



รายงานการวิจัย

๔
๒๕๓๕

การสอบสมิทธิภาพทั่วไปทางภาษาอังกฤษ ครั้งที่ ๒ ปี พ.ศ. ๒๕๓๕

ในโครงการ

การสร้างและพัฒนาคุณภาพแบบทดสอบ
สมิทธิภาพทั่วไปทางภาษาอังกฤษและระบบการทดสอบที่เกี่ยวข้อง

สถาบันวิจัยประชากร

จุฬาลงกรณ์มหาวิทยาลัย

จพ
กษ ๑๕
๐๐ ๘๑๘๙

สถาบันภาษาและคำขวัญการ
จุฬาลงกรณ์มหาวิทยาลัย

๒๕๓๕

บทคัดย่อ

งานวิจัยเรื่อง "การสอบสมรรถภาพทั่วไปทางภาษาอังกฤษ ครั้งที่ 2 ปี 2534" เป็นส่วนหนึ่งของโครงการ "การสร้างและพัฒนาคูณภาพแบบทดสอบสมรรถภาพทั่วไปทางภาษาอังกฤษและระบบการทดสอบที่เกี่ยวข้อง" ซึ่งเป็นโครงการที่ฝ่ายวิชาการและสถาบันภาษาของจุฬาลงกรณ์มหาวิทยาลัย จัดขึ้นร่วมกัน

วัตถุประสงค์ของการดำเนินงานระยะที่ 2 พ.ศ. 2534 มีดังนี้

1. เพื่อพัฒนาข้อทดสอบสมรรถภาพทั่วไปทางภาษาอังกฤษด้านการฟัง การอ่าน และเขียนเพิ่มเติมมากขึ้นจากข้อทดสอบที่ได้จากการสร้างขึ้นในการวิจัยระยะปีที่ 1 และให้ได้แบบทดสอบ 4 ฉบับ ที่คู่ขนานกันในด้านเนื้อหาและสถิติโดยใช้ตารางวิเคราะห์เดียวกัน
2. เพื่อวิเคราะห์และปรับปรุงแบบทดสอบทั้งฉบับและเป็นรายข้อ
3. เพื่อพัฒนาแบบทดสอบให้เป็นมาตรฐานและนำเก็บไว้ในคลังข้อทดสอบโดยใช้คอมพิวเตอร์

ประชากรเชิงนิยาม ในการวิจัยครั้งนี้คือนิสิตชั้นปีที่ 1 ถึงนิสิตบัณฑิตศึกษาและบุคคลภายนอกทั่วไปที่มีพื้นฐานความรู้ตั้งแต่ระดับปริญญาตรีขึ้นไปที่มีความสนใจและสมัครสอบ เพื่อวัดระดับสมรรถภาพทั่วไปทางภาษาอังกฤษด้วยตนเอง ซึ่งมีผู้มีสิทธิสมัครสอบจำนวน 2,334 คน พลวิจัยคือผู้ที่ได้เข้าทดสอบแบบทดสอบสมรรถภาพทั่วไปทางภาษาอังกฤษ จำนวน 481 คน คิดเป็นร้อยละ 20.57 ของผู้มีสิทธิสมัครสอบทั้งหมด และบุคคลภายนอกจำนวน 87 คน เข้ารับการทดสอบแบบทดสอบสมรรถภาพทั่วไปทางภาษาอังกฤษฉบับปรับปรุง

แบบทดสอบสมรรถภาพทั่วไปทางภาษาอังกฤษทั้ง 4 ฉบับ เป็นแบบทดสอบอิงเกณฑ์ประกอบด้วยข้อสอบปรนัย 480 ข้อ อัตนัย 12 ข้อ ผ่านการวิเคราะห์เป็นรายข้อและทั้งฉบับเพื่อหาค่าความยากง่าย อำนาจจำแนก ค่าความเที่ยง และค่าอำนาจในการทำนาย ผู้วิจัยใช้การวิเคราะห์แบบประเพณีนิยม เพื่อปรับปรุงคัดเลือก ตัวलग และใช้การวิเคราะห์รายข้อตามแบบทฤษฎีความสามารถแฝงด้วยวิธี Rasch Model เพื่อให้ได้ข้อทดสอบที่มีค่าความยากง่ายคงที่สำหรับเก็บไว้ในคลังข้อสอบ

ผลการวิจัยสรุปได้ว่า (1) ได้แบบทดสอบปรนัยวัดทักษะฟัง อ่าน และเขียน จำนวน 4 ฉบับ มีค่าความเที่ยงดีมาก 2 ฉบับ และดี 2 ฉบับ แบบทดสอบปรนัยมีค่าความเที่ยงแบบ $KR_{20} = .854-.910$ และ แบบ Phi = $.822-.896$ ค่าอำนาจในการทำนาย $48.00-58.50$ และได้แบบทดสอบอัตนัยวัดความสามารถทางการเขียนที่อยู่ในเกณฑ์พอใช้ได้ - ดี ($R_{xy} = .594-.840$) (2) ได้ข้อทดสอบแบบปรนัยวัดสมรรถภาพทั่วไปทางภาษาอังกฤษที่มีคุณภาพดีมากถึงพอใช้ได้ 471 ข้อ (3) ได้ข้อทดสอบแบบอัตนัยวัดสมรรถภาพทางการเขียนที่มีคุณภาพดีถึงพอใช้ได้ 12 ข้อ และ (4) ได้ข้อทดสอบแบบปรนัยวัดสมรรถภาพทั่วไปทางภาษาอังกฤษ ซึ่งได้มาตรฐานเข้าเกณฑ์ของ Rasch Model 414 ข้อ

ABSTRACT

The research project entitled "General English Proficiency Testing: Phase II, 1991" is part of a project coordinated by the Academic Affairs Division and Chulalongkorn University Language Institute to develop standardized tests of general English proficiency and related testing systems. The objectives of the research were as follows:

1. to generate more test items than those in the Phase I research by using the same test table of specifications
2. to analyze and improve the constructed test items item by item as well as by form
3. to standardize the tests and enter them into the computerized test-item banking system.

The population of the research numbering 2,334 were Chulalongkorn University students from first year to graduate level, and the general public with at least a Bachelor's degree who voluntarily enrolled for the tests. The subjects in the first administration of the four forms of the tests numbered 481 or 20.57 % of the population. The subjects in the second administration taking the revised version of the tests numbered 87.

The tests constructed, administered, analyzed, and revised are criterion-referenced tests of general English proficiency consisting of 480 objective-type items and 12 subjective-type items. The classical model of test item analysis was used for item revision while the Rasch model was for item banking.

The study led to (1) the development of 480 objective items testing the listening, reading and writing skills grouped in four parallel test forms with reliability indices ranging from acceptable to high (KR 20 = .854 - .910, Phi coefficients = .820 - .896), the efficiency of prediction $E = 48.00 - 58.50$; and 12 subjective items measuring the skills of writing with acceptable to high reliability indices ($r_{xy} = .554 - .840$); (2) the 471 standardized objective test items; (3) 12 standardized subjective test items; and (4) 414 objective test items which fit the Rasch model of item analysis.

คำนำและกิตติกรรมประกาศ

งานวิจัยเรื่อง "การสอบสมิทธิภาพทั่วไปทางภาษาอังกฤษครั้งที่ 2 ปี 2534" เป็นส่วนหนึ่งของโครงการวิจัยและพัฒนาเรื่อง "การสร้างและพัฒนาคณาจารย์ของแบบทดสอบสมิทธิภาพทั่วไปทางภาษาอังกฤษ และระบบการทดสอบที่เกี่ยวข้อง" ซึ่งเป็นโครงการร่วมกันระหว่างฝ่ายวิชาการและสถาบันภาษา ของจุฬาลงกรณ์มหาวิทยาลัย

รายงานวิจัยเรื่องนี้ต่อเนื่องจากรายงานการวิจัยเรื่อง "แนวโน้มระดับสมิทธิภาพทั่วไปทางภาษาอังกฤษของนิสิตจุฬาลงกรณ์มหาวิทยาลัยและบุคคลภายนอกที่สนใจ ปี พ.ศ. 2533" จัดทำโดย รองศาสตราจารย์ ดร. สุทัศน์ สุกมลลันต์ และคณะ

การดำเนินงานของโครงการในปีที่ 2 นี้ ประสบผลสำเร็จด้วยดี เป็นเพราะได้รับความอนุเคราะห์ในการดำเนินการสอบจากคณาจารย์และบุคลากรที่ช่วยในการจัดสถานที่และดำเนินการสอบ และวิเคราะห์ผล คณะดำเนินงานจึงใคร่ขอขอบพระคุณทุกท่านที่เกี่ยวข้องเป็นอย่างสูง รวมถึง รองศาสตราจารย์มาลินี จันทวิมล และรองศาสตราจารย์ ดร. สุทัศน์ สุกมลลันต์ ผู้สนับสนุนการวิจัยครั้งนี้ และได้ช่วยให้ข้อคิดเห็นและข้อเสนอแนะที่เป็นประโยชน์

รองศาสตราจารย์ ดร. อัจฉรา วงศ์โสธร
หัวหน้าโครงการ

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

สารบัญ

	หน้า
บทคัดย่อภาษาไทย	i
บทคัดย่อภาษาอังกฤษ	ii
คำนำและกิตติกรรมประกาศ	iii
สารบัญ	iv
สารบัญตาราง	vi
สารบัญแผนภูมิ	viii
บทที่ 1 บทนำ	1
ความเป็นมาและความสำคัญของปัญหา	1
วัตถุประสงค์ของการวิจัย	1
ขอบเขตของการวิจัย	2
ข้อตกลงเบื้องต้น	2
คำจำกัดความในการวิจัย	3
ประโยชน์ที่คาดว่าจะได้รับ	3
บทที่ 2 เอกสารทางวิชาการที่เกี่ยวข้อง	4
การวิจัยทางการทดสอบภาษา	4
การพัฒนาแบบทดสอบวัดสมรรถภาพทางภาษาแบบอิงเกณฑ์	7
การวิเคราะห์ข้อทดสอบแบบอิงเกณฑ์ - กลุ่ม และ	10
การวิเคราะห์ตามทฤษฎีความสามารถแฝงแบบ Rasch Model	11
การสร้างคลังข้อทดสอบ	14

	หน้า
บทที่ 3 วิธีดำเนินการวิจัย	17
ประชากรและผลวิจัย	17
การสร้างแบบทดสอบ	17
การบริหารการสอบ	19
การวิเคราะห์ข้อมูล	19
เกณฑ์ในการพิจารณาคุณภาพข้อสอบ	23
บทที่ 4 ผลการวิเคราะห์ข้อมูล	27
ผลการวิเคราะห์ข้อทดสอบสมิทริภาพทั่วไปทางภาษาอังกฤษ	27
เพื่อการพัฒนาปรับปรุงข้อกระทง	
ผลการวิเคราะห์แบบทดสอบสมิทริภาพทั่วไปทางภาษาอังกฤษฉบับ	54
ปรับปรุง	
ผลการวิเคราะห์ข้อทดสอบแบบ Rasch Model เพื่อพัฒนาเป็น	58
คลังข้อสอบ	
สรุปผลการวิเคราะห์	83
บทที่ 5 สรุป อภิปรายผล และข้อเสนอแนะ	89
สรุป	89
อภิปรายผล	90
ข้อเสนอแนะ	91
บรรณานุกรม	92
ภาคผนวก	86
ภาคผนวก ก.	86
ภาคผนวก ข.	101
ภาคผนวก ค.	106

เลขหมึก	๒๕๖๑
เลขทะเบียน	๐๐๕๑๙๙
วัน,เดือน,ปี	๒๕๖๓.๑๑.๑๗

สารบัญตาราง

		หน้า
ตารางที่ 1	ตัวอย่างตารางวิเคราะห์การใช้ภาษาในชีวิตประจำวัน	15
ตารางที่ 2	แสดงค่าสถิติต่าง ๆ ของแบบทดสอบสมิทธิภาพทั่วไปทางภาษาอังกฤษ ฉบับที่ 1-4	27
ตารางที่ 3	ผลการวิเคราะห์ข้อกระทบแบบทดสอบสมิทธิภาพทั่วไปทางภาษา อังกฤษ ฉบับที่ 1	28
ตารางที่ 4	ผลการวิเคราะห์ข้อกระทบแบบทดสอบสมิทธิภาพทั่วไปทางภาษา อังกฤษ ฉบับที่ 2	34
ตารางที่ 5	ผลการวิเคราะห์ข้อกระทบแบบทดสอบสมิทธิภาพทั่วไปทางภาษา อังกฤษ ฉบับที่ 3	40
ตารางที่ 6	ผลการวิเคราะห์ข้อกระทบแบบทดสอบสมิทธิภาพทั่วไปทางภาษา อังกฤษ ฉบับที่ 4	46
ตารางที่ 7	ผลการวิเคราะห์ข้อทดสอบสมิทธิภาพทั่วไปทางภาษาอังกฤษ ส่วนอัตนัย	52
ตารางที่ 8	ค่าความเที่ยงแบบ inter-rater	53
ตารางที่ 9	ผลการวิเคราะห์คุณภาพของแบบทดสอบสมิทธิภาพทั่วไปทางภาษา อังกฤษ ฉบับปรับปรุง	54
ตารางที่ 10	ผลการวิเคราะห์ข้อกระทบแบบทดสอบสมิทธิภาพทั่วไปทางภาษา อังกฤษ ฉบับปรับปรุง	55
ตารางที่ 11	ผลการวิเคราะห์ข้อกระทบโดยใช้ Rasch Model ของแบบทดสอบ สมิทธิภาพทั่วไปทางภาษาอังกฤษ ฉบับที่ 1	59
ตารางที่ 12	ผลการวิเคราะห์ข้อกระทบโดยใช้ Rasch Model ของแบบทดสอบ สมิทธิภาพทั่วไปทางภาษาอังกฤษ ฉบับที่ 2	65
ตารางที่ 13	ผลการวิเคราะห์ข้อกระทบโดยใช้ Rasch Model ของแบบทดสอบ สมิทธิภาพทั่วไปทางภาษาอังกฤษ ฉบับที่ 3	71

	หน้า	
ตารางที่ 14	ผลการวิเคราะห์ข้อกระทงโดยใช้ Rasch Model ของแบบทดสอบ สมิทธิภาพทั่วไปทางภาษาอังกฤษ ฉบับที่ 4	77
ตารางที่ 15	สรุปผลการวิเคราะห์	83
ตารางที่ 16	ค่าความยากง่ายและค่าอำนาจจำแนกของข้อสอบปรนัย ของ แบบทดสอบสมิทธิภาพทั่วไปทางภาษาอังกฤษ	85
ตารางที่ 17	จำนวนและร้อยละของข้อสอบของแบบทดสอบสมิทธิภาพทั่วไปทาง ภาษาอังกฤษที่เข้าเกณฑ์ตามการวิเคราะห์แบบ Classical Model และแบบ Rasch Model	87
ตารางที่ 18	ค่าความยากง่ายและค่าอำนาจจำแนกของข้อสอบอัตนัย ของ แบบทดสอบสมิทธิภาพทั่วไปทางภาษาอังกฤษ	88
ตารางที่ 13	สรุปค่าสถิติพื้นฐานของแบบทดสอบสมิทธิภาพทั่วไปทางภาษาอังกฤษ ฉบับที่ 1	101
ตารางที่ 20	สรุปค่าสถิติพื้นฐานของแบบทดสอบสมิทธิภาพทั่วไปทางภาษาอังกฤษ ฉบับที่ 2	102
ตารางที่ 21	สรุปค่าสถิติพื้นฐานของแบบทดสอบสมิทธิภาพทั่วไปทางภาษาอังกฤษ ฉบับที่ 3	103
ตารางที่ 22	สรุปค่าสถิติพื้นฐานของแบบทดสอบสมิทธิภาพทั่วไปทางภาษาอังกฤษ ฉบับที่ 4	104
ตารางที่ 23	สรุปค่าสถิติพื้นฐานของแบบทดสอบสมิทธิภาพทั่วไปทางภาษาอังกฤษ ฉบับปรับปรุง	105
ตารางที่ 24	แสดงค่าสหสัมพันธ์ระหว่างทักษะ การฟัง การอ่าน และการเขียน ของข้อสอบวัดสมิทธิภาพทั่วไปทางภาษาอังกฤษ	106

สารบัญแผนภูมิ

	หน้า
แผนภูมิที่ 1 สมรรถภาพทางภาษา	6
แผนภูมิที่ 2 กระบวนการคัดเลือกข้อทดสอบ (แบบอิง เกณฑ์)	9
แผนภูมิที่ 3 การกระจายของคะแนนจากแบบทดสอบวัดสมิทธิภาพทั่วไปทาง ภาษาอังกฤษ ฉบับที่ 1	96
แผนภูมิที่ 4 แสดงการกระจายของคะแนนจากแบบทดสอบวัดสมิทธิภาพทั่วไปทาง ภาษาอังกฤษ ฉบับที่ 2	97
แผนภูมิที่ 5 แสดงการกระจายของคะแนนจากแบบทดสอบวัดสมิทธิภาพทั่วไปทาง ภาษาอังกฤษ ฉบับที่ 3	98
แผนภูมิที่ 6 แสดงการกระจายของคะแนนจากแบบทดสอบวัดสมิทธิภาพทั่วไปทาง ภาษาอังกฤษ ฉบับที่ 4	99
แผนภูมิที่ 7 แสดงการกระจายของคะแนนจากแบบทดสอบวัดสมิทธิภาพทั่วไปทาง ภาษาอังกฤษ ฉบับปรับปรุง	100

