต็ม 20 คะแนน คะแนนสูงสุดที่ผู้เข้าสอบทำได้คือ 14.33 คะแนน หรือ 71.67 % คะแนนต่ำสุด คือ 5.33 คะแนน หรือ 26.67 % ค่าเฉลี่ยเท่ากับ 9.97 และส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 3.173



ตารางที่ 50 รายงานผลคะแนนการสอบวัดความสามารถภาษาอังกฤษโดยใช้แบบทดสอบ IELTS (International English Language Testing System) - การพูด

ที่	รหัสประจำตัว	Task A			รวม	%	Task B			รวม	%	Task C			รวม	%	Speaking			รวม	%
		1	2	3			1	2	3			1	2	3							
1	581011523	6	7	6	6.33	31.67	5	6	5	5.33	26.67	6	4	7	5.67	28.33	17	17	18	17.33	28.89
2	581011528	13	14	14	13.67	68.33	14	13	15	14.00	70.00	15	12	13	13.33	66.67	42	39	42	41.00	68.33
3	581011538	13	14	13	13.33	66.67	14	13	13	13.33	66.67	14	13	13	13.33	66.67	41	40	39	40.00	66.67
4	581011539	6	9	6	7.00	35.00	5	7	5	5.67	28.33	6	8	6	6.67	33.33	17	24	17	19.33	32.22
5	591011350	7	9	6	7.33	36.67	5	8	5	6.00	30.00	5	5	8	6.00	30.00	17	22	19	19.33	32.22
6	591011381	7	5	6	6.00	30.00	5	7	5	5.67	28.33	5	6	7	6.00	30.00	17	18	18	17.67	29.44
7	591011521	13	15	12	13.33	66.67	14	15	16	15.00	75.00	14	12	15	13.67	68.33	41	42	43	42.00	70.00
8	591081055	5	8	6	6.33	31.67	5	6	4	5.00	25.00	6	4	5	5.00	25.00	16	18	15	16.33	27.22
9	591081262	5	5	6	5.33	26.67	5	9	4	6.00	30.00	6	4	9	6.33	31.67	16	18	19	17.67	29.44
10	601011529	15	16	13	14.67	73.33	16	14	15	15.00	75.00	14	15	16	15.00	75.00	45	45	44	44.67	74.44
11	601011533	6	8	6	6.67	33.33	5	6	5	5.33	26.67	6	5	6	5.67	28.33	17	19	17	17.67	29.44
12	601011553	10	12	11	11.00	55.00	12	12	10	11.33	56.67	11	10	10	10.33	51.67	33	34	31	32.67	54.44
13	601011575	15	12	13	13.33	66.67	15	12	12	13.00	65.00	13	12	13	12.67	63.33	43	36	38	39.00	65.00
14	601031239	14	15	14	14.33	71.67	16	14	15	15.00	75.00	14	15	14	14.33	71.67	44	44	43	43.67	72.78
15	601031260	13	15	13	13.67	68.33	15	14	12	13.67	68.33	13	12	15	13.33	66.67	41	41	40	40.67	67.78

ตารางที่ 52 รายงานผลคะแนนการสอบวัดความสามารถภาษาอังกฤษโดยใช้แบบทดสอบ IELTS (International English Language Testing System) –การพูด (ต่อ)

ที่	รหัสประจำตัว	Task A			รวม	%	Task B			รวม	%	Task C			รวม	%	Speaking			รวม	%
		1	2	3			1	2	3			1	2	3							
16	601031481	10	12	11	11.00	55.00	12	10	10	20	100	20	20	20	20	100	60	60	60	60	100
17	611011553	13	14	13	13.33	66.67	14	13	13	13.33	66.67	14	13	13	13.33	66.67	41	40	39	40.00	66.67
18	611011559	10	14	13	12.33	61.67	11	13	13	12.33	61.67	12	13	13	12.67	63.33	33	40	39	37.33	62.22
19	611011576	11	12	10	11.00	55.00	11	11	11	11.00	55.00	11	11	11	11.00	55.00	33	34	32	33.00	55.00
20	611011580	15	14	15	14.67	73.33	15	13	12	13.33	66.67	13	13	15	13.67	68.33	43	40	42	41.67	69.44
21	611011665	10	12	11	11.00	55.00	12	10	10	10.67	53.33	11	12	10	11.00	55.00	33	34	31	32.67	54.44
	Max	15.00	16.00	15.00	14.67	73.33	16.00	15.00	16.00	15.00	75.00	15.00	15.00	16.00	15.00	75.00	45.00	45.00	44.00	44.67	74.44
	Min	5.00	5.00	6.00	5.33	26.67	5.00	6.00	4.00	5.00	25.00	5.00	4.00	5.00	5.00	25.00	16.00	17.00	15.00	16.33	27.22
	\bar{X}	10.33	11.52	10.38	10.75	53.73	10.76	10.76	10.00	10.51	52.54	10.48	9.95	10.94	10.94	52.22	31.57	32.24	31.29	31.70	52.83
	S.D.	3.540	3.400	3.369	3.341	16.707	4.403	3.408	4.171	3.799	18.993	3.642	3.773	3.434	3.511	17.554	11.439	9.929	10.710	10.605	17.675

ตารางที่ 52 แสดงรายงานผลคะแนนการสอบวัดความสามารถภาษาอังกฤษโดยใช้แบบทดสอบ IELTS (International English Language Testing System) เฉพาะทักษะการพูด จากจำนวนผู้เข้าสอบ 32 คน มีการคัดเลือกผู้เข้าสอบที่ได้คะแนนการสอบครั้งที่ 1 อยู่ในกลุ่มสูง กลาง และต่ำ กลุ่มละ 7 คน รวม 21 คน มาทดสอบการพูด พบว่า

จากคะแนน 60 คะแนน คะแนนสูงสุดที่ผู้เข้าสอบทำได้คือ 44.67 คะแนน หรือ 74.44 % คะแนนต่ำสุด คือ 16.33 คะแนน หรือ 27.22 % ค่าเฉลี่ยเท่ากับ 31.70 และส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 10.605

เมื่อพิจารณาค่าร้อยละของคะแนนเฉลี่ยรายการงาน พบว่า ผู้เข้าสอบสามารถทำ Task A ได้ดีกว่า Task B และ C โดย Task A มีคะแนนเต็ม 20 คะแนน คะแนนสูงสุดที่ผู้เข้าสอบทำได้คือ 14.67 คะแนน หรือ 73.33 % คะแนนต่ำสุด คือ 5.33 คะแนน หรือ 26.67 % ค่าเฉลี่ยเท่ากับ 10.75 และส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 3.341 สำหรับ Task B มีคะแนนเต็ม 20 คะแนน คะแนนสูงสุดที่ผู้เข้าสอบทำได้คือ 15.00 คะแนน หรือ 75.00 % คะแนนต่ำสุด คือ 5.00 คะแนน หรือ 25.00 % ค่าเฉลี่ยเท่ากับ 10.51 และส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 3.799 และ Task C มีคะแนนเต็ม 20 คะแนน คะแนนสูงสุดที่ผู้เข้าสอบทำได้คือ 15.00 คะแนน หรือ 75.00 % คะแนนต่ำสุด คือ 5.00 คะแนน หรือ 25.00 % ค่าเฉลี่ยเท่ากับ 10.94 และส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 3.511

ความสัมพันธ์ระหว่างคะแนนของแบบทดสอบด้านการพูด

ในการหาความสัมพันธ์ของคะแนนสอบการพูด ผู้วิจัยได้นำคะแนนรายบุคคลของผู้เข้าสอบที่ได้จากการทำแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test) และแบบทดสอบ IELTS (International English Language Testing System) มาหาความสัมพันธ์หรือเปรียบเทียบกัน เนื่องจากจำนวนผู้สอบที่ได้รับการคัดเลือกให้เข้าทดสอบการพูดของแบบทดสอบทั้ง 2 ชนิด ไม่ได้มีรายชื่อตรงกันทั้ง 21 คน มีผู้เข้าสอบ 1 คนที่เข้าสอบพูดโดยใช้แบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test) แต่ไม่ได้เข้าสอบพูดโดยใช้แบบทดสอบ IELTS (International English Language Testing System) และอีก 1 คน เข้าสอบพูดโดยใช้แบบทดสอบ IELTS (International English Language Testing System) แต่ไม่ได้เข้าสอบพูดโดยใช้แบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test) ดังนั้นจึงมีผลคะแนนการพูดที่สามารถนำมาเปรียบเทียบได้คือจากผู้เข้าสอบ 20 คน ดังรายละเอียดในตารางที่ 53

ตารางที่ 51 แสดงความตรงของคะแนนแบบทดสอบอัตนัย (การพูด) ระหว่างคะแนน
จากแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัยและแบบทดสอบ
IELTS (International English Language Testing System)

ที่	รหัส ประจำตัว	Task A		Task B		Task C		คะแนนรวม	
		UEPT	IELTS	UEPT	IELTS	UEPT	IELTS	UEPT	IELTS
		20	20	20	20	20	20	60	60
1	581011523	12.33	6.33	6.33	5.33	6.33	5.67	19.33	17.33
2	581011528	9.33	13.67	13.33	14.00	13.33	13.33	39.00	41.00
3	581011538	7.33	13.33	9.00	13.33	11.00	13.33	29.33	40.00
4	581011539	7.00	7.00	6.00	5.67	5.67	6.67	19.00	19.33
5	591011381	14.00	6.00	7.00	5.67	6.00	6.00	20.00	17.67
6	591011521	6.00	13.33	13.33	15.00	13.67	13.67	41.00	42.00
7	591081055	6.33	6.33	6.00	5.00	6.33	5.00	18.33	16.33
8	591081262	13.67	5.33	6.33	6.00	7.33	6.33	20.00	17.67
9	601011529	7.00	14.67	13.67	15.00	14.00	15.00	41.33	44.67
10	601011533	10.33	6.67	5.67	5.33	7.67	5.67	20.33	17.67
11	601011553	12.33	11.00	10.00	11.33	10.33	10.33	30.67	32.67
12	601011575	13.00	13.33	12.67	13.00	12.33	12.67	37.33	39.00
13	601031239	14.00	14.33	13.00	15.00	12.67	14.33	38.67	43.67
14	601031260	10.00	13.67	14.33	13.67	13.67	13.33	42.00	40.67
15	601031481	10.33	11.00	10.33	10.67	9.67	10.33	30.00	32.00
16	611011553	10.67	13.33	10.67	13.33	10.67	13.33	31.67	40.00
17	611011559	11.00	12.33	11.33	12.33	11.00	12.67	33.00	37.33
18	611011576	14.00	11.00	10.33	11.00	11.33	11.00	32.67	33.00
19	611011580	9.67	14.67	14.33	13.33	14.67	13.67	43.00	41.67
20	611011665	6.67	11.00	10.33	10.67	11.33	11.00	31.33	32.67
Correlation (r)		0.93		0.93		0.94		0.95	

หมายเหตุ: UEPT = University English Proficiency Test, IELTS = International English Language Testing System

ตารางที่ 53 แสดงความตรงของคะแนนแบบทดสอบอัตนัย (การพูด) ระหว่างคะแนนของ
แบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัยและแบบทดสอบ IELTS
(International English Language Testing System) ของผู้เข้าสอบ จำนวน 20 คน โดยภาพรวม
พบว่า คะแนนสอบในทักษะการพูดมีความสัมพันธ์กันในระดับสูง โดยมีค่าความสัมพันธ์เท่ากับ .95

เมื่อพิจารณารายภาระงาน พบว่าค่าความสัมพันธ์ของคะแนนการเขียนทั้ง 3 ภาระงานยังคงอยู่ในระดับสูง โดย Task A และ B มีค่าความสัมพันธ์ของคะแนนเท่ากับ .93 และ Task C มีค่าความสัมพันธ์ของคะแนนเท่ากับ .94

สรุปได้ว่า คะแนนทักษะการพูดของผู้เข้าสอบโดยใช้แบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test) และแบบทดสอบ IELTS (International English Language Testing System) มีความสัมพันธ์กันสูงกว่าเกณฑ์ที่กำหนด ทั้งภาพรวมและรายภาระงาน ($r_{xy} \geq .75$)



ตารางที่ 52 สรุปความสัมพันธ์ระหว่างคะแนนของแบบทดสอบวัดความสามารถภาษาอังฤษระดับมหาวิทยาลัยและแบบทดสอบมาตรฐานสากลในแต่ละด้าน

r	แบบทดสอบวัดความสามารถภาษาอังฤษระดับมหาวิทยาลัย																			
	ทักษะ	A	B	C	ฟัง	A	B	ใช้ภาษา	อ่าน	ปรนัย	A	B	เขียน	A	B	C	พูด	อัตนัย	ฟังฉบับ	
	A	.91																		
	B		.37																	
	C			.65																
	ฟัง				.93															
	A					.59														
	B						.71													
	ใช้ภาษา							.76												
	อ่าน								.87											
	ปรนัย									.85										
	A										.88									
	B											.90								
	เขียน												.95							
	A													.93						
	B														.93					
	C															.94				
	พูด																.95			
	อัตนัย																	.95		
	ฟังฉบับ																		.90	

คณะศึกษาศาสตร์ มหาวิทยาลัยขอนแก่น

รายการอ้างอิง

- Alderson, J. C., Clapham, C., and Wall, D. . (1995). *Language Test Constructed and Evaluation*. Cambridge: Cambridge University Press.
- Arkoudis, & Sophie et.al. (2009). The Impact of English Language Proficiency and Workplace Readiness on the Employment Outcomes of Tertiary International Students. Retrieved from https://internationaleducation.gov.au/research/Publications/Documents/ELP_Full_Report.pdf
- Arunee Wiriyachitra. (n.d.). English Language Teaching and Learning in Thailand in this Decade. Retrieved from <http://apecknowledgebank.org/resources/downloads/English>
- ASTV ผู้จัดการออนไลน์. (2558). ศธ.เตรียมปรับแนวทางการเรียนภาษาอังกฤษ เล็งให้บัณฑิตทำการทดสอบก่อนจบ. Retrieved from <http://www.manager.co.thHome/ViewNews.aspx?NewsID=9580000117933>
- Bachman, & Lyle F. (2004). *Statistical Analyses for Language Assessment*. Cambridge: New York: Cambridge University Press.
- Bachman, & Lyle F. & Palmer, A. S. (1996). *Language Testing in Practice: Designing and Developing Useful Language Tests pp. 86-89*. Oxford: Oxford University Press.
- British Council. (2015). IELTS สำหรับวีซ่าและการตรวจคนเข้าเมืองแห่งสหราชอาณาจักร. Retrieved from <https://www.britishcouncil.or.th/exam/ielts-uk-visa-immigration>
- Brown, H. D. (2004). *Language Assessment: Principles and Classroom Practices*. New York: Pearson Education, Inc.
- Brown, J. D. (2005). *Testing in Language Programs*. Upper Saddle River, N.J.: Prentice Hall Regents.
- Cohen, & Andrew D. (1994). *Assessing Language Ability in the Classroom*. Boston, Mass.: Heinle & Heinle Publishers.
- Council of Europe. (2003). Common European Framework of Reference for Languages: Learning, Teaching, Assessment. Retrieved from http://www.coe.int/t/dg4/linguistic/Source/Framework_EN.pdf

- Djiwandono, & Patrisius Istiarto. (2006). Cultural Bias in Language Testing. Retrieved from <http://teflin.org/journal/index.php/journal/article/download/61/257>
- Educational Testing Service. (2010). TOEFL iBT Research Insight: TOEFL iBT Test Framework and Test Development. Retrieved from https://www.ets.org/s/toefl/pdf/toefl_ibt_research_insight.pdf
- EF Education First. (2015). "EF Proficiency Index 2015.". Retrieved from <http://www.ef.co.th/epi/regions/asia/thailand/>
- Gao, F. (2010). "What's Wrong with Current Chinese College English Assessment System? Reform or Not?" in *International Education Studies*, 3(1).
- Genesee, & Fred & Upshur, J. A. (1996). *Classroom-Based Evaluation in Second Language Education*. Cambridge: Cambridge University Press.
- Heaton, J. B. (1975). *Writing English Language Tests*. London: Longman.
- Henning, G. (1987). *A Guide to Language Testing*. Massachusetts: Newbury House Publishers.
- Hughes, A. (1989). *Testing for Language Teachers*. Cambridge: Cambridge University Press.
- IELTS. (2015). Information for Candidates. Retrieved from <http://takeielts.britishcouncil.org/sites/default/files/Information%20for%20Candidates.pdf>
- IMET. (2555). ข้อสรุปการระดมความคิดเห็นการประชุมผู้บริหารสถาบันอุดมศึกษา เรื่อง ความสามารถในการแข่งขันของสถาบันอุดมศึกษาไทย เมื่อเข้าสู่ประชาคมเศรษฐกิจอาเซียน (AEC) ปี 2558. Retrieved from <http://www.imetthai.com/news-15.php>
- Ito, A. (2005). A Validation Study on the English language test in a Japanese Nationwide University Entrance Examination. *The Asian EFL Journal*. Retrieved from http://www.asian-efl-journal.com/June_05_ai.pdf
- Ivanova , V. (2011). Construction and Evaluation of Achievement Tests in English. Retrieved from <http://sci-gems.math.bas.bg/jspui/bitstream/10525/1554/1/adis-may-2011-276p-285p.pdf>
- Jin, Y. (2011). "Fundamental Concerns in High-Stakes Language Testing: The Case of the College English Test". in *Pan-Pacific Association of Applied Linguistics.*, 15(2), 71-

83.

Jobsdb.com. (2015). 3 เหตุผลที่คนทำงานไม่สนใจภาษาอังกฤษไม่ได้แล้ว. Retrieved from

<http://th.jobsdb.com/th-th/articles/%E0%B8%A0%E0%B8%B2%E0%B8%A9%E0%B8%B2%E0%B8%AD%E0%B8%B1%E0%B8%87%E0%B8%81%E0%B8%A4%E0%B8%A9%E0%B8%AA%E0%B8%B3%E0%B8%84%E0%B8%B1%E0%B8%8D%E0%B8%AD%E0%B8%A2%E0%B9%88%E0%B8%B2%E0%B8%87%E0%B9%84%E0%B8%A3>

Khoii, R. S., Nazli. (2012). A Fairness Issue: Test Method and the Validity of Grammar Sub-tests of High-Stakes Admissions Tests. Retrieved from

http://www.academia.edu/3658679/A_Fairness_Issue_Test_Method_Facet_and_the_Validity_of_Grammar_Subtests_of_High-Stakes_Admissions_Tests

LCRT Consulting. (2013). MELA Validation Project: Final Report. Retrieved from

<http://www.themela.com/assets/documents/MELAValidationProjectFinalReport2013.pdf>

Lopez, & Alexis Augusto & Janssen, G. (2010). Validation Study of Colombia's ECAES English Exam. *in Lenguaje*, 38(2), 423-448.

Madsen, & Harold S. (1983). *Techniques in Testing*. Oxford: Oxford University Press.

Malaysia Examination Council. (2006). Malaysia University English Test (MUET). Retrieved from <http://www.mpm.edu.my/documents/10156/c5c332ab-3d97-4959-83c0-09866eea0774>

McNamara Tim. (2000). *Language Testing*. Oxford: Oxford University Press.

Munson Robin, & Parton Cinda. (2013). Bias and Fairness in State Testing. Retrieved from <http://www.k12.wa.us/LegisGov/2013documents/BiasAndFairnessInStateTestingDec2013.pdf>

Paltridge, & Brian. (2007). "Beyond the text: A textography of Chinese College English writing" *in University of Sydney Papers in TESOL*, 2(2), 149-165.

Pan Yi-Ching. (2009a). "Evaluating the appropriateness and consequences of test use" *in Colomb. Appl. Linguist. J.* (11), Bogotá, Colombia. 93-105.

Pan Yi-Ching. (2009b). "Test Impact: English Certification Exit Requirements in Taiwan". *in TEFLIN Journal*(20).

Peterson Doug. (2012). Avoiding Bias and Stereotypes – Test Design & Delivery Part 5.

- Retrieved from <http://blog.questionmark.com/avoiding-bias-and-stereotypes-test-design-delivery-part-5>
- Pizorn, Karmen, & Moe Ali. (2012). A Validation Study of the National Assessment Instruments for Young English Language Learners in Norway and Slovenia. Retrieved from <http://www.dlib.si/stream/URN:NBN:SI:doc-BNLUTFZY/bd80a568-429d-415a-88a5-71a993368fa9/PDF>
- Riazi, & Mehdi. (2014). Research Note; Concurrent and Predictive Validating of Pearson Test of English Academic (PTE Academic). Retrieved from http://pearsonpte.com/wp-content/uploads/2014/07/Riazi_M_2014.pdf
- Saint John's for Education. (2555). AEC กับการปรับตัวของอุดมศึกษาไทย ในการก้าวสู่ประชาคมอาเซียน. Retrieved from <https://wiki.stjohn.ac.th/groups/aec/wiki/d2461/>
- Samaie, & Mahmoud. (2014). Achievement Test Development and Validation: A Measure of Reading Comprehension Strategies for Iranian Learners of English, in International Journal of Linguistics 6 (2). Retrieved from http://www.macrothink.org/journal/index.php/ijl/article/download/5355/pdf_78
- telc GmbH. (2012). telc Language Test: Handbook English. Retrieved from https://www.telc.net/fileadmin/user_upload/handbook_english_b1-b2.pdf
- The European Association for Quality Language Services. (2009). Calibrating tests to the CEFR: A simplified guide for EAQUALS Members on calibrating entry or progress tests to the CEFR. Retrieved from <http://clients.squareeye.net/uploads/eaquals2011/documents/>
- TNN Thailand 24 Hours. (2558). ศธ.จ่อเห็นศ.ป.ตรีสอบภาษาอังกฤษก่อนจบ. Retrieved from http://www.tnnthailand.com/news_detail.php?id=86322&t=news
- UCLES. (2014). Cambridge English First: Information for Candidates for Exams. Retrieved from <http://www.cambridgeenglish.org/images/181310-first-information-for-candidates-2015-document.pdf>
- Urmston Alan, Raquel, Michelle, & Tsang Carrie. (2013). "Diagnostic Testing of Hong Kong Tertiary Students' English Language Proficiency: The Development and Validation of DELTA". in *Hong Kong Journal of Applied Linguistics*, 14 (2)(2013), 60-82.
- Weyant, Katie, Chisholm, & Amanda. (2014). Safeguarding Fairness Principles through the

- Test Development Process: A Tale of Two Organizations Cambridge English: Research Notes 55 (3-6). Retrieved from <http://www.cambridgeenglish.org/images/164055-research-notes-55-document.pdf>
- Whiting, & Gilman. (2009). Cultural Bias in Testing. Retrieved from <http://www.education.com/reference/article/cultural-bias-in-testing/>
- Wu, & Mei. (2012). Comparing PETS and GEPT in China and Taiwan. Retrieved from <http://files.eric.ed.gov/fulltext/EJ1079501.pdf>
- Wu, & Rachel Yi-Fen. (2011). "Establishing the Validity of the General English Proficiency Test Reading Component through a Critical Evaluation on Alignment with the Common European Framework of Reference." A thesis submitted to the University of Bedfordshire in partial fulfillment of the requirement of the degree of Doctor of Philosophy. Retrieved from <http://uobrep.openrepository.com/uobrep/bitstream/10547/223000/1/wu>
- Zheng Ying, & Cheng Liying. (2008). Test Review: College English Test (CET) in China in Language Testing. pp 408-409. Retrieved from <http://ltj.sagepub.com/content/25/3/408.full.pdf>
- Zhou Yuemei. (2004). Comparability Study of Two National EFL Tests (CET-6 and TEM-4) in China. Retrieved from http://www.asiatefl.org/main/download_pdf.php?i=4&c=1391758451
- จุฬาลงกรณ์มหาวิทยาลัย. (2557). Chulalongkorn University Test of English Proficiency. Retrieved from http://www.atc.chula.ac.th/th_html/th_tep.html
- นันท์ตา เพชรภรณ์. (2553). ความคาดหวังของผู้ใช้บัณฑิตที่มีต่อคุณลักษณะของบัณฑิตที่สำเร็จการศึกษาระดับปริญญาตรี ปีการศึกษา 2553 สาขาวิชาการตลาด วิทยาลัยราชพฤกษ์. Retrieved from http://www.rc.ac.th/Library_web/doc/RC_RR/2553_Market_Nantita.pdf
- มหาวิทยาลัยธรรมศาสตร์. (2557). TU-GET (Thammasat University General English Test). Retrieved from <http://litu.tu.ac.th/2012/index.php/th/litu-tests>
- มหาวิทยาลัยศิลปากร. (2557). Silpakorn English Proficiency Test. . Retrieved from http://www.sept.su.ac.th/pdf/about_sept.pdf
- ศิริพร พงษ์สุรพิพัฒน์, กรองแก้ว กรรณสุด, & และสุภาณี ชินวงศ์. (2544). ความตรงร่วมสมัย: การเปรียบเทียบคะแนนสอบแบบทดสอบสมรรถภาพทั่วไปทางภาษาอังกฤษของจุฬาลงกรณ์มหาวิทยาลัย (CU-TEP) กับ

คะแนนแบบทดสอบ TOEFL. สถาบันภาษา จุฬาลงกรณ์มหาวิทยาลัย. วารสารภาษาปริทัศน์, 19(2544), 32-48.

สมเกียรติ อ่อนวิมล. (2554). การใช้ภาษาอังกฤษกับอนาคตของไทยในอาเซียน. Retrieved from http://www.englishexpress.ac.th/index.php?option=com_content&view=article&id=108:2012-03-11-02-09-50&catid=10:news-of-institute-of-english-express&Itemid=15

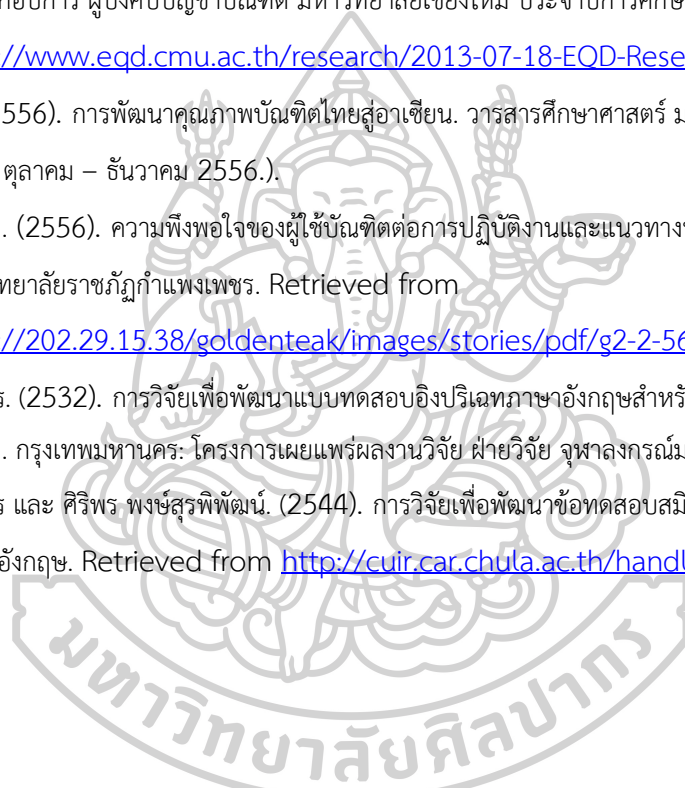
สำนักพัฒนาคุณภาพการศึกษา มหาวิทยาลัยเชียงใหม่. (2555). รายงานการวิจัยเรื่องความพึงพอใจของนายจ้างผู้ประกอบการ ผู้บังคับบัญชาบัณฑิต มหาวิทยาลัยเชียงใหม่ ประจำปีการศึกษา 2555. Retrieved from <http://www.eqd.cmu.ac.th/research/2013-07-18-EOD-Research.pdf.pdf>

สุบิน ยุระรัช. (2556). การพัฒนาคุณภาพบัณฑิตไทยสู่อาเซียน. วารสารศึกษาศาสตร์ มหาวิทยาลัยนเรศวร ปีที่ 15, 14(4 ตุลาคม – ธันวาคม 2556.).

สุรัชย์ โกศิยะกุล. (2556). ความพึงพอใจของผู้ใช้บัณฑิตต่อการปฏิบัติงานและแนวทางพัฒนาการผลิตบัณฑิต มหาวิทยาลัยราชภัฏกำแพงเพชร. Retrieved from <http://202.29.15.38/goldenteak/images/stories/pdf/g2-2-56>

อัจฉรา วงศ์โสธร. (2532). การวิจัยเพื่อพัฒนาแบบทดสอบอิงปริเฉทภาษาอังกฤษสำหรับใช้กับนักศึกษาไทยในระดับต่าง ๆ. กรุงเทพมหานคร: โครงการเผยแพร่ผลงานวิจัย ฝ่ายวิจัย จุฬาลงกรณ์มหาวิทยาลัย.