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย



ความเป็นมาและความสำคัญของปัญหา

จุฬาลงกรณ์มหาวิทยาลัยได้เล็งเห็นว่าภาษาอังกฤษเป็นภาษาสากลในการสื่อสาร และแสวงหาความรู้ ตลอดจนมีความสำคัญมากทั้งต่อการศึกษาในระดับอุดมศึกษาและในการ ประกอบอาชีพในวงการต่าง ๆ ทั้งภายในและภายนอกประเทศ ถ้าบัณฑิตของจุฬาลงกรณ์ มหาวิทยาลัยมีความสามารถในการใช้ภาษาอังกฤษในด้านเนื้อหาวิชาการ และในการใช้ภาษา ภาษาอังกฤษในระดับที่สามารถติดต่อทำความเข้าใจกับต่างประเทศได้ก็จะทำให้บัณฑิตทันต่อเหตุการณ์ มีคุณค่าเป็นที่ต้องการของตลาดแรงงานและสังคมมากขึ้น จุฬาลงกรณ์มหาวิทยาลัยจึงได้จัด โครงการการก่อสร้างและพัฒนาคุณภาพของแบบทดสอบวัดสมรรถภาพทั่วไปทางภาษาอังกฤษ และ ระบบการทดสอบที่เกี่ยวข้องขึ้น เพื่อกระตุ้นให้เกิดการตื่นตัวทั้งอาจารย์และนิสิตให้เกิดความ สนใจและให้ความสำคัญต่อการเรียนการสอนภาษาอังกฤษ โดยโครงการนี้มีระยะดำเนินการ 3 ปี ปี 2533-2535 ซึ่งโครงการนี้ได้ดำเนินการเสร็จสิ้นไปในปีที่ 1 พ.ศ. 2533 โดยมี วัตถุประสงค์ เพื่อศึกษาสำรวจเกี่ยวกับแนวโน้มระดับสมรรถภาพทั่วไปทางภาษาอังกฤษของนิสิต ชั้นปีที่ 1-6 บัณฑิตวิทยาลัย และบุคคลภายนอกที่สนใจ เมื่อจำแนกออกเป็นมิติต่าง ๆ คือ เพศ ชั้นปี คณะวิชา เชื้อชาติ สัญชาติ ศาสนา อายุ และรายได้ของครอบครัว ในการดำเนินการ ในระยะปีที่ 2 คณะผู้วิจัยได้ดำเนินการตามเป้าหมายของโครงการและจากข้อเสนอแนะของ คณะผู้วิจัยในระยะปีที่ 1 ได้ดำเนินการเพื่อสร้างและพัฒนาคุณภาพของแบบทดสอบสมรรถภาพ ทั่วไปทางภาษาอังกฤษที่ใช้แล้วในการวิจัยในระยะปีที่ 1 ให้มีจำนวนข้อเพิ่มมากขึ้นและจัดเก็บใน คลังข้อสอบโดยใช้ระบบคอมพิวเตอร์

วัตถุประสงค์ของการวิจัย

1. เพื่อพัฒนาข้อทดสอบสมรรถภาพทั่วไปทางภาษาอังกฤษด้านการฟัง การอ่านและ เขียน เพิ่มมากขึ้นจากข้อทดสอบที่ได้จากการสร้างขึ้นในการวิจัยระยะปีที่ 1 และให้ได้แบบ ทดสอบ 4 ฉบับที่คู่ขนานกันในด้านเนื้อหาและสถิติ
2. เพื่อวิเคราะห์และปรับปรุงแบบทดสอบที่สร้างขึ้นทั้งฉบับและเป็นรายชื่อ

3. เพื่อพัฒนาแบบทดสอบให้เป็นมาตรฐานและเก็บไว้ในคลังข้อทดสอบโดยใช้คอมพิวเตอร์ (Computerized item bank)

ขอบเขตของการวิจัย

1. การสร้างแบบทดสอบสมรรถภาพทั่วไปทางภาษาอังกฤษครั้งนี้สร้างขึ้นเพื่อวัดความรู้ความสามารถทั่วไปทางภาษาอังกฤษเฉพาะการวัดทักษะการฟัง การอ่าน การเขียน เท่านั้น

2. การวิจัยครั้งนี้ผู้วิจัยไม่คำนึงถึงความแตกต่างด้านเพศ อายุ เชื้อชาติ ปรากฏา ตลอดจนภูมิหลังของผู้เข้าทดสอบ

3. กลุ่มผลวิจัยของการวิจัยครั้งนี้มี 2 กลุ่ม คือ

3.1 ผลวิจัยกลุ่มที่ 1 คือ นิสิตจุฬาลงกรณ์มหาวิทยาลัยตั้งแต่ระดับชั้นปีที่ 1 ถึงบัณฑิตศึกษา รวมทั้งบุคคลภายนอกที่มีผลสอบตามประกาศของสถาบันภาษาและเข้าสอบในภาคปลายปีการศึกษา 2533 เมื่อวันที่ 16-31 มกราคม 2534 จำนวน 481 คน

3.2 ผลวิจัยกลุ่มที่ 2 คือ บุคคลภายนอกที่มีพื้นความรู้ระดับปริญญาตรีและปริญญาโท ซึ่งเข้าทดสอบในภาคปลายปีการศึกษา 2534 จำนวน 87 คน

ข้อตกลงเบื้องต้น

คณะผู้วิจัยได้ใช้ตารางกำหนดลักษณะเฉพาะของข้อสอบ (Test Table of Specifications) ที่คณะผู้วิจัยในโครงการระยะปีที่ 1 ได้จัดทำขึ้น ซึ่งเป็นการจัดทำของคณะผู้เชี่ยวชาญทางด้านการออกข้อสอบภาษาอังกฤษ มีความเชื่อถือได้เนื่องจากตารางนี้ได้ใช้เพื่อสร้างข้อทดสอบสมรรถภาพทั่วไปทางภาษาอังกฤษในระะยะปีที่ 1 แล้วมีความเที่ยงสูง ($KR_{20} = 0.90$) มีความตรงเชิงประจักษ์ (Face Validity) และความตรงเชิงเนื้อหา (Content Validity) สูงโดยพิจารณาจากเนื้อหาและความสอดคล้องของข้อทดสอบกับวัตถุประสงค์ของการวัดซึ่งระบุไว้ในตารางดังกล่าว

คำจำกัดความที่ใช้ในการวิจัย

ข้อทดสอบสมิทริภาพทั่วไป ทางภาษาอังกฤษ	หมายถึง	ความสามารถทั่วไปทางภาษาซึ่งไม่ขึ้นต่อหลักสูตรใดหลักสูตรหนึ่งอาจเรียกได้ว่าเป็นสมรรถวิสัยในการใช้ภาษา (Language Competence) ซึ่งมีความถนัดทางภาษา (Language Aptitude) และสัมฤทธิ์ผลทางการเรียนภาษา (Language Achievement) เป็นรากฐาน
ทักษะการฟัง	หมายถึง	ทักษะรับสารซึ่งสื่อด้วยเสียง ได้แก่ความสามารถในการฟังเพื่อความเข้าใจข้อความ ข่าว บทสนทนาและการบรรยาย
ทักษะการอ่าน	หมายถึง	ทักษะรับสารซึ่งสื่อด้วยตัวหนังสือ หรือความสามารถในการอ่านเพื่อความเข้าใจข้อความที่ใช้โวหารแบบต่าง ๆ ได้แก่การเล่าเรื่อง การแสดงความคิดเห็นเชิงวิพากษ์วิจารณ์ การบรรยายสถานที่ การโฆษณาจูงใจ และการอธิบาย
ทักษะการเขียน	หมายถึง	ทักษะสื่อสารด้วยตัวหนังสือ หรือความสามารถใช้ภาษาได้อย่างถูกต้องทั้งทางศัพท์และไวยากรณ์ ความสามารถเขียนข้อความในระดับอนุเจตตามหัวข้อและกรอบเนื้อความที่กำหนดและกรอบเนื้อหา และการเขียนแบบเสรีเขียนความเรียงตามหัวข้อเรื่องที่กำหนดให้

ประโยชน์ที่คาดว่าจะได้รับ

1. สถาบันภาษา จุฬาลงกรณ์มหาวิทยาลัยได้แบบทดสอบสมิทริภาพทั่วไปทางภาษาอังกฤษที่มีมาตรฐาน และเป็นที่ยอมรับทั่วไป มีจำนวน 4 ชุด และจำนวนข้อเพิ่มขึ้น
2. สถาบันภาษาได้มีคลังข้อสอบทางภาษา ซึ่งใช้ระบบคอมพิวเตอร์ในการจัดการเพื่อนำมาใช้และพัฒนาข้อสอบต่อไป

บทที่ 2

เอกสารทางวิชาการที่เกี่ยวข้อง

บทนี้แบ่งเป็น 4 ตอนดังนี้

- ตอนที่ 1: การวิจัยทางการทดสอบภาษา
- ตอนที่ 2: การพัฒนาแบบทดสอบวัดสมรรถภาพทางภาษาแบบอิงเกณฑ์
- ตอนที่ 3: การวิเคราะห์ข้อทดสอบแบบอิงกลุ่ม - อิงเกณฑ์และการวิเคราะห์ตามทฤษฎีความสามารถแฝงแบบ Rasch Model
- ตอนที่ 4: การสร้างคลังข้อทดสอบ

ตอนที่ 1: การวิจัยทางการทดสอบภาษา (Language Testing Research)

การวิจัยทางการทดสอบภาษามีความหมายโดย 3 นัย ดังนี้

นัยที่หนึ่ง การวิจัยทางการทดสอบภาษา หมายถึงการพัฒนาข้อทดสอบทางภาษาขึ้นมาให้เป็นมาตรฐาน โดยนักวิจัยใช้ขั้นตอนดำเนินงานดังนี้

- (1) ศึกษาเกี่ยวกับคุณลักษณะหรือภาวะสันนิษฐาน (Construct) ที่จะวัด และกำหนดว่าประกอบด้วยอะไร
- (2) สร้างเครื่องมือหรือข้อทดสอบที่มีความเที่ยง (Reliable) และความตรง (Valid) ขึ้น เพื่อวัดคุณลักษณะที่ประสงค์
- (3) นำข้อทดสอบไปทดลองใช้และวิเคราะห์คุณภาพของรายข้อและของทั้งฉบับ
- (4) ปรับปรุงข้อทดสอบโดยใช้สถิติวิเคราะห์ที่ได้จากการทดลองข้อทดสอบ
- (5) นำข้อทดสอบไปทดลองใช้อีกครั้งหนึ่ง แล้วตรวจสอบค่าความยากง่าย ค่าความเที่ยง และค่าความตรง

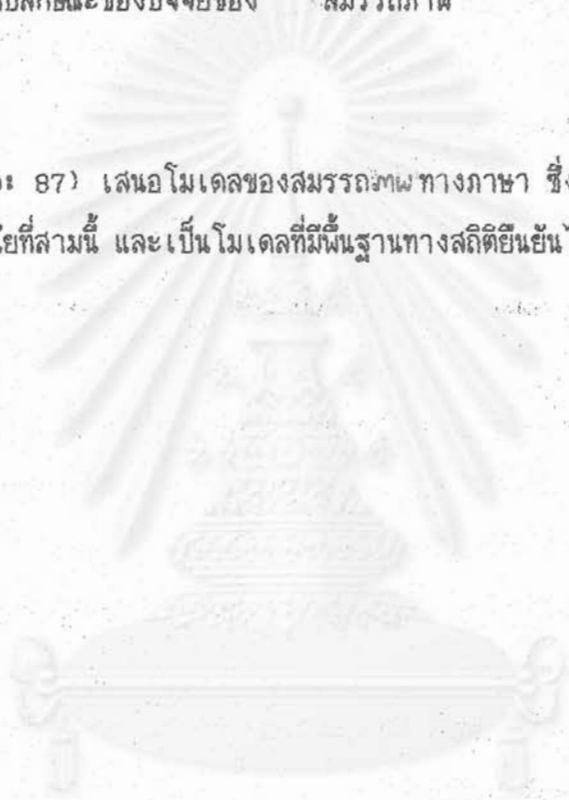
นัยที่สอง

การศึกษาเกี่ยวกับระดับความสามารถหรือสัมฤทธิ์ผลทางการใช้ภาษา โดยนำเครื่องมือที่เป็นแบบทดสอบทางภาษาไปใช้ในการเก็บข้อมูล

นัยที่สาม

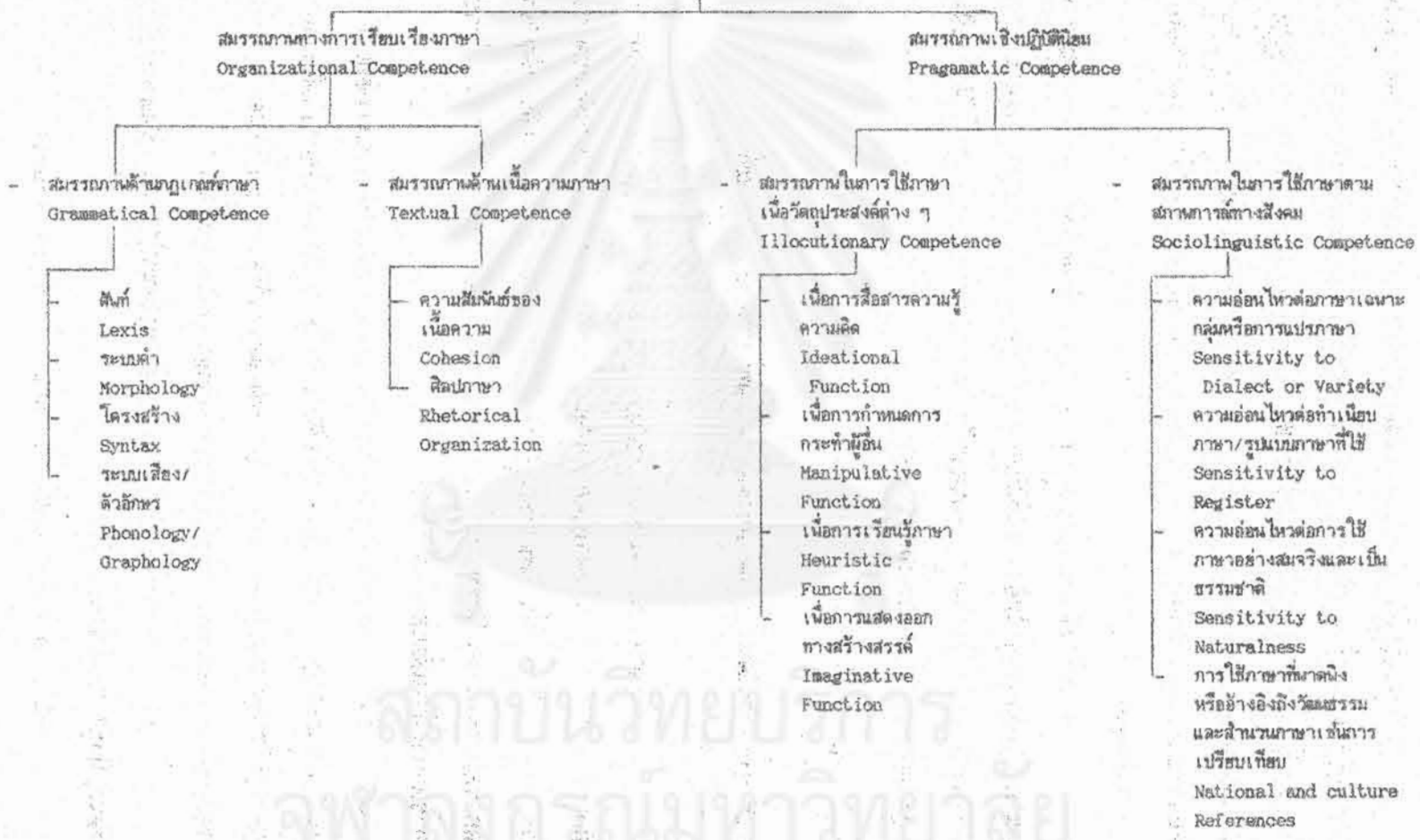
การศึกษาเกี่ยวกับองค์ประกอบหรือปัจจัย (Factor) ของทักษะ ความสามารถทางการใช้ภาษาว่ามีลักษณะเช่นใด เช่น งานวิจัยของ Oller และ Hinofotis (1980) Bachman และ Palmer (1981) [ทั้ง 2 เรื่องอยู่ใน Oller, J.W. Jr. (ed.)] ของ J.B. Carroll (1983) และงานวิจัยของ อัจฉรา วงศ์โสธร (1990 : 209-214) ซึ่งให้ข้อมูลเกี่ยวกับลักษณะของปัจจัยของ สมรรถภาพ

Bachman (1990 : 87) เสนอโมเดลของสมรรถภาพทางภาษา ซึ่งพัฒนาจากงานวิจัยทางการทดสอบภาษานัยที่สามนี้ และเป็นโมเดลที่มีพื้นฐานทางสถิติยืนยันไม่ใช่เป็นแต่เพียงโมเดลทางทฤษฎีเท่านั้น



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

สมรรถภาพทางภาษา
Language Competence



จากแผนภูมิสรุปได้ว่า โมเดลของสมรรถภาพทางการใช้ภาษาของ Bachman ซึ่งมีงานวิจัยยืนยันมีความคล้ายคลึงกับโมเดลทางทฤษฎีซึ่ง De Saussure (1916) ได้เสนอไว้ว่า ประกอบด้วยความสามารถใช้กฎเกณฑ์ภาษา (Langue) และความสามารถใช้ภาษาเพื่อการสื่อความหมาย (Parole) โดยที่ Langue ตามโมเดลของ De Saussure คือ Organizational Competence และ Parole คือ Pragmatic Competence ตามโมเดลของ Bachman

การสร้างข้อทดสอบวัดสมรรถภาพทางภาษาที่สมบูรณ์ต้องครอบคลุมสมรรถภาพทั้ง 2 ด้าน ดังนั้นการวัดจึงควรมุ่งไปยังการใช้ทักษะภาษา ได้แก่ การฟัง การพูด การอ่าน การเขียนตามสถานการณ์ในชีวิตจริง ซึ่งครอบคลุมการวัดเนื้อหาของภาษา ได้แก่ การใช้ศัพท์ เสียง โครงสร้าง ความสัมพันธ์ของเนื้อความและศิลปภาษา

ตอนที่ 2: การพัฒนาแบบทดสอบวัดสมรรถภาพทางภาษาแบบอิงเกณฑ์

แบบทดสอบอิงเกณฑ์มักจะใช้กับการสอบสัมฤทธิ์ผล (Achievement Test) ซึ่งอิงหลักสูตร เพราะเป็นการวัดตามวัตถุประสงค์ที่กำหนดไว้ชัดเจนและมีการกำหนดเกณฑ์การผ่านไว้ ส่วนการทดสอบอิงกลุ่มนั้น ให้นิยามโดยคำนี้ถึงตัวแปรเกี่ยวกับผู้เข้าสอบเป็นสำคัญ และมุ่งเปรียบเทียบผลการสอบระหว่างผู้เรียนมากกว่าเป็นการทดสอบความวัตถุประสงค์ และเกณฑ์การให้คะแนนก็นิยามมาจากค่าเฉลี่ยและค่าการกระจายของคะแนน

L.F. Bachman (1986 : 264) ได้นำเสนอแนวคิดในการพัฒนาแบบทดสอบวัดสมรรถภาพทางภาษาแบบอิงเกณฑ์ว่าเป็นการวัดที่มีประสิทธิภาพ มีอำนาจในการวัดดีกว่าการวัดแบบอิงกลุ่ม เนื่องจากมีการกำหนดวัตถุประสงค์ ขอบเขตเนื้อหาไว้ชัดเจน ข้อกำหนด 2 ประการที่จำเป็นสำหรับการพัฒนาแบบทดสอบวัดสมรรถภาพทางภาษาแบบอิงเกณฑ์ คือ

- (1) ข้อทดสอบที่วัดวัตถุประสงค์ตามเนื้อหาซึ่งลุ่มมา เป็นตัวแทนที่แท้จริงของเนื้อหาทั้งหมด ที่อยู่ในขอบเขตของเนื้อหา
- (2) การกำหนดจุดตัดของคะแนน (Cut-off Score)

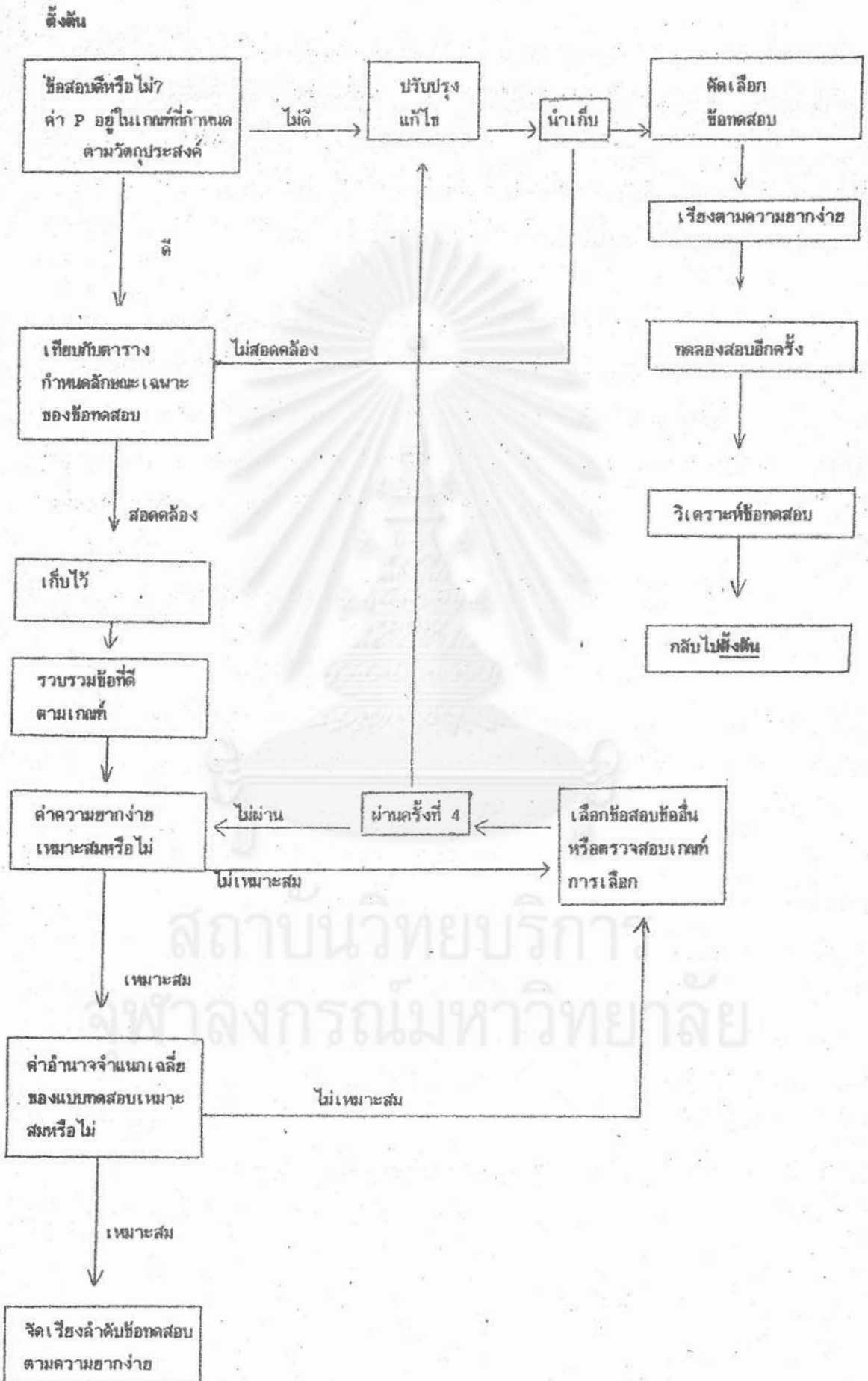
Geoff Brindley (1990: 6-7) ได้กล่าวถึงการพัฒนาแบบทดสอบวัดสมรรถภาพทางภาษาแบบอิงเกณฑ์ว่ามุ่งวัดกิจกรรมที่เป็นตัวแทนของการใช้ภาษาที่แท้จริงในชีวิตประจำวัน และการประเมินระดับความสามารถทางการใช้ภาษานั้นจะอิงเกณฑ์ซึ่งเป็นคำอธิบาย (Descriptor) ระดับความสามารถ (Band Scales) ซึ่งมีใช้ต่าง ๆ กันมากมาย และเกณฑ์หนึ่งที่นิยมใช้กันคือเกณฑ์ของสภายุโรป (The Council of Europe)

Bachman (1989: 256) ได้กล่าวถึงข้อจำกัดของการอธิบายระดับความสามารถและความยากของการกำหนดระดับจากคะแนนที่ได้ Bachman ได้เสนอว่าวิธีการเดียวที่จะใช้สำหรับการวัดสมรรถภาพทางการสื่อสารทางภาษา (Communicative Language Competence) ก็โดยการระบุองค์ประกอบของสมรรถวิสัยหรือสมรรถภาพทางภาษาและกำหนดระดับความสามารถที่ไม่ขึ้นต่อบริบทโดยบริบทหนึ่ง แต่ขึ้นกับความสามารถหรือพฤติกรรมทางภาษาของแต่ละบุคคลอย่างแท้จริง

Bachman (1985 : 3) ได้เสนอขั้นตอนการคัดเลือกข้อทดสอบไว้ตามแผนภูมิต่อไปนี้

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

แผนภูมิที่ 2: กระบวนการคัดเลือกข้อทดสอบ (แบบอิงเกณฑ์)



จากแผนภูมิจะเห็นได้ว่าการพัฒนาข้อทดสอบนั้นใช้การวิเคราะห์และปรับปรุงเป็นพื้นฐาน และค่าสถิติที่ใช้ได้แก่ค่าความยากง่ายตามแบบทดสอบอิงเกณฑ์นั้น จะกำหนดตามเกณฑ์ความสามารถ และค่าอำนาจจำแนกจะไม่เน้นเป็นรายข้อ แต่พิจารณาจากค่าเฉลี่ยของฉบับ การกำหนดค่าความยากง่ายจากเกณฑ์ .25-.75 หรือ .20-.80 นั้น เป็นจากเกณฑ์ความสามารถขั้นสูงที่ทำ ข้อสอบยากได้ค่าความยากคือ .20 หรือ .25 ไปตามลำดับถึงความสามารถขั้นต่ำ ค่าความยากง่าย จำแนกนั้นอย่างต่ำ ควรเป็น .25 หรือ .30

ตอนที่ 3: การวิเคราะห์ข้อทดสอบแบบอิงกลุ่ม - อิงเกณฑ์และการวิเคราะห์ตามทฤษฎี ความสามารถแฝงแบบ Rasch Model

การวิเคราะห์ข้อทดสอบแบบอิงกลุ่มและแบบอิงเกณฑ์

ศาสตร์ของการทดสอบทางภาษาในยุคปัจจุบันเป็นยุคของภาษาศาสตร์เชิงจิตวิทยาและสังคมวิทยาและเชิงการสื่อสาร ซึ่ง Morrow (1981) เรียกว่าเป็นยุค The Promised Land การทดสอบหันมาเน้นความสมจริงตามสถานการณ์ที่ใช้ภาษาเพื่อการสื่อสารจริง ๆ ในชีวิตประจำวัน ทิศทางของการวิจัยทางการทดสอบภาษาก็ได้หันเหไปจากการเน้นตัวค่าสถิติวิเคราะห์แบบอิงกลุ่ม เพราะสถิติวิเคราะห์ดังกล่าวยึดมั่นอยู่กับกลุ่มตัวอย่าง (Sample Group) ค่าสถิติจึงเปลี่ยนแปลงไปตามกลุ่มที่ทดสอบและการกำหนดมาตรฐานของข้อทดสอบคือค่าความยากง่ายและค่าอำนาจจำแนกแบบอิงกลุ่มตลอดจนการเทียบมาตรฐานของแบบทดสอบกับเกณฑ์ภายนอก ได้แก่ การหาความตรงร่วมสมัย (Concurrent Validity) โดยหาค่าสหสัมพันธ์ระหว่างคะแนนที่ผู้สอบได้จากแบบทดสอบที่สร้างขึ้นกับคะแนนที่ได้จากแบบทดสอบมาตรฐาน เพื่อพิสูจน์ความเป็นมาตรฐานของแบบทดสอบที่สร้างขึ้น ถูกแทนที่โดยการวิเคราะห์แบบอิงเกณฑ์ และการหาค่าความตรงเชิงเนื้อหา (Content Validity) และเชิงภาวะสันนิษฐาน (Construct Validity) แทน

Pilliner (1983: 3-4) กล่าวว่า การทดสอบซึ่งมีหลายวัตถุประสงค์ครอบคลุมเนื้อหา มาก แต่อยู่ในขอบเขตที่ระบุไว้ชัดเจน โดยใช้กรอบอ้างอิง (Frame of Reference) ที่แน่นอนในการกำหนดเนื้อหาของข้อทดสอบควรเป็นข้อทดสอบอิงเกณฑ์ และใช้ผลการตัดสินคะแนนในแบบต่อเนื่อง (Continuum Model) คือมีระดับได้-ตก มากกว่า 2 ระดับ แทนการใช้วิธี

การตัดสินคะแนนที่ใช้เกณฑ์ได้-ตก ที่กำหนดไว้ตายตัวของแบบ State Model ซึ่งแบบหลังนี้ เหมาะที่จะใช้ในการสอบวัดความก้าวหน้าระหว่างเรียน (Formative Test) ที่มีวัตถุประสงค์ และเนื้อหาบ่อย นอกจากนี้การทดสอบการสื่อความหมายทางภาษา (Productive) ซึ่งรวมถึง การรับสาร (Receptive) ด้วย จำเป็นต้องครอบคลุมภาวะสันนิษฐาน (Construct) ที่ หลากหลายเพื่อวัดการใช้ภาษาที่สมจริงตามสถานแวดล้อมทางสังคม เนื้อหาและทักษะที่วัดโดยแบบ ทดสอบจึงหลากหลาย (Heterogeneous) มากกว่าเป็นอันหนึ่งอันเดียว (Homogeneous) ดังนั้น การวิเคราะห์ข้อสอบจึงไม่สามารถใช้วิธีการแบบประเพณีนิยม (Classical Model) ซึ่งการ วิเคราะห์ตั้งอยู่บนรากฐานของความเป็นอันหนึ่งอันเดียวกันอย่าง เป็นกลุ่มก้อนหรือมีความเป็น เอกพันธ์ (Homogeneity)

การวิเคราะห์แบบทดสอบทั้งฉบับโดยความเที่ยง (Reliability) ของแบบทดสอบ จึงไม่ควรใช้วิธีการหาค่าความมั่นคงภายใน (Internal Consistency) เช่น Kuder Richardson Formula 20 และ 21 หรือ Split-half Pilliner (1983: 14) ได้ แนะนำวิธีการหาค่าความเที่ยงของแบบทดสอบอิง เกณฑ์ว่าควรใช้วิธีการหาค่าสัมประสิทธิ์สหสัมพันธ์ แบบ Phi (Phi Coefficient) ค่าสัมประสิทธิ์นี้บางครั้งเรียกว่าค่าสัมประสิทธิ์การอ้างอิง สรุปล้ำสำหรับความผิดพลาดที่สัมบูรณ์ (Generalizability for absolute error (Brown, 1990: 16-17) และสามารถใช้อธิบายปริเณท (Domain) หรือขอบเขตเนื้อหาทักษะของ การทดสอบ เมื่อข้อสอบที่วิเคราะห์ได้ลุ่มอย่างเป็นตัวแทนของปริเณท สูตรการวิเคราะห์ Phi นี้พัฒนามาจากสูตรของ Brennan (1983, 1984) และ Brown ได้เสนอไว้ดังนี้

$$\phi = \frac{\frac{nS_p^2}{n-1} [KR_{20}]}{\frac{nS_p^2}{n-1} [KR_{20}] + \frac{\bar{X}_p(1-\bar{X}_p) - s_p^2}{k-1}}$$

n = จำนวนผู้เข้าสอบ

k = จำนวนข้อกระทง

X_p = ค่ามัชฌิม เลขคณิตของคะแนนสัดส่วน (Proportion Score)

S_p = ค่าเบี่ยงเบนมาตรฐานของคะแนนสัดส่วน (ได้จากสูตรที่ใช้ N คิด ไม่ใช่ N-1)

KR_{20} = ค่าความเที่ยง Kuder Richardson Formula 20

$$\bar{X}_p = \frac{\bar{X}}{k}$$

$$S_p = \frac{SD}{k}$$

สูตรดังกล่าวใช้วิเคราะห์แบบทดสอบอิงเกณฑ์โดยใช้กับคะแนนจากการสอบเพียงครั้งเดียว

การหาค่าความตรง (Validity) ของแบบทดสอบใน Pilliner (1983, 14-15) แนะนำวิธีการวิเคราะห์ค่าความตรงเชิงเนื้อหา (Content Validity) โดยให้ผู้เชี่ยวชาญด้านเนื้อหาที่วัดมากกว่า 1 คน ประเมินความสอดคล้องของข้อทดสอบกับเนื้อหา/ทักษะ นอกจากนี้ Cronbach (1971 อ้างถึงใน Pilliner (1983: 15) กล่าวถึงขั้นตอนการสร้างแบบทดสอบให้มีความเที่ยงและความตรงว่า เริ่มตั้งแต่การที่คณะผู้สร้างข้อทดสอบเขียนข้อทดสอบจากตารางวิเคราะห์มีความคู่ขนานกันจริงทางสถิติ โดยให้ผู้สอบกลุ่มเดียวกันทดลองสอบทั้ง 2 ฉบับ คะแนนที่ได้ถ้าอยู่ในพิสัยความคลาดเคลื่อนเดียวกัน ถือว่าเป็นคะแนนที่เท่ากัน ความเที่ยงที่หาได้โดยวิธีนี้เป็นค่าความเที่ยงแบบ Test-retest Equivalent Form ส่วนค่าความตรงเป็นค่าความตรงเชิงเนื้อหาเพราะข้อทดสอบพัฒนาจากตารางวิเคราะห์เนื้อหาเดียวกัน

การวิเคราะห์ข้อทดสอบเป็นรายข้อ (Test-item Analysis) สำหรับแบบทดสอบอิงเกณฑ์นั้นจะพิจารณาให้ความสำคัญแก่ค่าความยากง่าย (Difficulty หรือ Facility Index) เท่านั้น ส่วนค่าอำนาจจำแนกไม่มีความสำคัญ ซึ่งต่างกับแบบทดสอบอิงกลุ่มที่ให้ความสำคัญต่อค่าความยากง่ายเท่าเทียมกับค่าอำนาจจำแนกทั้งนี้เนื่องจากว่าข้อทดสอบอิงเกณฑ์ไม่ได้สร้างขึ้นมาเพื่อจำแนกผู้สอบ (Pilliner, 1983: 12)

การวิเคราะห์ข้อสอบตามทฤษฎีความสามารถแฝงแบบ Rasch Model

การวิเคราะห์ข้อทดสอบโดยใช้ Classical Test Theory/CTS นั้นอาศัยทฤษฎีที่ไม่มีข้อตั้งตั้งเบื้องต้น (Assumption) เกี่ยวกับระดับความสามารถของแต่ละบุคคลกับผลของการทดสอบอีกทั้งข้อมูลที่ใช้ในการทำนายความสามารถของบุคคลในการทำสอบคือ β นั้น เป็นอัตราส่วนที่คนทั้งกลุ่มทำข้อทดสอบถูกและใช้พาดพิงถึงความสามารถของแต่ละบุคคลในการตอบข้อสอบข้อใดข้อหนึ่งได้โดยเฉลี่ยเท่านั้น จึงจำเป็นต้องมีการวิเคราะห์ที่สามารถชี้บ่งถึงความสามารถของบุคคลว่ามีความสามารถในระดับใดจึงจะทำข้อสอบได้ ความสัมพันธ์ของระดับความยากง่ายของข้อสอบกับระดับความสามารถของบุคคลในการทำข้อสอบได้ นำไปสู่การวิเคราะห์ข้อสอบตามทฤษฎีความสามารถแฝง (Latent Trait Model) หรือทฤษฎีการสนองตอบต่อข้อกระทง (Item Response Theory/IRT)

IRT มีรากฐานบนข้อตั้งตั้งเบื้องต้นที่มั่นคงและเฉพาะเจาะจงกว่า CTT เพราะสามารถใช้ทำนายเกี่ยวกับระดับความสามารถของบุคคลในการทำสอบที่มีความยากง่ายในระดับใดระดับหนึ่งได้ IRT มีข้อตั้งตั้งเบื้องต้นว่า (1) คุณลักษณะที่วัดเป็นมิติเดียว (Unidimensionality) (2) มาตรการที่ใช้วัดเป็นมิติเดียวและ (3) ความเป็นอิสระในตำแหน่งที่อยู่ของข้อกระทง (Local independence)

การคัดเลือกข้อกระทงแบบ IRT นั้น จะตัดข้อที่ไม่อยู่ในโค้งลักษณะข้อกระทง (Item Characteristic Curve/ICC) ICC เป็นรากฐานของ IRT (Bachman, 1990: 205-206) เพราะกำหนดข้อตั้งตั้งเบื้องต้นเกี่ยวกับความสัมพันธ์ของความเป็นไปได้ของความสามารถของบุคคลกับความเป็นไปได้ที่จะทำข้อกระทงได้ เกณฑ์ของการกำหนด ICC เต็มรูปแบบประกอบด้วย

- (1) ค่าอำนาจจำแนกของข้อทดสอบ (a)
- (2) ค่าความยากง่ายของข้อทดสอบ (b)
- (3) การเดาถูกของข้อทดสอบ (c)

ในการวิเคราะห์ข้อสอบโมเดล IRT นี้ อาจทำได้โดยใช้เกณฑ์ a และ b คือ Two parameter ซึ่งมีข้อตั้งตั้งเบื้องต้นว่าคนที่มีความสามารถต่ำจะไม่สามารถตอบข้อสอบถูกและค่าการเดา = 0 ในการทดสอบภาษาโมเดลที่นิยมใช้กันมากเป็นแบบพารามิเตอร์เดียวหรือ One parameter ของ Rasch ซึ่งกำหนดให้ค่าอำนาจจำแนกของข้อกระทงทุกข้อเท่ากัน และไม่