อัจฉรา วงศ์โสธร และ ศิริพร พงษ์สุรพิพัฒน์. (2544). การวิจัยเพื่อพัฒนาข้อทดสอบสมิทภาพทั่วไปทางภาษาอังกฤษ. Retrieved from <http://cuir.car.chula.ac.th/handle/123456789/757>



ภาคผนวก ก

รายชื่อผู้เชี่ยวชาญและหนังสือแต่งตั้งผู้เชี่ยวชาญตรวจเครื่องมือวิจัย

รายชื่อผู้เชี่ยวชาญ

- | | |
|---|--|
| 1. อาจารย์พรทิพย์ บดีพงษ์ | ผู้เชี่ยวชาญ
ข้าราชการบำนาญ
อดีตอาจารย์ประจำสาขาวิชาภาษาตะวันตก
คณะมนุษยศาสตร์และสังคมศาสตร์
มหาวิทยาลัยทักษิณ |
| 2. อาจารย์ ดร. ศิริรัตน์ สีนประจักษ์ผล | ผู้เชี่ยวชาญ
ข้าราชการบำนาญ
อดีตอาจารย์ประจำสาขาวิชาภาษาตะวันตก
คณะมนุษยศาสตร์และสังคมศาสตร์
มหาวิทยาลัยทักษิณ |
| 3. Mr. Neil C. Cuthbertson
อาจารย์ประจำโปรแกรมภาษาอังกฤษ | ผู้เชี่ยวชาญ
คณะศิลปศาสตร์ |
| มหาวิทยาลัยเทคโนโลยีราชมงคลศรีวิชัย | |



ประวัติอาจารย์พรทิพย์ บดีพงศ์

1. ชื่อ-สกุล นางสาวพรทิพย์ บดีพงศ์
2. ตำแหน่งทางวิชาการ อาจารย์
3. สาขาที่เชี่ยวชาญ ภาษาอังกฤษ การสอนภาษาอังกฤษเป็นภาษาต่างประเทศ การวัดและประเมินผลทางภาษาอังกฤษ การใช้เทคโนโลยีสารสนเทศในการสอนและการผลิตสื่อการสอนภาษาอังกฤษ การใช้คอมพิวเตอร์ในการเรียนการสอนภาษา
4. สังกัด สาขาวิชาภาษาตะวันตก คณะมนุษยศาสตร์และสังคมศาสตร์

5. การศึกษา

ชื่อวุฒิการศึกษา	สาขา/ วิชาเอก	สถาบันการศึกษา	ปีที่
ศศ.บ. (ศึกษาศาสตร์) เกียรตินิยมอันดับสอง	ภาษาอังกฤษ	มหาวิทยาลัยสงขลานครินทร์ วิทยาเขตปัตตานี	2518
MA in Media Technology for the Teaching of English as a Foreign Language with Distinction	การสอนภาษาอังกฤษ	University of Newcastle upon Tyne, United Kingdom	2544

อื่นๆ:

- ได้รับทุนจากรัฐบาลนิวซีแลนด์ ภายใต้แผนโคลัมโบ ให้ไปเรียนหลักสูตร Diploma in the Teaching of English as a Second Language (การสอนภาษาอังกฤษเป็นภาษาที่ 2) ที่ Victoria University of Wellington, เมือง Wellington, New Zealand ปี พ.ศ. 2527-2548
- ได้รับทุนจากรัฐบาลประเทศสหราชอาณาจักร ภายใต้แผนโคลัมโบ ให้ไปเรียนหลักสูตร Certificate in PRINSELT (The Principles and Practice of In-Service Training for English Language Teaching) จาก University College Plymouth St Mark & St John, เมือง Plymouth, United Kingdom, 2533
- ได้รับทุนจากบริติชเคาน์ซิล ประเทศไทย ให้ไปอบรม หลักสูตร Assessing Young Learners: Classroom-based Approaches ในโครงการ Regional Hornby School ที่จัดขึ้น ณ กรุงกัวลาลัมเปอร์ ประเทศมาเลเซีย ระหว่างวันที่ 14-18 พฤศจิกายน 2554

6. ผลงานทางวิชาการ

บทความทางวิชาการ

พรทิพย์ บดีพงศ์. Learning Diaries as a Tool for Reflection in a Short-Term Teacher Training Course for CALL (computer-assisted language learning) เสนอที่การประชุมวิชาการระดับนานาชาติประจำปี ครั้งที่ 25 ของสมาคมครูผู้สอนภาษาอังกฤษแห่งประเทศไทย กรุงเทพมหานคร, 2548

_____ . Training teachers to use CALL in the UK and Thailand: a reflective approach เสนอและตีพิมพ์ในรายงานการประชุมวิชาการระดับนานาชาติประจำปี ครั้งที่ 26 ของสมาคมครูผู้สอนภาษาอังกฤษแห่งประเทศไทย เชียงใหม่, 2549

_____ . 'Noticing' and a Self-Access Grammar Learning CALL Programme เสนอในการประชุมวิชาการระดับนานาชาติประจำปี ครั้งที่ 27 ของสมาคมครูผู้สอนภาษาอังกฤษแห่งประเทศไทยซึ่งจัดร่วมกับสมาคมครูผู้สอนภาษาอังกฤษในภูมิภาคเอเชียแปซิฟิก ขอนแก่น, 2550

7. ประสบการณ์ทางวิชาการ

7.1 ประสบการณ์การสอน

7.1.1 ระดับปริญญาตรี 12 ปี รายวิชาที่สอน

- 0000121 ภาษาอังกฤษทั่วไป 1
- 0000122 ภาษาอังกฤษทั่วไป 2
- 0115131 ไวยากรณ์ 1
- 0115231 ไวยากรณ์ 2
- 0115421 การอ่านเชิงวิชาการ
- 0115391 เทคนิคการสอนภาษาอังกฤษ 1
- 0115491 เทคนิคการสอนภาษาอังกฤษ 2
- 0115489 คอมพิวเตอร์เพื่อการเรียนภาษา
- 0115492 เทคโนโลยีและสื่อการสอนภาษาอังกฤษ
- 0115493 การวัดและประเมินผลทางภาษา
- 0115495 คอมพิวเตอร์ในการสอนภาษา

7.1.2 ระดับปริญญาโท 5 ปี รายวิชาที่สอน

- 0115500 การอ่าน
- 0115551 ระเบียบวิธีการสอนภาษา
- 0115653 คอมพิวเตอร์เพื่อการเรียนการสอนภาษา

7.2 ประวัติการทำงาน

ตำแหน่ง	ชื่อ/ที่อยู่ของสถานที่ทำงาน	ลักษณะงาน
อาจารย์ 2 ระดับ 7	โรงเรียนสุโขทัย-โก-ลก จังหวัด นราธิวาส และ โรงเรียนนราสิกขาลัย จังหวัด นราธิวาส	<ul style="list-style-type: none"> • สอนวิชาภาษาอังกฤษในระดับมัธยมศึกษาปลาย • วิทยากรประจำศูนย์พัฒนาการเรียนการสอนภาษา (ERIC) โรงเรียนสุโขทัย-โก-ลก • ครูผู้นิเทศ(Supervisory Teacher)สาขาวิชาภาษาอังกฤษของกรมสามัญศึกษา ประจำจังหวัดนราธิวาส • ผู้ประสานงานเอ เอฟ เอส จังหวัดนราธิวาส

ตำแหน่ง	ชื่อ/ที่อยู่ของสถานที่ทำงาน	ลักษณะงาน
อาจารย์	ภาควิชาภาษาตะวันตก คณะมนุษยศาสตร์และ สังคมศาสตร์ มหาวิทยาลัย ทักษิณ	<ul style="list-style-type: none"> สอนวิชาภาษาอังกฤษพื้นฐาน ภาษาอังกฤษเฉพาะสาขา การสอนภาษาอังกฤษ การวัดและประเมินผลทางภาษา และการใช้เทคโนโลยี และการใช้คอมพิวเตอร์ในการสอนภาษาอังกฤษ เป็นวิทยากรเกี่ยวกับการเรียนการสอนภาษาอังกฤษ และการใช้คอมพิวเตอร์ในการสอนภาษาอังกฤษ (รายละเอียดในข้อ 7.3)

7.3 ประสบการณ์อื่นๆ

การเป็นวิทยากรให้กับหน่วยงานภายนอก (พ.ศ. 2547 ถึงปัจจุบัน)

เรื่อง	หน่วยงานที่จัด	สถานที่จัด	ระยะเวลา
การอบรมเรื่อง Develop Reading and Writing Skills for Critical Thinking	ศูนย์พัฒนาการเรียนการสอนภาษาอังกฤษ (ERIC) สพม. สงขลา เขต 1	โรงเรียนมหาวชิราวุธ	21-22 มิถุนายน 2555
การอบรมเรื่องการพัฒนาการเรียนการสอนภาษาอังกฤษเพื่อเสริมสร้างคุณลักษณะเยาวชนอาเซียน	สำนักพิมพ์อักษรเจริญทัศน์ ร่วมกับสำนักงานการศึกษาเอกชนจังหวัดปัตตานี ยะลา และนราธิวาส	โรงแรมปาร์ควิว ปัตตานี หอประชุม ศอบต. ยะลา หอประชุมเทศบาลเมืองนราธิวาส	18-20 มีนาคม 2555
การอบรมครูผู้สอนไม่ตรงวุฒิ/วิชาเอก ในเขตพัฒนาพิเศษเฉพาะกิจจังหวัดชายแดนใต้ (เรื่องการประเมินผลทางภาษาอังกฤษ)	สำนักงานเขตพื้นที่การศึกษาประถมศึกษา ปัตตานี เขต 2	โรงแรมหาดแก้วรีสอร์ท สงขลา	16 กันยายน 2554
การอบรมครูผู้สอนไม่ตรงวุฒิ/วิชาเอก ในเขตพัฒนาพิเศษเฉพาะกิจจังหวัดชายแดนใต้ (เรื่องการประเมินผลทางภาษาอังกฤษ)	สำนักงานเขตพื้นที่การศึกษาประถมศึกษา ปัตตานี เขต 2	โรงแรมหาดแก้วรีสอร์ท สงขลา	29 กรกฎาคม 2554
การสอนอ่าน – การอบรมครูผู้สอนสาระการเรียนรู้ภาษาต่างประเทศ (ภาษาอังกฤษ) จากโรงเรียนในสังกัดกองกำกับการตชด. และโรงเรียนในมูลนิธิชัยพัฒนา	มหาวิทยาลัยทักษิณ	มหาวิทยาลัยทักษิณ	26 สิงหาคม 2554
การอบรมครูผู้สอนภาษาอังกฤษหลักสูตร TKT Module 2 และ 3	สำนักงานเขตพื้นที่การศึกษา สงขลา เขต 2	โรงเรียนหาดใหญ่ สมบูรณกุลกัลยา	20-21 และ 27 มิย. 2553
การอบรมครูผู้สอนภาษาอังกฤษหลักสูตร TKT Module 2 และ 3	สำนักงานเขตพื้นที่การศึกษา นครศรีธรรมราช เขต 3	โรงแรมแกรนด์พาร์ค นครศรีธรรมราช	11-15 พค. 2552
การอบรมครูผู้สอนภาษาอังกฤษหลักสูตร TKT Module 2 และ 3	สำนักงานเขตพื้นที่การศึกษา สงขลา เขต 1	โรงเรียนมหาวชิราวุธ	4-8 พ.ค. 2552

เรื่อง	หน่วยงานที่จัด	สถานที่จัด	ระยะเวลา
การอบรมครูผู้สอนภาษาอังกฤษ หลักสูตรการผลิตสื่อการเรียนการสอน สอนแบบ web-based โดยใช้ คอมพิวเตอร์	โรงเรียนกำแพงวิทยา อ.ละงู จ.สตูล	โรงเรียนกำแพงวิทยา	15-16 พ.ย. 2551
การอบรมครูผู้สอนภาษาอังกฤษ หลักสูตร TKT Module 2	สำนักงานเขตพื้นที่การศึกษา นครศรีธรรมราช เขต 3	โรงแรมแกรนด์พาร์ค นครศรีธรรมราช	12-15 พค. 2551
การอบรมการทำวิจัยในชั้นเรียน วิชาภาษาอังกฤษสำหรับผู้สอน ภาษาอังกฤษในช่วงชั้นที่ 1-4	ภาควิชาภาษาตะวันตก คณะมนุษยศาสตร์และ สังคมศาสตร์	คณะมนุษยศาสตร์และ สังคมศาสตร์ มหาวิทยาลัยทักษิณ	7-9 พฤษภาคม 2551
การอบรมครูผู้สอนภาษาอังกฤษ หลักสูตร TKT Module 1 และ 2	สำนักงานเขตพื้นที่การศึกษา สงขลา เขต 1	โรงเรียนมหาวิทยาลัยราชวูธ	28 เมย. -2 พ.ค. 2551
การอบรมเชิงปฏิบัติการวัดผลและ ประเมินผลทางภาษา	สำนักงานเทศบาลนครสงขลา	โรงเรียนเทศบาล 2 (อ่อนอุทิศ)	31 มีค. - 3 เมย. 2551
โครงการสัมมนาอาจารย์ผู้สอน ภาษาอังกฤษ Cambridge Four Corners	Cambridge University Press	โรงแรมลีการ์เดนส์ หาดใหญ่ สงขลา	9 ม.ค. 2551
การอบรมครูสอนภาษาอังกฤษ ระดับการศึกษาขั้นพื้นฐานสังกัด เทศบาลนครสงขลา	ภาควิชาภาษาตะวันตก มหาวิทยาลัยทักษิณ	มหาวิทยาลัยทักษิณ	30 มีค. - 4 เมย. 2550
การอบรมเชิงปฏิบัติการการใช้ คอมพิวเตอร์ช่วยสอนภาษาอังกฤษ	ภาควิชาภาษาตะวันตก มหาวิทยาลัยทักษิณ	มหาวิทยาลัยทักษิณ	18-22 เมษายน 2548
โครงการอบรมเพื่อพัฒนาศักยภาพ ครูผู้สอนภาษาอังกฤษใน 30 จังหวัดท่องเที่ยว	สำนักงานเขตพื้นที่การศึกษา สงขลา เขต 1	สถาบันราชภัฏสงขลา จังหวัดสงขลา	มีค. - เมย.2547 (6 สัปดาห์)
การเตรียมตัวสอบ TOEFL และ IELTS และประสบการณ์การศึกษา ต่อต่างประเทศ	วิทยาลัยพยาบาลบรมราช ชนนี สงขลา	วิทยาลัยพยาบาล บรมราชชนนี สงขลา	1-5 กรกฎาคม 2547

การเป็นวิทยากรให้กับหน่วยงานภายใน

- เป็นวิทยากรในโครงการอบรมเชิงปฏิบัติการทำวิจัยในชั้นเรียนให้แก่สิดหลักสูตร กศ.บ. วิชาเอก
ภาษาอังกฤษ ชั้นปีที่ 4 และโครงการอบรมเพื่อเตรียมความพร้อมให้แก่สิดชั้นปีที่ 1 วิชาเอก
ภาษาอังกฤษ ทั้งหลักสูตร ศศ.บ. และกศ.บ.

นอกจากนี้ ยังเป็นผู้รับผิดชอบโครงการบริการวิชาการต่างๆที่จัดขึ้นโดยภาควิชาภาษาตะวันตก และสาขาวิชา
ภาษาตะวันตก ตั้งแต่ พ.ศ. 2546 มาโดยตลอด

อ.ดร.ศิริรัตน์ ลินประจักษ์ผล

ประวัติการศึกษา

- ปริญญาตรี อักษรศาสตรบัณฑิต อบ.(เกียรตินิยมอันดับสอง)
2517 วิชาเอกภาษาอังกฤษ วิชาโทภาษาฝรั่งเศส
คณะอักษรศาสตร์ มหาวิทยาลัยศิลปากร ณ พระราชวังสนามจันทร์ นครปฐม
- ปริญญาโท MAE (Master of Arts in Education with Reading endorsement Ball State
2531 University
เมือง มันทซ์, รัฐ อินเดียนา ประเทศ สหรัฐอเมริกา
- ปริญญาเอก Ph.D. สาขา Foreign and Second Language Education
2547 State University of New York at Buffalo, NY, USA.

ประสบการณ์การอบรม

- 2519 อบรมวิชาการสอนภาษาฝรั่งเศส จัดโดยกรมสามัญศึกษา กระทรวงศึกษาธิการ
และสำนักพัฒนาวัฒนธรรม สถานเอกอัครราชทูตฝรั่งเศส ระหว่างวันที่ 3 พฤษภาคม – 19
พฤษภาคม 2519
- 2523 Audio-Visual French Teaching Training Program at CAVILAM (Centre Audio-
Visuel de Langue Moderne) for nine months at Vichy, France
- 2525 7th REGIONAL INTENSIVE –COURSE FOPR THAI TEACHERS OF ENGLISH at the
British Council Centre, Bangkok from 18 October 1982 to 17 December 1982
- 2528 Teacher Training Program at Department of English as a Foreign Language
Institute of Education, University of London, England

ประวัติการทำงาน

- 2517- 2535 ครูสอนภาษาอังกฤษและภาษาฝรั่งเศส โรงเรียนเดชะปัตตนยานุกูล จังหวัดปัตตานี
- 2536 - 2539 อาจารย์วิชาภาษาอังกฤษ สาขาวิชาภาษาตะวันตก มหาวิทยาลัยศรีนครินทรวิโรฒ ภาคใต้
สงขลา
- 2540 อาจารย์วิชาภาษาอังกฤษ สาขาวิชาภาษาตะวันตก มหาวิทยาลัยทักษิณ วิทยาเขตสงขลา
- 2541- 2545 ศึกษาค้นคว้าระดับปริญญาเอก
- 2546 – 2555 อาจารย์วิชาภาษาอังกฤษ สาขาวิชาภาษาตะวันตก มหาวิทยาลัยทักษิณ
- 2555 เกษียณอายุราชการ

ผลงานทางวิชาการ

หนังสือ

ศิริรัตน์ สิ้นประจักษ์ผล (2540) ศัพท์น่ารู้ขั้นพื้นฐานสำหรับภาษาอังกฤษระดับต้น. กรุงเทพฯ: โอเดียนสโตร์ (2549) ACTIVE-PASSIVE. กรุงเทพฯ: โอเดียนสโตร์

บทความวิชาการ

ศิริรัตน์ สิ้นประจักษ์ผล. (2537). “อ่านหนังสือให้เด็กฟัง เพิ่มพลังพัฒนาเด็ก”
วารสารทางวิชาการ มหาวิทยาลัยศรีนครินทรวิโรฒ ภาคใต้, 8(1), 23-25.
(2547). “K-W-L กับการพัฒนาการอ่าน”
วารสารสำนักหอสมุด มหาวิทยาลัยทักษิณ, 3(1), 9-13.
(2548). “มุมมองของการสอนภาษากับการพัฒนาการอ่าน”
วารสารสำนักหอสมุด มหาวิทยาลัยทักษิณ, 4(1), 5-9.
(2548). “ผลกระทบจากการตั้งคำถามของผู้สอน”
วารสารปริชาต, 18(1), 79-86.

ภาระงานด้านต่างๆที่เกี่ยวข้องกับการปฏิบัติงานของนางสาวศิริรัตน์ สิ้นประจักษ์ผล อาจารย์ประจำสาขาวิชาภาษาอังกฤษ คณะมนุษยศาสตร์และสังคมศาสตร์ ภาระงานที่ปฏิบัติใน ตุลาคม 2554- ถึงปัจจุบัน

คำสั่ง	ภาระงานที่รับผิดชอบหน่วยงานคณะที่สังกัดและหน่วยงานนอกคณะ
81/2554	กรรมการผู้ทรงคุณวุฒิประเมินหลักสูตรการศึกษาระดับมัธยมศึกษาตอนต้น (หลักสูตร 5 ปี)
142/2554	กรรมการโครงการอนุรักษ์วัฒนธรรมไทยและกิจกรรมทางพระพุทธศาสนากิจกรรมที่ 1 : วันลอยกระทง
148/2554	กรรมการโครงการอบรมการเรียนรู้วัฒนธรรมนานาชาติ
177/2554	กรรมการอำนวยการโครงการสัมมนาเชิงปฏิบัติการพัฒนาการสอนภาษาอังกฤษเป็นภาษาต่างประเทศโดยใช้วิจัยเป็นฐาน สำหรับครูเครือข่ายวิจัย ภาคใต้ตอนล่าง ครั้งที่ 1
0014/2555	กรรมการการจัดค่ายปลูกฝังคุณลักษณะความเป็นครูและเสริมประสบการณ์วิชาชีพครูของนิสิต/นักศึกษาในโครงการ สควค ค่ายที่ 2 (คำสั่งคณะศึกษาศาสตร์)
019/2555	ประธานกรรมการออกข้อสอบการสอบคัดเลือกนิสิตเข้าศึกษาวิชาโทภาษาอังกฤษ
022/2555	กรรมการโครงการพัฒนาเครื่องมือวัดทางภาษาของนิสิตตามหลักสูตร ศศ.บ. และ กศ.บ. ภาษาอังกฤษ
0025/2555	อนุกรรมการภาควิชาการการจัดค่ายปลูกฝังคุณลักษณะความเป็นครูและเสริมประสบการณ์วิชาชีพครูสำหรับนิสิต/นักศึกษาทุนโครงการ สควค ระดับประกาศนียบัตรบัณฑิตวิชาชีพครู ประจำปีการศึกษา 2554 ค่ายที่ 2
075/2555	กรรมการคุมสอบ การสอบประเมินผลความรู้ของนิสิตระดับปริญญาโท ประจำปีภาคเรียนฤดูร้อน

คำสั่ง	ภาระงานที่รับผิดชอบหน่วยงานคณะที่สังกัดและหน่วยงานนอกคณะ
0188/2555	กรรมการสอบสัมภาษณ์ การสอบคัดเลือกบุคคลเข้าเป็นนิสิตระดับปริญญาโท ประจำปีการศึกษา 2555
0189/2555	กรรมการออกข้อสอบ ตรวจสอบคะแนน หลักสูตร กศ.ม. สาขาวิชาภาษาอังกฤษ
0442/2555	คณะทำงานศูนย์พัฒนาทักษะภาษาอังกฤษ โครงการบูรณาการวิชาการเพื่อสังคม
095/2555	ประธานคัดเลือกผลงานวิชาการโครงการประชุมวิชาการงานวิจัยเพื่อพัฒนาการเรียนการสอนภาษาอังกฤษครั้งที่ 2
0068/2555	กองบรรณาธิการโครงการจัดทำวารสารวิชาการเทคโนโลยีและสื่อสารการศึกษา
118/2555	กรรมการและเลขานุการ English Pop Quiz Contest โครงการแข่งขันทักษะภาษาอังกฤษ
128/2555	กรรมการโครงการอนุรักษ์วัฒนธรรมไทยและกิจกรรมทางพระพุทธศาสนา กิจกรรมที่ 2 และ 3 วันอาสาฬหบูชาและวันเข้าพรรษา
	กรรมการโครงการเทศน์มหาชาติ เฉลิมพระเกียรติ 84 พรรษา

ภาระงานที่ได้รับเชิญจากหน่วยงานภายนอก

วัน เดือนปี	ภาระงาน	หน่วยงาน
4 เมษายน 55	ประธานคณะกรรมการสอบวิทยานิพนธ์หลักสูตรศิลปศาสตรมหาบัณฑิตเรื่อง “Impact of Journal Writing with Peer Feedback on EFL Students’ Writing Ability”	สาขาวิชาการสอน ภาษาอังกฤษเป็นภาษา นานาชาติ(ภาคปกติ) ภาควิชา ภาษาและภาษาศาสตร์ คณะ ศิลปศาสตร์ มหาวิทยาลัยสงขลานครินทร์ วิทยาเขตหาดใหญ่
26 เมษายน 2555	กรรมการสอบวิทยานิพนธ์หลักสูตรศึกษาศาสตรมหาบัณฑิต สาขาวิชาเทคโนโลยีและสื่อสารการศึกษา เรื่อง “ผลของการสอนตามแนวคิดวิธีธรรมชาติต่อความสามารถทางภาษาอังกฤษของนักเรียนระดับเตรียมความพร้อม”	สาขาวิชาเทคโนโลยีและ สื่อสารการศึกษา คณะ ศึกษาศาสตร์ มหาวิทยาลัย สงขลานครินทร์ วิทยาเขต ปัตตานี
3 พฤษภาคม 2555	กรรมการสอบวิทยานิพนธ์หลักสูตรศิลปศาสตรมหาบัณฑิตเรื่อง “Incidental Vocabulary Acquisition through Newspapers Reading: The Impact of Content Familiarity”	สาขาวิชาการสอน ภาษาอังกฤษเป็นภาษา นานาชาติ(ภาคปกติ) ภาควิชา ภาษาและภาษาศาสตร์ คณะ ศิลปศาสตร์ มหาวิทยาลัยสงขลานครินทร์ วิทยาเขตหาดใหญ่
2554	กรรมการสอบวิทยานิพนธ์ หลักสูตรศึกษาศาสตรมหาบัณฑิต สาขาวิชาจิตวิทยาการศึกษา แผน/แบบก2 เรื่อง ผลของการสอนแบบเพื่อนช่วยสอนที่มีต่อผลสัมฤทธิ์ทางการเรียนวิชาภาษาอังกฤษของนักเรียนชั้นมัธยมศึกษาปีที่ 2 ที่มีบริบทพหุวัฒนธรรมทางด้านภาษาต่างกัน	คณะศึกษาศาสตร์ มหาวิทยาลัยสงขลานครินทร์ วิทยาเขตปัตตานี
ตุลาคม 2554	ผู้เชี่ยวชาญพิจารณาหลักสูตรครุศาสตรบัณฑิต สาขาวิชาภาษาอังกฤษ	คณะครุศาสตร์ มหาวิทยาลัยราชภัฏสงขลา
4 พฤศจิกายน 2554	ผู้ทรงคุณวุฒิวิพากษ์หลักสูตรของหลักสูตรศิลปศาสตรบัณฑิต 4 ปี สาขาภาษาอังกฤษ หลักสูตรปรับปรุง พ.ศ. 2554	มหาวิทยาลัยราชภัฏยะลา
24-26 สิงหาคม 31-2 กันยายน	วิทยากรหลักสูตรภาษาอังกฤษระดับมัธยมศึกษาตอนต้นของโครงการพัฒนาครูกลุ่มสาระการเรียนรู้ภาษาอังกฤษสังกัดสำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน ปีงบประมาณ 2555	คณะศึกษาศาสตร์ มหาวิทยาลัยสงขลานครินทร์ วิทยาเขตปัตตานี ศูนย์ เครือข่ายระดับภูมิภาค: ภาคใต้

ผู้ทรงคุณวุฒิพิจารณา บทความของวารสาร / ตรวจสอบเนื้อหาข้อสอบ แบบสอบถาม
เพื่อการทำวิทยานิพนธ์

เมื่อ	เรื่อง	หน่วยงาน
	บทความ เรื่อง “ผลการใช้เทคนิคการแบ่งกลุ่มตามผลสัมฤทธิ์ทางการเรียนของวิธีการเรียนแบบร่วมมือต่อการอ่านภาษาอังกฤษเพื่อความเข้าใจของนิสิตชั้นปีที่ 2 มหาวิทยาลัยทักษิณ วิทยาเขตพัทลุง”	วารสาร “มนุษยศาสตร์และสังคมศาสตร์” มหาวิทยาลัยทักษิณ ปีที่ 5 ฉบับที่ 1(เมษายน-กันยายน 2553)
17/6/2554	ตรวจสอบเนื้อหา โครงสร้าง และการใช้ภาษาในประเด็นคำถามของวิทยานิพนธ์เรื่อง การศึกษาแรงจูงใจในการเรียนภาษาอังกฤษของนิสิตมหาวิทยาลัยทักษิณ	หลักสูตรครุทศ.สาขาวิชาภาษาอังกฤษ
26/10/2554	ตรวจสอบเนื้อหา โครงสร้าง และการใช้ภาษาในประเด็นคำถามของวิทยานิพนธ์เรื่อง กลวิธีการเรียนรู้ภาษาและทักษะการคิดอย่างมีวิจารณญาณของนิสิตวิชาเอกภาษาอังกฤษ มหาวิทยาลัยทักษิณ	หลักสูตรครุทศ.สาขาวิชาภาษาอังกฤษ
26/4/2555	บทความ ”กลวิธีการเรียนรู้ภาษาและทักษะการคิดอย่างมีวิจารณญาณของนิสิตวิชาเอกภาษาอังกฤษ มหาวิทยาลัยทักษิณ”	วารสารสำนักหอสมุด
มีนาคม 2555	การใช้วิธีการเขียนแบบเน้นกระบวนการเพื่อเพิ่มพูนความสามารถในการเขียนภาษาอังกฤษของนิสิตวิชาเอกภาษาอังกฤษชั้นปีที่ 3 มหาวิทยาลัยทักษิณ	การประชุมวิชาการระดับนานาชาติ มหาวิทยาลัยทักษิณ ครั้งที่ 22 ประจำปี 2555 ของสถาบันวิจัยและพัฒนา มหาวิทยาลัยทักษิณ
	ผู้เชี่ยวชาญในการตรวจสอบคุณภาพบทความวิจัยของวารสารสำนักหอสมุด มหาวิทยาลัยทักษิณเพื่อจัดพิมพ์ในรูปแบบอิเล็กทรอนิกส์ เมื่อ 26 เมษายน 2555	สำนักหอสมุด

หนังสือแต่งตั้งผู้เชี่ยวชาญตรวจเครื่องมือวิจัย



ที่ ศธ 6806 (ทศ)/8511

บัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร
พระราชวังสนามจันทร์
อ.เมือง จ.นครปฐม 73000

8 ธันวาคม 2560

เรื่อง ขอเชิญเป็นผู้ตรวจคุณภาพเครื่องมือวิจัย

เรียน อาจารย์พรทิพย์ บดีพงศ์

ด้วย นายจตุพงศ์ โมรา รหัสประจำตัว 56254904 นักศึกษาระดับปริญญาตรีบัณฑิต
สาขาวิชา หลักสูตรและการสอน บัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร กำลังทำวิทยานิพนธ์ เรื่อง " การพัฒนาแบบ
ทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย "

ในการนี้ บัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร มีความประสงค์ ขอเรียนเชิญท่านในฐานะผู้เชี่ยวชาญเป็นผู้
ตรวจคุณภาพเครื่องมือวิจัย ให้กับนักศึกษาดังกล่าว เพื่อประกอบการทำวิทยานิพนธ์

จึงเรียนมาเพื่อโปรดให้ความอนุเคราะห์ จักขอบพระคุณยิ่ง

ขอแสดงความนับถือ

(รองศาสตราจารย์ ดร.ไชยยศ ไพวิทยศิริธรรม)

รองคณบดีบัณฑิตวิทยาลัย ฝ่ายวิชาการและวิจัย

รักษาการแทน คณบดีบัณฑิตวิทยาลัย

สำนักงานบัณฑิตวิทยาลัย
นครปฐม โทร.034-218790



ที่ ศธ 6806/วท/8512

บัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร
พระราชวังสนามจันทร์
อ.เมือง จ.นครปฐม 73000

8 ธันวาคม 2560

เรื่อง ขอเชิญเป็นผู้ตรวจคุณภาพเครื่องมือวิจัย

เรียน อาจารย์ ดร.ศิริรัตน์ สีนประจักษ์ผล

ด้วย นายจตุพงศ์ โมรา รหัสประจำตัว 56254904 นักศึกษาระดับปริญญาตรีบัณฑิต
สาขาวิชา หลักสูตรและการสอน บัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร กำลังทำวิทยานิพนธ์ เรื่อง " การพัฒนาแบบ
ทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย "

ในการนี้ บัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร มีความประสงค์ ขอเรียนเชิญท่านในฐานะผู้เชี่ยวชาญเป็นผู้
ตรวจคุณภาพเครื่องมือวิจัยให้กับนักศึกษาดังกล่าว เพื่อประกอบการทำวิทยานิพนธ์

จึงเรียนมาเพื่อ โปรดให้ความอนุเคราะห์ จักขอบพระคุณยิ่ง

ขอแสดงความนับถือ

(รองศาสตราจารย์ ดร.ไชยยศ ไพบูลย์ศิริธรรม)

รองคณบดีบัณฑิตวิทยาลัย ฝ่ายวิชาการและวิจัย

รักษาการแทน คณบดีบัณฑิตวิทยาลัย

สำนักงานบัณฑิตวิทยาลัย
นครปฐม โทร.034-218790



ที่ ศธ 6806 (งค) / 8519

บัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร
พระราชวังสนามจันทร์
อ.เมือง จ.นครปฐม 73000