มีค่าการเดา มีการใช้ Rasch อย่างแพร่หลายและอย่างมีประสิทธิภาพในการวิเคราะห์ข้อกระทง เช่นกล่าวถึงใน Davidson และ Henning (1975) Griffin (1985) Henning et al (1985) Madsen และ Larson (1986) Pollitt และ Hutchinson (1987) Larson (1987) Madsen (1987) Adams et al (1987) Henning (1987) (อ้างถึงใน Bachman, 1990: 206) และ McNamara (1990) ได้กล่าวถึงการใช้ Rasch วิเคราะห์ข้อทดสอบภาษาเพื่อการสื่อสาร นอกจากนี้ Laurier (1990) ได้กล่าวถึงการใช้ IRT โดยระบบคอมพิวเตอร์ขนาดเล็ก (Micro-computer) ในการจัดทำข้อทดสอบจัดระดับภาษาฝรั่งเศส

ตอนที่ 4: การสร้างคลังข้อทดสอบ (Test-item Bank)

ข้อทดสอบที่ผ่านการวิเคราะห์ความเป็นมาตรฐานแล้ว ควรนำเข้ามาเก็บไว้ในคลังข้อทดสอบเพื่อการนำมาใช้ต่อไป การสร้างคลังข้อสอบโดยใช้คอมพิวเตอร์ (Computerized item banking) ช่วยในการคัดเลือกข้อสอบใช้แต่ละครั้งและในการจัดพิมพ์ ผลิตข้อสอบ ตลอดจนการพัฒนาข้อสอบขึ้นใหม่อย่างต่อเนื่อง โดยคอมพิวเตอร์แบบตั้งโต๊ะที่ใช้ในการจัดระบบตั้งแต่การวิเคราะห์ จัดเก็บ คัดเลือกในการนำออกมาใช้ในแต่ละครั้ง และจัดพิมพ์ ข้อทดสอบที่อยู่ในธนาคารต้องมีมากกว่าจำนวนที่จำเป็นที่จะต้องใช้ในแต่ละครั้งอย่างน้อยมากกว่า 1 เท่าขึ้นไป และควรมีปริมาณตามอัตราส่วนของความสำคัญ และความจำเป็นของเนื้อหา ทักษะทางภาษา

คลังข้อทดสอบที่สร้างขึ้นตามการใช้ภาษาในชีวิตจริงเหมาะสมกับการจัดเก็บข้อทดสอบวัดสมรรถภาพทั่วไปทางภาษาแบบอิงเกณฑ์ โดยการใช้ตารางวิเคราะห์ภาษาที่ใช้ในชีวิตประจำวัน โดยนักวิชาการทางภาษาและนักวัดผลทางภาษา ตารางต่อไปนี้เสนอตัวอย่างของตารางวิเคราะห์การใช้ภาษาในชีวิตประจำวัน

ศูนย์วิจัยและบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 1 ตัวอย่างตารางวิเคราะห์การใช้ภาษาในชีวิตประจำวัน

ทักษะ	เนื้อหา	หน้าที่/วัตถุประสงค์	ความสำคัญ
อ่าน	เอกสาร	จับใจความ	25%
	วารสาร	สรุปความ	
	หนังสือพิมพ์		
	บทความ		
	เรื่องสั้น		
	นวนิยาย		
ฟัง	บทสนทนา	เพื่อการติดต่อทางสังคม	25%
	เพลง	เพื่อความรู้ทันโลก	
	ข่าว	เพื่อความบันเทิง	
อ่าน-เขียน	จดหมาย	ถาม/ตอบ เพื่อให้ได้ข้อมูล รายละเอียด	20%
	บทความ	สรุปประเด็น	
เขียน	เรื่องสั้น	วิจารณ์	
	บันทึก	อ้างอิง	
	รายงาน	บันทึกความจำ ให้ข้อมูล	
ฟัง-พูด	เรื่องส่วนตัว	ถาม/ตอบ เพื่อให้ได้	30%
	กิจกรรม	ข้อมูล	
	โต้วาที	เสนอข้อคิดเห็น/โต้แย้ง	
พูด	บรรยาย/อธิบาย	ให้ข้อมูลความรู้	

คลังข้อทดสอบมีประโยชน์ทั้งทางด้านการบริหารและด้านวิชาการ

ประโยชน์ทางด้านการบริหาร คือช่วยไม่ให้เกิดการซ้ำซ้อนกันในการทำงาน การสิ้นเปลืองเวลาและทุนทรัพย์ที่ต้องเสียไปในการสร้าง พัฒนาและใช้ข้อทดสอบที่สร้างขึ้นเพื่อใช้ตามวัตถุประสงค์เดียวกันเพราะคลังข้อทดสอบเป็นแหล่งรวบรวมข้อสอบภาษาไว้ตามพารามิเตอร์

ประโยชน์ทางด้านวิชาการ มีดังนี้

- (1) ทำให้รักษามาตรฐานของการทดสอบและประเมินผลทางภาษา
- (2) ทำให้มีข้อทดสอบที่ได้มาตรฐานและสร้างตามพารามิเตอร์ที่พึงประสงค์ได้เป็นจำนวนมากมาย
- (3) ทำให้การทดสอบเป็นไปตามเอกัตภาพ (Individualized Testing) วัดความสามารถของบุคคลตามเกณฑ์และจัดสอบได้เป็นรายบุคคล
- (4) ทำให้การทดสอบไม่จำกัดระดับโดยมีข้อทดสอบที่มีระดับความยากง่ายต่าง ๆ กัน
- (5) ทำให้มีการพัฒนาข้อทดสอบให้เหมาะสมกับความสามารถของผู้เรียนแต่ละคน (Tailored testing)
- (6) ทำให้เพิ่มพูนศักยภาพในการเรียนรู้ภาษาและสมรรถวิสัยในการใช้ภาษาเพื่อการสื่อสารของผู้เรียน โดยการวัดความสามารถของตนอย่างสม่ำเสมอด้วยข้อสอบที่เป็นมาตรฐาน
- (7) ทำให้ผู้สอบสามารถฟื้นฟูปรับปรุงทักษะภาษาของตนได้อย่างสม่ำเสมอ โดยการวัดผลที่ต่อเนื่องแบบเอกัตบุคคล

การสร้างคลังข้อทดสอบสำหรับการวัดความสามารถในการใช้ทักษะรับสารทางภาษา ได้แก่ การฟัง อ่าน และทักษะสื่อสารทางภาษา ได้แก่ พูด (จัดเก็บตัวคำถาม บทบาทที่กำหนด สถานการณ์การพูด) และการเขียน (จัดเก็บ คำสั่ง หัวข้อการเขียน) มีประโยชน์มากมายทั้งทางด้านการบริหารและวิชาการ)

บทที่ 3.

วิธีดำเนินการวิจัย

ในการวิจัยครั้งนี้ ผู้วิจัยมุ่งสร้างพร้อมทั้งวิเคราะห์คุณภาพของแบบทดสอบสมิทริภาพทั่วไปทางภาษาอังกฤษ แบบทดสอบสมิทริภาพทั่วไปทางภาษาอังกฤษที่สร้างขึ้นนี้ทำการวิเคราะห์คัดเลือกข้อกระทงที่มีคุณภาพดี จำนวน 4 ฉบับ รวม 480 ข้อ เพื่อทำการวิเคราะห์หาคุณภาพของแบบทดสอบ แล้วนำมาจัดเก็บในคลังข้อทดสอบ เพื่อนำมาใช้และพัฒนาข้อสอบต่อไป

ประชากรและพลวิจัย

ประชากรเชิงนิยามที่ใช้ในการวิจัยครั้งนี้ คือ นิสิตชั้นปีที่ 1 ถึง นิสิตบัณฑิตศึกษา และบุคคลภายนอกทั่วไปที่มีพื้นฐานความรู้ตั้งแต่ระดับปริญญาตรีขึ้นไป ที่มีความสนใจและสมัครสอบเพื่อวัดระดับสมิทริภาพทั่วไปทางภาษาอังกฤษของตนเอง ซึ่งมีผู้มีสิทธิสมัครสอบจำนวน 2,334 คน

พลวิจัย (Subject) ที่ใช้ในการวิจัยครั้งนี้ คือผู้ที่เข้ารับการทดสอบแบบทดสอบสมิทริภาพทั่วไปทางภาษาอังกฤษเมื่อระหว่างวันที่ 16-21 มกราคม 2534 จำนวน 481 คน คิดเป็นร้อยละ 20.57 ของผู้มีสิทธิสมัครสอบทั้งหมด และบุคคลภายนอกจำนวน 87 คน เข้ารับการทดสอบแบบทดสอบสมิทริภาพทั่วไปทางภาษาอังกฤษฉบับปรับปรุงในเดือนมกราคม 2535

เนื่องจากการสร้างแบบทดสอบสมิทริภาพทั่วไปทางภาษาอังกฤษครั้งนี้คณะผู้วิจัยและผู้สร้างแบบทดสอบได้จัดสร้างแบบทดสอบเป็น 4 ฉบับ ผู้วิจัยจึงจัดแบ่งพลวิจัยออกเป็น 4 กลุ่มตามกลุ่มคณะต่าง ๆ ตามผลการวิจัยตามโครงการวิจัยและพัฒนาการสร้างและพัฒนาคุณภาพของแบบทดสอบสมิทริภาพ ปีที่ 1 (สุพันธ์์ สุขมลลันต์ และคณะ 2533) โดยจัดกลุ่มคณะต่าง ๆ เป็น 4 กลุ่ม ที่ค่าเฉลี่ยของผลการทดสอบวัดสมิทริภาพภาษา อังกฤษครั้งแรกไม่แตกต่างกันอย่างมีนัยสำคัญเข้าด้วยกันเป็น 4 กลุ่ม ($F = .111$ $df = 1,466$)

การสร้างแบบทดสอบ

สถาบันภาษา จุฬาลงกรณ์มหาวิทยาลัย ได้แต่งตั้งคณะกรรมการผู้เชี่ยวชาญทางด้านการออกข้อสอบและวัดผลทางภาษาอังกฤษขึ้นมาคณะหนึ่ง และคณะผู้ออกข้อสอบได้ดำเนินการสร้างแบบทดสอบดังนี้

1. ศึกษาตารางกำหนดลักษณะเฉพาะของข้อสอบ (Test table of specifications) ซึ่งได้ใช้ในการสร้างข้อสอบสมรรถภาพทั่วไปทางภาษาอังกฤษและนำมาวิจัยในปีที่ 1 ของโครงการวิจัยนี้แล้วเห็นว่ามีเหมาะสม

2. คณะผู้สร้างข้อทดสอบสมรรถภาพทั่วไปทางภาษาอังกฤษได้ดำเนินการสร้างตามตารางวิเคราะห์ดังกล่าว ได้จำนวนข้อกระทงดังนี้

2.1 แบบปรนัย (เลือกคำตอบจากตัวเลือก 4 ตัว) แบ่งเป็นข้อกระทงที่สร้างขึ้นตามทักษะต่าง ๆ ดังนี้

2.1.1	ทดสอบทักษะการฟังเข้าใจความ	105	ข้อ
2.1.2	ทดสอบทักษะการอ่านเข้าใจความ	204	ข้อ
2.1.3	ทดสอบการเขียน	75	ข้อ

2.2 แบบอัตนัย

2.2.1	วัดการเขียนระดับอนุเจต	6	ข้อ
2.2.2	การเขียนระดับเรียงความ	3	ข้อ

3. คณะผู้ออกข้อสอบได้นำข้อสอบที่พัฒนาและปรับปรุงแล้วจากการสอบระยะที่ 1 พ.ศ. 2533 คละกับข้อสอบที่สร้างขึ้นตามข้อ 2 แล้วดำเนินการคัดเลือกข้อสอบอีกครั้ง ได้จำนวนข้อกระทงดังนี้

3.1 แบบปรนัย แบ่งเป็นข้อกระทงที่คัดเลือกไว้ดังนี้

3.1.1	ทดสอบทักษะการฟังเข้าใจความ	140	ข้อ
3.1.2	ทดสอบทักษะการอ่านเข้าใจความ	240	ข้อ
3.1.3	ทดสอบการเขียน	100	ข้อ

3.2 แบบอัตนัย

3.2.1	วัดการเขียนระดับอนุเจต	8	ข้อ
3.2.2	การเขียนระดับเรียงความ	4	ข้อ

4. นำข้อทดสอบที่คละกันมาจัดเข้าเป็นฉบับตามตารางกำหนดลักษณะเฉพาะทางของข้อสอบให้มีเนื้อหา และความยากง่ายของเนื้อหาที่ตัดสิน โดยคณะกรรมการพัฒนาข้อทดสอบว่าไม่แตกต่างกัน ได้แบบทดสอบสมรรถภาพทั่วไปทางภาษาอังกฤษ 4 ฉบับ

การบริหารการสอบ

1. การผลิต ข้อคำถามและเขียนเข้าเล่มที่สถาบันภาษา จุฬาลงกรณ์มหาวิทยาลัย
2. การประชาสัมพันธ์การสอบ ผ่านที่ประชุมคณะดี เพื่อพิจารณาเกี่ยวกับเวลาสอบ และขอความร่วมมือในการเผยแพร่ ประชาสัมพันธ์ให้กับนิสิตในคณะ นอกจากนี้ยังใช้การติดประกาศตามคณะและที่ป้ายประกาศของสถาบันภาษาและบัณฑิตวิทยาลัย
3. การสมัครสอบ ผู้สนใจสมัครและกรอกใบสมัครที่สถาบันภาษา
4. การจัดสอบ พลวิชัย เข้าสอบเป็นกลุ่ม ที่เป็นนิสิตเข้าตามกลุ่มคณะ ส่วนที่เป็นบุคคลภายนอกเข้าสอบเป็นกลุ่มตามเวลาที่กำหนด สถานที่สอบอยู่ที่คณะและสถาบันภาษา ห้องสอบเป็นห้องเรียนและห้องประชุม การสอบการฟังให้ผ่านลำโพงขยายเสียง ผู้คุมสอบ เป็นอาจารย์และเจ้าหน้าที่ของคณะและของสถาบันภาษา ซึ่งได้ผ่านการอบรมชี้แจงวิธีการดำเนินการสอบ เป็นผู้นำแบบทดสอบและกระดาษคำตอบไปดำเนินการสอบ
5. การตรวจข้อสอบ ข้อทดสอบแบบปรนัยไว้ตรวจและวิเคราะห์โดยเครื่องคอมพิวเตอร์ของสถาบันบริการคอมพิวเตอร์ของจุฬาลงกรณ์มหาวิทยาลัย ข้อทดสอบแบบอัตนัยตรวจโดยอาจารย์สถาบันภาษา ฉบับละ 2 คน โดยใช้เกณฑ์การให้คะแนนเกณฑ์เดียวกัน

การวิเคราะห์ข้อมูล

1. ค่าสถิติพื้นฐาน
คะแนนต่ำสุด คะแนนสูงสุด คะแนนเฉลี่ย ค่าเบี่ยงเบนมาตรฐานของคะแนนของแบบทดสอบแต่ละฉบับ
2. ค่าสัมประสิทธิ์การกระจาย จากสูตร

$$C.V. = \frac{S.D.}{\bar{X}} \times 100$$

เมื่อ C.V. = สัมประสิทธิ์การกระจาย
 S.D. = ส่วนเบี่ยงเบนมาตรฐาน
 \bar{X} = คะแนนเฉลี่ย

3. ค่าความยากง่ายของข้อกระทง

$$3.1 \quad P = \frac{R}{T}$$

เมื่อ P = ความยากง่ายของข้อทดสอบ

R = จำนวนผู้ตอบถูก

T = จำนวนคนที่ตอบทั้งสิ้น

3.2 ค่าความยากง่ายแบบ Delta (Δ) เป็นค่าที่แปลงมาเป็นเส้นตรงจากค่า P โดยวิธี linear transformation ประโยชน์ของค่า Δ คือการหาค่าเฉลี่ยความยากง่ายของข้อทดสอบหลาย ๆ ข้อ ทั้งนี้เพราะค่า Δ มีลักษณะเป็นคะแนนมาตรฐานอย่างหนึ่งที่มีหน่วยเท่ากันคือ

ค่ามัธยัม เลขคณิต = 13

ค่าความเบี่ยงเบนมาตรฐาน = 4

z = ค่า S.D. ณ จุดค่า P หรือจุดที่บอกอัตราส่วนของผู้ตอบข้อกระทงได้และตอบไม่ได้ (Fan, 1952)

$$\text{สูตรการหาค่า } \Delta = 13 + 4z$$

4. ค่าอำนาจจำแนก ใช้สูตรสหสัมพันธ์แบบ Point biserial Correlation

$$r_{pb} = \frac{(\bar{X}_r - \bar{X}_w)}{s_t} \sqrt{p(1-p)}$$

เมื่อ \bar{X}_r = ค่าเฉลี่ยของคะแนนรวมของผู้ที่ตอบถูก

\bar{X}_w = ค่าเฉลี่ยของคะแนนรวมของผู้ที่ตอบผิด

s_t = ความเบี่ยงเบนมาตรฐานของคะแนนรวมทั้งหมด

p = ค่าความยากง่ายของข้อสอบ

5. ค่าความเที่ยงของข้อสอบ ใช้สูตร Kuder-Richardson 20 และค่าสัมประสิทธิ์สหสัมพันธ์ Phi

5.1 KR_{20}

$$r_{tt} = \left(\frac{k}{k-1} \right) \left(\frac{s_x^2 - \sum p_i q_i}{s_x^2} \right)$$

- เมื่อ r_{cc} = สัมประสิทธิ์ความเที่ยง
 k = จำนวนข้อสอบในแบบทดสอบที่ต้องการหาความเที่ยง
 p = สัดส่วนของคนที่ตอบแต่ละข้อถูก
 q = สัดส่วนของคนที่ตอบแต่ละข้อผิด ($1-p$)
 pq = ผลรวมของผลคูณระหว่างสัดส่วนของคนที่ตอบแต่ละข้อถูกและผิด
 S^2_x = ความแปรปรวนของคะแนนของผู้ถูกทดสอบทั้งหมด

5.2 การหาค่าความเที่ยงแบบ Phi (Brown, 1990 : 16-17)

$$\phi = \frac{\frac{nS_p^2}{n-1} [KR_{20}]}{\frac{nS_p^2}{n-1} [KR_{20}] + \frac{\bar{x}_p(1-\bar{x}_p) - S_p^2}{k-1}}$$

(ดูรายละเอียดในบทที่ 2)

สูตรการหาค่านัยสำคัญทางสถิติของค่า Phi (Anastasi, 1976 : 214)

ระดับนัยสำคัญที่ $p < .05 = \frac{1.96}{\sqrt{N}}$

$p < .01 = \frac{2.58}{\sqrt{N}}$

สูตรนี้ N คือจำนวนคนทั้งหมด ค่า phi ค่าที่ต่ำสุด ที่จะมีนัยสำคัญอยู่ที่ระดับ .05 จะเป็น $1.96/10 = .196$ ข้อสอบใด ๆ ที่มีค่า phi เท่ากับหรือมากกว่าค่า .258 จะมีนัยสำคัญที่ระดับ .01

5.3 ค่าความคลาดเคลื่อนมาตรฐานของการวัด SE_m

$$SE_m = S.D. \sqrt{1 - r_{tt}}$$

- เมื่อ r_{tt} = ค่าความเที่ยงของแบบทดสอบ

6. ค่าอำนาจในการทำนาย E (Efficiency of Prediction) (โกวิท ประวาลพกษ์ และสมศักดิ์ สีนระเวชญ์ 2527 : 280)

$$E = 100 (1 - \sqrt{1 - r^2})$$

E = ประสิทธิภาพในการทำนาย

r = ค่าความเที่ยง

7. การวิเคราะห์ข้อสอบอัตโนมัติวัดความสามารถทางการเขียน ใช้การวิเคราะห์ของ Whitney และ Saber (1970) อ้างถึงใน Scannell และ Tracy (1975)

7.1 การวิเคราะห์หาค่าความง่าย (Facility value /FV)

$$FV = \frac{S_u + S_L - (N_T) X_{min}}{N_T (X_{max} - X_{min})}$$

- เมื่อ
- S_u = คะแนนรวมของกลุ่มที่ได้คะแนนสูง
 - S_L = คะแนนรวมของกลุ่มที่ได้คะแนนต่ำ
 - N_T = จำนวนผู้สอบทั้งหมดที่ใช้ในการวิเคราะห์
 - X_{max} = คะแนนสูงสุด
 - X_{min} = คะแนนต่ำสุด

7.2 การวิเคราะห์ค่าอำนาจจำแนก

$$D = \frac{S_u - S_L}{N_u (S_{max} - X_{min})}$$

- เมื่อ
- S_u = คะแนนรวมของผู้สอบในกลุ่มคะแนนสูง
 - S_L = คะแนนรวมของผู้สอบในกลุ่มคะแนนต่ำ
 - N_u = จำนวนของผู้สอบในกลุ่มคะแนนสูง
 - X_{max} = คะแนนสูงสุด
 - X_{min} = คะแนนต่ำสุด

7.3 การวิเคราะห์ค่าความเที่ยง การวิเคราะห์หาค่าความเที่ยงใช้วิธีหาค่าสัมประสิทธิ์สหสัมพันธ์ Alpha (α) พัฒนาขึ้นโดย Cronbach (1951) อ้างถึงใน Pilliner (1985) สูตรนี้ใช้ในกรณีที่ข้อสอบไม่ได้ใช้วิธีการให้คะแนนแบบ 1 และ 0 สูตรนี้จะให้ค่าเท่ากับ KR_{20}

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

เมื่อ $\sum S_i^2 =$ ผลรวมของค่าความแปรปรวนของแต่ละข้อ
 $S_x^2 =$ ค่าความแปรปรวนของแบบสอบทั้งฉบับ

8. การวิเคราะห์ข้อทดสอบเป็นรายข้อกระทางโดยใช้วิธี Rasch Model เพื่อตรวจสอบคุณสมบัติของข้อสอบ และคัดเลือกนำมาบรรจุในคลังข้อสอบด้วยระบบคอมพิวเตอร์

สูตรในการหาค่าของความเป็นไปได้ที่คน ๆ หนึ่งจะตอบข้อทดสอบถูกหรือผิด คือ

$$P\{x_{ve} | B_v, \theta_e\} = \frac{\text{Exp}[x_{ve} (B_v - \theta_e)]}{[1 + \text{Exp}(B_v - \theta_e)]}$$

เมื่อ P = probability
 x = คะแนน
 ve = ถูกหรือผิด มีค่า 1 และ 0
 / = ทหาร
 Exp = Exponential
 | = given

เกณฑ์ในการพิจารณาคุณภาพข้อสอบ

1. คุณภาพข้อสอบทั้งฉบับ

1.1 ความยากง่ายของข้อสอบ (p) พิจารณาจากค่าความยากง่ายของข้อสอบทั้งฉบับ โดยที่ข้อสอบสมิทริกานทั่วไปทางภาษาอังกฤษเป็นข้อสอบที่วัดความสามารถทั่วไป จึงได้กำหนดค่าของความยากง่ายไว้ดังนี้

ความยากง่าย	ระดับ
.40 - .60	ดีมาก
.31 - .39 , .61-.69	ดี
.20-.30 , .70-.80	พอใช้ได้
.81 >	ง่ายเกินไป
<.19	ยากเกินไป

1.2 อำนาจจำแนก

ค่าอำนาจจำแนกของแบบทดสอบสมิทธิภาพทั่วไปทางภาษาอังกฤษ (r) หมายถึงความสามารถของข้อสอบสมิทธิภาพทั่วไปทางภาษาอังกฤษที่จะแยกคนเก่งออกจากคนอ่อน ค่าอำนาจจำแนกมีค่าตั้งแต่ +1.00 ถึง -1.00 ค่า r ที่ใช้ได้จะมีค่าตั้งแต่ +.20 ขึ้นไป และถ้าค่า r ยิ่งใกล้ + 1.00 ข้อสอบนั้นจะยิ่งมีอำนาจจำแนกมาก ถ้า r มีค่าเป็นลบ ข้อสอบนี้จะเป็นข้อสอบที่ไม่ดี เพราะเป็นข้อสอบที่ลวงคนเก่ง ซึ่งผิดจุดประสงค์ของข้อสอบเกณฑ์ที่ใช้ในการพิจารณาอำนาจจำแนกของข้อสอบดังนี้

อำนาจจำแนก	ระดับ
.60 >	ดีมาก
.35-.59	ดี
.20-.34	พอใช้ได้
>.20	ต่ำควรปรับปรุง
ติดลบ	ต้องตัดทิ้งหรือเปลี่ยนแปลง

สำหรับแบบทดสอบวัดสมิทธิภาพทั่วไปทางภาษาอังกฤษเป็นแบบทดสอบที่วัดแบบอิงเกณฑ์พิจารณาค่าเฉลี่ยของค่าอำนาจจำแนกทั้งฉบับ และแก้ไขข้อที่มีค่าอำนาจจำแนกติดลบ

1.3 ความเที่ยงของแบบทดสอบสมิทธิภาพทั่วไปทางภาษาอังกฤษ

ความเที่ยงของข้อสอบเป็นค่าที่แสดงถึงความคงเส้นคงวาหรือความสม่ำเสมอของข้อสอบที่ใช้วัด คือไม่ว่าผู้สอบสอบด้วยข้อสอบชุดเดิมกี่ครั้งก็ตาม ผู้สอบจะได้คะแนนเท่าเดิมหรือใกล้เคียงกัน ข้อสอบที่มีลักษณะนี้จัดว่าเป็นข้อสอบที่มีความเที่ยงสูง เกณฑ์ที่ใช้ในการพิจารณาค่าความเที่ยง คือ ค่า KR_{20} สำหรับข้อสอบปรนัย ค่า Cronbach Alpha สำหรับข้อสอบ



อัตร้าวัดความสามารถในการเขียน และค่าสัมประสิทธิ์ Phi สำหรับข้อสอบทั้งปรนัยและอัตนัย โดยพิจารณาความสำคัญทางสถิติเป็นเกณฑ์

1.4 คุณภาพของรายข้อกระทง

คุณภาพของรายข้อกระทงพิจารณาจากความยากง่าย และอำนาจจำแนกรายข้อทดสอบ เกณฑ์ที่ใช้ในการพิจารณาค่าคุณภาพของรายข้อกระทง แบ่งเป็น

1. ข้อกระทงที่มีคุณภาพทั้งความยากง่ายและอำนาจจำแนกเป็นข้อกระทงที่มีความยากง่ายระหว่าง .20 ถึง .80 และมีค่าอำนาจจำแนกไม่ติดลบ
2. ข้อกระทงที่สามารถปรับปรุงในด้านความยากง่ายหรืออำนาจจำแนก คือ ข้อกระทงที่
 - 2.1 มีค่าความยากง่ายระหว่าง .20 ถึง .80 และมีอำนาจจำแนกติดลบ
 - 2.2 มีค่าความยากง่ายน้อยกว่า .20 หรือมากกว่า .80 และมีค่าอำนาจจำแนกไม่ติดลบ
3. ข้อกระทงที่ควรตัดทิ้ง เป็นข้อกระทงที่มีค่าความยากง่ายน้อยกว่า .20 หรือมากกว่า .80 และมีอำนาจจำแนกติดลบ

สำหรับคุณภาพของตัวเลือกที่เป็นตัวลวง ต้องปรับปรุงตัวเลือกเมื่อตัวเลือกนั้นมีผู้เลือกตอบไม่ถึงร้อยละ 5 ($P < .05$) และตัวเลือกนั้นมีผู้ตอบในกลุ่มที่ได้คะแนนสูงเลือกตอบมากกว่ากลุ่มที่ได้คะแนนต่ำ ($N_H > N_L$)

ส่วนการตรวจสอบคุณภาพของข้อกระทงตามวิธีการของ Rasch Model มีขั้นตอนดังนี้

1. ค่าความยากง่าย (β) ว่าอยู่ในพิสัยของการยอมรับได้คือ +3.00 หรือไม่
2. ค่าความสอดคล้องของโค้ง ICC (Item Characteristic Curve) เป็นไปตามที่คาดหวังหรือว่า หรือสอดคล้องมากน้อยเพียงใด โดยพิจารณาค่า weighted mean square (weighted MNSQ) ว่ามีค่าใกล้เคียงกับ 1.00 หรือไม่ โดยนำเอาค่าความคาดเคลื่อนมาตรฐานของการวัด (SE_m) มาบวกหรือลบ ในการพิจารณาความสอดคล้องด้วย

บทที่ 4

ผลการวิเคราะห์ข้อมูล

การวิเคราะห์ข้อทดสอบสมิทริภาพทั่วไปทางภาษาอังกฤษ ผู้วิจัยเสนอผลการวิจัย

เป็น 3 ตอน ดังนี้

ตอนที่ 1 ผลการวิเคราะห์ข้อทดสอบสมิทริภาพทั่วไปทางภาษาอังกฤษเพื่อการพัฒนาปรับปรุงข้อกระทง แยกการวิเคราะห์ ดังนี้

1. ผลการวิเคราะห์ข้อทดสอบแบบปรนัย
2. ผลการวิเคราะห์ข้อทดสอบอัตนัย

ตอนที่ 2 ผลการวิเคราะห์แบบทดสอบสมิทริภาพทั่วไปทางภาษาอังกฤษฉบับปรับปรุง

1. ผลการวิเคราะห์ข้อทดสอบแบบปรนัย
2. ผลการวิเคราะห์ข้อทดสอบอัตนัย

ตอนที่ 3 ผลการวิเคราะห์ข้อทดสอบแบบ Rasch Model เพื่อพัฒนาเป็นคลังข้อสอบ

ตอนที่ 4 สรุปผลการวิเคราะห์

ตอนที่ 1 ผลการวิเคราะห์ข้อทดสอบสมิทริภาพทั่วไปทางภาษาอังกฤษเพื่อการพัฒนาปรับปรุงข้อสอบรายชื่อ

- 1.1 คุณภาพของแบบทดสอบ

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 2 แสดงค่าสถิติต่าง ๆ ของแบบทดสอบสมิทภาพทั่วไปทางภาษาอังกฤษฉบับที่ 1 - 4

ฉบับที่	N	\bar{X}	SD	CV	Max	Min	Range	ค่าความยากง่าย		อำนาจจำแนก		ค่าความเที่ยง		อำนาจทำนาย	SE _m	สรุปคุณภาพ
								ค่า p	Δ	r _{pb}	KR ₂₀	Phi	E			
1	121	57.909	13.974	24.131	103	27	76	0.488	13.142	0.258	0.878	.896**	52.10	4.881	ดีมาก	
2	103	59.320	15.996	26.966	111	31	80	0.494	13.041	0.288	0.910	.893**	58.50	4.799	ดี	
3	118	60.737	12.445	21.313	102	34	68	0.506	12.941	0.230	0.854	.822**	48.00	4.946	ดี	
4	139	62.029	13.788	22.228	106	33	73	0.517	12.831	0.250	0.877	.854**	52.00	4.836	ดีมาก	

แบบทดสอบสมิทภาพทั่วไปทางภาษาอังกฤษทั้ง 4 ฉบับ มีค่าเฉลี่ยความยากง่ายในเกณฑ์ดี - ดีมาก (Δ 12.831 ~ 13.219 , p 0.483 ~ 0.517)
 ค่าอำนาจจำแนกเฉลี่ยอยู่ในเกณฑ์ยอมรับได้ (r_{pb} 0.258 ~ 0.088) ค่าความเที่ยงสูง (KR₂₀ 0.854 ~ 0.910 Phi .822 ~ .893) ความแตกต่าง
 ของคะแนนอยู่ในระดับปานกลาง และได้เฉลี่ยกันโดยฉบับที่ 2 มีค่าสัมประสิทธิ์การกระจายสูงที่สุด (26.966) และฉบับที่ 3 ค่าที่สุด 21.313)

สรุปได้ว่าแบบทดสอบทั้ง 4 ฉบับมีคุณภาพในเกณฑ์ดี ความยากง่ายปานกลาง อำนาจจำแนกในการทำนายร้อยละ 48.00 ~ 58.50 ค่าความคลาดเคลื่อนมาตรฐานของ
 การวัด 4.799 ~ 4.946

1.2 ผลการวิเคราะห์ข้อทดสอบแบบปรนัย

ตารางที่ 3 ผลการวิเคราะห์ข้อกระทงแบบทดสอบสมิทริภาพทั่วไปทางภาษาอังกฤษ ฉบับที่ 1

ข้อที่	ความยากง่าย		อำนาจจำแนก		ตัวเลือกที่ต้องปรับปรุง	
	ขนาด	คุณภาพ	ขนาด	คุณภาพ		
	P	Δ				
1	.074	18.81	ยากไป	.281	พอใช้	4 $N_H > N_L$
2	.157	17.07	ยากไป	.351	พอใช้	3 $N_H > N_L$
3	.339	14.70	พอใช้	.345	พอใช้	
4	.595	12.00	ดี	.240	พอใช้	
5	.231	15.97	พอใช้	.210	พอใช้	
6	.215	16.20	พอใช้	.034	ต่ำ	
7	.545	12.50	ดีมาก	.094	ต่ำ	
8	.554	12.42	ดีมาก	.414	ดี	
9	.694	10.93	ดี	.322	พอใช้	
10	.405	14.00	ดีมาก	.384	พอใช้	
11	.496	13.04	ดีมาก	.158	ต่ำ	
12	.529	12.67	ดีมาก	.067	ต่ำ	
13	.215	16.20	พอใช้	.300	พอใช้	
14	.256	15.66	พอใช้	.049	ต่ำ	
15	.544	12.42	ดีมาก	.280	พอใช้	
16	.479	13.25	ดีมาก	.076	ต่ำ	
17	.471	13.33	ดีมาก	.107	ต่ำ	
18	.620	11.74	ดี	.217	พอใช้	
19	.256	15.66	พอใช้	.222	พอใช้	

ตารางที่ 3 (ต่อ)

ข้อที่	ความยากง่าย		อำนาจจำแนก		ตัวเลือกที่ต้องปรับปรุง
	ขนาด	คุณภาพ	ขนาด	คุณภาพ	
	P	Δ			
20	.521	12.75	ดีมาก	.189	ต่ำ
21	.471	13.33	ดีมาก	.157	ต่ำ
22	.488	13.16	ดีมาก	.426	ดี
23	.603	11.91	ดีมาก	.269	พอใช้
24	.612	11.82	ดีมาก	.338	พอใช้
25	.273	15.46	พอใช้	.359	พอใช้
26	.273	15.46	พอใช้	.300	พอใช้
27	.083	18.59	ยากไป	.030	ต่ำ
28	.512	12.84	ดีมาก	.224	พอใช้
29	.347	14.61	พอใช้	.242	พอใช้
30	.529	12.67	ดีมาก	.029	ต่ำ
31	.264	15.56	พอใช้	.043	ต่ำ
32	.512	12.84	ดีมาก	.216	พอใช้
33	.198	16.43	พอใช้	.239	พอใช้
34	.149	17.21	ยากไป	.280	พอใช้
35	.521	12.75	ดีมาก	.227	พอใช้
36	.289	15.26	พอใช้	.120	ต่ำ
37	.231	15.97	พอใช้	.140	ต่ำ
38	.760	10.13	พอใช้	.160	ต่ำ
39	.140	17.91	ยากไป	.244	พอใช้
40	.512	12.84	ดีมาก	.164	ต่ำ
41	.430	13.75	ดีมาก	.280	พอใช้

3 คนเลือกมากเกินไป

ตารางที่ 3 (ต่อ)

ข้อที่	ความยากง่าย		อำนาจจำแนก		ตัวเลือกที่ต้องปรับปรุง	
	ขนาด	คุณภาพ	ขนาด	คุณภาพ		
	P	Δ				
42	.331	14.79	พอใช้	.285	พอใช้	
43	.413	13.92	ดีมาก	.366	พอใช้	
44	.554	12.42	ดีมาก	.298	พอใช้	
45	.884	8.18	ง่ายไป	.236	พอใช้	2,3,4 ไม่มีประสิทธิภาพ
46	.281	15.36	พอใช้	.341	พอใช้	
47	.736	10.44	พอใช้	.358	พอใช้	
48	.322	14.88	พอใช้	.154	ต่ำ	
49	.529	12.67	ดีมาก	.276	พอใช้	
50	.860	8.09	ง่ายไป	.232	พอใช้	3,4 ไม่มีประสิทธิภาพ
51	.795	9.80	พอใช้	.358	พอใช้	
52	.748	15.76	พอใช้	.083	ต่ำ	
53	.645	11.48	ดีมาก	.293	พอใช้	
54	.636	11.57	พอใช้	.382	พอใช้	
55	.496	13.04	ดี	.409	ต่ำ	
56	.306	15.07	พอใช้	.234	พอใช้	
57	.223	16.09	พอใช้	.300	พอใช้	
58	.793	9.69	พอใช้	.340	พอใช้	
59	.231	15.97	พอใช้	.239	พอใช้	
60	.595	12.00	ดีมาก	.442	ดี	
61	.322	14.88	พอใช้	.390	พอใช้	
62	.744	10.34	พอใช้	.412	ดี	
63	.736	10.44	พอใช้	.263	พอใช้	

ตารางที่ 3 (ต่อ)