8 ธันวาคม 2560

เรื่อง ขอเชิญเป็นผู้ตรวจคุณภาพเครื่องมือวิจัย

เรียน Mr.Neil C.Cuthbertson

ด้วย นายจตุพงศ์ โมรา รหัสประจำตัว 56254904 นักศึกษาระดับปริญญาตรีบัณฑิต
สาขาวิชา หลักสูตรและการสอน บัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร กำลังทำวิทยานิพนธ์ เรื่อง " การพัฒนาแบบ
ทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย "

ในการนี้ บัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร มีความประสงค์ ขอเรียนเชิญท่านในฐานะผู้เชี่ยวชาญเป็นผู้
ตรวจคุณภาพเครื่องมือวิจัยให้กับนักศึกษาดังกล่าว เพื่อประกอบการทำวิทยานิพนธ์

จึงเรียนมาเพื่อโปรดให้ความอนุเคราะห์ จักขอบพระคุณยิ่ง

ขอแสดงความนับถือ

(รองศาสตราจารย์ ดร.ไชยยศ ไพวิทยศิริธรรม)

รองคณบดีบัณฑิตวิทยาลัย ฝ่ายวิชาการและวิจัย
รักษาการแทน คณบดีบัณฑิตวิทยาลัย

สำนักงานบัณฑิตวิทยาลัย
นครปฐม โทร.034-218790

ภาคผนวก ข

หนังสือขอทดลองเครื่องมือวิจัย



๕๙๙ อ.ศตพวงศ์ (งานวิจัย)



ที่ ศษ ๕๘๐๖(๙๖๙)/๐๐๐๔

มหาวิทยาลัยทักษิณ
รับ ๕๗๑๖ วันที่ 3 ธ.ค. 2561
เวลา 19.31

บัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร
พระราชวังสนามจันทร์
อ.เมือง จ.นครปฐม 73000

29 ตุลาคม 2561

คณะมนุษยศาสตร์และสังคมศาสตร์
รับ 1960 วันที่ 04 ธ.ค. 2561
เวลา 19.16

เรื่อง ขอลดลงเครื่องมือวิจัย

เรียน อธิการบดีมหาวิทยาลัยทักษิณ

ด้วย นายจตุพงศ์ โมรา รหัสประจำตัว 56254904 นักศึกษาระดับปริญญาตรีบัณฑิต
สาขาวิชา หลักสูตรและการสอน บัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร กำลังทำวิทยานิพนธ์ เรื่อง " การพัฒนา
แบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย " มีความประสงค์ จะขอลดลงเครื่องมือวิจัยกับ
นิติศัณปีที 4 จำนวน 900 คน เพื่อประกอบการทำวิทยานิพนธ์

ในการนี้ บัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร จึงขอความอนุเคราะห์จากท่าน โปรดอนุญาตให้
นักศึกษาดังกล่าวได้ทดลองเครื่องมือวิจัยด้วย

จึงเรียนมาเพื่อโปรดให้ความอนุเคราะห์ จักขอบพระคุณยิ่ง

เรียน อธิการบดี
เพื่อโปรดทราบและพิจารณา
มอบหมายให้ดำเนินการ
ติดต่อขอใช้เครื่องมือวิจัย
ตามที่ขอมา

3 ธ.ค. ๖1
3 ธ.ค. ๖1

3/18/61

ขอแสดงความนับถือ

Om 21

(อาจารย์ ดร. อธิภมร มากชัย)

รองคณบดีบัณฑิตวิทยาลัย ฝ่ายวิชาการและวิจัย
รักษาการแทน คณบดีบัณฑิตวิทยาลัย

3 ธ.ค. ๖1

เรียน คณบดี
เพื่อโปรดทราบและพิจารณา
เนน ดร มอช. อ.ศตพวงศ์ โมรา
ปร. สำนักร่วมเป็นนายเพื่อทดลอง
ได้ออวมือวิจัย

4 ธ.ค. ๖1
4 ธ.ค. ๖1

ขอ
อ.ศตพวงศ์ (เนนตี่)
3 ธ.ค. ๖๑

สำนักงานบัณฑิตวิทยาลัย
นครปฐม โทร.034-218790

ภาคผนวก ค

หนังสือขอความอนุเคราะห์ในการเก็บรวบรวมข้อมูล



ภาคผนวก ง

รายชื่อผู้ตอบแบบสำรวจความคิดเห็นอาจารย์มหาวิทยาลัย
ต่อโครงสร้างของแบบทดสอบ

- | | |
|--|--|
| 1. อาจารย์พรทิพย์ บดีพงศ์ | ผู้เชี่ยวชาญ
ข้าราชการบำนาญ
อดีตอาจารย์ประจำสาขาวิชาภาษาตะวันตก
คณะมนุษยศาสตร์และสังคมศาสตร์
มหาวิทยาลัยทักษิณ |
| 2. อาจารย์ ดร. ศิริรัตน์ สิ้นประจักษ์ผล | ผู้เชี่ยวชาญ
ข้าราชการบำนาญ
อดีตอาจารย์ประจำสาขาวิชาภาษาตะวันตก
คณะมนุษยศาสตร์และสังคมศาสตร์
มหาวิทยาลัยทักษิณ |
| 3. Mr. Neil C. Cuthbertson
อาจารย์ประจำโปรแกรมภาษาอังกฤษ
มหาวิทยาลัยเทคโนโลยีราชมงคลศรีวิชัย | ผู้เชี่ยวชาญ
คณะศิลปศาสตร์ |
| 4. ผู้ช่วยศาสตราจารย์ ดร. เกียรติทราย วุฒิพงษ์ | อาจารย์ประจำสาขาวิชาภาษาตะวันตก
คณะมนุษยศาสตร์และสังคมศาสตร์
มหาวิทยาลัยทักษิณ |
| 5. อาจารย์ ดร. ซูไฮลา อีแอต | อาจารย์ประจำสาขาวิชาภาษาตะวันตก
คณะมนุษยศาสตร์และสังคมศาสตร์
มหาวิทยาลัยทักษิณ |
| 6. ผู้ช่วยศาสตราจารย์มลฤดี สิทธิชัย | อาจารย์ประจำสาขาวิชาภาษาตะวันตก
คณะมนุษยศาสตร์และสังคมศาสตร์
มหาวิทยาลัยทักษิณ |
| 7. อาจารย์ดิญะพร วิสะมิตนันท์ | อาจารย์ประจำสาขาวิชาภาษาตะวันตก
คณะมนุษยศาสตร์และสังคมศาสตร์
มหาวิทยาลัยทักษิณ |
| 8. อาจารย์นวลทิพย์ มหามงคล | อาจารย์ประจำสาขาวิชาภาษาตะวันตก
คณะมนุษยศาสตร์และสังคมศาสตร์
มหาวิทยาลัยทักษิณ |

9. อาจารย์จุรีภรณ์ มะเล่โลหิต
อาจารย์ประจำสาขาวิชาภาษาตะวันตก
คณะมนุษยศาสตร์และสังคมศาสตร์
มหาวิทยาลัยทักษิณ
10. Miss Irish Chan Seosan
อาจารย์ประจำสาขาวิชาภาษาตะวันตก
คณะมนุษยศาสตร์และสังคมศาสตร์
มหาวิทยาลัยทักษิณ



ภาคผนวก จ

รายชื่อคณะกรรมการตรวจให้คะแนนแบบทดสอบอัตนัย

ทักษะการเขียน

- | | |
|--|--|
| 1. ผู้ช่วยศาสตราจารย์ ดร. เกร็ดทราย วุฒิพงษ์ | อาจารย์ประจำสาขาวิชาภาษาตะวันตก
คณะมนุษยศาสตร์และสังคมศาสตร์
มหาวิทยาลัยทักษิณ |
| 2. ผู้ช่วยศาสตราจารย์มลฤดี สิทธิชัย | อาจารย์ประจำสาขาวิชาภาษาตะวันตก
คณะมนุษยศาสตร์และสังคมศาสตร์
มหาวิทยาลัยทักษิณ |
| 3. อาจารย์จรีภรณ์ มะเลโล่หิต | อาจารย์ประจำสาขาวิชาภาษาตะวันตก
คณะมนุษยศาสตร์และสังคมศาสตร์
มหาวิทยาลัยทักษิณ |

ทักษะการพูด

- | | |
|-------------------------------|--|
| 1. อาจารย์ ดร. ชูไฮลา อีแต | อาจารย์ประจำสาขาวิชาภาษาตะวันตก
คณะมนุษยศาสตร์และสังคมศาสตร์
มหาวิทยาลัยทักษิณ |
| 2. อาจารย์ดิฐะพร วิสะมิตนันท์ | อาจารย์ประจำสาขาวิชาภาษาตะวันตก
คณะมนุษยศาสตร์และสังคมศาสตร์
มหาวิทยาลัยทักษิณ |
| 3. อาจารย์นวลทิพย์ มหามงคล | อาจารย์ประจำสาขาวิชาภาษาตะวันตก
คณะมนุษยศาสตร์และสังคมศาสตร์
มหาวิทยาลัยทักษิณ |

ภาคผนวก ฉ
แบบประเมินค่า IOC โดยผู้เชี่ยวชาญ
IOC Evaluation Form by Expert Reviewers

Please put a tick (✓) under the number corresponding to your judgment on each of the test items.

-1 = you disagree that the test item exactly tests the sub-skill and the skill level identified.

0 = you are not sure if the test item exactly tests the sub-skill and the skill level identified.

1 = you agree that the test item exactly tests the sub-skill and the skill level identified.

Section 1

Listening Test (University English Proficiency Test)

Part/ Item	Test Objectives/Listening Sub-skills	Skill Level	Reviewer's Rating			Suggestions
			-1	0	1	
A: Short Conversations						
Q-1	Focus on the second line. (S-1)	Comprehension				
Q-2	Focus on the second line. (S-1)	Comprehension				
Q-3	Choose answers with synonyms. (S-2)	Comprehension				
Q-4	Choose answers with synonyms. (S-2)	Comprehension				
Q-5	Choose answers with synonyms. (S-2)	Comprehension				
Q-6	Choose answers with synonyms. (S-2)	Comprehension				
Q-7	Choose answers with synonyms. (S-2)	Comprehension				
Q-8	Avoid similar sounds. (S-3)	Comprehension				
Q-9	Avoid similar sounds. (S-3)	Comprehension				
Q-10	Draw conclusions about Who . (S-4)	Analysis				
Q-11	Draw conclusions about Where . (S-4)	Analysis				
Q-12	Draw conclusions about What . (S-4)	Analysis				
Q-13	Listen for Who and What in passives. (S-5)	Comprehension				
Q-14	Listen for Who and What in passives. (S-5)	Comprehension				
Q-15	Listen for Who and What in passives. (S-5)	Comprehension				
Q-16	Listen for Who and What with multiple nouns. (S-6)	Comprehension				
Q-17	Listen for Who and What with multiple nouns. (S-6)	Comprehension				
Q-18	Listen for negative expressions. (S-7)	Comprehension				
Q-19	Listen for negative expressions. (S-7)	Comprehension				
Q-20	Listen for negative expressions. (S-7)	Comprehension				
Q-21	Listen for double negative expressions. (S-8)	Comprehension				
Q-22	Listen for double negative expressions. (S-8)	Comprehension				
Q-23	Listen for double negative expressions. (S-8)	Comprehension				
Q-24	Listen for "almost negative" expressions. (S-9)	Comprehension				
Q-25	Listen for "almost negative" expressions. (S-9)	Comprehension				
Q-26	Listen for "almost negative" expressions. (S-9)	Comprehension				
Q-27	Listen for negatives with comparatives. (S-10)	Comprehension				
Q-28	Listen for negatives with comparatives. (S-10)	Comprehension				

Part/ Item	Test Objectives/Listening Sub-skills	Skill Level	Reviewer's Rating			Suggestions
			-1	0	1	
Q-29	Listen for negatives with comparatives. (S-10)	Comprehension				
Q-30	Listen for expressions of agreement. (S-11)	Comprehension				
Q-31	Listen for expressions of agreement. (S-11)	Comprehension				
Q-32	Listen for expressions of uncertainty and suggestion. (S-12)	Comprehension				
Q-33	Listen for expressions of uncertainty and suggestion. (S-12)	Comprehension				
Q-34	Listen for expressions of uncertainty and suggestion. (S-12)	Comprehension				
Q-35	Listen for emphatic expressions of surprise. (S-13)	Comprehension				
Q-36	Listen for emphatic expressions of surprise. (S-13)	Comprehension				
Q-37	Listen for emphatic expressions of surprise. (S-13)	Comprehension				
Q-38	Listen for wishes. (S-14)	Comprehension				
Q-39	Listen for wishes. (S-14)	Comprehension				
Q-40	Listen for untrue conditions. (S-15)	Analysis				
Q-41	Listen for untrue conditions. (S-15)	Analysis				
Q-42	Listen for untrue conditions. (S-15)	Analysis				
Q-43	Listen for two-and three-part verbs. (S-16)	Comprehension				
Q-44	Listen for two-and three-part verbs. (S-16)	Comprehension				
Q-45	Listen for two-and three-part verbs. (S-16)	Comprehension				
Q-46	Listen for idioms. (S-17)	Analysis				
Q-47	Listen for idioms. (S-17)	Analysis				
Q-48	Listen for idioms. (S-17)	Analysis				
Q-49	Listen for idioms. (S-17)	Analysis				
B: Long Conversations						
Q-50	Answer stated detail questions correctly. (B-2)	Comprehension				
Q-51	Answer implied detail questions correctly. (B-3)	Analysis				
Q-52	Answer stated detail questions correctly. (B-2)	Comprehension				
Q-53	Answer stated detail questions correctly. (B-2)	Comprehension				
Q-54	Answer stated detail questions correctly. (B-2)	Comprehension				
Q-55	Answer stated detail questions correctly. (B-2)	Comprehension				
Q-56	Answer implied detail questions correctly. (B-3)	Analysis				
Q-57	Answer stated detail questions correctly. (B-2)	Comprehension				
Q-58	Identify who the speakers are (B-1)	Comprehension				
Q-59	Answer stated detail questions correctly. (B-2)	Comprehension				
Q-60	Answer stated detail questions correctly. (B-2)	Comprehension				
Q-61	Answer stated detail questions correctly. (B-2)	Comprehension				

Part/ Item	Test Objectives/Listening Sub-skills	Skill Level	Reviewer's Rating			Suggestions
			-1	0	1	
Q-62	Answer stated detail questions correctly. (B-2)	Comprehension				
Q-63	Answer stated detail questions correctly. (B-2)	Comprehension				
Q-64	Find "Unstated" details. (B-4)	Comprehension				
Q-65	Answer stated detail questions correctly. (B-2)	Comprehension				
Q-66	Answer stated detail questions correctly. (B-2)	Comprehension				
C: Talks and Lectures						
Q-67	Identify the topic of the talk. (C-1)	Comprehension				
Q-68	Answer stated detail questions correctly. (C-2)	Comprehension				
Q-69	Answer stated detail questions correctly. (C-2)	Comprehension				
Q-70	Answer stated detail questions correctly. (C-2)	Comprehension				
Q-71	Answer stated detail questions correctly. (C-2)	Comprehension				
Q-72	Find "Unstated" details. (C-3)	Comprehension				
Q-73	Identify the topic of the talk. (C-1)	Comprehension				
Q-74	Answer implied detail questions correctly. (C-4)	Analysis				
Q-75	Answer stated detail questions correctly. (C-2)	Comprehension				
Q-76	Find "Unstated" details. (C-3)	Comprehension				
Q-77	Answer stated detail questions correctly. (C-2)	Comprehension				
Q-78	Answer stated detail questions correctly. (C-2)	Comprehension				
Q-79	Find "Unstated" details. (C-3)	Comprehension				
Q-80	Answer implied detail questions correctly. (C-4)	Analysis				
Q-81	Answer implied detail questions correctly. (C-4)	Analysis				
Q-82	Answer stated detail questions correctly. (C-2)	Comprehension				
Q-83	Answer stated detail questions correctly. (C-2)	Comprehension				
Q-84	Find "Unstated" details. (C-3)	Comprehension				
Q-85	Answer stated detail questions correctly. (C-2)	Comprehension				
Q-86	Find "Unstated" details. (C-3)	Comprehension				
Q-87	Answer implied detail questions correctly. (C-4)	Analysis				

Section 2

Language Use (University English Proficiency Test)

Part/ Item	Test Objectives/Listening Sub-skills	Skill Level	Reviewer's Rating			Suggestions
			-1	0	1	
Structure						
Q-1	Be sure the sentence has a subject and a verb. (S-1)	Comprehension				
Q-2	Be sure the sentence has a subject and a verb. (S-1)	Comprehension				
Q-3	Be sure the sentence has a subject and a verb. (S-1)	Comprehension				
Q-4	Be careful of objects of prepositions. (S-2)	Comprehension				
Q-5	Be careful of objects of prepositions. (S-2)	Comprehension				
Q-6	Be careful of present participles. (S-4)	Comprehension				
Q-7	Be careful of present participles. (S-4)	Comprehension				
Q-8	Use reduced adjective clause correctly. (S-13)	Comprehension				
Q-9	Use reduced adjective clause correctly. (S-13)	Comprehension				
Q-10	Use adverb <i>time</i> and <i>cause</i> connectors correctly. (S-7)	Comprehension				
Q-11	Use adverb <i>time</i> and <i>cause</i> connectors correctly. (S-7)	Comprehension				
Q-12	Use noun clause connector / subjects correctly. (S-10)	Comprehension				
Q-13	Use other adverb connectors correctly. (S-8)	Comprehension				
Q-14	Use other adverb connectors correctly. (S-8)	Comprehension				
Q-15	Inverted the subject and verb with place expressions. (S-16)	Comprehension				
Q-16	Be careful of past participles. (S-5)	Comprehension				
Q-17	Use noun clause connectors correctly. (S-9)	Comprehension				
Q-18	Invert the subject and verb with conditionals. (S-18)	Comprehension				
Q-19	Use reduced adverb clauses correctly. (S-14)	Comprehension				
Q-20	Invert the subject and verb with negatives. (S-17)	Comprehension				
Q-21	Use adjective clause connectors correctly. (S-11)	Comprehension				
Q-22	Invert the subject and verb with comparisons. (S-19)	Comprehension				
Written Expression						
Q-23	After <i>be</i> , use the present participle or the past participle. (S-31)	Comprehension				
Q-24	After <i>be</i> , use the present participle or the past participle. (S-31)	Comprehension				
Q-25	Form comparatives and superlatives correctly. (S-27)	Comprehension				
Q-26	Form comparatives and superlatives correctly. (S-27)	Comprehension				
Q-27	Make verbs agree after certain words. (S-23)	Comprehension				

Part/ Item	Test Objectives/Listening Sub-skills	Skill Level	Reviewer's Rating			Suggestions
			-1	0	1	
Q-28	Make verbs agree after certain words. (S-23)	Comprehension				
Q-29	After will, would, or other modals, use the base form of the verb. (S-32)	Comprehension				
Q-30	After will, would, or other modals, use the base form of the verb. (S-32)	Comprehension				
Q-31	Use the correct tense with time expressions. (S-35)	Comprehension				
Q-32	Use the correct tense with time expressions. (S-35)	Comprehension				
Q-33	Distinguish subject and object pronouns. (S-43)	Comprehension				
Q-34	Distinguish subject and object pronouns. (S-43)	Comprehension				
Q-35	Distinguish possessive adjectives and pronoun. (S-44)	Comprehension				
Q-36	Distinguish possessive adjectives and pronoun. (S-44)	Comprehension				
Q-37	Use parallel structure with paired conjunctions. (S-25)	Comprehension				
Q-38	Use parallel structure with paired conjunctions. (S-25)	Comprehension				
Q-39	Make verbs agree after prepositional phrases. (S-20)	Comprehension				
Q-40	Use comparatives and superlatives correctly. (S-28)	Comprehension				
Q-41	Use comparatives and superlatives correctly. (S-28)	Comprehension				
Q-42	Distinguish a and an. (S-53)	Comprehension				
Q-43	Distinguish a and an. (S-53)	Comprehension				
Q-44	Use the correct tense with will and would. (S-36)	Comprehension				
Q-45	Distinguish like, alike, unlike, and dislike. (S-59)	Comprehension				
Q-46	Distinguish like, alike, unlike, and dislike. (S-59)	Comprehension				
Q-47	Distinguish make and do. (S-58)	Comprehension				
Q-48	Make inverted verbs agree. (S-22)	Comprehension				
Q-49	Distinguish countable and uncountable nouns. (S-40)	Comprehension				
Q-50	Distinguish countable and uncountable nouns. (S-40)	Comprehension				
Q-51	Use articles with singular nouns. (S-52)	Comprehension				
Q-52	Position adjectives and adverbs correctly. (S-48)	Comprehension				
Q-53	Position adjectives and adverbs correctly. (S-48)	Comprehension				
Q-54	Recognize active and passive meanings. (S-38)	Comprehension				
Q-55	Recognize active and passive meanings. (S-38)	Comprehension				
Q-56	Use articles with singular nouns. (S-52)	Comprehension				
Q-57	Use predicate adjectives correctly. (S-50)	Comprehension				

Part/ Item	Test Objectives/Listening Sub-skills	Skill Level	Reviewer's Rating			Suggestions
			-1	0	1	
Q-58	Use the correct form of the passive. (S-37)	Comprehension				
Q-59	Use of prepositions after certain words.	Comprehension				
Q-60	Use of prepositions after certain words.	Comprehension				
Q-61	Distinguish the person from the thing. (S-42)	Comprehension				
Q-62	Use adjectives after linking verbs. (S-47)	Comprehension				

Comments for Section 2

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....



Section 3

Reading Comprehension (University English Proficiency Test)

Part/ Item	Test Objectives/Reading Sub-Skills	Skill Level	Reviewer's Rating			Suggestions
			-1	0	1	
A: English in Everyday Use (EEU)						
Q-1	To answer an implied detail question correctly. (A-3)	Analysis				
Q-2	To answer an implied detail question correctly. (A-3)	Analysis				
Q-3	To answer stated detail questions correctly. (A-1)	Comprehension				
Q-4	To answer an implied detail question correctly. (A-3)	Analysis				
Q-5	To answer stated detail questions correctly. (A-1)	Comprehension				
Q-6	To find "unstated" details. (A-2)	Comprehension				
Q-7	To answer stated detail questions correctly. (A-1)	Comprehension				
Q-8	To answer stated detail questions correctly. (A-1)	Comprehension				
Q-9	To answer stated detail questions correctly. (A-1)	Comprehension				
Q-10	To answer stated detail questions correctly. (A-1)	Comprehension				
Q-11	To choose a general synonym for some specific words in the reading texts. (A-5)	Comprehension				
Q-12	To find "unstated" details. (A-2)	Comprehension				
Q-13	To determine meanings from word parts. (A-4)	Comprehension				
Q-14	To answer an implied detail question correctly. (A-3)	Analysis				
Q-15	To determine meanings from word parts. (A-4)	Comprehension				
Q-16	To find "unstated" details. (A-2)	Comprehension				
Q-17	To answer stated detail questions correctly. (A-1)	Comprehension				
Q-18	To choose a general synonym for some specific words in the reading texts. (A-5)	Comprehension				
Q-19	To answer an implied detail question correctly. (A-3)	Analysis				
Q-20	To answer stated detail questions correctly. (A-1)	Comprehension				
Q-21	To answer stated detail questions correctly. (A-1)	Comprehension				
Q-22	To answer stated detail questions correctly. (A-1)	Comprehension				
Q-23	To choose a general synonym for some specific words in the reading texts. (A-5)	Comprehension				
Q-24	To choose a general synonym for some specific words in the reading texts. (A-5)	Comprehension				
Q-25	To find "unstated" details. (A-2)	Comprehension				

Part/ Item	Test Objectives/Reading Sub-Skills	Skill Level	Reviewer's Rating			Suggestions
			-1	0	1	
Q-26	To answer stated detail questions correctly. (A-1)	Comprehension				
Q-27	To find "unstated" details. (A-2)	Comprehension				
Q-28	To use context to determine meanings of simple words. (A-6)	Comprehension				
Q-29	To use context to determine meanings of simple words. (A-6)	Comprehension				
Q-30	To use context to determine meanings of simple words. (A-6)	Comprehension				
Q-31	To use context to determine meanings of simple words. (A-6)	Comprehension				
Q-32	To use context to determine meanings of simple words. (A-6)	Comprehension				
Q-33	To answer stated detail questions correctly. (A-1)	Comprehension				
Q-34	To answer stated detail questions correctly. (A-1)	Comprehension				
Q-35	To answer an implied detail question correctly. (A-3)	Analysis				
Q-36	To answer stated detail questions correctly. (A-1)	Comprehension				
Q-37	To answer stated detail questions correctly. (A-1)	Comprehension				
Q-38	To answer stated detail questions correctly. (A-1)	Comprehension				
Q-39	To answer stated detail questions correctly. (A-1)	Comprehension				
Q-40	To answer stated detail questions correctly. (A-1)	Comprehension				
Q-41	To use context to determine meanings of simple words. (A-6)	Comprehension				
Q-42	To use context to determine meanings of simple words. (A-6)	Comprehension				
Q-43	To use context to determine meanings of simple words. (A-6)	Comprehension				
Q-44	To use context to determine meanings of simple words. (A-6)	Comprehension				
Q-45	To use context to determine meanings of simple words. (A-6)	Comprehension				
Q-46	To use context to determine meanings of simple words. (A-6)	Comprehension				
Q-47	To use context to determine meanings of simple words. (A-6)	Comprehension				
Q-48	To use context to determine meanings of simple words. (A-6)	Comprehension				
Q-49	To use context to determine meanings of simple words. (A-6)	Comprehension				
Q-50	To use context to determine meanings of simple words. (A-6)	Comprehension				

Part/ Item	Test Objectives/Reading Sub-Skills	Skill Level	Reviewer's Rating			Suggestions
			-1	0	1	
B: English for Academic Purpose (EEU)						
Q-51	To recognize the organization of the ideas in the passage. (B-9)	Analysis				
Q-52	To answer main idea question correctly. (B-1)	Comprehension				
Q-53	To answer stated detail questions correctly. (B-2)	Comprehension				
Q-54	To find "unstated" details. (B-3)	Comprehension				
Q-55	To determine where specific information is found. (B-8)	Comprehension				
Q-56	To answer stated detail questions correctly. (B-2)	Comprehension				
Q-57	To find "unstated" details. (B-3)	Comprehension				
Q-58	To answer stated detail questions correctly. (B-2)	Comprehension				
Q-59	To determine where specific information is found. (B-8)	Comprehension				
Q-60	To determine where specific information is found. (B-8)	Comprehension				
Q-61	To answer an implied detail question correctly. (B-5)	Analysis				
Q-62	To determine the purpose of the author. (B-12)	Analysis				
Q-63	To choose a general synonym for some specific words in the reading texts. (B-14)	Comprehension				
Q-64	To choose a general synonym for some specific words in the reading texts. (B-14)	Comprehension				
Q-65	To choose a general synonym for some specific words in the reading texts. (B-14)	Comprehension				
Q-66	To fine pronoun referents. (B-6)	Comprehension				
Q-67	To answer an implied detail question correctly. (B-5)	Analysis				
Q-68	To answer an implied detail question correctly. (B-5)	Analysis				
Q-69	To answer main idea question correctly. (B-1)	Comprehension				
Q-70	To find "unstated" details. (B-3)	Comprehension				
Q-71	To answer stated detail questions correctly. (B-2)	Comprehension				
Q-72	To answer stated detail questions correctly. (B-2)	Comprehension				
Q-73	To answer stated detail questions correctly. (B-2)	Comprehension				
Q-74	To answer stated detail questions correctly. (B-2)	Comprehension				
Q-75	To answer stated detail questions correctly. (B-2)	Comprehension				
Q-76	To find "unstated" details. (B-3)	Comprehension				
Q-77	To answer stated detail questions correctly. (B-2)	Comprehension				

Part/ Item	Test Objectives/Reading Sub-Skills	Skill Level	Reviewer's Rating			Suggestions
			-1	0	1	
Q-78	To choose a general synonym for some specific words in the reading texts. (B-14)	Comprehension				
Q-79	To answer main idea question correctly. (B-1)	Comprehension				
Q-80	To find "unstated" details. (B-3)	Comprehension				
Q-81	To answer stated detail questions correctly. (B-2)	Comprehension				
Q-82	To answer stated detail questions correctly. (B-2)	Comprehension				
Q-83	To choose a general synonym for some specific words in the reading texts. (B-14)	Comprehension				
Q-84	To find pronoun referents. (B-6)	Comprehension				
Q-85	To choose a general synonym for some specific words in the reading texts. (B-14)	Comprehension				
Q-86	To draw a conclusion by using main points or contexts in the passage. (B-11)	Comprehension				
Q-87	To answer transition questions correctly. (B-16)	Analysis				
Q-88	To answer stated detail questions correctly. (B-2)	Comprehension				
Q-89	To answer stated detail questions correctly. (B-2)	Comprehension				
Q-90	To answer stated detail questions correctly. (B-2)	Comprehension				
Q-91	To answer stated detail questions correctly. (B-2)	Comprehension				
Q-92	To answer stated detail questions correctly. (B-2)	Comprehension				
Q-93	To answer stated detail questions correctly. (B-2)	Comprehension				
Q-94	To choose a general synonym for some specific words in the reading texts. (B-14)	Comprehension				
Q-95	To choose a general synonym for some specific words in the reading texts. (B-14)	Comprehension				
Q-96	To find "unstated" details. (B-3)	Comprehension				
Q-97	To answer stated detail questions correctly. (B-2)	Comprehension				
Q-98	To answer an implied detail question correctly. (B-5)	Analysis				
Q-99	To determine the tone and mood of the author. (B-15)	Analysis				
Q-100	To answer an implied detail question correctly. (B-5)	Analysis				