ข้อที่	ความยากง่าย		อำนาจจำแนก		ตัวเลือกที่ต้องปรับปรุง	
	ขนาด	คุณภาพ	ขนาด	คุณภาพ		
	p	Δ				
64	.413	13.92	ดีมาก	.349	พอใช้	
65	.851	8.79	ง่ายไป	.262	พอใช้	
66	.488	13.16	ดี	.277	พอใช้	
67	.380	14.26	พอใช้	.075	ต่ำ	
68	.810	9.45	ง่ายไป	.253	พอใช้	1 $N_H > N_L$
69	.760	10.13	พอใช้	.373	พอใช้	
70	.570	12.25	ดีมาก	.254	พอใช้	
71	.372	14.35	พอใช้	.012	ต่ำ	
72	.579	12.17	ดีมาก	.423	ดี	
73	.397	14.09	พอใช้	.499	ดี	
74	.364	14.43	พอใช้	.547	ดี	
75	.521	12.75	ดีมาก	.017	ต่ำ	
76	.678	11.12	พอใช้	.486	ดี	
77	.570	12.25	ดีมาก	.329	พอใช้	
78	.669	11.21	ดี	.093	ต่ำ	
79	.620	11.74	ดี	.529	ดี	
80	.653	11.39	ดี	.211	พอใช้	
81	.281	15.36	ดี	.286	พอใช้	
82	.463	13.41	ดีมาก	.157	ต่ำ	
83	.603	11.91	ดีมาก	.357	พอใช้	
84	.413	13.92	ดีมาก	.267	พอใช้	
85	.843	8.93	ง่ายไป	.228	พอใช้	4 $N_H > N_L$

ตารางที่ 3 (ต่อ)

ข้อที่	ความยากง่าย			อำนาจจำแนก		ตัวเลือกที่ต้องปรับปรุง
	ขนาด	คุณภาพ	ขนาด	คุณภาพ		
	P	Δ				
86	.719	10.64	พอใช้	.172	ต่ำ	
87	.917	7.41	ง่ายไป	.056	ต่ำ	1 $N_H > N_L$
88	.083	18.59	ยากไป	.277	พอใช้	4 $N_H > N_L$
89	.397	14.09	พอใช้	.347	พอใช้	
90	.579	12.17	ดีมาก	.406	ดี	
91	.446	13.58	ดีมาก	.235	พอใช้	
92	.331	17.79	พอใช้	.223	พอใช้	
93	.529	12.67	ดีมาก	.251	พอใช้	
94	.496	13.04	ดีมาก	.377	พอใช้	
95	.645	11.48	ดี	.158	ต่ำ	
96	.661	11.30	ดี	.269	พอใช้	
97	.372	14.35	พอใช้	.283	พอใช้	
98	.421	13.83	ดีมาก	.367	พอใช้	
99	.612	11.83	ดี	.353	พอใช้	
100	.496	13.04	ดีมาก	.334	พอใช้	
101	.570	12.25	ดีมาก	.269	พอใช้	
102	.537	12.59	ดีมาก	.388	พอใช้	
103	.537	12.59	ดีมาก	.301	พอใช้	
104	.579	12.17	ดีมาก	.239	พอใช้	
105	.438	13.66	ดีมาก	.385	พอใช้	
106	.471	13.33	ดีมาก	.358	พอใช้	
107	.603	11.91	ดีมาก	.267	พอใช้	

ตารางที่ 3 (ต่อ)

ข้อที่	ความยากง่าย		อำนาจจำแนก		ตัวเลือกที่ต้องปรับปรุง	
	ขนาด	คุณภาพ	ขนาด	คุณภาพ		
	p	Δ				
108	.719	10.64	พอใช้	.256	พอใช้	
109	.860	8.09	ง่ายไป	.312	พอใช้	1 ไม่มีประสิทธิภาพ
110	.669	11.21	ดี	.100	ต่ำ	
111	.537	12.59	ดีมาก	.251	พอใช้	
112	.471	13.33	ดีมาก	.237	พอใช้	
113	.661	11.30	ดี	.396	ดี	
114	.207	16.31	พอใช้	.000	ต่ำ	
115	.777	9.91	พอใช้	.291	พอใช้	
116	.190	16.56	ยากไป	.286	พอใช้	3 $N_H > N_L$
117	.587	12.08	ดีมาก	.236	พอใช้	
118	.388	14.17	พอใช้	.64	พอใช้	
119	.620	11.74	ดี	.366	พอใช้	
120	.264	15.56	พอใช้	.357	พอใช้	

* ค่าอำนาจจำแนกที่ใช้ในเกณฑ์ปรับปรุง คือข้อที่มีค่าอำนาจจำแนกติดลบ (ต่ำมาก)

สรุป แบบทดสอบสมิทริภาพทั่วไปทางภาษาอังกฤษฉบับที่ 1 มีข้อสอบที่ใช้ได้ตามเกณฑ์ความยากง่าย 109 ข้อ และไม่มีข้อใดที่มีค่าอำนาจจำแนกติดลบ ข้อสอบที่ต้องปรับปรุง 11 ข้อ เป็นข้อสอบที่ง่ายเกินไป 6 ข้อ และยากเกินไป 5 ข้อ

ตารางที่ 4 ผลการวิเคราะห์ข้อกระทบแบบทดสอบสมิทริภาพทั่วไปทางภาษาอังกฤษ ฉบับที่ 2

ข้อที่	ความยากง่าย		อำนาจจำแนก		ตัวเลือกที่ต้องปรับปรุง	
	ขนาด	คุณภาพ	ขนาด	คุณภาพ		
	P	Δ				
1	.864	8.57	ง่ายไป	.281	พอใช้	3 ไม่มีประสิทธิภาพ
2	.485	13.19	ดีมาก	.272	พอใช้	
3	.660	11.31	ดี	.307	พอใช้	
4	.621	11.72	ดี	.338	พอใช้	
5	.563	12.32	ดีมาก	.406	ดี	
6	.320	14.91	ดี	.473	ดี	
7	.709	10.76	พอใช้	.375	ดี	
8	.340	14.69	ดี	.252	พอใช้	
9	.243	15.83	พอใช้	.446	ดี	
10	.379	14.28	ดี	.321	พอใช้	
11	.272	15.47	พอใช้	.214	พอใช้	
12	.282	15.35	พอใช้	.373	ดี	
13	.117	17.81	ยากไป	.333	พอใช้	
14	.233	15.59	พอใช้	.072	ต่ำ	
15	.427	13.77	ดีมาก	.130	ต่ำ	4 $N_H > N_L$
16	.291	15.24	พอใช้	.317	พอใช้	
17	.155	17.09	ยากไป	.358	ดี	
18	.447	13.54	ดีมาก	.395	ดี	
19	.738	10.41	พอใช้	.382	ดี	
20	.534	12.62	ดีมาก	.475	ดี	
21	.301	15.13	พอใช้	.122	ต่ำ	
22	.223	16.08	พอใช้	.206	พอใช้	

ตารางที่ 4 (ต่อ)

ข้อที่	ความยากง่าย		อำนาจจำแนก		ตัวเลือกที่ต้องปรับปรุง	
	ขนาด	คุณภาพ	ขนาด	คุณภาพ		
	p	Δ				
23	.388	14.17	ดี	.313	พอใช้	
24	.204	16.35	พอใช้	.330	พอใช้	
25	.262	15.59	พอใช้	-.022	ต่ำ	$2 N_H = N_L, 4 N_H > N_L$
26	.427	13.77	ดีมาก	.380	ดี	
27	.515	12.81	ดีมาก	.350	ดี	
28	.262	15.59	พอใช้	-.099	ต่ำ	$2 N_H > N_L, 3$ คนเลือกมากไป
29	.631	11.62	ดี	.283	พอใช้	
30	.515	12.81	ดีมาก	.165	ต่ำ	
31	.408	13.97	ดีมาก	.200	พอใช้	
32	.583	12.13	ดีมาก	.074	ต่ำ	
33	.505	12.96	ดีมาก	.210	พอใช้	
34	.786	9.79	พอใช้	.299	พอใช้	
35	.359	14.48	ดี	.540	ดี	
36	.680	11.09	ดี	.279	พอใช้	
37	.544	12.52	ดีมาก	.238	พอใช้	
38	.660	11.31	ดี	.300	พอใช้	
39	.573	12.23	ดีมาก	.257	พอใช้	
40	.728	10.53	พอใช้	.180	ต่ำ	
41	.689	10.98	ดี	.123	ต่ำ	
42	.282	15.35	พอใช้	.434	ดี	
43	.583	12.13	ดีมาก	.309	พอใช้	
44	.864	8.57	หายไป	.105	ต่ำ	1 ไม่มีประสิทธิภาพ

ตารางที่ 4 (ต่อ)

ข้อที่	ความยากง่าย		อำนาจจำแนก		ตัวเลือกที่ต้องปรับปรุง	
	ขนาด	คุณภาพ	ขนาด	คุณภาพ		
	p	Δ				
45	.660	11.31	ดี	.360	ดี	
46	.204	14.28	พอใช้	.321	พอใช้	
47	.417	13.87	ดีมาก	.267	พอใช้	
48	.505	12.96	ดีมาก	.408	ดี	
49	.524	12.72	ดีมาก	.322	พอใช้	
50	.709	10.76	พอใช้	.209	พอใช้	
51	.126	12.13	ยากไป	-.223	ต่ำ	1 $N_H > N_L$, แก้ตัวเลือกที่ถูก
52	.806	9.51	ง่ายไป	.331	พอใช้	4 $N_H > N_L$
53	.417	13.87	ดีมาก	.447	ดี	
54	.505	12.96	ดีมาก	.444	ดี	
55	.175	16.78	ยากไป	.518	ดี	4 คนเลือกมากไป
56	.427	13.77	ดีมาก	.224	พอใช้	
57	.272	15.47	พอใช้	.513	ดี	
58	.408	13.97	ดีมาก	.392	ดี	
59	.379	14.28	ดี	.427	ดี	
60	.825	9.22	ง่ายไป	.246	พอใช้	3 ไม่มีประสิทธิภาพ
61	.602	11.93	ดีมาก	.170	ต่ำ	
62	.883	8.19	ง่ายไป	.215	พอใช้	1 ไม่มีประสิทธิภาพ
63	.835	9.07	ง่ายไป	.298	พอใช้	2,3 ไม่มีประสิทธิภาพ
64	.796	9.65	พอใช้	.397	ดี	
65	.854	5.81	ง่ายไป	.242	พอใช้	

ตารางที่ 4 (ต่อ)

ข้อที่	ความยากง่าย		อำนาจจำแนก		ตัวเลือกที่ต้องปรับปรุง	
	ขนาด	คุณภาพ	ขนาด	คุณภาพ		
	P	Δ				
66	.796	9.65	พอใช้	.304	พอใช้	
67	.204	16.35	พอใช้	.285	พอใช้	
68	.165	16.93	ยากไป	.375	ดี	1 $N_H > N_L$
69	.709	10.76	พอใช้	.375	ดี	
70	.456	13.48	ดีมาก	.559	ดี	
71	.650	11.41	ดี	.451	ดี	
72	.845	8.91	ง่ายไป	.285	พอใช้	2,3 ไม่มีประสิทธิภาพ
73	.602	11.93	ดีมาก	.314	พอใช้	
74	.544	12.52	ดีมาก	.530	ดี	
75	.291	15.24	พอใช้	.469	ดี	
76	.505	12.96	ดีมาก	.331	พอใช้	
77	.097	10.41	ยากไป	.020	ต่ำ	1 $N_H > N_L$, 1 ไม่มีประสิทธิภาพ
78	.796	9.65	พอใช้	.376	ดี	
79	.680	11.09	ดี	.344	พอใช้	
80	.874	8.38	ง่ายไป	.185	ต่ำ	2 $N_H > N_L$, 3 $N_H > N_L$
81	.350	14.59	ดี	.365	ดี	
82	.408	13.97	ดีมาก	.150	ต่ำ	
83	.126	17.62	ยากไป	.159	ต่ำ	2 คนเลือกมากไป, แก้ตัวเลือกที่ถูกลง
84	.485	13.91	ดีมาก	.352	ดี	

ตารางที่ 4 (ต่อ)

ข้อที่	ความยากง่าย		อำนาจจำแนก		ตัวเลือกที่ต้องปรับปรุง	
	ขนาด	คุณภาพ	ขนาด	คุณภาพ		
	p	Δ				
85	.903	7.71	ง่ายไป	.173	ต่ำ	1,4 ไม่มีประสิทธิภาพ
86	.388	14.71	พอใช้	.176	ต่ำ	
87	.369	14.38	พอใช้	.388	ดี	
88	.350	14.59	พอใช้	.371	ดี	
89	.456	13.48	ดีมาก	.447	ดี	
90	.369	14.38	ดี	.463	ดี	
91	.689	10.98	พอใช้	.326	พอใช้	
92	.563	12.32	ดีมาก	.340	พอใช้	
93	.777	9.92	พอใช้	.339	พอใช้	
94	.330	14.80	ดี	.341	พอใช้	
95	.689	10.98	พอใช้	.250	พอใช้	
96	.864	8.57	ง่ายไป	.219	พอใช้	4 ไม่มีประสิทธิภาพ
97	.583	12.13	ดีมาก	.349	พอใช้	
98	.204	16.35	พอใช้	.342	พอใช้	
99	.505	12.96	ดีมาก	.325	พอใช้	
100	.388	14.17	ดี	.028	ต่ำ	
101	.631	11.62	ดี	.271	พอใช้	
102	.583	12.13	ดีมาก	.213	พอใช้	
103	.456	13.48	ดีมาก	.245	พอใช้	
104	.146	13.87	ดีมาก	.430	ดี	
105	.214	16.21	พอใช้	.404	ดี	

ตารางที่ 4 (ต่อ)

ข้อที่	ความยากง่าย		อำนาจจำแนก		ตัวเลือกที่ต้องปรับปรุง	
	ขนาด	คุณภาพ	ขนาด	คุณภาพ		
	P	Δ				
106	.515	12.81	ดีมาก	.155	ต่ำ	
107	.301	15.13	พอใช้	.249	พอใช้	
108	.553	12.42	ดีมาก	.350	ดี	
109	.592	12.03	ดีมาก	.528	ดี	
110	.447	13.58	ดีมาก	.218	พอใช้	
111	.320	14.91	ดี	.308	พอใช้	
112	.485	13.19	ดีมาก	.324	พอใช้	
113	.573	12.23	ดีมาก	.394	ดี	
114	.282	15.35	พอใช้	.156	ต่ำ	
115	.252	15.71	พอใช้	.237	พอใช้	
116	.214	16.21	พอใช้	-.040	ต่ำ	2,3 $N_H > N_L$
117	.427	13.77	ดีมาก	.024	ต่ำ	
118	.340	14.69	ดี	.325	พอใช้	
119	.243	15.83	พอใช้	.150	ต่ำ	
120	.670	11.20	ดี	.541	ดี	

สรุป แบบทดสอบสมิทธิภาพทางภาษาอังกฤษฉบับที่ 2 มีข้อสอบที่ใช้ได้ตามเกณฑ์ 102 ข้อ ข้อสอบที่ต้องปรับปรุง 18 ข้อ เป็นข้อที่ง่ายเกินไป 11 ข้อ ยากเกินไป 7 ข้อ มีค่าอำนาจจำแนกติดลบ 4 ข้อ

ตารางที่ 5 ผลการวิเคราะห์ข้อกระถงแบบทดสอบสมมติฐานทั่วไปทางภาษาอังกฤษ ฉบับที่ 3

ข้อที่	ความยากง่าย		อำนาจจำแนก		ตัวเลือกที่ต้องปรับปรุง	
	ขนาด	คุณภาพ	ขนาด	คุณภาพ		
	p	Δ				
1	.924	7.24	ง่ายไป	.159	ต่ำ	1,3 ไม่มีประสิทธิภาพ
2	.559	12.36	ดีมาก	.397	ดี	
3	.568	12.28	ดีมาก	.022	ต่ำ	
4	.322	18.76	ดี	.147	ต่ำ	3,4 $N_H > N_L$
5	.602	11.93	ดีมาก	.161	ต่ำ	
6	.144	17.62	ยากไป	.016	ต่ำ	
7	.568	12.28	ดีมาก	.276	พอใช้	
8	.254	15.68	พอใช้	.242	พอใช้	
9	.542	12.53	ดีมาก	.190	ต่ำ	
10	.441	13.64	ดีมาก	.334	พอใช้	
11	.415	13.90	ดีมาก	.117	ต่ำ	
12	.500	13.00	ดีมาก	.310	พอใช้	
13	.347	14.61	ดี	.196	ต่ำ	
14	.678	11.11	ดี	.231	พอใช้	
15	.517	12.79	ดีมาก	.381	ดี	
16	.458	13.47	ดีมาก	.108	ต่ำ	
17	.712	10.72	พอใช้	-.097	ต่ำ	
18	.364	14.43	ดี	-.057	ต่ำ	
19	.220	16.12	พอใช้	-.010	ต่ำ	
20	.585	12.10	ดีมาก	-.105	ต่ำ	
21	.551	12.45	ดีมาก	.141	ต่ำ	
22	.703	10.82	พอใช้	.056	ต่ำ	

ตารางที่ 5 (ต่อ)

ข้อที่	ความยากง่าย		อำนาจจำแนก		ตัวเลือกที่ต้องปรับปรุง	
	ขนาด	คุณภาพ	ขนาด	คุณภาพ		
	p	Δ				
23	.364	14.43	ดี	.180	ต่ำ	
24	.678	11.11	ดี	.167	ต่ำ	
25	.169	16.86	ยากไป	.365	ดี	3 มีคนเลือกมากไป, แก้ตัวเลือกที่ถูก
26	.585	12.10	ดีมาก	.322	พอใช้	
27	.754	10.21	พอใช้	.306	พอใช้	
28	.602	11.93	ดีมาก	.355	ดี	
29	.534	12.62	ดีมาก	.175	ต่ำ	
30	.898	7.87	ง่ายไป	.192	ต่ำ	1,4 ไม่มีประสิทธิภาพ
31	.339	14.70	ดี	-.017	ต่ำ	1 $N_H > N_L$
32	.415	13.90	ดีมาก	.222	พอใช้	
33	.263	15.58	พอใช้	.247	พอใช้	
34	.314	14.98	ดี	.159	ต่ำ	
35	.475	13.30	ดีมาก	.430	ดี	
36	.661	11.30	ดี	.251	พอใช้	
37	.763	10.10	พอใช้	.224	พอใช้	
38	.517	12.79	ดีมาก	.274	พอใช้	
39	.483	13.21	ดีมาก	.203	พอใช้	
40	.610	11.84	ดี	.224	พอใช้	
41	.331	14.79	ดี	.393	ดี	
42	.644	11.48	ดี	.321	พอใช้	
43	.534	12.62	ดีมาก	.368	ดี	