ภาคผนวก ช

แบบประเมินการให้คะแนนแบบทดสอบอัตนัย



Grading Rubric for Writing Tasks

Criteria/Levels	5	4	3	2	1	Points
Content	Writing addresses the assigned topic; the ideas are concrete and thoroughly developed; no extraneous material.	Writing addresses the issues but misses some points; ideas could be more fully developed; some extraneous material is present.	Development of ideas not complete or the writing is somewhat off the topic; paragraphs aren't divided exactly right.	Ideas incomplete; essay does not reflect careful thinking or was hurriedly written; inadequate effort in area of content.	Writing is completely inadequate and does not reflect college-level work; no apparent effort to consider the topic carefully.	
Organization	Writing shows high degree of attention to logic and reasoning of points. Unity clearly leads the reader to the conclusion and stirs thought regarding the topic.	Writing is coherent and logically organized with transitions used between ideas and paragraphs to create coherence. Overall unity of ideas is present.	Writing is coherent and logically organized. A few points remain misplaced and stray from the topic. Some transitions evident but not used throughout the writing.	Writing is coherent and logically organized. Some points remain misplaced and stray from the topic. A few transitions are evident.	Writing lacks logical organization. It shows some coherence but ideas lack unity. Serious errors.	
Language	Writing contains a range of vocabulary, including less-common lexis. Writing has a range of simple and complex grammatical forms, with control and flexibility. Occasional errors may be present but do not impede communication.	Writing mostly contains a range of vocabulary, including less-common lexis. Writing has a range of simple and many complex grammatical forms, with control and flexibility. A few errors may be present but do not impede communication.	Writing contains a range of everyday vocabulary, with occasional inappropriate use of less common lexis. Writing has a range of simple and some complex grammatical forms with a good degree of control. Errors do not impede communication.	Writing contains some everyday vocabulary, while sometimes overusing certain lexis. Writing has a few complex grammatical forms with a good degree of control. While errors are noticeable, meaning can still be understandable.	Writing contains everyday vocabulary, while occasionally overusing certain lexis. Writing has simple grammatical forms with a good degree of control. While errors are noticeable, meaning can still be determined.	
Mechanics	Writing is free of distracting spelling, punctuation, and grammatical errors; absent of fragments, comma splices, and run-ons.	Writing is few spelling, punctuation, and grammatical errors allowing reader to follow ideas clearly. Very few fragments or run-ons.	Writing has some spelling, punctuation, and grammatical errors; some evidences of fragments, comma splices, and run-ons.	Most spelling punctuation, and grammar is correct allowing reader to progress through writing. Some errors remain.	Spelling, punctuation and grammatical errors create distraction, making reading difficult; fragments, comma splices, run-ons evident. Errors are frequent./20

Adapted from Brown (2004) and Cambridge English (2015)

Grading Rubric for Speaking Tasks

Criteria	5	4	3	2	1	Points
Communication and Size	Understand all questions, topics of talking, and the task description very well. Can give correct answers to all questions, excellent level of description and additional details beyond the required, and reasons or express opinions to support the idea logically.	One or two questions doesn't get the correct answers. Can give level of description; all required information included. Mostly understand the topic and can give many reasons to support the ideas expressed.	Quite understand the questions but cannot give right answers to a few questions. Can give adequate description; some additional details should be provided. Quite understand the topic and can give some ideas together with some opinions to support the ideas.	Can understand only some questions and can give right answers to one or two questions. Description lacks some critical details that make it difficult for listener to understand. Unsure to understand the topic, but some ideas expressed can support the topic.	Mostly cannot understand what is asked, so most of the answers doesn't match the questions. Description is so lacking that the listener cannot understand. Mostly cannot understand the topic, so doesn't have any idea to express.	
Fluency	Smooth and fluid speech; few to no hesitations; no attempts to search for words; volume is excellent.	Smooth and fluid speech; few hesitations; a slight search for words; inaudible word or two.	Speech is relatively smooth, some hesitation and unevenness caused by rephrasing and searching for words, volume wavers.	Speech is frequently hesitant with some sentences left uncompleted; volume very soft.	Speech is slow, hesitant & strained except for short memorized phrases; difficult to perceive continuity in speech; inaudible.	
Accuracy	Excellent control of language features; a wide range of well-chosen vocabulary. Adequacy and variety of grammatical structures.	Good language control; good range of relatively well-chosen vocabulary. Some errors in grammatical structures possibly caused by attempts to include a variety.	Adequate language control; vocabulary range is lacking. Fluent grammatical errors that do not obscure meaning; little variety in structures.	Weak language control; basic vocabulary choice with some words clearly lacking. Frequent grammatical errors even in simple structures that at times obscure meaning.	Weak language control; vocabulary that is used does not match the task. Frequent grammatical errors even in simple structures; meaning is obscured.	
Comprehensibility	Stay in target language and comprehensibility not affected by errors. Rhythm and intonation are very good.	Stay in target language but there are minor errors in pronunciation and comprehensibility sometimes affected.	Stay in target language but there are some errors in pronunciation and comprehensibility sometimes affected.	Sometimes uses English or another language and/or pronunciation is weak and impedes or affect comprehensibility.	Use less target language and comprehensibility is an issue and/or major pronunciation errors are excessive for this level.	
Adapted from Brown (2004), Carroll (1983), and Cambridge English (2015)					/20

ภาคผนวก ข

แบบประเมินความพึงพอใจของผู้เกี่ยวข้อง
ต่อแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย
(University English Proficiency Test)

วัตถุประสงค์ของแบบประเมิน คือ เพื่อศึกษาความพึงพอใจของผู้เกี่ยวข้องต่อแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย (University English Proficiency Test)

คำชี้แจง แบบประเมินฉบับนี้มี 3 ตอน จำนวน 22 ข้อ ให้ผู้ตอบแบบประเมินทำเครื่องหมาย ✓ ลงในช่องความคิดเห็นของข้อคำถาม และเขียนข้อเสนอแนะในช่องความคิดเห็นหรือข้อเสนอแนะที่กำหนดให้ สำหรับผู้ตอบที่เป็นนิสิต แบบประเมินฉบับนี้ไม่มีผลต่อผลการเรียนหรือคะแนนสอบของนิสิตแต่ประการใด โดยมีค่าระดับความพึงพอใจ ดังนี้

5	หมายถึง	พึงพอใจมากที่สุด
4	หมายถึง	พึงพอใจมาก
3	หมายถึง	พึงพอใจปานกลาง
2	หมายถึง	พึงพอใจน้อย
1	หมายถึง	พึงพอใจน้อยที่สุด

ตอนที่ 1 สถานภาพของผู้ตอบแบบประเมิน

- 1.1 ผู้เชี่ยวชาญด้านการวัดและประเมินผลการสอนภาษาอังกฤษ
- 1.2 อาจารย์ผู้สอนภาษาอังกฤษระดับมหาวิทยาลัย
- 1.3 ผู้เข้าสอบ

ตอนที่ 2 ความพึงพอใจของผู้ตอบแบบประเมิน

ข้อคำถาม	ระดับความพึงพอใจ				
	มากที่สุด	มาก	ปานกลาง	น้อย	น้อยที่สุด
คำสั่งของแบบทดสอบ					
1) คำสั่งรวมและคำสั่งสำหรับแต่ละตอนให้ไว้อย่างชัดเจน ใช้ภาษาเข้าใจง่าย					
2) ตัวอย่างแสดงการตอบในแต่ละตอนที่เข้าใจได้ง่าย					
3) การระบุน้ำหนักคะแนนของแบบทดสอบแต่ละตอน					
4) การกำหนดเวลาในการทำแบบทดสอบแต่ละตอน					
5) แบบทดสอบการเขียนตอบได้มีการกำหนดเวลาในการทำและความยาวของคำตอบให้ไว้อย่างชัดเจน					
6) ผู้เข้าสอบได้รับคำแนะนำการทำสอบอย่างชัดเจน ครอบคลุม และเพียงพอ					

ข้อคำถาม	ระดับความคิดเห็น				
	มากที่สุด	มาก	ปานกลาง	น้อย	น้อยที่สุด
เนื้อหาสาระของแบบทดสอบ					
1) ความยากง่ายของเนื้อหาในแบบทดสอบเหมาะสมกับระดับของผู้เข้าสอบ					
2) การจัดลำดับเนื้อหาของแบบทดสอบแต่ละตอนมีความเหมาะสม					
3) จำนวนข้อของแบบทดสอบเหมาะสมกับเวลาสอบ					
4) เนื้อหาสาระของแบบทดสอบน่าสนใจ มีประโยชน์ต่อผู้เข้าสอบ					
1) เนื้อหาของแบบทดสอบเหมาะสมกับการสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย					
2) แบบทดสอบที่ใช้สามารถวัดและประเมินวัตถุประสงค์ของการสอบแต่ละทักษะได้อย่างเหมาะสมและเพียงพอ					
3) แบบทดสอบวัดทักษะทางภาษาอย่างหลากหลาย วัดทั้งระเบียบวิธี/กฎเกณฑ์ทางภาษา และการใช้ภาษาอย่างสมจริง					
4) ระดับภาษาที่ใช้เหมาะสม มีลีลาที่สอดคล้องกับสภาพการใช้ภาษาอย่างสมจริง เช่น การใช้ภาษาเป็นกันเองในการสนทนา และการใช้ภาษาเป็นทางการในการอ่าน และการบรรยาย ฯลฯ หรือไม่					
5) แบบทดสอบใช้ภาษาที่มีลีลาสุภาพ ไม่ขัดต่อจริยธรรมและศีลธรรมอันดีหรือไม่					
6) ชื่อของแบบทดสอบสอดคล้องกับเนื้อหาที่วัด ไม่ทำให้ผู้เข้าสอบไขว้เขวหรือเข้าใจผิดได้					
7) แบบทดสอบไม่ลำเอียงต่อเชื้อชาติ พื้นภูมิลัทธิ เพศ และสภาพทางร่างกายของผู้เข้าสอบหรือไม่ และ					
8) มีการใช้ภาษาโดยไม่คำนึงถึงความรู้สึกนึกคิด จิตใจของผู้เข้าสอบบางกลุ่มหรือไม่					
รูปแบบและการจัดลำดับความ					
1) แบบทดสอบมีความยาวเหมาะสมกับระดับผู้เข้าสอบ ไม่สั้นหรือยาวจนเกินไป					
2) มีการวัดวัตถุประสงค์แต่ละข้อเหมาะสม ไม่มากเกินไปหรือน้อยเกินไป					
3) แบบทดสอบไม่มีการชี้แนะคำตอบซึ่งกันและกัน					
4) ไม่มีการใช้ภาษาที่ชี้แนะคำตอบที่ถูกต้อง					
5) ทำสอบโดยใช้ความรู้ความสามารถในการเลือกคำตอบที่ถูกต้อง และไม่ใช้การเดาหรือการคัดคำตอบที่คิดว่าไม่ใช่คำตอบออกไปที่ละข้อ					
6) มีการกำหนดเวลาที่เหมาะสมโดยให้ผู้เข้าสอบที่อ่อนสามารถทำได้ด้วย					
7) มีการใช้วิธีการวัดเหมาะสมกับคำตอบที่ต้องการหรือไม่ เช่น แบบจับคู่ มีการใช้ภาษาในตัวแบบทดสอบและคำตอบในระดับความยากง่ายเดียวกัน					

ข้อคำถาม	ระดับความคิดเห็น				
	มากที่สุด	มาก	ปานกลาง	น้อย	น้อยที่สุด
8) มีการใช้วิธีการวัดที่หลากหลาย เพื่อวัดครอบคลุมทุกทักษะ หรือวัดทักษะย่อยที่ประสงค์					
9) มีการใช้ตัวกระตุ้น/ตัวแบบทดสอบ เพื่อให้ได้คำตอบที่ต้องการอย่างสมจริง เช่น มีการใช้เทปบันทึกเสียง วิดีทัศน์ ภาพประกอบ ฯลฯ หรือไม่					
10) การวางรูปแบบแบบทดสอบชัดเจน สะดวกต่อการทำสอบ ไม่ทำให้ผู้เข้าสอบพลาดพลั้ง เช่น การเรียงข้อ และตัวเลือก a b c d ชัดเจน และถ้าเป็นแบบโคลงหรือแบบเติมคำ มีการวางตัวเลือกในหน้าเดียวกันกับตัวแบบทดสอบ เพื่อให้ผู้เข้าสอบไม่ต้องสับสน หรือเสียเวลาพลิกหน้ากระดาษหรือไม่					
11) การเว้นช่องไฟและช่องว่างระหว่างบรรทัดเหมาะสม ชัดเจนหรือไม่					
การตรวจให้คะแนน					
1) การตรวจให้คะแนนแบบทดสอบอัตนัย มีเกณฑ์ให้คะแนนที่เหมาะสมกับภาระงานสอบ					
ประโยชน์ของแบบทดสอบ					
1) ผลการสอบมีประโยชน์ในการพัฒนาและปรับปรุงทักษะภาษาอังกฤษของผู้เข้าสอบ					
2) ผู้เข้าสอบได้รับประสบการณ์ที่จะเป็นประโยชน์ต่อการเข้าสอบวัดความสามารถภาษาอังกฤษในอนาคต					
การจัดสอบ					
1) การสมัครและลงทะเบียนสอบทันสมัย สะดวก และเข้าถึงได้รวดเร็ว					
2) มีช่องทางในการติดต่อสื่อสารระหว่างผู้จัดสอบและผู้เข้าสอบ					
3) ที่นั่งสะดวกสบายเหมาะกับการนั่งทำแบบทดสอบ					
4) อากาศในห้องสอบเหมาะสม ไม่ร้อนหรือเย็นจนเกินไป					
5) ไม่มีเสียงรบกวนอื่นที่เป็นการทำลายสมาธิในการทำแบบทดสอบ					
6) แสงสว่างเพียงพอ ไม่เป็นอุปสรรคต่อการทำแบบทดสอบ					
7) เสียงจากลำโพงในการทดสอบฟังชัดเจน ไม่เป็นอุปสรรคต่อการได้ยิน					

ตอนที่ 3 ความคิดเห็นและข้อเสนอแนะของผู้ตอบแบบประเมิน

.....

.....

.....

.....

.....

.....

ขอขอบคุณในความร่วมมือ

ภาคผนวก ฅ

แบบประเมินคุณประโยชน์ของแบบทดสอบ



Evaluation Form for Test Usefulness (University English Proficiency Test-UEPT)

To what extent do you agree with each statement? Put a tick (✓) in a column under the number which best indicates your agreement to the statement based on the specified criteria.

- 5 strongly agree
4 agree
3 moderately agree
2 disagree
1 strongly disagree

Listening Test

Test Usefulness	Levels of agreement				
	5	4	3	2	1
Construct Validity					
1. The test objective of each listening part is clearly defined.					
2. The test objective of each listening part is in accordance with the purpose of the listening test.					
3. All listening tasks are in accordance with the listening test objectives.					
4. The scoring methods are appropriate, and they can reflect the test takers' listening abilities.					
5. The test takers' scores can predict their listening abilities.					
Reliability					
6. The expected response of each test item is sufficiently specific for one correct answer.					
7. The scores of the tests are reliable, irrespective of any raters.					
8. The response to each test item is reliable, irrespective of any test-takers, any time, and any place for test-taking.					
9. The test input in each part of the test does not contribute to understanding or knowledge in doing other items of other parts.					
Authenticity					
10. The listening tasks are related to the test takers' future real-life tasks.					
11. The listening texts are authentic, and they reflect real-life situations.					
Interactiveness					
12. The listening tasks are familiar topics for all test takers.					
13. The listening tasks are suitable for the English knowledge of all the test takers.					
Impact					
14. The listening tasks provide the test takers with experience in doing future listening tasks.					
15. The listening tasks are similar to the listening activities that the test takers have experienced in class.					
16. The purpose of the listening test is consistent with the goals and the values mentioned in the instructional program.					
Test Usefulness	Levels of agreement				

Language Use Test

Test Usefulness	Levels of agreement				
	5	4	3	2	1
Construct Validity					
1. The test objective of each Language Use part is clearly defined.					
2. The test objective of each Language Use part is in accordance with the purpose of the Language Use test.					
3. All Language Use tasks are in accordance with the Language Use test objectives.					
4. The scoring methods are appropriate, and they can reflect the test takers' Language Use abilities.					
5. The test takers' scores can predict their Language Use abilities.					
Reliability					
6. The expected response of each test item is sufficiently specific for one correct answer.					
7. The scores of the tests are reliable, irrespective of any raters.					
8. The response to each test item is reliable, irrespective of any test-takers, any time, and any place for test-taking.					
9. The test input in each part of the test does not contribute to understanding or knowledge in doing other items of other parts.					
Authenticity					
10. The test tasks are related to the test takers' future real-life tasks.					
11. The input texts are authentic and reflect real-life situations.					
Interactiveness					
12. The test tasks are familiar topics for all test takers.					
13. The test tasks are suitable for the English knowledge of all the test takers.					
Impact					
14. The test tasks provide the test takers with experience in doing future language use tasks.					
15. The test tasks are similar to the language use activities that the test takers have experienced in class.					
16. The purpose of the Language Use test is consistent with the goals and the values mentioned in the instructional program.					
17. The test tasks can reflect test takers' Language Use abilities currently demanded by Thai society and educational system					
Practicality					
18. The test tasks can possibly be constructed with existing resources.					
19. The Language Use test can be administered and scored without much difficulty.					

Reading Test

Test Usefulness	Levels of agreement				
	5	4	3	2	1
Construct Validity					
1. The test objective of the reading section is clearly defined.					
2. The test objectives of the reading section are in accordance with the purpose of the reading test.					
3. The reading tasks are all in accordance with the reading test objectives.					
4. The scoring methods are appropriate, and they can reflect all reading abilities of the test takers.					
5. The test takers' reading scores can predict their reading abilities.					
Reliability					
6. The expected response of each test item is sufficiently specific for one correct answer.					
7. The scores of the tests are reliable, irrespective of any raters.					
8. The response to each test item is reliable, irrespective of any test-takers, any time, and any place for test-taking.					
9. The test input in each part of the test does not contribute to understanding or knowledge in doing other items of other parts.					
Authenticity					
10. The reading tasks are related to the test takers' future real-life tasks.					
11. The reading texts are authentic, and they reflect real-life situations.					
Interactiveness					
12. The reading tasks are familiar topics for all test takers.					
13. The reading tasks are suitable for the English knowledge of all the test takers.					
Impact					
14. The reading tasks provide the test takers with experience in doing future reading tasks.					
15. The reading tasks are similar to the reading activities that the test takers have experienced in class.					
16. The purpose of the reading test is consistent with the goals and the values mentioned in the instructional program.					
17. Reading tasks can reflect test takers' reading abilities currently demanded by Thai society and educational system					
Practicality					
18. The reading test can possibly be constructed with existing resources.					
19. The reading test can be administered and scored without much difficulty.					

Writing Test

Test Usefulness	Levels of agreement				
	5	4	3	2	1
Construct Validity					
1. The test objectives of each writing task are clearly defined.					
2. The test objectives of each writing task are in accordance with the purpose of the writing test.					
3. The writing tasks are all in accordance with the writing test objectives.					
4. The scoring methods are appropriate, and they can reflect the test takers' writing abilities.					
5. The test takers' scores can predict their writing abilities.					
Reliability					
6. The scoring criteria is specific enough to make scoring writing tasks generally fair for all test takers.					
7. The scores of the tests are reliable, irrespective of any raters.					
8. The response to each test task is reliable, irrespective of any test-takers, any time, and any place for test-taking.					
9. The test input in each part of the test does not contribute to understanding or knowledge in doing other items of other parts.					
Authenticity					
10. The writing tasks are related to the test takers' future real-life tasks.					
11. The writing tasks are authentic and reflect real-life situations.					
Interactiveness					
12. The writing tasks are familiar topics for all test takers.					
13. The writing tasks are suitable for the English knowledge of all the test takers.					
Impact					
14. The writing tasks provide the test takers with experience in doing future language use tasks.					
15. The writing tasks are similar to the writing activities that the test takers have experienced in class.					
16. The purpose of the writing test is consistent with the goals and the values mentioned in the instructional program.					
17. The writing tasks can reflect test takers' writing abilities currently demanded by Thai society and educational system					
Practicality					
18. The writing tasks can possibly be constructed with existing resources.					
19. The writing test can be administered and scored without much difficulty.					

Speaking Test

Test Usefulness	Levels of agreement				
	5	4	3	2	1
Construct Validity					
1. The test objectives of each speaking part are clearly defined.					
2. The test objectives of each speaking part are in accordance with the purpose of the speaking test.					
3. The speaking tasks are all in accordance with the speaking test objectives.					
4. The scoring methods are appropriate, and they can reflect the test takers' speaking abilities.					
5. The test takers' scores can predict their speaking abilities.					
Reliability					
6. The scoring criteria are specific enough to make scoring speaking tasks generally fair for all test takers.					
7. The scores of the tests are reliable, irrespective of any raters.					
8. The response to each test item is reliable, irrespective of any test-takers, any time, and any place for test-taking.					
9. The test input in each part of the test does not contribute to understanding or knowledge in doing other items of other parts.					
Authenticity					
10. The speaking tasks are related to the test takers' future real-life tasks.					
11. The speaking tasks are authentic and reflect real-life situations.					
Interactiveness					
12. The speaking tasks are familiar topics for all test takers.					
13. The speaking tasks are suitable for the English knowledge of all the test takers.					
Impact					
14. The speaking tasks provide the test takers with experience in doing future language use tasks.					
15. The speaking tasks are similar to the speaking activities that the test takers have experienced in class.					
16. The purpose of the speaking test is consistent with the goals and the values mentioned in the instructional program.					
17. The speaking tasks can reflect test takers' speaking abilities currently demanded by Thai society and educational system					
Practicality					
18. The speaking tasks can possibly be constructed with existing resources.					
19. The speaking test can be administered and scored without much difficulty.					

ภาคผนวก ญ
สรุปค่าดัชนีความสอดคล้องของความตรงเชิงเนื้อหา (IOC)
ของแบบทดสอบวัดความสามารถระดับมหาวิทยาลัย: ด้านการฟัง

Part/ Item	Test Objectives/Listening Sub-skills	Skill Level	Rater No.			Sum	Mean	IOC	สรุป
			1	2	3				
Q-1	Focus on the second line. (S-1)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-2	Focus on the second line. (S-1)	Comprehension	1	0	1	2	0.67	0.83	เก็บไว้
Q-3	Choose answers with synonyms. (S-2)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-4	Choose answers with synonyms. (S-2)	Comprehension	1	1	0	2	0.67	0.83	เก็บไว้
Q-5	Choose answers with synonyms. (S-2)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-6	Choose answers with synonyms. (S-2)	Comprehension	1	1	0	2	0.67	0.83	เก็บไว้
Q-7	Choose answers with synonyms. (S-2)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-8	Avoid similar sounds. (S-3)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-9	Avoid similar sounds. (S-3)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-10	Draw conclusions about Who . (S-4)	Analysis	1	1	1	3	1	1.00	เก็บไว้
Q-11	Draw conclusions about Where . (S-4)	Analysis	1	1	1	3	1	1.00	เก็บไว้
Q-12	Draw conclusions about What . (S-4)	Analysis	1	1	1	3		1.00	เก็บไว้
Q-13	Listen for Who and What in passives. (S-5)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-14	Listen for Who and What in passives. (S-5)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-15	Listen for Who and What in passives. (S-5)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-16	Listen for Who and What with multiple nouns. (S-6)	Comprehension	1	1	0	2	0.67	0.83	เก็บไว้
Q-17	Listen for Who and What with multiple nouns. (S-6)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-18	Listen for negative expressions. (S-7)	Comprehension	1	1	0	2	0.67	0.83	เก็บไว้
Q-19	Listen for negative expressions. (S-7)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-20	Listen for negative expressions. (S-7)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-21	Listen for double negative expressions. (S-8)	Comprehension	1	1	0	2	0.67	0.83	เก็บไว้
Q-22	Listen for double negative expressions. (S-8)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-23	Listen for double negative expressions. (S-8)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-24	Listen for “almost negative” expressions. (S-9)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-25	Listen for “almost negative” expressions. (S-9)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-26	Listen for “almost negative” expressions. (S-9)	Comprehension	1	1	1	3	1	1.00	เก็บไว้

Part/ Item	Test Objectives/Listening Sub-skills	Skill Level	Rater No.			Sum	Mean	IOC	สรุป
			1	2	3				
Q-27	Listen for negatives with comparatives. (S-10)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-28	Listen for negatives with comparatives. (S-10)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-29	Listen for negatives with comparatives. (S-10)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-30	Listen for expressions of agreement. (S-11)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-31	Listen for expressions of agreement. (S-11)	Comprehension	1	1	0	2	0.67	0.83	เก็บไว้
Q-32	Listen for expressions of uncertainty and suggestion. (S-12)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-33	Listen for expressions of uncertainty and suggestion. (S-12)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-34	Listen for expressions of uncertainty and suggestion. (S-12)	Comprehension	1	1	0	2	0.67	0.83	เก็บไว้
Q-35	Listen for emphatic expressions of surprise. (S-13)	Comprehension	1	1	0	2	0.67	0.83	เก็บไว้
Q-36	Listen for emphatic expressions of surprise. (S-13)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-37	Listen for emphatic expressions of surprise. (S-13)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-38	Listen for wishes. (S-14)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-39	Listen for wishes. (S-14)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-40	Listen for untrue conditions. (S-15)	Analysis	1	1	1	3	1	1.00	เก็บไว้
Q-41	Listen for untrue conditions. (S-15)	Analysis	1	1	1	3	1	1.00	เก็บไว้
Q-42	Listen for untrue conditions. (S-15)	Analysis	1	1	1	3	1	1.00	เก็บไว้
Q-43	Listen for two-and three-part verbs. (S-16)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-44	Listen for two-and three-part verbs. (S-16)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-45	Listen for two-and three-part verbs. (S-16)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-46	Listen for idioms. (S-17)	Analysis	1	1	1	3	1	1.00	เก็บไว้
Q-47	Listen for idioms. (S-17)	Analysis	1	1	1	3	1	1.00	เก็บไว้
Q-48	Listen for idioms. (S-17)	Analysis	1	1	1	3	1	1.00	เก็บไว้
Q-49	Listen for idioms. (S-17)	Analysis	1	1	1	3	1	1.00	เก็บไว้
B: Long Conversations									
Q-50	Answer stated detail questions correctly. (B-2)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-51	Answer implied detail questions correctly. (B-3)	Analysis	1	1	1	3	1	1.00	เก็บไว้

Part/ Item	Test Objectives/Listening Sub-skills	Skill Level	Rater No.			Sum	Mean	IOC	สรุป
			1	2	3				
Q-52	Answer stated detail questions correctly. (B-2)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-53	Answer stated detail questions correctly. (B-2)	Comprehension	1	1	0	2	0.67	0.83	เก็บไว้
Q-54	Answer stated detail questions correctly. (B-2)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-55	Answer stated detail questions correctly. (B-2)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-56	Answer implied detail questions correctly. (B-3)	Analysis	1	1	0	2	0.67	0.83	เก็บไว้
Q-57	Answer stated detail questions correctly. (B-2)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-58	Identify who the speakers are (B-1)	Comprehension	1	1	0	2	0.67	0.83	เก็บไว้
Q-59	Answer stated detail questions correctly. (B-2)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-60	Answer stated detail questions correctly. (B-2)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-61	Answer stated detail questions correctly. (B-2)	Comprehension	1	1	0	2	0.67	0.83	เก็บไว้
Q-62	Answer stated detail questions correctly. (B-2)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-63	Answer stated detail questions correctly. (B-2)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-64	Find "Unstated" details. (B-4)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-65	Answer stated detail questions correctly. (B-2)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-66	Answer stated detail questions correctly. (B-2)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
C: Talks and Lectures									
Q-67	Identify the topic of the talk. (C-1)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-68	Answer stated detail questions correctly. (C-2)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-69	Answer stated detail questions correctly. (C-2)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-70	Answer stated detail questions correctly. (C-2)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-71	Answer stated detail questions correctly. (C-2)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-72	Find "Unstated" details. (C-3)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-73	Identify the topic of the talk. (C-1)	Comprehension	1	1	0	2	0.67	0.83	เก็บไว้
Q-74	Answer implied detail questions correctly. (C-4)	Analysis	1	1	0	2	0.67	0.83	เก็บไว้

Part/ Item	Test Objectives/Listening Sub-skills	Skill Level	Rater No.			Sum	Mean	IOC	สรุป
			1	2	3				
Q-75	Answer stated detail questions correctly. (C-2)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-76	Find “Unstated” details. (C-3)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-77	Answer stated detail questions correctly. (C-2)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-78	Answer stated detail questions correctly. (C-2)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-79	Find “Unstated” details. (C-3)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-80	Answer implied detail questions correctly. (C-4)	Analysis	1	1	-1	1	0.33	0.50	เก็บไว้
Q-81	Answer implied detail questions correctly. (C-4)	Analysis	1	1	0	2	0.67	0.83	เก็บไว้
Q-82	Answer stated detail questions correctly. (C-2)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-83	Answer stated detail questions correctly. (C-2)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-84	Find “Unstated” details. (C-3)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-85	Answer stated detail questions correctly. (C-2)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-86	Find “Unstated” details. (C-3)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-87	Answer implied detail questions correctly. (C-4)	Analysis	1	1	-1	1	0.33	0.50	เก็บไว้

ข้อเสนอแนะ/ความคิดเห็น

- ผู้เชี่ยวชาญคนที่ 1
- ข้อที่ 4 ให้แก้ไขคำตอบถูก
 - ข้อ 17 น่าจะใช้ advice แทนคำว่า orders
- ผู้เชี่ยวชาญคนที่ 2
- ข้อ 2 go น่าจะเป็น get
 - ข้อ 9 ควรปรับคำถาม
 - ข้อ 11 finish เติม -ed
 - ข้อ 22 you เป็น yours และเติม the หน้า tickets
 - ข้อ 48 เติม I หลัง after
- ผู้เชี่ยวชาญคนที่ 3
- ข้อ 6 may not be synonyms but inference
 - ข้อ 11 finish เติม -ed
 - ข้อ 26 option C and D have almost the same meaning
 - ข้อ 53 from the idiom “under the weather” – may not be consistent with “stated detail”
 - ข้อ 56 “That’s a relief” may be “stated detail”, not implied detailed”

ข้อ 74 “economics professor” may not be “implied detail”

ข้อ 80 This question is identifying author’s/speaker’s purpose as skill and not answering “implied detail”

ข้อ 87 This question seems to be identifying author’s/speaker’s attitude and may not be “implied detail”

สรุปได้ว่าค่า IOC รายข้อของแบบทดสอบด้านการฟัง อยู่ระหว่าง 0.50 -1.00 หลังจากปรับปรุงข้อผิดพลาดด้านการพิมพ์และไวยากรณ์แล้ว มีข้อคำถามที่สามารถนำไปใช้ได้ 87 ข้อ



สรุปค่าดัชนีความสอดคล้องของความตรงเชิงเนื้อหา (IOC)
ของแบบทดสอบวัดความสามารถระดับมหาวิทยาลัย: ด้านการใช้ภาษา

Part/ Item	Test Objectives/Listening Sub-skills	Skill Level	Rater No.			Sum	Mean	IOC	สรุป
			1	2	3				
A: Structure									
Q-1	Be sure the sentence has a subject and a verb. (S-1)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-2	Be sure the sentence has a subject and a verb. (S-1)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-3	Be sure the sentence has a subject and a verb. (S-1)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-4	Be careful of objects of prepositions. (S-2)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-5	Be careful of objects of prepositions. (S-2)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-6	Be careful of present participles. (S-4)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-7	Be careful of present participles. (S-4)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-8	Use reduced adjective clause correctly. (S-13)	Comprehension	0	1	-1	0	0	0.50	ปรับปรุง
Q-9	Use reduced adjective clause correctly. (S-13)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-10	Use adverb <i>time</i> and <i>cause</i> connectors correctly. (S-7)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-11	Use adverb <i>time</i> and <i>cause</i> connectors correctly. (S-7)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-12	Use noun clause connector / subjects correctly. (S-10)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-13	Use other adverb connectors correctly. (S-8)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-14	Use other adverb connectors correctly. (S-8)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-15	Inverted the subject and verb with place expressions. (S-16)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-16	Be careful of past participles. (S-5)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-17	Use noun clause connectors correctly. (S-9)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-18	Invert the subject and verb with conditionals. (S-18)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-19	Use reduced adverb clauses correctly. (S-14)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-20	Invert the subject and verb with negatives. (S-17)	Comprehension	1	1	1	3	1	1.00	เก็บไว้

Part/ Item	Test Objectives/Listening Sub-skills	Skill Level	Rater No.			Sum	Mean	IOC	สรุป
			1	2	3				
Q-21	Use adjective clause connectors correctly. (S-11)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-22	Invert the subject and verb with comparisons. (S-19)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
B. Written Expression									
Q-23	After <i>be</i> , use the present participle or the past participle. (S-31)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-24	After <i>be</i> , use the present participle or the past participle. (S-31)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-25	Form comparatives and superlatives correctly. (S-27)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-26	Form comparatives and superlatives correctly. (S-27)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-27	Make verbs agree after certain words. (S-23)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-28	Make verbs agree after certain words. (S-23)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-29	After <i>will</i> , <i>would</i> , or other modals, use the base form of the verb. (S-32)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-30	After <i>will</i> , <i>would</i> , or other modals, use the base form of the verb. (S-32)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-31	Use the correct tense with time expressions. (S-35)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-32	Use the correct tense with time expressions. (S-35)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-33	Distinguish subject and object pronouns. (S-43)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-34	Distinguish subject and object pronouns. (S-43)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-35	Distinguish possessive adjectives and pronoun. (S-44)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-36	Distinguish possessive adjectives and pronoun. (S-44)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-37	Use parallel structure with paired conjunctions. (S-25)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-38	Use parallel structure with paired conjunctions. (S-25)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-39	Make verbs agree after prepositional phrases. (S-20)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-40	Use comparatives and superlatives correctly. (S-28)	Comprehension	1	1	1	3		1.00	เก็บไว้

Part/ Item	Test Objectives/Listening Sub-skills	Skill Level	Rater No.			Sum	Mean	IOC	สรุป
			1	2	3				
Q-41	Use comparatives and superlatives correctly. (S-28)	Comprehension	1	1	0	2	0.67	0.83	เก็บไว้
Q-42	Distinguish a and an. (S-53)	Comprehension	1	1	0	2	0.67	0.83	เก็บไว้
Q-43	Distinguish a and an. (S-53)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-44	Use the correct tense with will and would. (S-36)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-45	Distinguish like, alike, unlike, and dislike. (S-59)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-46	Distinguish like, alike, unlike, and dislike. (S-59)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-47	Distinguish make and do. (S-58)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-48	Make inverted verbs agree. (S-22)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-49	Distinguish countable and uncountable nouns. (S-40)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-50	Distinguish countable and uncountable nouns. (S-40)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-51	Use articles with singular nouns. (S-52)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-52	Position adjectives and adverbs correctly. (S-48)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-53	Position adjectives and adverbs correctly. (S-48)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-54	Recognize active and passive meanings. (S-38)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-55	Recognize active and passive meanings. (S-38)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-56	Use articles with singular nouns. (S-52)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-57	Use predicate adjectives correctly. (S-50)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-58	Use the correct form of the passive. (S-37)	Comprehension	1	1	1	3	1	1.00	ซ้ำ (ตัดออก)
Q-59	Use of prepositions after certain words.	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-60	Use of prepositions after certain words.	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-61	Distinguish the person from the thing. (S-42)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-62	Use adjectives after linking verbs. (S-47)	Comprehension	1	1	0	2	0.67	0.83	เก็บไว้

ข้อเสนอแนะ/ความคิดเห็น

ผู้เชี่ยวชาญคนที่ 1

ผู้เชี่ยวชาญคนที่ 2

ข้อที่ 54 ซ้ำกับข้อ 58

ข้อที่ 20 โจทย์ไม่ได้ใส่ subskill ใน ()

แบบทดสอบในส่วนนี้ ออกข้อสอบได้ครอบคลุมตามที่ระบุไว้ใน Blue Print

ผู้เชี่ยวชาญคนที่ 3

ข้อ 8 “Where’s the supposed reduced adjective clause?”

ข้อ 42 This sentence seems unclear. How is the first clause related to the second one?

ข้อ 48 “a bunch” may be singular or plural → maybe change the noun, instead of using this collective noun.

ข้อที่ 58 same as 54

ข้อที่ 62 unclear if the item really tests this.

สรุปได้ว่าค่า IOC รายข้อของแบบทดสอบด้านการใช้ภาษา อยู่ระหว่าง 0.50 -1.00
หลังจากปรับปรุงข้อผิดพลาดด้านการพิมพ์ หลักไวยากรณ์และคำแนะนำจากผู้เชี่ยวชาญแล้ว



สรุปค่าดัชนีความสอดคล้องของความตรงเชิงเนื้อหา (IOC)
ของแบบทดสอบวัดความสามารถระดับมหาวิทยาลัย: ด้านการอ่าน

Part/ Item	Test Objectives/Listening Sub-skills	Skill Level	Rater No.			Sum	Mean	IOC	สรุป
			1	2	3				
A: English in Everyday Use (EEU)									
Q-1	To answer an implied detail question correctly. (A-3)	Analysis	1	1	1	3	1	1.00	เก็บไว้
Q-2	To answer an implied detail question correctly. (A-3)	Analysis	0	1	1	2	0.67	0.83	เก็บไว้
Q-3	To answer stated detail questions correctly. (A-1)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-4	To answer an implied detail question correctly. (A-3)	Analysis	1	1	1	3	1	1.00	เก็บไว้
Q-5	To answer stated detail questions correctly. (A-1)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-6	To find “unstated” details. (A-2)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-7	To answer stated detail questions correctly. (A-1)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-8	To answer stated detail questions correctly. (A-1)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-9	To answer stated detail questions correctly. (A-1)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-10	To answer stated detail questions correctly. (A-1)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-11	To choose a general synonym for some specific words in the reading texts. (A-5)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-12	To find “unstated” details. (A-2)	Comprehension	0	1	-1	0	0	0.50	ตัดออก
Q-13	To determine meanings from word parts. (A-4)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-14	To answer an implied detail question correctly. (A-3)	Analysis	1	1	1	3	1	1.00	เก็บไว้
Q-15	To determine meanings from word parts. (A-4)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-16	To find “unstated” details. (A-2)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-17	To answer stated detail questions correctly. (A-1)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-18	To choose a general synonym for some specific words in the reading texts. (A-5)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-19	To answer an implied detail question correctly. (A-3)	Analysis	1	1	1	3	1	1.00	เก็บไว้
Q-20	To answer stated detail questions correctly. (A-1)	Comprehension	1	1	1	3	1	1.00	เก็บไว้

Part/ Item	Test Objectives/Listening Sub-skills	Skill Level	Rater No.			Sum	Mean	IOC	สรุป
			1	2	3				
Q-21	To answer stated detail questions correctly. (A-1)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-22	To answer stated detail questions correctly. (A-1)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-23	To choose a general synonym for some specific words in the reading texts. (A-5)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-24	To choose a general synonym for some specific words in the reading texts. (A-5)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-25	To find “unstated” details. (A-2)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-26	To answer stated detail questions correctly. (A-1)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-27	To find “unstated” details. (A-2)	Comprehension	1	1	0	2	0.67	0.83	เก็บไว้
Q-28	To use context to determine meanings of simple words. (A-6)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-29	To use context to determine meanings of simple words. (A-6)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-30	To use context to determine meanings of simple words. (A-6)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-31	To use context to determine meanings of simple words. (A-6)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-32	To use context to determine meanings of simple words. (A-6)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-33	To answer stated detail questions correctly. (A-1)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-34	To answer stated detail questions correctly. (A-1)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-35	To answer an implied detail question correctly. (A-3)	Analysis	1	1	1	3	1	1.00	เก็บไว้
Q-36	To answer stated detail questions correctly. (A-1)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-37	To answer stated detail questions correctly. (A-1)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-38	To answer stated detail questions correctly. (A-1)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-39	To answer stated detail questions correctly. (A-1)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-40	To answer stated detail questions correctly. (A-1)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-41	To use context to determine meanings of simple words. (A-6)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-42	To use context to determine meanings of simple words. (A-6)	Comprehension	1	1	1	3	1	1.00	เก็บไว้

Part/ Item	Test Objectives/Listening Sub-skills	Skill Level	Rater No.			Sum	Mean	IOC	สรุป
			1	2	3				
Q-43	To use context to determine meanings of simple words. (A-6)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-44	To use context to determine meanings of simple words. (A-6)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-45	To use context to determine meanings of simple words. (A-6)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-46	To use context to determine meanings of simple words. (A-6)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-47	To use context to determine meanings of simple words. (A-6)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-48	To use context to determine meanings of simple words. (A-6)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-49	To use context to determine meanings of simple words. (A-6)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-50	To use context to determine meanings of simple words. (A-6)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
B: English for Academic Purpose (EAP)									
Q-51	To recognize the organization of the ideas in the passage. (B-9)	Analysis	1	1	1	3	1	1.00	เก็บไว้
Q-52	To answer main idea question correctly. (B-1)	Comprehension	1	0	1	2	0.67	0.83	เก็บไว้
Q-53	To answer stated detail questions correctly. (B-2)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-54	To find “unstated” details. (B-3)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-55	To determine where specific information is found. (B-8)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-56	To answer stated detail questions correctly. (B-2)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-57	To find “unstated” details. (B-3)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-58	To answer stated detail questions correctly. (B-2)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-59	To determine where specific information is found. (B-8)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-60	To determine where specific information is found. (B-8)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-61	To answer an implied detail question correctly. (B-5)	Analysis	1	1	0	2	0.67	0.83	ตัดออก
Q-62	To determine the purpose of the author. (B-12)	Analysis	1	1	1	3	1	1.00	เก็บไว้
Q-63	To choose a general synonym for some specific words in the reading texts. (B-14)	Comprehension	1	1	1	3	1	1.00	เก็บไว้

Part/ Item	Test Objectives/Listening Sub-skills	Skill Level	Rater No.			Sum	Mean	IOC	สรุป
			1	2	3				
Q-64	To choose a general synonym for some specific words in the reading texts. (B-14)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-65	To choose a general synonym for some specific words in the reading texts. (B-14)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-66	To find pronoun referents. (B-6)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-67	To answer an implied detail question correctly. (B-5)	Analysis	1	1	1	3	1	1.00	เก็บไว้
Q-68	To answer an implied detail question correctly. (B-5)	Analysis	1	1	0	2	0.67	0.83	ตัดออก
Q-69	To answer main idea question correctly. (B-1)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-70	To find “unstated” details. (B-3)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-71	To answer stated detail questions correctly. (B-2)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-72	To answer stated detail questions correctly. (B-2)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-73	To answer stated detail questions correctly. (B-2)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-74	To answer stated detail questions correctly. (B-2)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-75	To answer stated detail questions correctly. (B-2)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-76	To find “unstated” details. (B-3)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-77	To answer stated detail questions correctly. (B-2)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-78	To choose a general synonym for some specific words in the reading texts. (B-14)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-79	To answer main idea question correctly. (B-1)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-80	To find “unstated” details. (B-3)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-81	To answer stated detail questions correctly. (B-2)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-82	To answer stated detail questions correctly. (B-2)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-83	To choose a general synonym for some specific words in the reading texts. (B-14)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-84	To find pronoun referents. (B-6)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-85	To choose a general synonym for some specific words in the reading texts. (B-14)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-86	To draw a conclusion by using main points or contexts in the passage. (B-11)	Comprehension	1	1	1	3	1	1.00	เก็บไว้

Part/ Item	Test Objectives/Listening Sub-skills	Skill Level	Rater No.			Sum	Mean	IOC	สรุป
			1	2	3				
Q-87	To answer transition questions correctly. (B-16)	Analysis	1	1	1	3	1	1.00	เก็บไว้
Q-88	To answer stated detail questions correctly. (B-2)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-89	To answer stated detail questions correctly. (B-2)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-90	To answer stated detail questions correctly. (B-2)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-91	To answer stated detail questions correctly. (B-2)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-92	To answer stated detail questions correctly. (B-2)	Comprehension	1	1	0	2	0.67	0.83	เก็บไว้
Q-93	To answer stated detail questions correctly. (B-2)	Comprehension	1	1	0	2	0.67	0.83	เก็บไว้
Q-94	To choose a general synonym for some specific words in the reading texts. (B-14)	Comprehension	1	1	0	2	0.67	0.83	เก็บไว้
Q-95	To choose a general synonym for some specific words in the reading texts. (B-14)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-96	To find “unstated” details. (B-3)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-97	To answer stated detail questions correctly. (B-2)	Comprehension	1	1	1	3	1	1.00	เก็บไว้
Q-98	To answer an implied detail question correctly. (B-5)	Analysis	1	1	-1	1	0.33	0.67	ปรับปรุง
Q-99	To determine the tone and mood of the author. (B-15)	Analysis	1	1	0	2	0.67	0.83	ปรับปรุง
Q-100	To answer an implied detail question correctly. (B-5)	Analysis	1	1	1	3	1	1.00	เก็บไว้

ข้อเสนอแนะ/ความคิดเห็น

ผู้เชี่ยวชาญคนที่ 2 ข้อที่ 52 ปรับ word order
 ข้อที่ 65 เปลี่ยน with เป็น to
 ข้อที่ 79 Internet ใช้ “I”
 ข้อที่ 96 เติม “statement” หลังคำว่า which
 ออกข้อสอบในส่วนนี้ได้ครอบคลุมตาม Blue Print การอ่านตาราง แผนภูมิในหน้า 48 ควรจะเป็น
 ภาพขาว-ดำ ภาพสีอาจจะไม่ชัดเจน
 ผู้เชี่ยวชาญคนที่ 3 ข้อที่ 34 Maybe use full dates (e.g. 12 January 2016)
 ข้อที่ 41-50 Authentic use of skills
 ข้อที่ 61 In which part of the passage does it suggest
 this?

ข้อที่ 94 Option A: protect (would be more appropriate for the context/usage)

ข้อที่ 98 Line 1 (night) stated → not implied detail

ข้อที่ 99 Option A, B : very close/related meaning, please replace this with another word

สรุปได้ว่าค่า IOC รายชื่อของแบบทดสอบสอบด้านการอ่าน อยู่ระหว่าง 0.67 -1.00 หลังจากปรับปรุงข้อผิดพลาดด้านการพิมพ์ หลักไวยากรณ์และคำแนะนำจากผู้เชี่ยวชาญแล้ว มีข้อคำถามที่สามารถนำไปใช้ได้ 97 ข้อ



สรุปค่าดัชนีความสอดคล้องของความตรงเชิงเนื้อหา (IOC)
ของแบบทดสอบวัดความสามารถระดับมหาวิทยาลัย: ด้านการเขียน

Part/ Item	Test Objectives/Listening Sub-skills	Skill Level	Rater No.			Sum	Mean	IOC	สรุป
			1	2	3				
Writing Task 1									
	The ability to describe the information objectively; The ability to contrast and compare the information; The ability to use specialized graphic language, not to express your own opinion and don't evaluate the data.	Application	1	1	1	3	1	1.00	เก็บไว้
Writing Task 2									
	The ability to make solutions; The ability to express your opinions; The ability to contrast and compare the opinions and facts; The ability to evaluate facts, ideas or arguments.	Application	1	1	1	3	1	1.00	เก็บไว้

ข้อเสนอแนะ/ความคิดเห็น

ผู้เชี่ยวชาญคนที่ 2

ภาพชัดเจน คำสั่งไม่ซับซ้อน ผู้เข้าสอบสามารถปฏิบัติตามได้

Rubric ที่ใช้ประเมิน ชัดเจนรายละเอียดได้ดีมาก

ผู้เชี่ยวชาญคนที่ 3

เกณฑ์การให้คะแนนไม่สอดคล้องกับจุดประสงค์ แนะนำให้เพิ่มจุดประสงค์

สรุป เพิ่มจุดประสงค์และเกณฑ์การให้คะแนนให้สอดคล้องกัน ได้แบบประเมินการเขียน

สรุปค่าดัชนีความสอดคล้องของความตรงเชิงเนื้อหา (IOC)
ของแบบทดสอบวัดความสามารถระดับมหาวิทยาลัย: ด้านการพูด

Part/ Item	Test Objectives/Listening Sub-skills	Skill Level	Rater No.			Sum	Mean	IOC	สรุป
			1	2	3				
Short Talk									
	To answer personal questions and give information on familiar topics	Application	1	1	1	3	1	1.00	เก็บไว้
Individual Long Run									
	To give a 2-minute-talk about how to solve a problem specified in the task card.	Application	1	1	1	3	1	1.00	เก็บไว้
Two-way Discussion									
	To express opinions and discuss some possible ways to solve the problem in Task 2 effectively by giving reasons and asking the examiner to share the ideas.	Application	1	1	1	3	1	1.00	เก็บไว้

ข้อเสนอแนะ/ความคิดเห็น

ผู้เชี่ยวชาญคนที่ 2

ควรระบุเวลาใน Task Card เพื่อให้ผู้เข้าสอบได้บริหารเวลาในการเตรียมหัวข้อที่ได้รับได้เหมาะสม Rubric ที่ใช้ประเมิน แยกประเภทและชี้แจงรายละเอียดของค่าระดับคะแนนที่ผู้เข้าสอบสามารถจะได้รับในการสอบอย่างชัดเจน

ผู้เชี่ยวชาญคนที่ 3

ควรเปลี่ยนเกณฑ์การประเมิน จาก Pronunciation and Accent เป็น Intelligibility หรือ Comprehensibility

สรุป ปรับเกณฑ์การประเมินตามคำแนะนำของผู้เชี่ยวชาญ ได้แบบประเมินการพูด

ภาคผนวก ก**การวิเคราะห์ค่าความลำเอียงของแบบทดสอบ**

(เป็นการแยกกลุ่มผู้เข้าสอบตามสายวิชา คือสายวิทย์และสายศิลป์
คัดเลือกกระดาษคำตอบของผู้เข้าสอบมาวิเคราะห์ สายละ 20 ชุด)



การวิเคราะห์ค่าความลำเอียงของแบบทดสอบด้านการฟัง

Item	P_S	P_A	Z_S	Z_A	r_S	r_A	D_i	แปลผล
1	0.81	0.30	-0.88	0.53	9.50	15.10	-1.10	ลำเอียง
2	0.10	0.70	1.28	-0.53	18.10	10.90	1.03	ลำเอียง
3	0.37	0.30	0.33	0.53	14.30	15.10	-0.42	
4	0.25	0.10	0.68	1.28	15.70	18.10	-0.88	ลำเอียง
5	0.10	0.70	1.28	-0.53	18.10	10.90	1.03	ลำเอียง
6	0.19	0.50	0.88	0.00	16.50	13.00	0.34	
7	0.10	0.30	1.28	0.53	18.10	15.10	0.11	
8	0.81	0.81	-0.88	-0.88	9.50	9.50	0.12	
9	0.25	0.19	0.68	0.88	15.70	16.50	-0.53	
10	0.19	0.70	0.88	-0.53	16.50	10.90	0.80	ลำเอียง
11	0.10	0.19	1.28	0.88	18.10	16.50	-0.19	
12	0.00	0.10	-3.25	1.28	0.00	18.10	-3.10	ลำเอียง
13	0.25	0.50	0.68	0.00	15.70	13.00	0.23	
14	0.10	0.10	1.28	1.28	18.10	18.10	-0.54	
15	0.19	0.63	0.88	-0.33	16.50	11.70	0.63	
16	0.00	0.10	-3.25	1.28	0.00	18.10	-3.10	ลำเอียง
17	0.10	0.10	1.28	1.28	18.10	18.10	-0.54	
18	0.19	0.19	0.88	0.88	16.50	16.50	-0.42	
19	0.81	0.50	-0.88	0.00	9.50	13.00	-0.65	
20	0.37	0.19	0.33	0.88	14.30	16.50	-0.73	
21	0.19	0.25	0.88	0.68	16.50	15.70	-0.24	
22	0.37	0.19	0.33	0.88	14.30	16.50	-0.73	
23	0.50	0.37	0.00	0.33	13.00	14.30	-0.43	
24	0.25	0.50	0.68	0.00	15.70	13.00	0.23	
25	0.00	0.10	-3.25	1.28	0.00	18.10	-3.10	ลำเอียง
26	0.25	0.63	0.68	-0.33	15.70	11.70	0.51	
27	0.00	0.30	-3.25	0.53	0.00	15.10	-2.45	ลำเอียง
28	0.00	0.70	-3.25	-0.53	0.00	10.90	-1.53	ลำเอียง
29	0.00	0.10	-3.25	1.28	0.00	18.10	-3.10	ลำเอียง
30	0.10	0.30	1.28	0.53	18.10	15.10	0.11	
31	0.50	0.50	0.00	0.00	13.00	13.00	-0.15	
32	0.25	0.63	0.68	-0.33	15.70	11.70	0.51	
33	0.10	0.10	1.28	1.28	18.10	18.10	-0.54	
34	0.37	0.70	0.33	-0.53	14.30	10.90	0.49	
35	0.19	0.81	0.88	-0.88	16.50	9.50	1.11	ลำเอียง
36	0.63	0.50	-0.33	0.00	11.70	13.00	-0.33	
37	0.63	0.50	-0.33	0.00	11.70	13.00	-0.33	
38	0.25	0.50	0.68	0.00	15.70	13.00	0.23	
39	0.50	0.10	0.00	1.28	13.00	18.10	-1.26	ลำเอียง

การวิเคราะห์ค่าความลำเอียงของแบบทดสอบด้านการฟัง

Item	P_S	P_A	Z_S	Z_A	r_S	r_A	D_i	แปลผล
40	0.50	0.50	0.00	0.00	13.00	13.00	-0.15	
41	0.25	0.50	0.68	0.00	15.70	13.00	0.23	
42	0.50	0.75	0.00	-0.68	13.00	10.30	0.44	
43	0.00	0.19	-3.25	0.88	0.00	16.50	-2.75	ลำเอียง
44	0.19	0.81	0.88	-0.88	16.50	9.50	1.11	ลำเอียง
45	0.30	0.50	0.53	0.00	15.10	13.00	0.15	
46	0.10	0.19	1.28	0.88	18.10	16.50	-0.19	
47	0.50	0.50	0.00	0.00	13.00	13.00	-0.15	
48	0.19	0.25	0.88	0.68	16.50	15.70	-0.24	
49	0.50	0.50	0.00	0.00	13.00	13.00	-0.15	
50	0.63	0.50	-0.33	0.00	11.70	13.00	-0.33	
51	0.75	0.75	-0.68	-0.68	10.30	10.30	0.06	
52	0.10	0.19	1.28	0.88	18.10	16.50	-0.19	
53	0.63	0.90	-0.33	-1.28	11.70	7.90	0.78	ลำเอียง
54	0.19	0.25	0.88	0.68	16.50	15.70	-0.24	
55	0.00	0.19	-3.25	0.88	0.00	16.50	-2.75	ลำเอียง
56	0.50	0.81	0.00	-0.88	13.00	9.50	0.61	
57	0.10	0.19	1.28	0.88	18.10	16.50	-0.19	
58	0.63	0.50	-0.33	0.00	11.70	13.00	-0.33	
59	0.50	0.00	0.00	-3.25	13.00	0.00	2.68	ลำเอียง
60	0.10	0.37	1.28	0.33	18.10	14.30	0.29	
61	0.50	0.70	0.00	-0.53	13.00	10.90	0.31	
62	0.10	0.19	1.28	0.88	18.10	16.50	-0.19	
63	0.37	0.50	0.33	0.00	14.30	13.00	0.03	
64	0.19	0.63	0.88	-0.33	16.50	11.70	0.63	
65	0.50	0.50	0.00	0.00	13.00	13.00	-0.15	
X			0.03	0.19			จำนวนข้อที่ลำเอียง	18
S.D.			1.375	0.808			ร้อยละ	27.69
r			-0.13					
a			0.5655					
b			0.173035					

เกณฑ์ ค่า D_i มากกว่า .75 หรือ น้อยกว่า -.75 เท่ากับ ข้อคำถามนั้นมีความลำเอียง

การวิเคราะห์ค่าความลำเอียงของแบบทดสอบด้านการใช้ภาษา

Item	P_S	P_A	Z_S	Z_A	r_S	r_A	D_i	แปลผล
1	0.90	0.50	-1.28	0.00	7.90	13.00	0.40	
2	0.90	0.50	-1.28	0.00	7.90	13.00	0.40	
3	0.25	0.25	0.68	0.68	15.70	15.70	-0.54	
4	0.30	0.50	0.53	0.00	15.10	13.00	0.15	
5	0.10	0.37	1.28	0.33	18.10	14.30	-0.28	
6	0.00	0.81	-3.25	-0.88	0.00	9.50	1.55	ลำเอียง
7	0.25	0.50	0.68	0.00	15.70	13.00	0.13	
8	0.75	0.63	-0.68	-0.33	10.30	11.70	0.64	
9	0.63	0.63	-0.33	-0.33	11.70	11.70	0.59	
10	0.70	0.81	-0.53	-0.88	10.90	9.50	1.16	ลำเอียง
11	0.81	0.70	-0.88	-0.53	9.50	10.90	0.87	ลำเอียง
12	0.10	0.37	1.28	0.33	18.10	14.30	-0.28	
13	0.50	0.75	0.00	-0.68	13.00	10.30	0.89	
14	0.37	0.50	0.33	0.00	14.30	13.00	0.18	
15	0.37	0.50	0.33	0.00	14.30	13.00	0.18	
16	0.81	0.30	-0.88	0.53	9.50	15.10	-0.17	
17	0.10	0.37	1.28	0.33	18.10	14.30	-0.28	
18	0.63	0.70	-0.33	-0.53	11.70	10.90	0.79	ลำเอียง
19	0.50	0.81	0.00	-0.88	13.00	9.50	1.09	ลำเอียง
20	0.00	0.00	-3.25	-3.25	0.00	0.00	3.90	ลำเอียง
21	0.37	0.50	0.33	0.00	14.30	13.00	0.18	
22	0.50	0.81	0.00	-0.88	13.00	9.50	1.09	ลำเอียง
23	0.63	0.81	-0.33	-0.88	11.70	9.50	1.14	ลำเอียง
24	0.10	0.19	1.28	0.88	18.10	16.50	-0.82	ลำเอียง
25	0.37	0.50	0.33	0.00	14.30	13.00	0.18	
26	0.37	0.70	0.33	-0.53	14.30	10.90	0.70	
27	0.81	0.50	-0.88	0.00	9.50	13.00	0.35	
28	0.50	0.63	0.00	-0.33	13.00	11.70	0.55	
29	0.25	0.25	0.68	0.68	15.70	15.70	-0.54	
30	0.25	0.81	0.68	-0.88	15.70	9.50	1.00	ลำเอียง
31	0.10	0.63	1.28	-0.33	18.10	11.70	0.37	
32	0.19	0.30	0.88	0.53	16.50	15.10	-0.42	
33	0.25	0.37	0.68	0.33	15.70	14.30	-0.19	
34	0.37	0.50	0.33	0.00	14.30	13.00	0.18	
35	0.30	0.81	0.53	-0.88	15.10	9.50	1.02	ลำเอียง
36	0.19	0.90	0.88	-1.28	16.50	7.90	1.36	ลำเอียง
37	0.50	0.10	0.00	1.28	13.00	18.10	-1.04	ลำเอียง
38	0.10	0.00	1.28	-3.25	18.10	0.00	3.26	ลำเอียง
39	0.50	0.37	0.00	0.33	13.00	14.30	-0.10	

การวิเคราะห์ค่าความลำเอียงของแบบทดสอบด้านการใช้ภาษา

Item	P_S	P_A	Z_S	Z_A	r_S	r_A	D_i	แปลผล
40	0.30	0.63	0.53	-0.33	15.10	11.70	0.47	
41	0.50	0.70	0.00	-0.53	13.00	10.90	0.74	
42	0.50	0.70	0.00	-0.53	13.00	10.90	0.74	
43	0.50	0.81	0.00	-0.88	13.00	9.50	1.09	ลำเอียง
44	0.37	0.50	0.33	0.00	14.30	13.00	0.18	
45	0.50	0.70	0.00	-0.53	13.00	10.90	0.74	
46	0.37	0.70	0.33	-0.53	14.30	10.90	0.70	
47	0.00	0.37	-3.25	0.33	0.00	14.30	0.36	
48	0.10	0.10	1.28	1.28	18.10	18.10	-1.22	ลำเอียง
49	0.50	0.19	0.00	0.88	13.00	16.50	-0.64	
50	0.25	0.19	0.68	0.88	15.70	16.50	-0.74	
X			0.04	-0.22			จำนวนข้อที่ลำเอียง	16
S.D.			1.065	0.868			ร้อยละ	32.00
r			0.29					
a			-0.1404					
b			-0.225616					

เกณฑ์ ค่า D_i มากกว่า .75 หรือ น้อยกว่า -.75 เท่ากับ ข้อคำถามนั้นมีความลำเอียง



การวิเคราะห์ค่าความลำเอียงของแบบทดสอบด้านการอ่าน (การอ่านในชีวิตประจำวัน)