ตารางที่ 5 (ต่อ)

ข้อที่	ความยากง่าย		อำนาจจำแนก		ตัวเลือกที่ต้องปรับปรุง	
	ขนาด	คุณภาพ	ขนาด	คุณภาพ		
	p	Δ				
44	.203	16.36	พอใช้	.025	ต่ำ	
45	.805	9.52	ง่ายไป	.211	พอใช้	3 ไม่มีประสิทธิภาพ
46	.390	14.16	ดี	.328	พอใช้	
47	.712	10.72	พอใช้	.064	ต่ำ	
48	.483	13.21	ดีมาก	.265	พอใช้	
49	.720	10.62	พอใช้	.241	พอใช้	
50	.712	10.72	พอใช้	.088	ต่ำ	
51	.898	7.87	ง่ายไป	.260	พอใช้	2,3 ไม่มีประสิทธิภาพ
52	.602	11.93	ดี	.371	ดี	
53	.636	11.57	ดี	.039	ต่ำ	
54	.661	11.30	ดี	.315	พอใช้	
55	.332	14.79	ดี	.098	ต่ำ	
56	.424	13.81	ดีมาก	-.175	ต่ำ	3 $N_H > N_L$
57	.246	15.79	พอใช้	.264	พอใช้	
58	.508	12.96	ดีมาก	.301	พอใช้	
59	.593	12.02	ดีมาก	.359	ดี	
60	.839	9.00	ง่ายไป	.308	พอใช้	1 ไม่มีประสิทธิภาพ, แก้ไขตัวเลือกที่ถูก
61	.466	13.38	ดีมาก	.187	ต่ำ	
62	.703	10.82	พอใช้	.137	ต่ำ	
63	.771	9.99	พอใช้	.215	พอใช้	
64	.314	14.98	ดี	.237	พอใช้	

ตารางที่ 5 (ต่อ)

ข้อที่	ความยากง่าย		อำนาจจำแนก		ตัวเลือกที่ต้องปรับปรุง	
	ขนาด	คุณภาพ	ขนาด	คุณภาพ		
	p	Δ				
65	.432	13.72	ดีมาก	.331	พอใช้	
66	.644	11.48	ดี	.242	พอใช้	
67	.424	13.81	ดีมาก	.341	พอใช้	
68	.127	17.60	ยากไป	.122	ต่ำ	$2 N_H > N_L$
69	.678	11.11	ดี	.363	ดี	
70	.475	13.30	ดีมาก	.486	ดี	
71	.254	15.68	พอใช้	.389	ดี	
72	.763	10.10	พอใช้	.320	พอใช้	
73	.695	10.92	พอใช้	.356	ดี	
74	.619	11.75	ดี	.275	พอใช้	
75	.263	15.58	พอใช้	.185	ต่ำ	
76	.110	15.90	ยากไป	-.277	ต่ำ	$1 N_H > N_L$, แก้ตัวเลือกที่ถูกลง
77	.559	12.36	ดีมาก	.293	พอใช้	
78	.203	16.36	พอใช้	-.143	ต่ำ	$3 N_H = N_L$, $4 N_H > N_L$
79	.432	13.72	ดีมาก	.377	ดี	
80	.669	11.21	ดี	.366	ดี	
81	.551	12.45	ดีมาก	.044	ต่ำ	
82	.398	14.07	ดี	.374	ดี	
83	.280	15.38	พอใช้	.392	ดี	
84	.534	12.62	ดีมาก	.278	พอใช้	
85	.364	14.43	ดี	.296	พอใช้	

ตารางที่ 5 (ต่อ)

ข้อที่	ความยากง่าย		อำนาจจำแนก		ตัวเลือกที่ต้องปรับปรุง
	ขนาด	คุณภาพ	ขนาด	คุณภาพ	
	p	Δ			
86	.246	15.79	พอใช้	.307	พอใช้
87	.746	10.32	พอใช้	.347	พอใช้
88	.644	11.48	ดี	.154	ต่ำ
89	.407	13.98	ดีมาก	.269	พอใช้
90	.636	11.57	ดี	.259	พอใช้
91	.229	16.01	พอใช้	.072	ต่ำ
92	.500	13.00	ดีมาก	.383	ดี
93	.424	13.81	ดีมาก	.363	ดี
94	.653	11.39	ดี	.365	ดี
95	.636	11.57	ดี	.342	พอใช้
96	.703	10.82	พอใช้	.192	ต่ำ
97	.890	8.06	ง่ายไป	.288	พอใช้
98	.508	12.96	ดีมาก	.301	พอใช้
99	.746	10.32	พอใช้	.035	ต่ำ
100	.585	12.10	ดีมาก	.156	ต่ำ
101	.763	10.10	พอใช้	.209	พอใช้
102	.534	12.62	ดีมาก	.450	ดี
103	.381	14.25	ดี	.300	พอใช้
104	.517	12.79	ดีมาก	.190	ต่ำ
105	.653	11.39	ดี	.286	พอใช้
106	.576	12.19	ดีมาก	.196	ต่ำ
107	.534	12.62	ดีมาก	.078	ต่ำ

3 ไม่มีประสิทธิภาพ

ตารางที่ 5 (ต่อ)

ข้อที่	ความยากง่าย		อำนาจจำแนก		ตัวเลือกที่ต้องปรับปรุง	
	ขนาด	คุณภาพ	ขนาด	คุณภาพ		
	p	Δ				
108	.415	13.90	ดีมาก	.158	ต่ำ	
109	.195	16.48	ยากไป	.403	ดี	1 คนเลือกมากไป, แก้ตัวเลือกที่ถูก
110	.085	18.53	ยากไป	.192	ต่ำ	3 $N_H > N_L$, 4 $N_H = N_L$
111	.458	13.47	ดีมาก	.431	ดี	
112	.729	10.52	พอใช้	.300	พอใช้	
113	.220	16.12	พอใช้	.145	พอใช้	
114	.458	13.47	ดีมาก	.370	ดี	
115	.441	13.64	ดีมาก	.030	ต่ำ	
116	.076	18.76	ยากไป	.008	ต่ำ	3 คนเลือกมากไป, 4 $N_H > N_L$
117	.449	13.55	ดีมาก	.566	ดี	
118	.347	14.61	ดี	.326	พอใช้	
119	.212	16.24	พอใช้	.052	ต่ำ	
120	.619	11.75	ดีมาก	.080	ต่ำ	

สรุป แบบทดสอบสมิทริภาพทั่วไปภาษาอังกฤษฉบับที่ 3 มีข้อกระทงที่ใช้ได้ตามเกณฑ์ 104 ข้อ ต้องปรับปรุง 16 ข้อ เป็นข้อที่ง่ายไป 6 ข้อ ยากไป 6 ข้อ อำนาจจำแนกดีดลบ 8 ข้อ

ตารางที่ 6 ผลการวิเคราะห์ข้อกระทงแบบทดสอบสมิทธิภาพทั่วไปทางภาษาอังกฤษ ฉบับที่ 4

ข้อที่	ความยากง่าย			อำนาจจำแนก		ตัวเลือกที่ต้องปรับปรุง
	ขนาด		คุณภาพ	ขนาด	คุณภาพ	
	p	Δ				
1	.410	13.95	ดีมาก	.203	พอใช้	
2	.770	10.01	พอใช้	.030	ต่ำ	
3	.309	15.03	พอใช้	.363	ดี	
4	.266	15.54	พอใช้	.257	พอใช้	
5	.295	15.20	พอใช้	.316	พอใช้	
6	.417	13.88	ดีมาก	.030	ต่ำ	
7	.317	14.95	ดี	.239	พอใช้	
8	.460	13.44	ดีมาก	.313	พอใช้	
9	.324	14.87	ดี	.272	พอใช้	
10	.317	14.95	ดี	.340	พอใช้	
11	.187	16.59	ยากไป	.173	ต่ำ	$4 N_H > N_L$
12	.540	12.56	ดีมาก	.085	ต่ำ	
13	.511	12.96	ดีมาก	.317	พอใช้	
14	.554	12.42	ดีมาก	.193	ต่ำ	
15	.576	12.20	ดีมาก	.232	พอใช้	
16	.360	14.48	ดี	.188	ต่ำ	
17	.561	12.34	ดีมาก	.022	ต่ำ	
18	.288	15.28	พอใช้	.085	ต่ำ	
19	.295	15.20	พอใช้	.305	พอใช้	
20	.374	14.32	ดี	.406	ดี	
21	.561	12.34	ดีมาก	.248	พอใช้	
22	.496	13.04	ดีมาก	.122	ต่ำ	

ตารางที่ 6 (ต่อ)

ข้อที่	ความยากง่าย		อำนาจจำแนก		ตัวเลือกที่ต้องปรับปรุง	
	ขนาด	คุณภาพ	ขนาด	คุณภาพ		
	p	Δ				
23	.568	14.95	ดีมาก	.330	พอใช้	
24	.432	13.73	ดีมาก	.275	พอใช้	
25	.266	15.54	พอใช้	.268	พอใช้	
26	.806	9.51	ง่ายไป	.169	ต่ำ	2 $N_H = N_L$ 4 ไม่มีประสิทธิภาพ
27	.158	17.05	ยากไป	.029	ต่ำ	2 $N_H > N_L$
28	.266	15.54	พอใช้	.243	พอใช้	
29	.532	12.64	ดีมาก	.246	พอใช้	
30	.410	13.95	ดีมาก	.128	ต่ำ	
31	.353	14.55	ดี	.159	ต่ำ	
32	.568	12.27	ดีมาก	.483	ดี	
33	.453	13.51	ดีมาก	.199	ต่ำ	
34	.540	12.56	ดีมาก	.286	พอใช้	
35	.482	13.22	ดีมาก	.334	พอใช้	
36	.014	21.78	ยากไป	-.092	ต่ำ	1,3 $N_H > N_L$
37	.612	11.83	ดี	.208	พอใช้	
38	.230	15.99	พอใช้	.246	พอใช้	
39	.741	10.37	พอใช้	.294	พอใช้	
40	.662	11.29	ดี	.321	พอใช้	
41	.158	17.05	ยากไป	-.037	ต่ำ	3,4 $N_H > N_L$ แก้ตัวเลือกที่ถูก
42	.647	11.45	ดี	.245	พอใช้	

ตารางที่ 6 (ต่อ)

ข้อที่	ความยากง่าย		อำนาจจำแนก		ตัวเลือกที่ต้องปรับปรุง	
	ขนาด	คุณภาพ	ขนาด	คุณภาพ		
	P	Δ				
43	.612	11.83	ดี	.406	ดี	
44	.626	11.68	ดี	.563	ดี	
45	.439	13.66	ดีมาก	.176	พอใช้	
46	.540	12.56	ดีมาก	.351	ดี	
47	.554	12.42	ดีมาก	.223	พอใช้	
48	.403	14.02	ดีมาก	.331	พอใช้	
49	.777	9.91	พอใช้	.368	ดี	
50	.058	19.34	ยากไป	.136	ต่ำ	$2 N_H > N_L$, แก้ตัวเลือกที่ถูก
51	.597	11.98	ดีมาก	.312	พอใช้	
52	.705	10.80	พอใช้	.438	ดี	
53	.381	14.25	ดี	.109	ต่ำ	
54	.835	9.07	ง่ายไป	.322	พอใช้	1 ไม่มีประสิทธิภาพ
55	.849	8.83	ง่ายไป	.263	พอใช้	3,4 ไม่มีประสิทธิภาพ
56	.914	7.51	ง่ายไป	.279	พอใช้	2,4 ไม่มีประสิทธิภาพ, แก้ตัวเลือกที่ถูก
57	.493	13.04	ดีมาก	.341	พอใช้	
58	.727	10.55	พอใช้	.262	พอใช้	
59	.755	10.19	พอใช้	.266	พอใช้	
60	.360	14.48	ดี	-.071	ต่ำ	
61	.835	9.07	ง่ายไป	.381	ดี	
62	.345	14.63	ดี	.222	พอใช้	

ตารางที่ 6 (ต่อ)

ข้อที่	ความยากง่าย		อำนาจจำแนก		ตัวเลือกที่ต้องปรับปรุง	
	ขนาด	คุณภาพ	ขนาด	คุณภาพ		
	p	Δ				
63	.863	8.58	ง่ายไป	.317	พอใช้	2 ไม่มีประสิทธิภาพ
64	.791	9.72	พอใช้	.267	พอใช้	
65	.403	14.02	ดีมาก	.109	ต่ำ	
66	.029	20.64	ยากไป	-.144	ต่ำ	4 $N_H > N_L$
67	.727	10.55	พอใช้	.312	พอใช้	
68	.849	8.83	ง่ายไป	.330	พอใช้	1 ไม่มีประสิทธิภาพ, 4 $N_M > N_L$
69	.554	12.42	ดีมาก	.012	ต่ำ	
70	.158	17.05	ยากไป	.288	พอใช้	1 $N_H > N_L$
71	.468	13.36	ดีมาก	.177	ต่ำ	
72	.734	10.46	พอใช้	.468	ดี	
73	.626	11.68	ดี	.304	พอใช้	
74	.655	11.37	ดี	.488	ดี	
75	.273	15.45	พอใช้	.245	พอใช้	
76	.547	12.49	ดีมาก	.324	พอใช้	
77	.367	14.40	ดี	.315	พอใช้	
78	.367	12.49	ดี	.267	พอใช้	
79	.827	9.19	ง่ายไป	.316	พอใช้	4 ไม่มีประสิทธิภาพ
80	.576	12.20	ดีมาก	.227	พอใช้	
81	.561	12.34	ดีมาก	.258	พอใช้	
82	.252	15.71	พอใช้	.169	ต่ำ	

ตารางที่ 6 (ต่อ)

ข้อที่	ความยากง่าย		อำนาจจำแนก		ตัวเลือกที่ต้องปรับปรุง	
	ขนาด	คุณภาพ	ขนาด	คุณภาพ		
	P	Δ				
83	.518	12.78	ดีมาก	.112	ต่ำ	
84	.914	7.51	ง่ายไป	.315	พอใช้	1,2,3 ไม่มีประสิทธิภาพ, แก้ตัวเลือกที่ถูก
85	.745	14.32	พอใช้	.384	ดี	
86	.302	15.11	พอใช้	.054	ต่ำ	
87	.288	15.28	พอใช้	.210	พอใช้	
88	.482	13.22	ดีมาก	.300	พอใช้	
89	.460	13.44	ดีมาก	.320	พอใช้	
90	.482	13.22	ดีมาก	.349	พอใช้	
91	.604	11.90	ดีมาก	.533	ดี	
92	.763	10.10	พอใช้	.352	ดี	
93	.532	12.64	ดีมาก	.442	ดี	
94	.504	12.96	ดีมาก	.205	พอใช้	
95	.813	9.41	ง่ายไป	.321	พอใช้	1 ไม่มีประสิทธิภาพ
96	.705	10.80	พอใช้	.192	ต่ำ	
97	.835	9.07	ง่ายไป	.315	พอใช้	1,2 ไม่มีประสิทธิภาพ
98	.302	15.11	พอใช้	.303	พอใช้	
99	.885	8.16	ง่ายไป	.244	พอใช้	3 ไม่มีประสิทธิภาพ
100	.683	11.05	ดี	.329	พอใช้	
101	.871	8.51	ง่ายไป	.211	พอใช้	4 ไม่มีประสิทธิภาพ
102	.237	15.90	พอใช้	.123	ต่ำ	

ตารางที่ 6 (ต่อ)

ข้อที่	ความยากง่าย		อำนาจจำแนก		ตัวเลือกที่ต้องปรับปรุง	
	ขนาด	คุณภาพ	ขนาด	คุณภาพ		
	p	Δ				
103	.827	9.19	ง่ายไป	.352	ดี	$3 N_M > N_L$
104	.741	10.37	พอใช้	.364	ดี	
105	.281	15.36	พอใช้	.368	ดี	
106	.662	11.29	ดี	.217	พอใช้	
107	.777	9.91	พอใช้	.342	พอใช้	
108	.799	9.61	พอใช้	.308	พอใช้	
109	.540	12.56	ดีมาก	.287	พอใช้	
110	.597	11.98	ดีมาก	.269	พอใช้	
111	.432	13.73	ดีมาก	.153	ต่ำ	
112	.511	12.96	ดีมาก	.406	ดี	
113	.374	14.32	ดี	.240	พอใช้	
114	.496	13.04	ดีมาก	.355	ดี	
115	.770	10.01	พอใช้	.305	พอใช้	
116	.439	13.66	ดีมาก	.041	ต่ำ	
117	.410	13.95	ดีมาก	.261	พอใช้	
118	.525	12.71	ดีมาก	.308	พอใช้	
119	.540	12.56	ดีมาก	.269	พอใช้	
120	.309	15.03	พอใช้	.119	ต่ำ	

สรุป แบบทดสอบสมิทธิภาพทั่วไปภาษาอังกฤษฉบับที่ 4 มีข้อกระทงที่อยู่ในเกณฑ์ 100 ข้อ ต้องปรับปรุง 20 ข้อ ยากไป 7 ข้อ ง่ายไป 13 ข้อ ค่าอำนาจจำแนกติดลบ 3 ข้อ

1.3 ผลการวิเคราะห์ข้อทดสอบอัตนัย

ข้อทดสอบสมิทธิภาพทั่วไปภาษาอังกฤษ อัตนัย วัดความสามารถในการเขียน มี 3 ตอน คือวัดความสามารถในการเขียนระดับอนุเจต 2 ตอน และระดับเรียงความ 1 ตอน การวิเคราะห์ใช้วิธีของ Whitney และ Sabers (ตอนที่ 3) ซึ่งได้ผลการวิเคราะห์ดังนี้ ตารางที่ 7 ผลการวิเคราะห์ข้อทดสอบสมิทธิภาพทั่วไปภาษาอังกฤษ ส่วนอัตนัย

แบบทดสอบฉบับที่		ค่าความยากง่าย		คุณภาพ	ค่าอำนาจจำแนก	คุณภาพ
		p	Δ			
1	ข้อที่ 1	.716	15.28	พอใช้	.287	พอใช้
	ข้อที่ 2	.504	13.04	ดีมาก	.370	ดี
	ข้อที่ 3	.520	13.24	ดี	.504	ดี
	ค่าเฉลี่ย	.580	13.85	ดีมาก	.387	ดี
2	ข้อที่ 1	.490	13.12	ดี	.340	พอใช้
	ข้อที่ 2	.548	13.30	ดี	.336	พอใช้
	ข้อที่ 3	.228	15.96	พอใช้	.360	ดี
	ค่าเฉลี่ย	.422	14.13	ดีมาก	.345	พอใช้
3	ข้อที่ 1	.477	13.24	ดีมาก	.327	พอใช้
	ข้อที่ 2	.353	14.52	ดี	.437	ดี
	ข้อที่ 3	.111	17.88	ยากไป	.210	พอใช้
	ค่าเฉลี่ย	.314	15.21	ดี	.325	พอใช้
4	ข้อที่ 1	.622	14.24	ดี	.287	พอใช้
	ข้อที่ 2	.604	16.38	ดีมาก	.308	พอใช้
	ข้อที่ 3	.412	13.92	ดีมาก	.487	ดี
	ค่าเฉลี่ย	.546	14.85	ดีมาก	.361	ดี