Item	P_S	P_A	Z_S	Z_A	r_S	r_A	D_i	แปลผล
1	0.30	0.37	0.53	0.33	15.10	14.30	0.29	
2	0.46	0.70	0.10	-0.53	13.40	10.90	0.60	
3	0.19	0.50	0.88	0.00	16.50	13.00	0.83	ลำเอียง
4	0.10	0.30	1.28	0.53	18.10	15.10	0.76	ลำเอียง
5	0.10	0.25	1.28	0.68	18.10	15.70	0.64	
6	0.10	0.37	1.28	0.33	18.10	14.30	0.91	ลำเอียง
7	0.37	0.30	0.33	0.53	14.30	15.10	-0.02	
8	0.37	0.50	0.33	0.00	14.30	13.00	0.38	
9	0.37	0.50	0.33	0.00	14.30	13.00	0.38	
10	0.25	0.63	0.68	-0.33	15.70	11.70	0.92	ลำเอียง
11	0.00	0.70	-3.25	-0.53	0.00	10.90	-2.15	ลำเอียง
12	0.25	0.50	0.68	0.00	15.70	13.00	0.67	
13	0.37	0.37	0.33	0.33	14.30	14.30	0.13	
14	0.10	0.10	1.28	1.28	18.10	18.10	0.18	
15	0.63	0.63	-0.33	-0.33	11.70	11.70	0.10	
16	0.37	0.53	0.33	-0.08	14.30	12.70	0.44	
17	0.50	0.75	0.00	-0.68	13.00	10.30	0.64	
18	0.30	0.70	0.53	-0.53	15.10	10.90	0.95	ลำเอียง
19	0.50	0.90	0.00	-1.28	13.00	7.90	1.10	ลำเอียง
20	0.37	0.37	0.33	0.33	14.30	14.30	0.13	
21	0.19	0.63	0.88	-0.33	16.50	11.70	1.08	ลำเอียง
22	0.10	0.50	1.28	0.00	18.10	13.00	1.16	ลำเอียง
23	0.10	0.81	1.28	-0.88	18.10	9.50	1.84	ลำเอียง
24	0.00	0.30	-3.25	0.53	0.00	15.10	-2.96	ลำเอียง
25	0.00	0.37	-3.25	0.33	0.00	14.30	-2.81	ลำเอียง
26	0.50	0.90	0.00	-1.28	13.00	7.90	1.10	ลำเอียง
27	0.10	0.30	1.28	0.53	18.10	15.10	0.76	ลำเอียง
28	0.10	0.19	1.28	0.88	18.10	16.50	0.49	
29	0.00	0.37	-3.25	0.33	0.00	14.30	-2.81	ลำเอียง
30	0.19	0.37	0.88	0.33	16.50	14.30	0.58	
31	0.00	0.10	-3.25	1.28	0.00	18.10	-3.54	ลำเอียง
32	0.10	0.50	1.28	0.00	18.10	13.00	1.16	ลำเอียง
33	0.19	0.50	0.88	0.00	16.50	13.00	0.83	
34	0.10	0.50	1.28	0.00	18.10	13.00	1.16	ลำเอียง
35	0.25	0.00	0.68	-3.25	15.70	0.00	3.18	ลำเอียง
36	0.50	0.70	0.00	-0.53	13.00	10.90	0.52	
37	0.00	0.81	-3.25	-0.88	0.00	9.50	-1.88	ลำเอียง
38	0.50	0.81	0.00	-0.88	13.00	9.50	0.79	ลำเอียง
39	0.00	0.19	-3.25	0.88	0.00	16.50	-3.23	ลำเอียง

การวิเคราะห์ค่าความลำเอียงของแบบทดสอบด้านการอ่าน (การอ่านในชีวิตประจำวัน)

Item	P_S	P_A	Z_S	Z_A	r_S	r_A	D_i	แปลผล
40	0.00	0.19	-3.25	0.88	0.00	16.50	-3.23	ลำเอียง
X			-0.12	-0.05			จำนวนข้อที่ลำเอียง	23
S.D.			1.648	0.812			ร้อยละ	46.00
r			-0.14					
a			0.8214					
b			-0.1485					

เกณฑ์ ค่า D_i มากกว่า .75 หรือ น้อยกว่า -.75 เท่ากับ ข้อคำถามนั้นมีความลำเอียง



การวิเคราะห์ค่าความลำเอียงของแบบทดสอบด้านการอ่าน (การอ่านเชิงวิชาการ)

Item	P_S	P_A	Z_S	Z_A	r_S	r_A	D_I	แปลผล
1	0.25	0.75	0.68	-0.68	15.70	10.30	-0.10	
2	0.19	0.00	0.88	-3.25	16.50	0.00	1.43	ลำเอียง
3	0.19	0.70	0.88	-0.53	16.50	10.90	-0.42	
4	0.10	0.10	1.28	1.28	18.10	18.10	-2.07	ลำเอียง
5	0.50	0.50	0.00	0.00	13.00	13.00	0.17	
6	0.10	0.63	1.28	-0.33	18.10	11.70	-0.99	
7	0.10	0.50	1.28	0.00	18.10	13.00	-1.21	ลำเอียง
8	0.25	0.30	0.68	0.53	15.70	15.10	-0.91	ลำเอียง
9	0.19	0.50	0.88	0.00	16.50	13.00	-0.77	ลำเอียง
10	0.10	0.30	1.28	0.53	18.10	15.10	-1.56	ลำเอียง
11	0.00	0.50	-3.25	0.00	0.00	13.00	3.69	ลำเอียง
12	0.50	0.19	0.00	0.88	13.00	16.50	-0.42	
13	0.10	0.30	1.28	0.53	18.10	15.10	-1.56	ลำเอียง
14	0.25	0.50	0.68	0.00	15.70	13.00	-0.56	
15	0.19	0.50	0.88	0.00	16.50	13.00	-0.77	ลำเอียง
16	0.19	0.63	0.88	-0.33	16.50	11.70	-0.55	
17	0.19	0.70	0.88	-0.53	16.50	10.90	-0.42	
18	0.25	0.50	0.68	0.00	15.70	13.00	-0.56	
19	0.30	0.37	0.53	0.33	15.10	14.30	-0.62	
20	0.19	0.50	0.88	0.00	16.50	13.00	-0.77	ลำเอียง
21	0.19	0.81	0.88	-0.88	16.50	9.50	-0.18	
22	0.37	0.90	0.33	-1.28	14.30	7.90	0.69	
23	0.25	0.50	0.68	0.00	15.70	13.00	-0.56	
24	0.19	0.37	0.88	0.33	16.50	14.30	-1.00	ลำเอียง
25	0.30	0.50	0.53	0.00	15.10	13.00	-0.40	
26	0.00	0.30	-3.25	0.53	0.00	15.10	3.33	ลำเอียง
27	0.50	0.19	0.00	0.88	13.00	16.50	-0.42	
28	0.50	0.70	0.00	-0.53	13.00	10.90	0.53	
29	0.63	0.30	-0.33	0.53	11.70	15.10	0.17	
30	0.19	0.30	0.88	0.53	16.50	15.10	-1.13	ลำเอียง
31	0.37	0.30	0.33	0.53	14.30	15.10	-0.54	
32	0.37	0.10	0.33	1.28	14.30	18.10	-1.05	ลำเอียง
33	0.25	0.70	0.68	-0.53	15.70	10.90	-0.20	
34	0.00	0.19	-3.25	0.88	0.00	16.50	3.09	ลำเอียง
35	0.50	0.70	0.00	-0.53	13.00	10.90	0.53	
36	0.25	0.10	0.68	1.28	15.70	18.10	-1.42	ลำเอียง
37	0.10	0.10	1.28	1.28	18.10	18.10	-2.07	ลำเอียง
38	0.00	0.37	-3.25	0.33	0.00	14.30	3.47	ลำเอียง
39	0.25	0.50	0.43	0.00	14.70	13.00	-0.29	

การวิเคราะห์ค่าความลำเอียงของแบบทดสอบด้านการอ่าน (การอ่านเชิงวิชาการ)

Item	P_S	P_A	Z_S	Z_A	r_S	r_A	D_i	แปลผล
40	0.19	0.50	0.88	0.00	16.50	13.00	-0.77	ลำเอียง
41	0.19	0.70	0.88	-0.53	16.50	10.90	-0.42	
42	0.63	0.70	-0.33	-0.53	11.70	10.90	0.88	ลำเอียง
X			0.28	0.05			จำนวนข้อที่ลำเอียง	20
S.D.			1.234	0.804			ร้อยละ	47.61
r			-0.14					
a			-1.0817					
b			-0.2528					

เกณฑ์ ค่า D_i มากกว่า .75 หรือ น้อยกว่า -.75 เท่ากับ ข้อคำถามนั้นมีความลำเอียง



ภาคผนวก ก

ผลการสำรวจความคิดเห็นอาจารย์มหาวิทยาลัย
ต่อโครงสร้างของแบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย



Survey Report

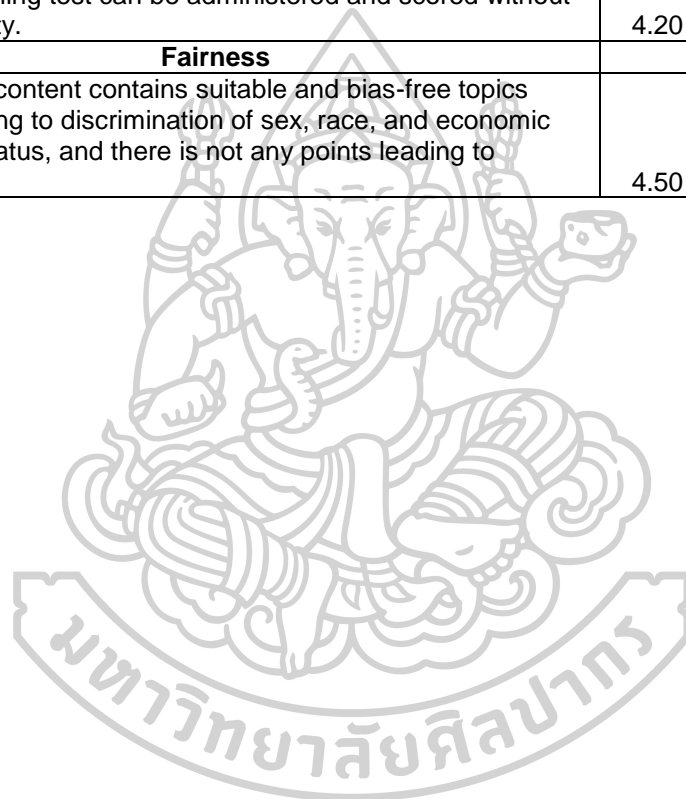
A Survey of University English Lecturers' Opinions towards the Test Usefulness of the University English Proficiency Test

4.51 – 5.00	means	Strongly agree
3.51 – 4.50	means	Agree
2.51 – 3.50	means	Moderately agree
1.51 – 2.50	means	Disagree
1.00 – 1.50	means	Strongly disagree

Listening Test

Test Usefulness	X	S.D.	Meaning
Construct Validity	4.52	0.110	Strongly agree
1. The test objective of each listening part is clearly defined.	4.60	0.699	Strongly agree
2. The test objective of each listening part is in accordance with the purpose of the listening test.	4.40	0.516	Agree
3. All listening tasks are in accordance with the listening test objectives.	4.60	0.516	Strongly agree
4. The scoring methods are appropriate, and they can reflect the test takers' listening abilities.	4.40	0.516	Agree
5. The test takers' scores can predict their listening abilities.	4.60	0.516	Strongly agree
Reliability	4.25	0.191	Agree
6. The expected response of each test item is sufficiently specific for one correct answer.	4.40	0.516	Agree
7. The scores of the tests are reliable, irrespective of any raters.	4.20	0.789	Agree
8. The response to each test item is reliable, irrespective of any test-takers, any time, and any place for test-taking.	4.00	0.000	Agree
9. The test input in each part of the test does not contribute to understanding or knowledge in doing other items of other parts.	4.40	0.516	Agree
Authenticity	4.40	0.000	Agree
10. The listening tasks are related to the test takers' future real-life tasks.	4.40	0.516	Agree
11. The listening texts are authentic, and they reflect real-life situations.	4.40	0.516	Agree
Interactiveness	4.30	0.224	Agree
12. The listening tasks are familiar topics for all test takers.	4.00	0.000	Agree
13. The listening tasks are suitable for the English knowledge of all the test takers.	4.30	0.483	Agree

Test Usefulness	X	S.D.	Meaning
Impact	4.40	0.432	Agree
14. The listening tasks provide the test takers with experience in doing future listening tasks.	4.00	0.000	Agree
15. The listening tasks are similar to the listening activities that the test takers have experienced in class.	5.00	0.000	Strongly agree
16. The purpose of the listening test is consistent with the goals and the values mentioned in the instructional program.	4.40	0.516	Agree
17. Listening tasks can reflect test takers' listening abilities currently demanded by Thai society and educational system.	4.20	0.422	Agree
Practicality	4.30	0.141	Agree
18. The listening test can possibly be constructed with existing resources.	4.40	0.516	Agree
19. The listening test can be administered and scored without much difficulty.	4.20	0.422	Agree
Fairness			
20. The test content contains suitable and bias-free topics without leading to discrimination of sex, race, and economic and social status, and there is not any points leading to violence.	4.50	0.527	Agree



Language Use Test

Test Usefulness	X	S.D.	Meaning
Construct Validity	4.20	0.100	Agree
1. The test objective of each Language Use part is clearly defined.	4.30	0.675	Agree
2. The test objective of each Language Use part is in accordance with the purpose of the Language Use test.	4.20	0.789	Agree
3. All Language Use tasks are in accordance with the Language Use test objectives.	4.10	0.738	Agree
4. The scoring methods are appropriate, and they can reflect the test takers' Language Use abilities.	4.10	0.738	Agree
5. The test takers' scores can predict their Language Use abilities.	4.30	0.675	Agree
Reliability	4.18	0.096	Agree
6. The expected response of each test item is sufficiently specific for one correct answer.	4.20	0.789	Agree
7. The scores of the tests are reliable, irrespective of any raters.	4.10	0.738	Agree
8. The response to each test item is reliable, irrespective of any test-takers, any time, and any place for test-taking.	4.10	0.738	Agree
9. The test input in each part of the test does not contribute to understanding or knowledge in doing other items of other parts.	4.30	0.675	Agree
Authenticity	4.20	0.141	Agree
10. The test tasks are related to the test takers' future real-life tasks.	4.10	0.738	Agree
11. The input texts are authentic and reflect real-life situations.	4.30	0.675	Agree
Interactiveness	4.16	0.089	Agree
12. The test tasks are familiar topics for all test takers.	4.20	0.789	Agree
13. The test tasks are suitable for the English knowledge of all the test takers.	4.10	0.738	Agree
Impact	4.25	0.100	Agree
14. The test tasks provide the test takers with experience in doing future language use tasks.	4.30	0.675	Agree
15. The test tasks are similar to the language use activities that the test takers have experienced in class.	4.30	0.675	Agree
16. The purpose of the Language Use test is consistent with the goals and the values mentioned in the instructional program.	4.10	0.738	Agree
17. The test tasks can reflect test takers' Language Use abilities currently demanded by Thai society and educational system	4.30	0.675	Agree
Practicality	4.15	0.071	Agree
18. The test tasks can possibly be constructed with existing resources.	4.20	0.789	Agree
19. The Language Use test can be administered and scored without much difficulty.	4.10	0.738	Agree
Fairness			
20. The test content contains suitable and bias-free topics without leading to discrimination of sex, race, and economic and social status, and there is not any points leading to violence.	4.70	0.483	Strongly agree

Reading Test

Test Usefulness	X	S.D.	Meaning
Construct Validity	4.14	0.207	Agree
1. The test objective of the reading section is clearly defined.	4.30	0.675	Agree
2. The test objectives of the reading section are in accordance with the purpose of the reading test.	4.10	0.568	Agree
3. The reading tasks are all in accordance with the reading test objectives.	4.20	0.789	Agree
4. The scoring methods are appropriate, and they can reflect all reading abilities of the test takers.	4.30	0.675	Agree
5. The test takers' reading scores can predict their reading abilities.	3.80	0.789	Agree
Reliability	4.10	0.216	Agree
6. The expected response of each test item is sufficiently specific for one correct answer.	4.00	0.667	Agree
7. The scores of the tests are reliable, irrespective of any raters.	4.40	0.516	Agree
8. The response to each test item is reliable, irrespective of any test-takers, any time, and any place for test-taking.	4.10	0.876	Agree
9. The test input in each part of the test does not contribute to understanding or knowledge in doing other items of other parts.	3.90	0.876	Agree
Authenticity	4.10	0.141	Agree
10. The reading tasks are related to the test takers' future real-life tasks.	4.00	0.471	Agree
11. The reading texts are authentic, and they reflect real-life situations.	4.20	0.789	Agree
Interactiveness	4.14	0.055	Agree
12. The reading tasks are familiar topics for all test takers.	4.20	0.632	Agree
13. The reading tasks are suitable for the English knowledge of all the test takers.	4.20	0.789	Agree
Impact	4.23	0.096	Agree
14. The reading tasks provide the test takers with experience in doing future reading tasks.	4.30	0.675	Agree
15. The reading tasks are similar to the reading activities that the test takers have experienced in class.	4.30	0.675	Agree
16. The purpose of the reading test is consistent with the goals and the values mentioned in the instructional program.	4.20	0.789	Agree
17. Reading tasks can reflect test takers' reading abilities currently demanded by Thai society and educational system	4.10	0.568	Agree
Practicality	3.95	0.212	Agree
18. The reading test can possibly be constructed with existing resources.	3.80	0.789	Agree
19. The reading test can be administered and scored without much difficulty.	4.10	0.738	Agree
Fairness			
20. The test content contains suitable and bias-free topics without leading to discrimination of sex, race, and economic and social status, and there is not any points leading to violence.	4.70	0.483	Strongly agree

Writing Test

Test Usefulness	X	S.D.	Meaning
Construct Validity	4.20	0.100	Agree
1. The test objectives of each writing task are clearly defined.	4.10	0.738	Agree
2. The test objectives of each writing task are in accordance with the purpose of the writing test.	4.30	0.675	Agree
3. The writing tasks are all in accordance with the writing test objectives.	4.10	0.738	Agree
4. The scoring methods are appropriate, and they can reflect the test takers' writing abilities.	4.30	0.675	Agree
5. The test takers' scores can predict their writing abilities.	4.20	0.789	Agree
Reliability	4.20	0.115	Agree
6. The scoring criteria is specific enough to make scoring writing tasks generally fair for all test takers.	4.10	0.738	Agree
7. The scores of the tests are reliable, irrespective of any raters.	4.30	0.675	Agree
8. The response to each test task is reliable, irrespective of any test-takers, any time, and any place for test-taking.	4.30	0.675	Agree
9. The test input in each part of the test does not contribute to understanding or knowledge in doing other items of other parts.	4.10	0.738	Agree
Authenticity	4.15	0.071	Agree
10. The writing tasks are related to the test takers' future real-life tasks.	4.10	0.738	Agree
11. The writing tasks are authentic and reflect real-life situations.	4.20	0.789	Agree
Interactiveness	4.16	0.089	Agree
12. The writing tasks are familiar topics for all test takers.	4.10	0.738	Agree
13. The writing tasks are suitable for the English knowledge of all the test takers.	4.10	0.738	Agree
Impact	4.15	0.100	Agree
14. The writing tasks provide the test takers with experience in doing future language use tasks.	4.10	0.738	Agree
15. The writing tasks are similar to the writing activities that the test takers have experienced in class.	4.10	0.738	Agree
16. The purpose of the writing test is consistent with the goals and the values mentioned in the instructional program.	4.10	0.738	Agree
17. The writing tasks can reflect test takers' writing abilities currently demanded by Thai society and educational system	4.30	0.675	Agree
Practicality	4.15	0.071	Agree
18. The writing tasks can possibly be constructed with existing resources.	4.20	0.789	Agree
19. The writing test can be administered and scored without much difficulty.	4.10	0.738	Agree
Fairness			
20. The test content contains suitable and bias-free topics without leading to discrimination of sex, race, and economic and social status, and there is not any points leading to violence.	4.70	0.483	Strongly agree

Speaking Test

Test Usefulness	X	S.D.	Meaning
Construct Validity	4.16	0.089	Agree
1. The test objectives of each speaking part are clearly defined.	4.20	0.789	Agree
2. The test objectives of each speaking part are in accordance with the purpose of the speaking test.	4.10	0.738	Agree
3. The speaking tasks are all in accordance with the speaking test objectives.	4.10	0.738	Agree
4. The scoring methods are appropriate, and they can reflect the test takers' speaking abilities.	4.30	0.675	Agree
5. The test takers' scores can predict their speaking abilities.	4.10	0.738	Agree
Reliability	4.25	0.100	Agree
6. The scoring criteria are specific enough to make scoring speaking tasks generally fair for all test takers.	4.30	0.675	Agree
7. The scores of the tests are reliable, irrespective of any raters.	4.30	0.675	Agree
8. The response to each test item is reliable, irrespective of any test-takers, any time, and any place for test-taking.	4.10	0.738	Agree
9. The test input in each part of the test does not contribute to understanding or knowledge in doing other items of other parts.	4.30	0.675	Agree
Authenticity	4.15	0.071	Agree
10. The speaking tasks are related to the test takers' future real-life tasks.	4.20	0.789	Agree
11. The speaking tasks are authentic and reflect real-life situations.	4.10	0.738	Agree
Interactiveness	4.16	0.089	Agree
12. The speaking tasks are familiar topics for all test takers.	4.20	0.789	Agree
13. The speaking tasks are suitable for the English knowledge of all the test takers.	4.10	0.738	Agree
Impact	4.18	0.096	Agree
14. The speaking tasks provide the test takers with experience in doing future language use tasks.	4.10	0.738	Agree
15. The speaking tasks are similar to the speaking activities that the test takers have experienced in class.	4.30	0.675	Agree
16. The purpose of the speaking test is consistent with the goals and the values mentioned in the instructional program.	4.20	0.789	Agree
17. The speaking tasks can reflect test takers' speaking abilities currently demanded by Thai society and educational system	4.10	0.738	Agree
Practicality	4.15	0.071	Agree
18. The speaking tasks can possibly be constructed with existing resources.	4.20	0.789	Agree
19. The speaking test can be administered and scored without much difficulty.	4.10	0.738	Agree
Fairness			
20. The test content contains suitable and bias-free topics without leading to discrimination of sex, race, and economic and social status, and there is not any points leading to violence.	4.80	0.422	Strongly agree

สรุปค่า P ค่า R รายข้อ แบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย Listening Part A

ข้อ	P	R	คุณภาพ	ข้อ	P	R	คุณภาพ
1	.39	.22	ค่อนข้างยาก จำนวนจำกัด	21	.25	.78	ค่อนข้างยาก จำนวนจำกัด
2	.50	.50	ยากง่ายปานกลาง จำนวนจำกัด	22	.19	.18	ยากมากหรือง่ายมาก จำนวนจำกัด
3	.20	-.23	จำนวนจำกัดแบบเป็นลบ	23	.19	.18	ยากมากหรือง่ายมาก จำนวนจำกัด
4	.19	.18	ยากมากหรือง่ายมาก จำนวนจำกัด	24	.25	.32	ค่อนข้างง่าย จำนวนจำกัด
5	.30	.43	ค่อนข้างยาก จำนวนจำกัด	25	.65	.54	ค่อนข้างง่าย จำนวนจำกัด
6	.58	.63	ยากง่ายปานกลาง จำนวนจำกัด	26	.00	.00	ไม่มีจำนวนจำกัด (จำนวนจำกัดเป็น 0)
7	.37	.27	ค่อนข้างยาก จำนวนจำกัด	27	.75	.00	ไม่มีจำนวนจำกัด (จำนวนจำกัดเป็น 0)
8	.25	.32	ค่อนข้างยาก จำนวนจำกัด	28	.44	.12	ความยากปานกลาง จำนวนจำกัด
9	.57	.38	ยากง่ายปานกลาง จำนวนจำกัด	29	.30	.43	ค่อนข้างยาก จำนวนจำกัด
10	.06	.45	ยากมาก จำนวนจำกัด	30	.25	.00	ไม่มีจำนวนจำกัด (จำนวนจำกัดเป็น 0)
11	.19	.72	ยากมาก จำนวนจำกัด	31	.19	.18	ยากมากหรือง่ายมาก จำนวนจำกัด
12	.25	.78	ค่อนข้างยาก จำนวนจำกัด	32	.46	.07	ความยากปานกลาง จำนวนจำกัด
13	.19	-.18	จำนวนจำกัดแบบเป็นลบ	33	.31	.15	ความยากปานกลาง จำนวนจำกัด
14	.19	.72	ยากมาก จำนวนจำกัด	34	.19	.18	ยากมากหรือง่ายมาก จำนวนจำกัด
15	.25	.32	ค่อนข้างยาก จำนวนจำกัด	35	.42	.62	ยากง่ายปานกลาง จำนวนจำกัด
16	.25	.00	ไม่มีจำนวนจำกัด (จำนวนจำกัดเป็น 0)	36	.30	.43	ค่อนข้างง่าย จำนวนจำกัด
17	.06	.45	ยากมาก จำนวนจำกัด	37	.58	.63	ยากง่ายปานกลาง จำนวนจำกัด
18	.51	.73	ยากง่ายปานกลาง จำนวนจำกัด	38	.10	.57	ยากมาก จำนวนจำกัด
19	.10	.57	ยากมาก จำนวนจำกัด	39	.37	.27	ค่อนข้างยาก จำนวนจำกัด
20	.00	.00	ไม่มีจำนวนจำกัด (จำนวนจำกัดเป็น 0)	40	.10	.57	ยากมาก จำนวนจำกัด

สรุปค่า P ค่า R รายข้อ แบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย Listening Part A

ข้อ	P	R	คุณภาพ	สรุปผลการวิเคราะห์
41	.19	.72	ยากมาก อำนาจจำแนกดี	จำนวนข้อสอบที่วิเคราะห์ 49
42	.57	.13	ความยากปานกลาง อำนาจจำแนกไม่ดี	จำนวนกระดาษคำตอบ 28
43	.58	.63	ยากง่ายปานกลาง อำนาจจำแนกดี	คะแนนเฉลี่ย 14.0357
44	.00	.00	ไม่มีอำนาจจำแนก (อำนาจจำแนกเป็น 0)	ส่วนเบี่ยงเบนมาตรฐาน 5.5773
45	.00	.00	ไม่มีอำนาจจำแนก (อำนาจจำแนกเป็น 0)	ความเชื่อมั่น KR-20 .7351
46	.19	.18	ยากมากหรือง่ายมาก อำนาจจำแนกไม่ดี	ความคลาดเคลื่อนมาตรฐาน 2.8708
47	.36	-.53	อำนาจจำแนกเป็นลบ	สรุปคุณภาพของข้อสอบ
48	.58	.63	ยากง่ายปานกลาง อำนาจจำแนกดี	ข้อสอบที่ดี ควรเก็บไว้ได้แก่
49	.36	.53	ค่อนข้างยาก อำนาจจำแนกดี	1) ข้อที่ค่อนข้างง่าย อำนาจจำแนกดี มี 1 ข้อ ได้แก่ข้อ 25
เฉลี่ย	.30	.39	ค่อนข้างยาก อำนาจจำแนกดี	2) ข้อที่ยากปานกลาง อำนาจจำแนกดี มี 8 ข้อ ได้แก่ข้อ 2, 6, 9, 18, 35, 37, 43, 48
				3) ข้อที่ค่อนข้างยาก อำนาจจำแนกดี มี 12 ข้อ ได้แก่ข้อ 1, 5, 7, 8, 12, 15, 21, 24, 29, 36, 39, 49
				ข้อที่ควรปรับปรุง ได้แก่
				1) ข้อที่ง่ายมาก อำนาจจำแนกดี ไม่มี
				2) ข้อที่ยากมาก อำนาจจำแนกดี มี 8 ข้อ ได้แก่ข้อ 10, 11, 14, 17, 19, 38, 40, 41
				3) ข้อที่ความยากปานกลาง อำนาจจำแนกไม่ดี มี 4 ข้อ ได้แก่ข้อ 28, 32, 33, 42
				ข้อสอบที่ควรตัดทิ้ง ได้แก่
				1) ข้อที่ยากมากหรือง่ายมาก อำนาจจำแนกไม่ดี มี 6 ข้อ ได้แก่ข้อ 4, 22, 23, 31, 34, 46
				2) ข้อที่ไม่มีอำนาจจำแนก (อำนาจจำแนกเป็น 0) มี 7 ข้อ ได้แก่ข้อ 16, 20, 26, 27, 30, 44, 45
				3) ข้อที่อำนาจจำแนกเป็นลบ มี 3 ข้อ ได้แก่ข้อ 3, 13, 47

สรุปค่า P ค่า R รายข้อ แบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย Listening Part B

ข้อ	P	R	คุณภาพ	สรุปผลการวิเคราะห์
1	.78	.75	ค่อนข้างง่าย อำนาจจำแนกดี	จำนวนข้อสอบที่วิเคราะห์ 17
2	.83	.69	ง่ายมาก อำนาจจำแนกดี	จำนวนกระดาษคำตอบ 28
3	.51	.73	ยากง่ายปานกลาง อำนาจจำแนกดี	คะแนนเฉลี่ย 6.3571
4	.06	.45	ยากมาก อำนาจจำแนกดี	ส่วนเบี่ยงเบนมาตรฐาน 2.9303
5	.44	.39	ยากง่ายปานกลาง อำนาจจำแนกดี	ความเชื่อมั่น KR-20 .6604
6	.42	.62	ยากง่ายปานกลาง อำนาจจำแนกดี	ความคลาดเคลื่อนมาตรฐาน 1.7078
7	.31	.15	ความยากปานกลาง อำนาจจำแนกไม่ดี	สรุปคุณภาพของข้อสอบ
8	.65	.54	ค่อนข้างง่าย อำนาจจำแนกดี	ข้อสอบที่ดี ควรเก็บไว้ได้แก่
9	.30	.82	ค่อนข้างยาก อำนาจจำแนกดี	1) ข้อที่ค่อนข้างง่าย อำนาจจำแนกดี มี 2 ข้อ ได้แก่ข้อ 1, 8
10	.42	.62	ยากง่ายปานกลาง อำนาจจำแนกดี	2) ข้อที่ยากง่ายปานกลาง อำนาจจำแนกดี มี 5 ข้อ ได้แก่ข้อ 3, 5, 6, 10, 17
11	.30	.43	ค่อนข้างยาก อำนาจจำแนกดี	3) ข้อที่ค่อนข้างยาก อำนาจจำแนกดี มี 3 ข้อ ได้แก่ข้อ 9, 11, 13
12	.19	.72	ยากมาก อำนาจจำแนกดี	ข้อที่ควรปรับปรุง ได้แก่
13	.25	.32	ค่อนข้างยาก อำนาจจำแนกดี	1) ข้อที่ง่ายมาก อำนาจจำแนกดี มี 1 ข้อ ได้แก่ข้อ 2
14	.00	.00	ไม่มีอำนาจจำแนก (อำนาจจำแนกเป็น 0)	2) ข้อที่ยากมาก อำนาจจำแนกดี มี 2 ข้อ ได้แก่ข้อ 4, 12
15	.19	.18	ยากมากหรือง่ายมาก อำนาจจำแนกไม่ดี	3) ข้อที่ความยากปานกลาง อำนาจจำแนกไม่ดี มี 1 ข้อ ได้แก่ข้อ 7
16	.34	-.20	อำนาจจำแนกเป็นลบ	ข้อสอบที่ควรตัดทิ้ง ได้แก่
17	.44	.39	ยากง่ายปานกลาง อำนาจจำแนกดี	1) ข้อที่ยากมาหรือง่ายมาก อำนาจจำแนกไม่ดี มี 1 ข้อ ได้แก่ข้อ 15
เฉลี่ย	.38	.51	ค่อนข้างยาก อำนาจจำแนกดี	2) ข้อที่ไม่มีอำนาจจำแนก (อำนาจจำแนกเป็น 0) มี 1 ข้อ ได้แก่ข้อ 14
				3) ข้อที่อำนาจจำแนกเป็นลบ มี 1 ข้อ ได้แก่ข้อ 16

สรุปค่า P ค่า R รายข้อ แบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย Listening Part C

ข้อ	P	R	คุณภาพ	สรุปผลการวิเคราะห์
1	.13	.00	ไม่มีอำนาจจำแนก (อำนาจจำแนกเป็น 0)	จำนวนข้อสอบที่วิเคราะห์ 21
2	.58	.63	ยากง่ายปานกลาง อำนาจจำแนกดี	จำนวนกระดาษคำตอบ 28
3	.69	.14	ความยากปานกลาง อำนาจจำแนกไม่ดี	คะแนนเฉลี่ย 8.5000
4	.50	.50	ยากง่ายปานกลาง อำนาจจำแนกดี	ส่วนเบี่ยงเบนมาตรฐาน 1.5924
5	.58	.63	ยากง่ายปานกลาง อำนาจจำแนกดี	ความเชื่อมั่น KR-20 -.5599
6	.31	.15	ความยากปานกลาง อำนาจจำแนกไม่ดี	ความคลาดเคลื่อนมาตรฐาน 1.9888
7	.13	.02	ยากมากหรือง่ายมาก อำนาจจำแนกไม่ดี	สรุปคุณภาพของข้อสอบ
8	.90	.57	ง่ายมาก อำนาจจำแนกดี	ข้อสอบที่ดี ควรเก็บไว้ได้แก่
9	.66	-.20	อำนาจจำแนกเป็นลบ	1) ข้อที่ค่อนข้างง่าย อำนาจจำแนกดี ไม่มี
10	.06	-.45	อำนาจจำแนกเป็นลบ	2) ข้อที่ยากง่ายปานกลาง อำนาจจำแนกดี มี 6 ข้อ ได้แก่ข้อ 2, 4, 5, 11, 15, 21
11	.56	.34	ยากง่ายปานกลาง อำนาจจำแนกดี	3) ข้อที่ค่อนข้างยาก อำนาจจำแนกดี มี 1 ข้อ ได้แก่ข้อ 16
12	.17	.69	ยากมาก อำนาจจำแนกดี	ข้อที่ควรปรับปรุง ได้แก่
13	.00	.00	ไม่มีอำนาจจำแนก (อำนาจจำแนกเป็น 0)	1) ข้อที่ง่ายมาก อำนาจจำแนกดี มี 1 ข้อ ได้แก่ข้อ 8
14	.19	-.18	อำนาจจำแนกเป็นลบ	2) ข้อที่ยากมาก อำนาจจำแนกดี มี 1 ข้อ ได้แก่ข้อ 12
15	.58	.63	ยากง่ายปานกลาง อำนาจจำแนกดี	3) ข้อที่ความยากปานกลาง อำนาจจำแนกไม่ดี มี 3 ข้อ ได้แก่ข้อ 3, 6, 18
16	.30	.43	ค่อนข้างยาก อำนาจจำแนกดี	ข้อสอบที่ควรตัดทิ้ง ได้แก่
17	.95	-.43	อำนาจจำแนกเป็นลบ	1) ข้อที่ยากมากหรือง่ายมาก อำนาจจำแนกไม่ดี มี 1 ข้อ ได้แก่ข้อ 7
18	.44	.12	ความยากปานกลาง อำนาจจำแนกไม่ดี	2) ข้อที่ไม่มีอำนาจจำแนก (อำนาจจำแนกเป็น 0) มี 4 ข้อ ได้แก่ข้อ 1, 13, 19, 20
19	.13	.00	ไม่มีอำนาจจำแนก (อำนาจจำแนกเป็น 0)	3) ข้อที่อำนาจจำแนกเป็นลบ มี 4 ข้อ ได้แก่ข้อ 9, 10, 14, 17
20	.13	.00	ไม่มีอำนาจจำแนก (อำนาจจำแนกเป็น 0)	
21	.58	.63	ยากง่ายปานกลาง อำนาจจำแนกดี	
เฉลี่ย	.41	.35	ความยากปานกลาง อำนาจจำแนกดี	

สรุปค่า P ค่า R รายข้อ แบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย Sentence Structure

ข้อ	P	R	คุณภาพ	สรุปผลการวิเคราะห์
1	.63	.27	ค่อนข้างง่าย อำนาจจำแนกดี	จำนวนข้อสอบที่วิเคราะห์ 22
2	.86	.66	ง่ายมาก อำนาจจำแนกดี	จำนวนกระดาษคำตอบ 28
3	.36	.53	ค่อนข้างยาก อำนาจจำแนกดี	คะแนนเฉลี่ย 9.1786
4	.19	.72	ยากมาก อำนาจจำแนกดี	ส่วนเบี่ยงเบนมาตรฐาน 3.6258
5	.30	.43	ค่อนข้างยาก อำนาจจำแนกดี	ความเชื่อมั่น KR-20 .6806
6	.37	.27	ค่อนข้างยาก อำนาจจำแนกดี	ความคลาดเคลื่อนมาตรฐาน 2.0492
7	.44	.39	ง่ายปานกลาง อำนาจจำแนกดี	สรุปคุณภาพของข้อสอบ
8	.57	.38	ง่ายปานกลาง อำนาจจำแนกดี	ข้อสอบที่ดี ควรเก็บไว้ได้แก่
9	.58	.63	ง่ายปานกลาง อำนาจจำแนกดี	1) ข้อที่ค่อนข้างง่าย อำนาจจำแนกดี มี 3 ข้อ ได้แก่ข้อ 1, 10, 18
10	.70	.82	ค่อนข้างง่าย อำนาจจำแนกดี	2) ข้อที่ยากปานกลาง อำนาจจำแนกดี มี 6 ข้อ ได้แก่ข้อ 7, 8, 9, 12, 17, 21
11	.86	.66	ง่ายมาก อำนาจจำแนกดี	3) ข้อที่ค่อนข้างยาก อำนาจจำแนกดี มี 5 ข้อ ได้แก่ข้อ 3, 5, 6, 13, 20
12	.42	.62	ง่ายปานกลาง อำนาจจำแนกดี	ข้อที่ควรปรับปรุง ได้แก่
13	.30	.43	ค่อนข้างยาก อำนาจจำแนกดี	1) ข้อที่ง่ายมาก อำนาจจำแนกดี มี 2 ข้อ ได้แก่ข้อ 2, 11
14	.25	.00	ไม่มีอำนาจจำแนก (อำนาจจำแนกเป็น 0)	2) ข้อที่ยากมาก อำนาจจำแนกดี มี 4 ข้อ ได้แก่ข้อ 4, 15, 19, 22
15	.15	.66	ยากมาก อำนาจจำแนกดี	3) ข้อที่ความยากปานกลาง อำนาจจำแนกไม่ดี ไม่มี
16	.31	-.15	อำนาจจำแนกเป็นลบ	ข้อสอบที่ควรตัดทิ้ง ได้แก่
17	.44	.39	ง่ายปานกลาง อำนาจจำแนกดี	1) ข้อที่ยากมากหรือง่ายมาก อำนาจจำแนกไม่ดี ไม่มี
18	.76	.33	ค่อนข้างง่าย อำนาจจำแนกดี	2) ข้อที่ไม่มีอำนาจจำแนก (อำนาจจำแนกเป็น 0) มี 1 ข้อ ได้แก่ข้อ 14
19	.15	.66	ยากมาก อำนาจจำแนกดี	3) ข้อที่อำนาจจำแนกเป็นลบ มี 1 ข้อ ได้แก่ข้อ 16
20	.30	.82	ค่อนข้างยาก อำนาจจำแนกดี	
21	.44	.39	ง่ายปานกลาง อำนาจจำแนกดี	
22	.15	.66	ยากมาก อำนาจจำแนกดี	
เฉลี่ย	.43	.53	ง่ายปานกลาง อำนาจจำแนกดี	

สรุปค่า P ค่า R รายข้อ แบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย Written Expression

ข้อ	P	R	คุณภาพ	ข้อ	P	R	คุณภาพ
1	.50	.00	ไม่มีอำนาจจำแนก (อำนาจจำแนกเป็น 0)	21	.81	.72	ง่ายมาก อำนาจจำแนกดี
2	.30	.43	ค่อนข้างยาก อำนาจจำแนกดี	22	.57	.61	ง่ายปานกลาง อำนาจจำแนกดี
3	.50	.00	ไม่มีอำนาจจำแนก (อำนาจจำแนกเป็น 0)	23	.19	.18	ยากมากหรือง่ายมาก อำนาจจำแนกไม่ได้
4	.81	.72	ง่ายมาก อำนาจจำแนกดี	24	.25	.32	ค่อนข้างยาก อำนาจจำแนกดี
5	.76	-.33	อำนาจจำแนกเป็นลบ	25	.30	.43	ค่อนข้างยาก อำนาจจำแนกดี
6	.25	.00	ไม่มีอำนาจจำแนก (อำนาจจำแนกเป็น 0)	26	.37	.27	ค่อนข้างยาก อำนาจจำแนกดี
7	.13	.00	ไม่มีอำนาจจำแนก (อำนาจจำแนกเป็น 0)	27	.36	.53	ค่อนข้างยาก อำนาจจำแนกดี
8	.15	.66	ยากมาก อำนาจจำแนกดี	28	.57	.38	ง่ายปานกลาง อำนาจจำแนกดี
9	.50	.00	ไม่มีอำนาจจำแนก (อำนาจจำแนกเป็น 0)	29	.63	.00	ไม่มีอำนาจจำแนก (อำนาจจำแนกเป็น 0)
10	.13	.00	ไม่มีอำนาจจำแนก (อำนาจจำแนกเป็น 0)	30	.58	.63	ง่ายปานกลาง อำนาจจำแนกดี
11	.25	.78	ค่อนข้างยาก อำนาจจำแนกดี	31	.51	.25	ง่ายปานกลาง อำนาจจำแนกดี
12	.25	-.32	อำนาจจำแนกเป็นลบ	32	.30	.82	ค่อนข้างยาก อำนาจจำแนกดี
13	.51	.73	ง่ายปานกลาง อำนาจจำแนกดี	33	.15	.66	ยากมาก อำนาจจำแนกดี
14	.69	-.14	อำนาจจำแนกเป็นลบ	34	.37	.87	ค่อนข้างยาก อำนาจจำแนกดี
15	.36	.53	ค่อนข้างยาก อำนาจจำแนกดี	35	.25	.00	ไม่มีอำนาจจำแนก (อำนาจจำแนกเป็น 0)
16	.76	.77	ค่อนข้างง่าย อำนาจจำแนกดี	36	.19	.18	ยากมากหรือง่ายมาก อำนาจจำแนกไม่ได้
17	.37	.27	ค่อนข้างยาก อำนาจจำแนกดี	37	.43	.61	ง่ายปานกลาง อำนาจจำแนกดี
18	.19	.72	ยากมาก อำนาจจำแนกดี	38	.60	.60	ค่อนข้างง่าย อำนาจจำแนกดี
19	.30	.82	ค่อนข้างยาก อำนาจจำแนกดี	เฉลี่ย	.42	.46	ง่ายปานกลาง อำนาจจำแนกดี
20	.69	.14	ความยากปานกลาง อำนาจจำแนกไม่ได้				

สรุปคำ P คำ R รายชื่อ แบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย Written Expression

สรุปผลการวิเคราะห์	
จำนวนข้อสอบที่วิเคราะห์	38
จำนวนกระดาษคำตอบ	28
คะแนนเฉลี่ย	16.1786
ส่วนเบี่ยงเบนมาตรฐาน	5.0854
ความเชื่อมั่น KR-20	.7016
ความคลาดเคลื่อนมาตรฐาน	2.7781
สรุปคุณภาพของข้อสอบ	
ข้อสอบที่ดี ควรเก็บไว้ ได้แก่	
1) ข้อที่ค่อนข้างง่าย จำนวนจำแนกดี มี 2 ข้อ ได้แก่ข้อ 16, 38	
2) ข้อที่ยากปานกลาง จำนวนจำแนกดี มี 6 ข้อ ได้แก่ข้อ 13, 22, 28, 30, 31, 37	
3) ข้อที่ค่อนข้างยาก จำนวนจำแนกดี มี 11 ข้อ ได้แก่ข้อ 2, 11, 15, 17, 19, 24, 25, 26, 27, 32, 34	
ข้อที่ควรปรับปรุง ได้แก่	
1) ข้อที่ง่ายมาก จำนวนจำแนกดี มี 2 ข้อ ได้แก่ข้อ 4, 21	
2) ข้อที่ยากมาก จำนวนจำแนกดี มี 3 ข้อ ได้แก่ข้อ 8, 18, 33	
3) ข้อที่ความยากปานกลาง จำนวนจำแนกไม่ดี มี 1 ข้อ ได้แก่ข้อ 20	
ข้อสอบที่ควรตัดทิ้ง ได้แก่	
1) ข้อที่ยากหรือง่ายมาก จำนวนจำแนกไม่ดี มี 2 ข้อ ได้แก่ข้อ 23, 36	
2) ข้อที่ไม่มีอำนาจจำแนก (อำนาจจำแนกเป็น 0) มี 8 ข้อ ได้แก่ข้อ 1, 3, 6, 7, 9, 10, 29, 35	
3) ข้อที่อำนาจจำแนกเป็นลบ มี 3 ข้อ ได้แก่ข้อ 5, 12, 14	

สรุปค่า P ค่า R รายข้อ แบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย Reading for Everyday Use

ข้อ	P	R	คุณภาพ	ข้อ	P	R	คุณภาพ
1	.69	.14	ความยากปานกลาง อำนาจจำแนกไม่ได้	21	.82	.20	ง่ายมาก อำนาจจำแนกดี
2	.90	.57	ง่ายมาก อำนาจจำแนกดี	22	.50	.00	ไม่มีอำนาจจำแนก (อำนาจจำแนกเป็น 0)
3	.30	.43	ค่อนข้างยากอำนาจจำแนกดี	23	.76	.33	ค่อนข้างง่าย อำนาจจำแนกดี
4	.86	.66	ง่ายมาก อำนาจจำแนกดี	24	.63	.27	ค่อนข้างง่าย อำนาจจำแนกดี
5	.51	.73	ง่ายปานกลาง อำนาจจำแนกดี	25	.13	.00	ไม่มีอำนาจจำแนก (อำนาจจำแนกเป็น 0)
6	.81	.72	ง่ายมาก อำนาจจำแนกดี	26	.46	.35	ยากง่ายปานกลาง อำนาจจำแนกดี
7	.57	.38	ง่ายปานกลาง อำนาจจำแนกดี	27	.82	.20	ง่ายมาก อำนาจจำแนกดี
8	.44	.12	ความยากปานกลาง อำนาจจำแนกไม่ได้	28	.31	.15	ความยากปานกลาง อำนาจจำแนกไม่ได้
9	.76	.77	ค่อนข้างง่าย อำนาจจำแนกดี	29	.71	.44	ค่อนข้างง่าย อำนาจจำแนกดี
10	.81	.72	ง่ายมาก อำนาจจำแนกดี	30	.76	.33	ค่อนข้างง่าย อำนาจจำแนกดี
11	.58	.63	ง่ายปานกลาง อำนาจจำแนกดี	31	.25	.32	ค่อนข้างยากอำนาจจำแนกดี
12	.44	-.12	อำนาจจำแนกเป็นลบ	32	.95	.43	ง่ายมาก อำนาจจำแนกดี
13	.25	-.31	อำนาจจำแนกเป็นลบ	33	.69	.14	ความยากปานกลาง อำนาจจำแนกไม่ได้
14	.36	.53	ค่อนข้างยากอำนาจจำแนกดี	34	.76	.33	ค่อนข้างง่าย อำนาจจำแนกดี
15	.63	.27	ค่อนข้างง่าย อำนาจจำแนกดี	35	.25	.78	ค่อนข้างยากอำนาจจำแนกดี
16	.06	.45	ยากมาก อำนาจจำแนกดี	36	.95	.43	ง่ายมาก อำนาจจำแนกดี
17	.36	-.53	อำนาจจำแนกเป็นลบ	37	.86	.66	ง่ายมาก อำนาจจำแนกดี
18	.81	.72	ง่ายมาก อำนาจจำแนกดี	38	.50	.93	ยากง่ายปานกลาง อำนาจจำแนกดี
19	.65	.54	ค่อนข้างง่าย อำนาจจำแนกดี	39	.36	.53	ค่อนข้างยากอำนาจจำแนกดี
20	.69	.14	ความยากปานกลาง อำนาจจำแนกไม่ได้	40	.00	.00	ไม่มีอำนาจจำแนก (อำนาจจำแนกเป็น 0)

สรุปค่า P ค่า R รายข้อ แบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย Reading for Everyday Use

ข้อ	P	R	คุณภาพ	สรุปผลการวิเคราะห์
41	.19	.72	ยากมาก อ่านจําแนกดี	จำนวนข้อสอบที่วิเคราะห์ 49
42	.44	.39	ยากง่ายปานกลาง อ่านจําแนกดี	จำนวนกระดาษคำตอบ 28
43	.42	.62	ยากง่ายปานกลาง อ่านจําแนกดี	คะแนนเฉลี่ย 26.8571
44	.25	-.32	อ่านจําแนกเป็นลบ	ส่วนเบี่ยงเบนมาตรฐาน 5.7366
45	.42	.62	ยากง่ายปานกลาง อ่านจําแนกดี	ความเชื่อมั่น KR-20 .7299
46	.63	.27	ค่อนข้างง่าย อ่านจําแนกดี	ความคลาดเคลื่อนมาตรฐาน 2.9816
47	.82	.20	ง่ายมาก อ่านจําแนกดี	สรุปคุณภาพของข้อสอบ
48	.31	.15	ความยากปานกลาง อ่านจําแนกไม่ดี	ข้อสอบที่ดี ควรเก็บไว้ได้แก่
49	.65	.54	ค่อนข้างง่าย อ่านจําแนกดี	1) ข้อที่ค่อนข้างง่าย อ่านจําแนกดี มี 10 ข้อ ได้แก่ข้อ 9, 15, 19, 23, 24, 29, 30, 34, 46, 49
เฉลี่ย	.55	.45	ยากง่ายปานกลาง อ่านจําแนกดี	2) ข้อที่ยากปานกลาง อ่านจําแนกดี มี 8 ข้อ ได้แก่ข้อ 5, 7, 11, 26, 38, 42, 43, 45
				3) ข้อที่ค่อนข้างยาก อ่านจําแนกดี มี 5 ข้อ ได้แก่ข้อ 3, 14, 31, 35, 39
				ข้อที่ควรปรับปรุง ได้แก่
				1) ข้อที่ง่ายมาก อ่านจําแนกดี มี 11 ข้อ ได้แก่ข้อ 2, 4, 6, 10, 18, 21, 27, 32, 36, 37, 47
				2) ข้อที่ยากมาก อ่านจําแนกดี มี 2 ข้อ ได้แก่ข้อ 16, 41
				3) ข้อที่ความยากปานกลาง อ่านจําแนกไม่ดี มี 6 ข้อ ได้แก่ข้อ 1, 8, 20, 28, 33, 48
				ข้อสอบที่ควรตัดทิ้ง ได้แก่
				1) ข้อที่ยากมากหรือง่ายมาก อ่านจําแนกไม่ดี ไม่มี
				2) ข้อที่ไม่มีอ่านจําแนก (อ่านจําแนกเป็น 0) มี 3 ข้อ ได้แก่ข้อ 22, 25, 40
				3) ข้อที่อ่านจําแนกเป็นลบ มี 4 ข้อ ได้แก่ข้อ 12, 13, 17, 44

สรุปค่า P ค่า R รายข้อ แบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย Reading for Academic Purpose

ข้อ	P	R	คุณภาพ	ข้อ	P	R	คุณภาพ
1	.38	.00	ไม่มีอำนาจจำแนก (อำนาจจำแนกเป็น 0)	21	.20	-.23	อำนาจจำแนกเป็นลบ
2	.25	.32	ค่อนข้างยาก อำนาจจำแนกดี	22	.06	-.46	อำนาจจำแนกเป็นลบ
3	.65	.54	ค่อนข้างง่าย อำนาจจำแนกดี	23	.46	.35	ยกง่ายปานกลาง อำนาจจำแนกดี
4	.30	.82	ค่อนข้างยาก อำนาจจำแนกดี	24	.34	-.20	อำนาจจำแนกเป็นลบ
5	.69	.14	ความยากปานกลาง อำนาจจำแนกไม่ได้	25	.80	.25	ง่ายมาก อำนาจจำแนกดี
6	.52	.46	ยกง่ายปานกลาง อำนาจจำแนกดี	26	.52	.46	ยกง่ายปานกลาง อำนาจจำแนกดี
7	.20	-.23	อำนาจจำแนกเป็นลบ	27	.33	.37	ค่อนข้างยาก อำนาจจำแนกดี
8	.57	.13	ความยากปานกลาง อำนาจจำแนกไม่ได้	28	.54	.43	ยกง่ายปานกลาง อำนาจจำแนกดี
9	.63	.27	ค่อนข้างง่าย อำนาจจำแนกดี	29	.41	.18	ความยากปานกลาง อำนาจจำแนกไม่ได้
10	.37	.87	ค่อนข้างยาก อำนาจจำแนกดี	30	.33	.37	ค่อนข้างยาก อำนาจจำแนกดี
11	.39	.48	ค่อนข้างยาก อำนาจจำแนกดี	31	.87	.63	ง่ายมาก อำนาจจำแนกดี
12	.53	.69	ยกง่ายปานกลาง อำนาจจำแนกดี	32	.87	.63	ง่ายมาก อำนาจจำแนกดี
13	.48	.30	ยกง่ายปานกลาง อำนาจจำแนกดี	33	.54	.43	ยกง่ายปานกลาง อำนาจจำแนกดี
14	.39	.48	ค่อนข้างยาก อำนาจจำแนกดี	34	.27	.26	ค่อนข้างยาก อำนาจจำแนกดี
15	.53	.69	ยกง่ายปานกลาง อำนาจจำแนกดี	35	.29	-.10	อำนาจจำแนกเป็นลบ
16	.52	.46	ยกง่ายปานกลาง อำนาจจำแนกดี	36	.29	.22	ค่อนข้างยาก อำนาจจำแนกดี
17	.37	.27	ค่อนข้างยาก อำนาจจำแนกดี	37	.42	-.36	อำนาจจำแนกเป็นลบ
18	.67	.50	ค่อนข้างง่าย อำนาจจำแนกดี	38	.22	.07	ความยากปานกลาง อำนาจจำแนกไม่ได้
19	.67	.50	ค่อนข้างง่าย อำนาจจำแนกดี	39	.45	.10	ความยากปานกลาง อำนาจจำแนกไม่ได้
20	.39	.22	ค่อนข้างยาก อำนาจจำแนกดี	40	.15	.66	ยกมาก อำนาจจำแนกดี

สรุปค่า P ค่า R รายข้อ แบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย Reading for Everyday Use

ข้อ	P	R	คุณภาพ	สรุปผลการวิเคราะห์
41	.10	.57	ยากมาก อ่านจําแนกดี	จำนวนข้อสอบที่วิเคราะห์ 48
42	.25	.78	ค่อนข้างยาก อ่านจําแนกดี	จำนวนกระดาษคำตอบ 28
43	.47	.55	ง่ายปานกลาง อ่านจําแนกดี	คะแนนเฉลี่ย 20.4643
44	.39	-.02	อ่านจําแนกเป็นลบ	ส่วนเบี่ยงเบนมาตรฐาน 7.1966
45	.35	.06	ความยากปานกลาง อ่านจําแนกไม่ได้	ความเชื่อมั่น KR-20 .8188
46	.62	.57	ค่อนข้างง่าย อ่านจําแนกดี	ความคลาดเคลื่อนมาตรฐาน 3.0639
47	.93	.50	ง่ายมาก อ่านจําแนกดี	สรุปคุณภาพของข้อสอบ
48	.66	.85	ค่อนข้างง่าย อ่านจําแนกดี	ข้อสอบที่ดี ควรเก็บไว้ได้แก่
เฉลี่ย	.45	.43	ง่ายปานกลาง อ่านจําแนกดี	1) ข้อที่ค่อนข้างง่าย อ่านจําแนกดี มี 6 ข้อ ได้แก่ข้อ 3, 9, 18, 19, 46, 48 2) ข้อที่ยากปานกลาง อ่านจําแนกดี มี 10 ข้อ ได้แก่ข้อ 6, 12, 13, 15, 16, 23, 26, 28, 33, 43 3) ข้อที่ค่อนข้างยาก อ่านจําแนกดี มี 12 ข้อ ได้แก่ข้อ 2, 4, 10, 11, 14, 17, 20, 27, 30, 34, 36, 42
				ข้อที่ควรปรับปรุง ได้แก่
				1) ข้อที่ง่ายมาก อ่านจําแนกดี มี 4 ข้อ ได้แก่ข้อ 25, 31, 32, 47
				2) ข้อที่ยากมาก อ่านจําแนกดี มี 2 ข้อ ได้แก่ข้อ 40, 41
				3) ข้อที่ความยากปานกลาง อ่านจําแนกไม่ได้ มี 6 ข้อ ได้แก่ข้อ 5, 8, 29, 38, 39, 45
				ข้อสอบที่ควรตัดทิ้ง ได้แก่
				1) ข้อที่ยากมากหรือง่ายมาก อ่านจําแนกไม่ได้ ไม่มี
				2) ข้อที่ไม่มีอ่านจําแนก (อ่านจําแนกเป็น 0) มี 1 ข้อ ได้แก่ข้อ 1
				3) ข้อที่อ่านจําแนกเป็นลบ มี 7 ข้อ ได้แก่ข้อ 7, 21, 22, 24, 35, 37, 44

รายงานผลการวิเคราะห์แบบทดสอบรายข้อ
แบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย
(University English Proficiency Test)

การทดลองใช้ครั้งที่ 2

ทดสอบวันพุธที่ 30 มกราคม 2562

เวลา 13.00 -17.00 น.
ณ ห้อง 15403 อาคาร 15

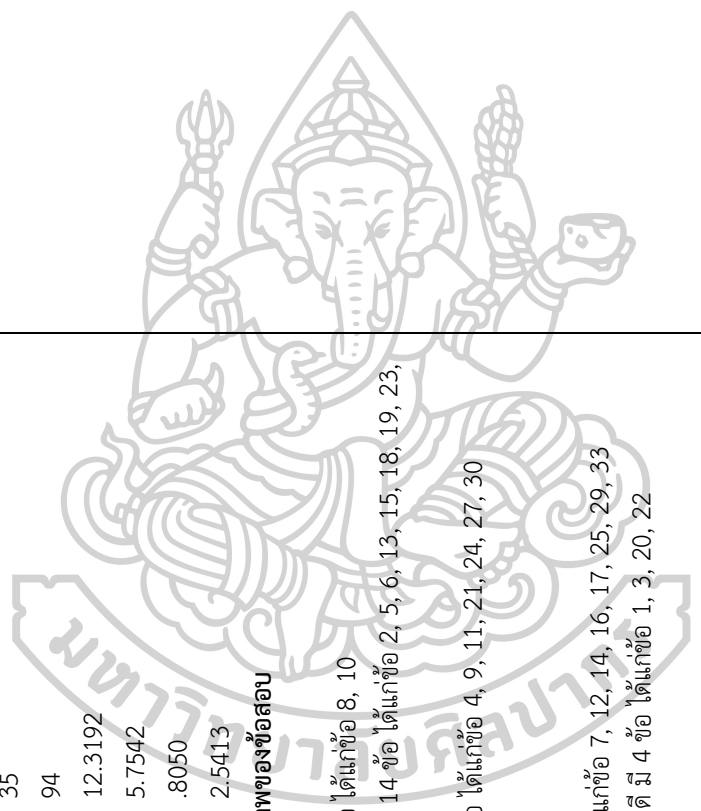


สรุปค่า P ค่า R รายข้อ แบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย Listening Part A

ข้อ	P	R	คุณภาพ	ข้อ	P	R	คุณภาพ
1	.48	.16	ความยากปานกลาง อ่านจจำแนกไม่ได้	21	.24	.21	ค่อนข้างยาก อ่านจจำแนกดี
2	.45	.60	ยากง่ายปานกลาง อ่านจจำแนกดี	22	.32	.18	ความยากปานกลาง อ่านจจำแนกไม่ได้
3	.28	.10	ความยากปานกลาง อ่านจจำแนกไม่ได้	23	.44	.32	ยากง่ายปานกลาง อ่านจจำแนกดี
4	.29	.43	ค่อนข้างยาก อ่านจจำแนกดี	24	.38	.30	ค่อนข้างยาก อ่านจจำแนกดี
5	.58	.64	ยากง่ายปานกลาง อ่านจจำแนกดี	25	.17	.32	ยากมาก อ่านจจำแนกดี
6	.44	.74	ยากง่ายปานกลาง อ่านจจำแนกดี	26	.40	.33	ยากง่ายปานกลาง อ่านจจำแนกดี
7	.19	.36	ยากมาก อ่านจจำแนกดี	27	.24	.47	ค่อนข้างยาก อ่านจจำแนกดี
8	.63	.47	ค่อนข้างง่าย อ่านจจำแนกดี	28	.55	.67	ยากง่ายปานกลาง อ่านจจำแนกดี
9	.35	.75	ค่อนข้างยาก อ่านจจำแนกดี	29	.19	.54	ยากมาก อ่านจจำแนกดี
10	.61	.69	ค่อนข้างง่าย อ่านจจำแนกดี	30	.36	.55	ค่อนข้างยาก อ่านจจำแนกดี
11	.28	.53	ค่อนข้างยาก อ่านจจำแนกดี	31	.48	.48	ยากง่ายปานกลาง อ่านจจำแนกดี
12	.08	.53	ยากมาก อ่านจจำแนกดี	32	.50	.28	ยากง่ายปานกลาง อ่านจจำแนกดี
13	.58	.64	ยากง่ายปานกลาง อ่านจจำแนกดี	33	.17	.32	ยากมาก อ่านจจำแนกดี
14	.14	.65	ยากมาก อ่านจจำแนกดี	34	.48	.56	ยากง่ายปานกลาง อ่านจจำแนกดี
15	.45	.52	ยากง่ายปานกลาง อ่านจจำแนกดี	35	.50	.59	ยากง่ายปานกลาง อ่านจจำแนกดี
16	.17	.32	ยากมาก อ่านจจำแนกดี	เฉลี่ย	.37	.47	ค่อนข้างยาก อ่านจจำแนกดี
17	.17	.51	ยากมาก อ่านจจำแนกดี				
18	.43	.56	ยากง่ายปานกลาง อ่านจจำแนกดี				
19	.50	.44	ยากง่ายปานกลาง อ่านจจำแนกดี				
20	.48	.08	ความยากปานกลาง อ่านจจำแนกไม่ได้				

สรุปคำ P คำ R รายชื่อ แบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย Listening Part A

สรุปผลการวิเคราะห์	
จำนวนข้อสอบที่วิเคราะห์	35
จำนวนกระดาษคำตอบ	94
คะแนนเฉลี่ย	12.3192
ส่วนเบี่ยงเบนมาตรฐาน	5.7542
ความเชื่อมั่น KR-20	.8050
ความคลาดเคลื่อนมาตรฐาน	2.5413
สรุปคุณภาพของข้อสอบ	
ข้อสอบที่ดี ควรเก็บไว้ ได้แก่	
1) ข้อที่ค่อนข้างง่าย จำนวนจำแนกดี มี 2 ข้อ ได้แก่ข้อ 8, 10	
2) ข้อที่ยากปานกลาง จำนวนจำแนกดี มี 14 ข้อ ได้แก่ข้อ 2, 5, 6, 13, 15, 18, 19, 23, 26, 28, 31, 32, 34, 35	
3) ข้อที่ค่อนข้างยาก จำนวนจำแนกดี มี 7 ข้อ ได้แก่ข้อ 4, 9, 11, 21, 24, 27, 30	
ข้อที่ควรปรับปรุง ได้แก่	
1) ข้อที่ง่ายมาก จำนวนจำแนกดี ไม่มี	
2) ข้อที่ยากมาก จำนวนจำแนกดี มี 8 ข้อ ได้แก่ข้อ 7, 12, 14, 16, 17, 25, 29, 33	
3) ข้อที่ความยากปานกลาง จำนวนจำแนกดี มี 4 ข้อ ได้แก่ข้อ 1, 3, 20, 22	
ข้อสอบที่ควรตัดทิ้ง ได้แก่	
1) ข้อที่ยากหรือง่ายมาก จำนวนจำแนกไม่ได้ ไม่มี	
2) ข้อที่ไม่มีอำนาจจำแนก (อำนาจจำแนกเป็น 0) ไม่มี	
3) ข้อที่อำนาจจำแนกเป็นลบ ไม่มี	

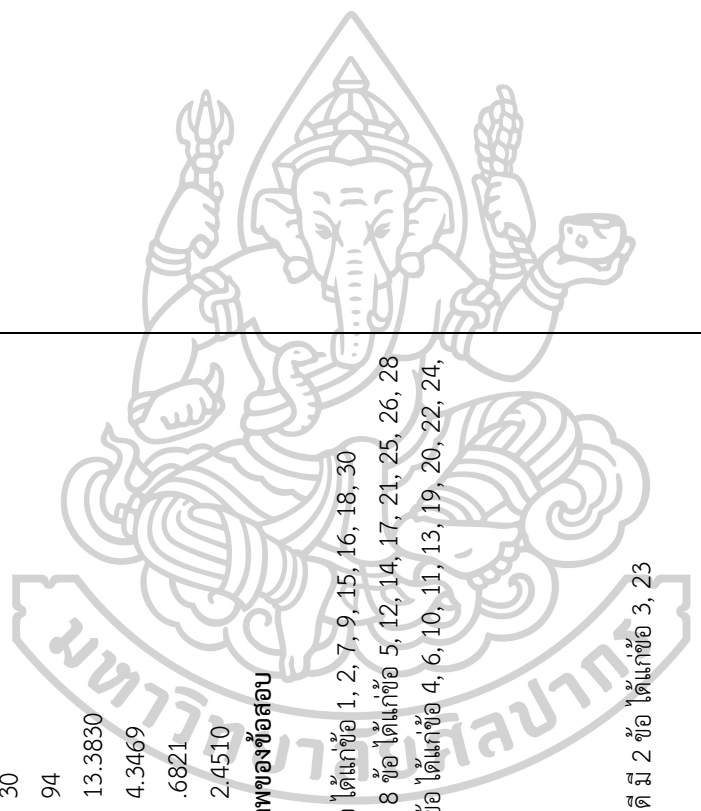


สรุปค่า P ค่า R รายข้อ แบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย Listening Part B-C

ข้อ	P	R	คุณภาพ	ข้อ	P	R	คุณภาพ
1	.64	.55	ค่อนข้างง่าย อ่านจําแนกดี	21	.52	.40	ยกง่ายปานกลาง อ่านจําแนกดี
2	.69	.58	ค่อนข้างง่าย อ่านจําแนกดี	22	.29	.43	ค่อนข้างยาก อ่านจําแนกดี
3	.44	.17	ความยากปานกลาง อ่านจําแนกไม่ได้	23	.50	.12	ความยากปานกลาง อ่านจําแนกไม่ได้
4	.37	.47	ค่อนข้างยาก อ่านจําแนกดี	24	.34	.22	ค่อนข้างยาก อ่านจําแนกดี
5	.59	.45	ยกง่ายปานกลาง อ่านจําแนกดี	25	.42	.29	ยกง่ายปานกลาง อ่านจําแนกดี
6	.25	.36	ค่อนข้างยาก อ่านจําแนกดี	26	.53	.62	ยกง่ายปานกลาง อ่านจําแนกดี
7	.78	.59	ค่อนข้างง่าย อ่านจําแนกดี	27	.36	.55	ค่อนข้างยาก อ่านจําแนกดี
8	.06	-.14	อ่านจําแนกเป็นลบ	28	.40	.33	ยกง่ายปานกลาง อ่านจําแนกดี
9	.65	.64	ค่อนข้างง่าย อ่านจําแนกดี	29	.36	.26	ค่อนข้างยาก อ่านจําแนกดี
10	.37	.47	ค่อนข้างยาก อ่านจําแนกดี	30	.66	.52	ค่อนข้างง่าย อ่านจําแนกดี
11	.21	.40	ค่อนข้างยาก อ่านจําแนกดี	เฉลี่ย	.46	.42	ยกง่ายปานกลาง อ่านจําแนกดี
12	.44	.41	ยกง่ายปานกลาง อ่านจําแนกดี				
13	.20	.24	ค่อนข้างยาก อ่านจําแนกดี				
14	.48	.56	ยกง่ายปานกลาง อ่านจําแนกดี				
15	.79	.40	ค่อนข้างง่าย อ่านจําแนกดี				
16	.68	.28	ค่อนข้างง่าย อ่านจําแนกดี				
17	.50	.52	ยกง่ายปานกลาง อ่านจําแนกดี				
18	.78	.59	ค่อนข้างง่าย อ่านจําแนกดี				
19	.27	.39	ค่อนข้างยาก อ่านจําแนกดี				
20	.21	.40	ค่อนข้างยาก อ่านจําแนกดี				

สรุปคำ P คำ R รายชื่อ แบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย Listening Part B-C

สรุปผลการวิเคราะห์	
จำนวนข้อสอบที่วิเคราะห์	30
จำนวนกระดาษคำตอบ	94
คะแนนเฉลี่ย	13.3830
ส่วนเบี่ยงเบนมาตรฐาน	4.3469
ความเชื่อมั่น KR-20	.6821
ความคลาดเคลื่อนมาตรฐาน	2.4510
สรุปคุณภาพของข้อสอบ	
ข้อสอบที่ดี ควรเก็บไว้ ได้แก่	
1) ข้อที่ค่อนข้างง่าย จำนวนจำแนกดี มี 8 ข้อ ได้แก่ข้อ 1, 2, 7, 9, 15, 16, 18, 30	
2) ข้อที่ยากปานกลาง จำนวนจำแนกดี มี 8 ข้อ ได้แก่ข้อ 5, 12, 14, 17, 21, 25, 26, 28	
3) ข้อที่ค่อนข้างยาก จำนวนจำแนกดี มี 11 ข้อ ได้แก่ข้อ 4, 6, 10, 11, 13, 19, 20, 22, 24, 27, 29	
ข้อที่ควรปรับปรุง ได้แก่	
1) ข้อที่ง่ายมาก จำนวนจำแนกดี ไม่มี	
2) ข้อที่ยากมาก จำนวนจำแนกดี ไม่มี	
3) ข้อที่ความยากปานกลาง จำนวนจำแนกดี มี 2 ข้อ ได้แก่ข้อ 3, 23	
ข้อสอบที่ควรตัดทิ้ง ได้แก่	
1) ข้อที่ยากหรือง่ายมาก จำนวนจำแนกไม่ดี ไม่มี	
2) ข้อที่ไม่มีอำนาจจำแนก (อำนาจจำแนกเป็น 0) ไม่มี	
3) ข้อที่อำนาจจำแนกเป็นลบ มี 1 ข้อ ได้แก่ข้อ 8	



สรุปค่า P ค่า R รายข้อ แบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย Sentence Structure

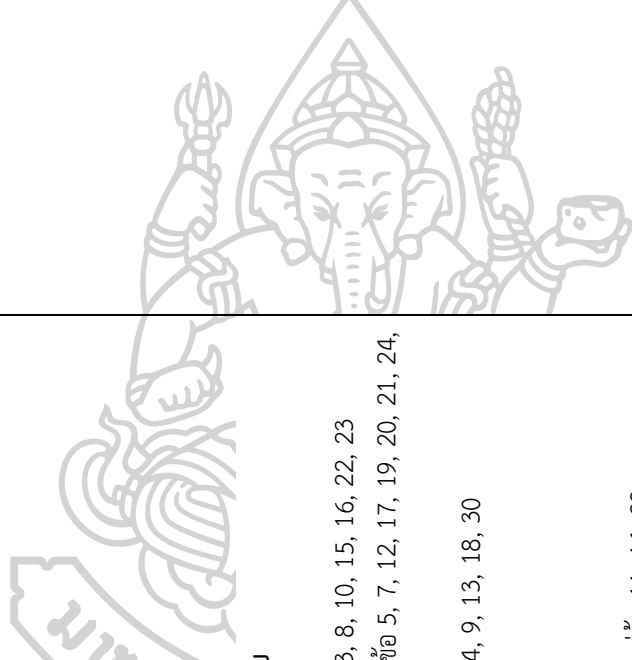
ข้อ	P	R	คุณภาพ	สรุปผลการวิเคราะห์
1	.72	.53	ค่อนข้างง่าย อำนาจจำแนกดี	จำนวนข้อสอบที่วิเคราะห์ 20
2	.83	.69	ง่ายมาก อำนาจจำแนกดี	จำนวนกระดาษคำตอบ 94
3	.30	.06	ความยากปานกลาง อำนาจจำแนกไม่ดี	คะแนนเฉลี่ย 9.9043
4	.55	.60	ยากง่ายปานกลาง อำนาจจำแนกดี	ส่วนเบี่ยงเบนมาตรฐาน 3.5252
5	.35	.43	ค่อนข้างยาก อำนาจจำแนกดี	ความเชื่อมั่น KR-20 .7014
6	.38	.58	ค่อนข้างยาก อำนาจจำแนกดี	ความคลาดเคลื่อนมาตรฐาน 1.9265
7	.36	.55	ค่อนข้างยาก อำนาจจำแนกดี	สรุปคุณภาพของข้อสอบ
8	.68	.18	ความยากปานกลาง อำนาจจำแนกไม่ดี	
9	.63	.38	ค่อนข้างง่าย อำนาจจำแนกดี	ข้อสอบที่ดี ควรเก็บไว้ได้แก่
10	.61	.69	ค่อนข้างง่าย อำนาจจำแนกดี	1) ข้อที่ค่อนข้างง่าย อำนาจจำแนกดี มี 6 ข้อ ได้แก่ข้อ 1, 9, 10, 11, 18, 19
11	.66	.52	ค่อนข้างง่าย อำนาจจำแนกดี	2) ข้อที่ค่อนข้างยาก อำนาจจำแนกดี มี 5 ข้อ ได้แก่ข้อ 4, 12, 13, 15, 16
12	.48	.40	ยากง่ายปานกลาง อำนาจจำแนกดี	3) ข้อที่ค่อนข้างยาก อำนาจจำแนกดี มี 4 ข้อ ได้แก่ข้อ 5, 6, 7, 17
13	.59	.45	ยากง่ายปานกลาง อำนาจจำแนกดี	ข้อที่ควรปรับปรุง ได้แก่
14	.36	.17	ความยากปานกลาง อำนาจจำแนกไม่ดี	1) ข้อที่ง่ายมาก อำนาจจำแนกดี มี 1 ข้อ ได้แก่ข้อ 2
15	.56	.74	ยากง่ายปานกลาง อำนาจจำแนกดี	2) ข้อที่ยากมาก อำนาจจำแนกดี มี 1 ข้อ ได้แก่ข้อ 20
16	.55	.67	ยากง่ายปานกลาง อำนาจจำแนกดี	3) ข้อที่ความยากปานกลาง อำนาจจำแนกดี มี 3 ข้อ ได้แก่ข้อ 3, 8, 14
17	.30	.24	ค่อนข้างยาก อำนาจจำแนกดี	ข้อสอบที่ควรตัดทิ้ง ได้แก่
18	.69	.71	ค่อนข้างง่าย อำนาจจำแนกดี	1) ข้อที่ยากมากหรือง่ายมาก อำนาจจำแนกไม่ดี ไม่มี
19	.69	.71	ค่อนข้างง่าย อำนาจจำแนกดี	2) ข้อที่ไม่มีอำนาจจำแนก (อำนาจจำแนกเป็น 0) ไม่มี
20	.08	.23	ยากมาก อำนาจจำแนกดี	3) ข้อที่อำนาจจำแนกเป็นลบ ไม่มี
เฉลี่ย	.52	.50	ยากง่ายปานกลาง อำนาจจำแนกดี	

สรุปค่า P ค่า R รายข้อ แบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย Written Expression

ข้อ	P	R	คุณภาพ	ข้อ	P	R	คุณภาพ
1	.33	.51	ค่อนข้างยาก อำนาจจำแนกดี	21	.52	.48	ยากง่ายปานกลาง อำนาจจำแนกดี
2	.75	.78	ค่อนข้างง่าย อำนาจจำแนกดี	22	.79	.57	ค่อนข้างง่าย อำนาจจำแนกดี
3	.69	.37	ค่อนข้างง่าย อำนาจจำแนกดี	23	.63	.66	ค่อนข้างง่าย อำนาจจำแนกดี
4	.25	.25	ค่อนข้างยาก อำนาจจำแนกดี	24	.50	.59	ยากง่ายปานกลาง อำนาจจำแนกดี
5	.52	.32	ยากง่ายปานกลาง อำนาจจำแนกดี	25	.46	.84	ยากง่ายปานกลาง อำนาจจำแนกดี
6	.36	-.08	อำนาจจำแนกเป็นลบ	26	.47	.63	ยากง่ายปานกลาง อำนาจจำแนกดี
7	.48	.40	ยากง่ายปานกลาง อำนาจจำแนกดี	27	.42	.64	ยากง่ายปานกลาง อำนาจจำแนกดี
8	.61	.50	ค่อนข้างง่าย อำนาจจำแนกดี	28	.14	.44	ยากมาก อำนาจจำแนกดี
9	.37	.47	ค่อนข้างยาก อำนาจจำแนกดี	29	.32	.18	ความยากปานกลาง อำนาจจำแนกไม่ได้
10	.69	.71	ค่อนข้างง่าย อำนาจจำแนกดี	30	.31	.26	ค่อนข้างง่าย อำนาจจำแนกดี
11	.37	.19	ความยากปานกลาง อำนาจจำแนกไม่ได้	เฉลี่ย	.48	.47	
12	.40	.25	ยากง่ายปานกลาง อำนาจจำแนกดี				
13	.39	.50	ค่อนข้างยาก อำนาจจำแนกดี				
14	.52	.08	ความยากปานกลาง อำนาจจำแนกไม่ได้				
15	.71	.69	ค่อนข้างง่าย อำนาจจำแนกดี				
16	.66	.30	ค่อนข้างง่าย อำนาจจำแนกดี				
17	.45	.36	ยากง่ายปานกลาง อำนาจจำแนกดี				
18	.21	.28	ค่อนข้างยาก อำนาจจำแนกดี				
19	.52	.56	ยากง่ายปานกลาง อำนาจจำแนกดี				
20	.50	.52	ยากง่ายปานกลาง อำนาจจำแนกดี				

สรุปคำ P คำ R รายชื่อ แบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย Written Expression

สรุปผลการวิเคราะห์	
จำนวนข้อสอบที่วิเคราะห์	30
จำนวนกระดาษคำตอบ	94
คะแนนเฉลี่ย	13.7553
ส่วนเบี่ยงเบนมาตรฐาน	4.8393
ความเชื่อมั่น KR-20	.7361
ความคลาดเคลื่อนมาตรฐาน	2.4862
สรุปคุณภาพของข้อสอบ	
ข้อสอบที่ดี ควรเก็บไว้ ได้แก่	
1) ข้อที่ค่อนข้างง่าย อำนาจจำแนกดี มี 8 ข้อ ได้แก่ข้อ 2, 3, 8, 10, 15, 16, 22, 23	
2) ข้อที่ยากปานกลาง อำนาจจำแนกดี มี 11 ข้อ ได้แก่ข้อ 5, 7, 12, 17, 19, 20, 21, 24, 25, 26, 27	
3) ข้อที่ค่อนข้างยาก อำนาจจำแนกดี มี 6 ข้อ ได้แก่ข้อ 1, 4, 9, 13, 18, 30	
ข้อที่ควรปรับปรุง ได้แก่	
1) ข้อที่ง่ายมาก อำนาจจำแนกดี ไม่มี	
2) ข้อที่ยากมาก อำนาจจำแนกดี มี 1 ข้อ ได้แก่ข้อ 28	
3) ข้อที่ความยากปานกลาง อำนาจจำแนกดี มี 3 ข้อ ได้แก่ข้อ 11, 14, 29	
ข้อสอบที่ควรตัดทิ้ง ได้แก่	
1) ข้อที่ยากหรือง่ายมาก อำนาจจำแนกไม่ดี ไม่มี	
2) ข้อที่ไม่มีอำนาจจำแนก (อำนาจจำแนกเป็น 0) ไม่มี	
3) ข้อที่อำนาจจำแนกเป็นลบ มี 1 ข้อ ได้แก่ข้อ 6	



สรุปค่า P ค่า R รายข้อ แบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย Reading for Everyday Use

ข้อ	P	R	คุณภาพ	ข้อ	P	R	คุณภาพ
1	.38	.30	ค่อนข้างยาก อ่านจําแนกดี	21	.43	.49	ง่ายปานกลาง อ่านจําแนกดี
2	.59	.45	ง่ายปานกลาง อ่านจําแนกดี	22	.37	.77	ค่อนข้างยาก อ่านจําแนกดี
3	.73	.80	ค่อนข้างง่าย อ่านจําแนกดี	23	.58	.71	ง่ายปานกลาง อ่านจําแนกดี
4	.67	.84	ค่อนข้างง่าย อ่านจําแนกดี	24	.50	.87	ง่ายปานกลาง อ่านจําแนกดี
5	.29	.33	ค่อนข้างยาก อ่านจําแนกดี	25	.34	.52	ค่อนข้างยาก อ่านจําแนกดี
6	.36	.55	ค่อนข้างยาก อ่านจําแนกดี	26	.73	.66	ค่อนข้างง่าย อ่านจําแนกดี
7	.69	.71	ค่อนข้างง่าย อ่านจําแนกดี	27	.35	.64	ค่อนข้างยาก อ่านจําแนกดี
8	.52	.56	ง่ายปานกลาง อ่านจําแนกดี	28	.19	-.36	อ่านจําแนกเป็นลบ
9	.56	.74	ง่ายปานกลาง อ่านจําแนกดี	29	.38	.58	ค่อนข้างยาก อ่านจําแนกดี
10	.44	.33	ง่ายปานกลาง อ่านจําแนกดี	30	.43	.49	ง่ายปานกลาง อ่านจําแนกดี
11	.47	.77	ง่ายปานกลาง อ่านจําแนกดี	31	.11	.35	ยากมาก อ่านจําแนกดี
12	.43	.23	ง่ายปานกลาง อ่านจําแนกดี	32	.35	.43	ค่อนข้างยาก อ่านจําแนกดี
13	.24	.21	ค่อนข้างยาก อ่านจําแนกดี	33	.46	.84	ง่ายปานกลาง อ่านจําแนกดี
14	.24	.10	ความยากปานกลาง อ่านจําแนกไม่ได้	34	.59	.71	ง่ายปานกลาง อ่านจําแนกดี
15	.61	.69	ค่อนข้างง่าย อ่านจําแนกดี	35	.38	.13	ความยากปานกลาง อ่านจําแนกไม่ได้
16	.71	.56	ค่อนข้างง่าย อ่านจําแนกดี	36	.76	.62	ค่อนข้างง่าย อ่านจําแนกดี
17	.54	.44	ง่ายปานกลาง อ่านจําแนกดี	37	.42	.64	ง่ายปานกลาง อ่านจําแนกดี
18	.50	.59	ง่ายปานกลาง อ่านจําแนกดี	38	.74	.64	ค่อนข้างง่าย อ่านจําแนกดี
19	.60	.33	ค่อนข้างง่าย อ่านจําแนกดี	39	.14	.22	ยากมาก อ่านจําแนกดี
20	.44	.08	ความยากปานกลาง อ่านจําแนกไม่ได้	40	.39	.50	ค่อนข้างยาก อ่านจําแนกดี
				เฉลี่ย	.47	.56	ง่ายปานกลาง อ่านจําแนกดี

สรุปคำ P คำ R รายชื่อ แบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย Reading for Everyday Use

สรุปผลการวิเคราะห์	
จำนวนข้อสอบที่วิเคราะห์	40
จำนวนกระดาษคำตอบ	94
คะแนนเฉลี่ย	17.0638
ส่วนเบี่ยงเบนมาตรฐาน	7.5917
ความเชื่อมั่น KR-20	.8665
ความคลาดเคลื่อนมาตรฐาน	2.7742
สรุปคุณภาพของข้อสอบ	
ข้อสอบที่ดี ควรเก็บไว้ ได้แก่	
1) ข้อที่ค่อนข้างง่าย จำนวนจำแนกดี มี 9 ข้อ ได้แก่ข้อ 3, 4, 7, 15, 16, 19, 26, 36, 38	
2) ข้อที่ยากง่ายปานกลาง จำนวนจำแนกดี มี 15 ข้อ ได้แก่ข้อ 2, 8, 9, 10, 11, 12, 17, 18, 21, 23, 24, 30, 33, 34, 37	
3) ข้อที่ค่อนข้างยาก จำนวนจำแนกดี มี 10 ข้อ ได้แก่ข้อ 1, 5, 6, 13, 22, 25, 27, 29, 32, 40	
ข้อที่ควรปรับปรุง ได้แก่	
1) ข้อที่ง่ายมาก จำนวนจำแนกดี ไม่มี	
2) ข้อที่ยากมาก จำนวนจำแนกดี มี 2 ข้อ ได้แก่ข้อ 31, 39	
3) ข้อที่ความยากปานกลาง จำนวนจำแนกไม่ดี มี 3 ข้อ ได้แก่ข้อ 14, 20, 35	
ข้อสอบที่ควรตัดทิ้ง ได้แก่	
1) ข้อที่ยากหรือง่ายมาก จำนวนจำแนกไม่ดี ไม่มี	
2) ข้อที่ไม่มีอำนาจจำแนก (อำนาจจำแนกเป็น 0) ไม่มี	
3) ข้อที่อำนาจจำแนกเป็นลบ มี 1 ข้อ ได้แก่ข้อ 28	

สรุปค่า P ค่า R รายข้อ แบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย Reading for Academic Purpose

ข้อ	P	R	คุณภาพ	ข้อ	P	R	คุณภาพ
1	.50	.20	ยากง่ายปานกลาง อำนาจจำแนกดี	21	.69	.71	ค่อนข้างง่าย อำนาจจำแนกดี
2	.36	.26	ค่อนข้างยาก อำนาจจำแนกดี	22	.71	.69	ค่อนข้างง่าย อำนาจจำแนกดี
3	.43	.56	ยากง่ายปานกลาง อำนาจจำแนกดี	23	.39	.50	ค่อนข้างยาก อำนาจจำแนกดี
4	.40	.61	ยากง่ายปานกลาง อำนาจจำแนกดี	24	.46	.58	ยากง่ายปานกลาง อำนาจจำแนกดี
5	.50	.59	ยากง่ายปานกลาง อำนาจจำแนกดี	25	.19	.36	ยากมาก อำนาจจำแนกดี
6	.37	.47	ค่อนข้างยาก อำนาจจำแนกดี	26	.50	.73	ยากง่ายปานกลาง อำนาจจำแนกดี
7	.43	.56	ยากง่ายปานกลาง อำนาจจำแนกดี	27	.55	.67	ยากง่ายปานกลาง อำนาจจำแนกดี
8	.47	.63	ยากง่ายปานกลาง อำนาจจำแนกดี	28	.69	.71	ค่อนข้างง่าย อำนาจจำแนกดี
9	.58	.71	ยากง่ายปานกลาง อำนาจจำแนกดี	29	.43	.49	ยากง่ายปานกลาง อำนาจจำแนกดี
10	.63	.77	ค่อนข้างง่าย อำนาจจำแนกดี	30	.33	.61	ค่อนข้างยาก อำนาจจำแนกดี
11	.47	.63	ยากง่ายปานกลาง อำนาจจำแนกดี	31	.37	.38	ค่อนข้างยาก อำนาจจำแนกดี
12	.42	.29	ยากง่ายปานกลาง อำนาจจำแนกดี	32	.17	.32	ยากมาก อำนาจจำแนกดี
13	.26	.50	ค่อนข้างยาก อำนาจจำแนกดี	33	.34	.22	ค่อนข้างยาก อำนาจจำแนกดี
14	.59	.45	ยากง่ายปานกลาง อำนาจจำแนกดี	34	.22	-.05	อำนาจจำแนกเป็นลบ
15	.40	.79	ยากง่ายปานกลาง อำนาจจำแนกดี	35	.24	.20	ค่อนข้างยาก อำนาจจำแนกดี
16	.40	.25	ยากง่ายปานกลาง อำนาจจำแนกดี	36	.32	.38	ค่อนข้างยาก อำนาจจำแนกดี
17	.60	.79	ค่อนข้างง่าย อำนาจจำแนกดี	37	.28	.53	ค่อนข้างยาก อำนาจจำแนกดี
18	.42	.37	ยากง่ายปานกลาง อำนาจจำแนกดี	38	.39	.50	ค่อนข้างยาก อำนาจจำแนกดี
19	.26	-.05	อำนาจจำแนกเป็นลบ	39	.32	.18	ความยากปานกลาง อำนาจจำแนกไม่ได้
20	.38	.21	ค่อนข้างยาก อำนาจจำแนกดี	40	.43	.56	ยากง่ายปานกลาง อำนาจจำแนกดี

สรุปค่า P ค่า R รายข้อ แบบทดสอบวัดความสามารถภาษาอังกฤษระดับมหาวิทยาลัย Reading for Everyday Use

ข้อ	P	r	คุณภาพ	สรุปผลการวิเคราะห์
41	.60	.79	ค่อนข้างง่าย อำนาจจำแนกดี	จำนวนข้อสอบที่วิเคราะห์ 42
42	.63	.77	ค่อนข้างง่าย อำนาจจำแนกดี	จำนวนกระดาษคำตอบ 94
เฉลี่ย	.43	.52		คะแนนเฉลี่ย 16.5958
				ส่วนเบี่ยงเบนมาตรฐาน 7.8987
				ความเชื่อมั่น KR-20 .8673
				ความคลาดเคลื่อนมาตรฐาน 2.8778
				สรุปคุณภาพของข้อสอบ
				ข้อสอบที่ดี ควรเก็บไว้ได้แก่
				1) ข้อที่ค่อนข้างง่าย อำนาจจำแนกดี มี 7 ข้อ ได้แก่ข้อ 10, 17, 21, 22, 28, 41, 42
				2) ข้อที่ยากง่ายปานกลาง อำนาจจำแนกดี มี 18 ข้อ ได้แก่ข้อ 1, 3, 4, 5, 7, 8, 9, 11, 12, 14, 15, 16, 18, 24, 26, 27, 29, 40
				3) ข้อที่ค่อนข้างยาก อำนาจจำแนกดี มี 12 ข้อ ได้แก่ข้อ 2, 6, 13, 20, 23, 30, 31, 33, 35, 36, 37, 38
				ข้อที่ควรปรับปรุง ได้แก่
				1) ข้อที่ง่ายมาก อำนาจจำแนกดี ไม่มี
				2) ข้อที่ยากมาก อำนาจจำแนกดี มี 2 ข้อ ได้แก่ข้อ 25, 32
				3) ข้อที่ความยากปานกลาง อำนาจจำแนกดี มี 1 ข้อ ได้แก่ข้อ 39
				ข้อสอบที่ควรตัดทิ้ง ได้แก่
				1) ข้อที่ยากหรือง่ายมาก อำนาจจำแนกดีไม่มี
				2) ข้อที่ไม่มีอำนาจจำแนก (อำนาจจำแนกเป็น 0) ไม่มี
				3) ข้อที่อำนาจจำแนกเป็นลบ มี 2 ข้อ ได้แก่ข้อ 19, 34

ประวัติผู้เขียน

ชื่อ-สกุล	จตุพงศ์ โมรา
วัน เดือน ปี เกิด	15 พฤษภาคม 2516
สถานที่เกิด	นครศรีธรรมราช
วุฒิการศึกษา	ประวัติการศึกษา พ.ศ. 2539 ศิลปศาสตรบัณฑิต สาขาวิชาภาษาอังกฤษ คณะมนุษยศาสตร์มหาวิทยาลัยศรีนครินทรวิโรฒภาคใต้ พ.ศ. 2545 ศิลปศาสตรมหาบัณฑิต สาขาวิชาภาษาศาสตร์ประยุกต์ คณะศิลปศาสตร์ มหาวิทยาลัยสงขลานครินทร์ วิทยาเขตหาดใหญ่ พ.ศ. 2562 ปรัชญาดุษฎีบัณฑิต สาขาวิชาหลักสูตรและการสอน (การสอนภาษาอังกฤษ) คณะศึกษาศาสตร์ มหาวิทยาลัยศิลปากร วิทยาเขตพระราชวังสนามจันทร์ ประวัติการทำงาน พ.ศ. 2539-2547 อาจารย์สอนวิชาภาษาอังกฤษ หัวหน้าแผนกวิชาภาษาอังกฤษ หัวหน้าฝ่ายวิชาการ โรงเรียนบริหารธุรกิจวิทยา สงขลา ตำบลบ่อยาง อำเภอเมืองสงขลา จังหวัดสงขลา พ.ศ. 2548-ปัจจุบัน อาจารย์สอนวิชาภาษาอังกฤษ สาขาวิชาภาษาตะวันตก คณะมนุษยศาสตร์และสังคมศาสตร์ มหาวิทยาลัยทักษิณ วิทยาเขตสงขลา อำเภอเมืองสงขลา จังหวัดสงขลา
ที่อยู่ปัจจุบัน	ที่อยู่ 140/104 หมู่ 4 ตำบลเขารูปช้าง อำเภอเมืองสงขลา จังหวัดสงขลา 90000 ที่ทำงาน สาขาวิชาภาษาตะวันตก คณะมนุษยศาสตร์และสังคมศาสตร์ มหาวิทยาลัยทักษิณ วิทยาเขตสงขลา อำเภอเมืองสงขลา จังหวัดสงขลา