จากตารางที่ 7 เห็นได้ว่าข้อสอบสมิทริภาพทั่วไปภาษาอังกฤษ ในส่วนของข้อสอบอัตนัย ของทั้ง 4 ฉบับ รวม 12 ข้อ มีเพียง 1 ข้อ คือ การเขียนแบบเรียงความของแบบทดสอบฉบับ ที่ 3 ที่ค่าความยากง่ายต่ำกว่าเกณฑ์ คือมีความยากเกินไปต้องปรับปรุง ($p = .111$, $\Delta = 17.88$)

ค่าความเที่ยงของข้อสอบสมิทริภาพทั่วไปภาษาอังกฤษ ในส่วนของข้อสอบอัตนัย วัดความสามารถทางการเขียน ใช้วิธีหาค่าความเที่ยงแบบ inter-rater หรือค่า สหสัมพันธ์ของคะแนนที่ผู้ตรวจ 2 คน ให้คะแนนงานเขียนแต่ละฉบับ เป็นดังนี้

ตารางที่ 8 ค่าความเที่ยงแบบ inter-rater

แบบทดสอบฉบับที่	ข้อสอบอัตนัยข้อที่		
	1	2	3
1	.741	.829	.834
2	.594	.730	.800
3	.727	.808	.684
4	.778	.685	.840

จากตารางที่ 8 เห็นได้ว่าค่าความเที่ยงของข้อสอบอัตนัย มีตั้งแต่ระดับสูง - ปานกลาง ($r_{xx} = .84 - .594$) โดยที่ข้อสอบอัตนัยข้อที่ 3 ของแบบทดสอบฉบับที่ 4 มี ค่าความเที่ยงระดับปานกลาง

สรุป ข้อสอบสมิทริภาพทั่วไปทางภาษาอังกฤษ ในส่วนของข้อสอบอัตนัยวัดความ สามารถในการเขียน ที่ต้องปรับปรุง คือข้อที่ 3 ของฉบับที่ 3 อีก 11 ข้ออยู่ในเกณฑ์ยอมรับได้

ตอนที่ 2 ผลการวิเคราะห์ข้อทดสอบสมมติภาพทั่วไปทางภาษาอังกฤษฉบับปรับปรุง

จากผลการวิเคราะห์ในตอนที่ 1 ข้อทดสอบสมมติภาพทั่วไปทางภาษาอังกฤษได้ปรับปรุงและพัฒนาขึ้นจากข้อทดสอบที่มีความยากง่ายเกินเกณฑ์ .20-.80 และมีค่าอำนาจจำแนกต่ำมากจนคิดลบ รวมเป็นข้อสอบจำนวน 63 ข้อ ได้ทดลองใช้กับผลวิจัยในกลุ่มที่ 2 จำนวน 75 คน ในจำนวนนี้มีผู้ทำข้อสอบอัตโนมัติ 12 คน

2.1 คุณภาพของแบบทดสอบ

ตารางที่ 9 ผลการวิเคราะห์คุณภาพของแบบทดสอบสมมติภาพทั่วไปทางภาษาอังกฤษฉบับปรับปรุง

ข้อสอบ	N	\bar{X}	SD	CV	Max	Min	Range	ค่าความยากง่าย		อำนาจจำแนก		ค่าความเที่ยง		อำนาจทำนาย	SE _m	สรุปคุณภาพ
								ค่า p	Δ	r_{pb}	KR ₂₀	Phi	E			
ปรนัย	75	33.102	6.572	19.854	52	19	33	0.487	13.235	.677	.727	.676**	31.30	3.434	ดี	
อัตนัย	12	6.502	1.878	28.883	8	3	5	0.503	13.04	D .360	.826	—	43.60	.783	ดีมาก	

** นัยสำคัญทางสถิติที่ $p < .01$

สรุปได้ว่าข้อทดสอบสมมติภาพทั่วไปทางภาษาอังกฤษฉบับปรับปรุง เฉพาะปรนัยมีค่าความยากอยู่ในเกณฑ์ดี ค่าอำนาจจำแนกอยู่ในเกณฑ์ดี ค่าความเที่ยง Phi ของข้อสอบอิงเกณฑ์ อยู่ในเกณฑ์สูง มีนัยสำคัญที่ระดับ .01 ส่วนข้อทดสอบแบบอัตนัยมีค่าความเที่ยง α .826 ซึ่งอยู่ในเกณฑ์ดี

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

2.2 ผลการวิเคราะห์ข้อทดสอบแบบปรนัย

ตารางที่ 10 ผลการวิเคราะห์ข้อกระทงแบบทดสอบสมิทธิภาพทั่วไปทางภาษาอังกฤษฉบับปรับปรุง

ข้อที่	ความยากง่าย		อำนาจจำแนก		ตัวเลือกที่ต้องปรับปรุง	
	ขนาด p	คุณภาพ Δ	ขนาด	คุณภาพ		
1	.227	16.04	พอใช้	.169	ต่ำ	
2	.307	15.06	พอใช้	.833	ดีมาก	
3	.253	15.70	พอใช้	.484	ดี	
4	.747	10.30	พอใช้	.892	ดีมาก	
5	.640	11.53	ดี	.551	ดี	
6	.040	20.04	ยากไป	.328	พอใช้	
7	.293	15.21	พอใช้	.572	ดี	
8	.680	11.09	ดี	.200	พอใช้	
9	.147	17.40	ยากไป	.411	ดี	1,2 $N_H > N_L$
10	.307	15.06	พอใช้	.599	ดี	
11	.400	14.05	ดีมาก	.809	ดีมาก	
12	.733	10.47	พอใช้	.837	ดีมาก	
13	.560	12.36	ดีมาก	.809	ดีมาก	
14	.640	11.53	ดี	.939	ดีมาก	
15	.747	10.30	พอใช้	.510	ดี	
16	.493	13.04	ดีมาก	.980	ดีมาก	
17	.120	17.74	ยากไป	.287	พอใช้	1,4 $N_H > N_L$
18	.200	16.41	พอใช้	.432	ดี	
19	.120	17.74	ยากไป	.462	ดี	1,4 $N_H > N_L$
20	.427	13.78	ดี	.370	ดี	

ตารางที่ 10 (ต่อ)

ข้อที่	ความยากง่าย		อำนาจจำแนก		ตัวเลือกที่ต้องปรับปรุง	
	ขนาด	คุณภาพ	ขนาด	คุณภาพ		
	p	Δ				
21	.227	16.04	พอใช้	.557	ดี	
22	.160	17.02	ยากไป	.568	ดี	แก้ตัวเลือกที่ถูกต้อง
23	.800	9.59	พอใช้	.927	ดีมาก	
24	.967	8.52	ง่ายไป	.034	ต่ำ	4 ไม่มีประสิทธิภาพ
25	.040	20.04	ยากไป	.027	ต่ำ	1 $N_H > N_L$, 3 ไม่มีประสิทธิภาพ
26	.253	15.07	พอใช้	.503	ดี	
27	.440	13.64	ดี	.352	ดี	
28	.693	10.94	พอใช้	.624	ดีมาก	
29	.413	13.92	ดีมาก	.388	ดี	
30	.773	9.96	พอใช้	.795	ดีมาก	
31	.773	9.96	พอใช้	.440	ดีมาก	
32	.440	13.64	ดีมาก	.289	พอใช้	
33	.147	17.40	ยากไป	.417	ดี	3 $N_H > N_L$
34	.707	10.79	พอใช้	.783	ดีมาก	
35	.587	12.08	ดีมาก	.617	ดีมาก	
36	.213	16.22	พอใช้	.649	ดีมาก	
37	.640	11.53	ดี	.466	ดี	
38	.373	14.33	ดี	.510	ดี	
39	.587	12.08	ดีมาก	.967	ดีมาก	
40	.560	12.36	ดีมาก	.216	พอใช้	
41	.707	10.79	พอใช้	.961	ดีมาก	
42	.707	10.79	พอใช้	.081	ต่ำ	

ตารางที่ 10 (ต่อ)

ข้อที่	ความยากง่าย		อำนาจจำแนก		ตัวเลือกที่ต้องปรับปรุง	
	ขนาด	คุณภาพ	ขนาด	คุณภาพ		
	p	Δ				
43	.827	9.20	ง่ายไป	.054	ต่ำ	
44	.693	10.94	ดี	.270	พอใช้	
45	.813	9.40	ง่ายไป	.027	ต่ำ	1 ไม่มีประสิทธิภาพ
46	.560	12.36	ดีมาก	.503	ดี	
47	.360	14.47	ดี	.798	ดีมาก	
48	.667	11.24	ดี	.783	ดีมาก	
49	.387	14.19	ดี	.881	ดีมาก	
50	.480	13.24	ดีมาก	.847	ดีมาก	
51	.453	13.51	ดี	.778	ดีมาก	
52	.307	15.06	พอใช้	.498	ดีมาก	
53	.827	9.20	ง่ายไป	.075	ดี	1,3 ไม่มีประสิทธิภาพ
54	.227	16.04	พอใช้	.232	ต่ำ	
55	.347	14.62	พอใช้	.473	พอใช้	
56	.373	14.33	พอใช้	.133	ดี	
57	.773	9.96	พอใช้	.834	ต่ำ	
58	.773	9.96	พอใช้	.897	ดีมาก	
59	.307	15.06	พอใช้	.806	ดีมาก	
60	.253	15.70	พอใช้	.684	ดีมาก	
61	.773	9.96	พอใช้	.115	ต่ำ	
62	.667	11.24	ดี	.603	ดีมาก	
63	.600	11.95	ดี	.179	ต่ำ	

จากตารางที่ 10 เห็นได้ว่าแบบทดสอบสมิทริภาพทั่วไปทางภาษาอังกฤษฉบับปรับปรุง มีข้อทดสอบที่ใช้ได้ตามเกณฑ์ .20 - .80 และค่าอำนาจจำแนกไม่ติดลบ 55 ข้อ จะเห็นได้ว่า ข้อกระทงทุกข้อไม่มีค่าอำนาจจำแนกติดลบ ยังมีข้อกระทงที่ต้องแก้ไขปรับปรุงอีก 9 ข้อ ยากไป 6 ข้อ ง่ายไป 3 ข้อ ข้อสอบที่ควรปรับปรุงเป็นข้อสอบ การฟัง 3 ข้อ การอ่าน 6 ข้อ สำหรับ ข้อสอบการเขียนแบบปรนัย (ข้อกระทงที่ 57-63) ไม่ต้องปรับปรุงอีก เพราะค่าความยากง่าย อยู่ในเกณฑ์ ดี - พอใช้ และค่าเฉลี่ยอำนาจจำแนก .588 ซึ่งอยู่ในเกณฑ์ดีมาก

ตอนที่ 3 ผลการวิเคราะห์ข้อทดสอบแบบ Rasch Model เพื่อพัฒนาเป็นข้อทดสอบเก็บไว้ในคลังข้อสอบ



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 11 ผลการวิเคราะห์ข้อกระทงโดยใช้ Rasch Model ของแบบทดสอบสมิทธิภาพทั่วไป
ทางภาษาอังกฤษ ฉบับที่ 1

ITEM NAME	ITEM DIFF	ERR IMPAC	FIT T-TEST		WEIGHTED	
			BETWN	TOTAL	MNSQ	DF
A74	-0.15	0.00	2.50	-0.92	0.86	88
A73	-0.33	0.00	2.48	-0.91	0.86	91
B05	-0.62	0.00	2.66	-0.84	0.87	92
A79	-1.27	0.00	3.04	-0.84	0.87	83
A72	-0.96	0.00	1.77	-0.77	0.88	90
A83	-0.96	0.00	1.63	-0.63	0.90	90
A60	-1.05	0.00	2.57	-0.63	0.90	88
B06	-0.75	0.00	1.19	-0.62	0.90	82
A61	-0.07	0.00	1.64	-0.55	0.91	83
B13	-1.46	0.00	0.89	-0.53	0.91	78
A77	-1.05	0.00	-0.03	-0.48	0.92	88
A22	-0.83	0.00	1.90	-0.47	0.93	91
A81	0.32	0.00	1.40	-0.45	0.92	77
A54	-1.41	0.00	1.64	-0.44	0.92	80
A90	-0.87	0.00	0.38	-0.44	0.93	91
A76	-1.72	0.00	1.71	-0.43	0.92	70
A22	-0.62	0.00	1.16	-0.42	0.93	92
B20	0.37	0.00	1.32	-0.39	0.93	75
A69	-1.84	0.00	0.26	-0.38	0.93	66
B16	0.82	0.00	-0.24	-0.38	0.92	60
A3	-0.02	0.00	-0.65	-0.37	0.94	86
A62	-1.78	0.00	2.06	-0.36	0.93	68
A2	1.03	0.00	0.13	-0.35	0.92	53

ตารางที่ 11 (ต่อ)

ITEM NAME	ITEM DIFF	ERR IMPAC	FIT T-TEST		WEIGHTED	
			BETWN	TOTAL	MNSQ	DF
A8	-1.09	0.00	0.67	-0.35	0.94	87
A88	1.54	0.00	-0.30	-0.34	0.91	37
B08	-1.56	0.00	0.69	-0.33	0.94	75
B12	-0.45	0.00	-0.24	-0.32	0.95	92
B04	-0.87	0.00	0.27	-0.32	0.95	91
A55	-0.92	0.00	0.83	-0.31	0.95	90
B00	-0.49	0.00	0.90	-0.31	0.95	92
B18	-0.37	0.00	-0.56	-0.31	0.95	91
A1	1.77	0.00	0.63	-0.31	0.90	31
A68	-2.15	0.00	0.71	-0.30	0.93	56
B07	-1.05	0.00	0.02	-0.30	0.95	88
A66	-0.54	0.00	0.12	-0.28	0.95	92
A98	-0.45	0.00	0.35	-0.27	0.95	92
B09	-2.64	0.00	0.45	-0.26	0.93	40
A38	0.70	0.00	-2.16	-0.23	0.95	64
B18	-1.56	0.00	1.29	-0.22	0.96	75
A80	-1.27	0.00	-1.05	-0.21	0.96	83
A92	0.12	0.00	-0.35	-0.20	0.96	82
A44	-0.92	0.00	-0.06	-0.19	0.96	90
A89	-0.33	0.00	-0.30	-0.19	0.96	91
A91	-0.41	0.00	-0.18	-0.19	0.97	91
A34	0.96	0.00	0.36	-0.17	0.96	56
A93	-0.96	0.00	-0.75	-0.16	0.97	90

ตารางที่ 11 (ต่อ)

ITEM NAME	ITEM DIFF	ERR IMPAC	FIT T-TEST		WEIGHTED	
			BETWN	TOTAL	MNSQ	DF
A82	-0.49	0.00	0.17	-0.14	0.97	92
B03	-0.87	0.00	-2.25	-0.13	0.97	91
B15	-2.22	0.00	0.32	-0.12	0.96	53
A21	-0.92	0.00	-0.86	-0.12	0.97	90
A65	-2.46	0.00	0.20	-0.10	0.96	45
A99	-1.37	0.00	2.07	-0.10	0.98	81
A70	-1.05	0.00	0.48	-0.09	0.98	88
B17	-1.05	0.00	0.78	-0.09	0.98	88
A63	-1.78	0.00	0.55	-0.07	0.98	68
B11	-0.96	0.00	-1.40	-0.07	0.98	90
A34	-0.20	0.00	-0.80	-0.06	0.98	89
A30	1.35	0.00	0.34	-0.05	0.98	43
A12	0.32	0.00	-0.70	-0.02	0.99	77
A37	0.70	0.00	-0.33	-0.02	0.99	64
A96	-1.46	0.00	-0.62	0.01	0.99	78
A6	0.89	0.00	-1.16	0.02	0.99	58
A29	2.40	0.00	1.64	0.08	0.99	18
A85	-2.46	0.00	-1.50	0.08	1.00	45
A64	-0.33	0.00	-0.10	0.09	1.01	91
A50	2.64	0.00	1.07	0.10	0.99	14
A97	-0.37	0.0	-1.24	0.10	1.01	91
A47	2.94	0.00	1.24	0.13	1.00	10
A35	3.35	0.00	1.26	0.14	0.98	6

ตารางที่ 11 (ต่อ)

ITEM NAME	ITEM DIFF	ERR IMPAC	FIT T-TEST		WEIGHTED	
			BETWN	TOTAL	MNSQ	DF
A43	2.64	0.00	0.52	0.14	1.01	14
A38	2.40	0.00	-0.06	0.14	1.01	18
A51	3.35	0.00	1.26	0.14	0.98	6
A40	1.65	0.01	0.95	0.17	1.02	34
A5	0.96	0.01	1.76	0.17	1.02	56
A36	0.86	0.01	-0.41	0.17	1.02	58
A24	2.21	0.01	-0.39	0.18	1.02	21
A4	-1.09	0.01	-0.84	0.19	1.02	87
A67	-0.02	0.01	-1.87	0.19	1.02	86
A18	1.10	0.01	-0.10	0.20	1.03	51
A87	-3.24	0.01	3.91	0.21	1.03	24
A45	4.05	0.02	1.12	0.23	0.97	3
A53	1.77	0.02	-0.10	0.24	1.04	31
A38	4.05	0.00	1.12	0.24	0.98	3
B14	0.70	0.02	0.44	0.25	1.03	64
A86	-1.62	0.00	0.29	0.26	1.03	74
A31	-0.28	0.02	-1.66	0.27	1.03	90
A28	1.54	0.02	1.65	0.28	1.05	37
A14	1.90	0.03	2.42	0.28	1.05	28
A21	1.54	0.02	1.32	0.28	1.05	37
A17	-0.58	0.02	0.85	0.30	1.04	92
A10	1.10	0.02	1.16	0.30	1.05	51
A78	-1.37	0.02	1.05	0.31	1.04	81

ตารางที่ 11 (ต่อ)

ITEM NAME	ITEM DIFF	ERR IMPAC	FIT T-TEST		WEIGHTED	
			BETWN	TOTAL	MNSQ	DF
A58	1.77	0.03	1.69	0.33	1.06	31
A20	0.64	0.02	-0.56	0.33	1.05	66
B10	-1.32	0.02	0.56	0.33	1.04	82
A95	-1.18	0.02	0.65	0.34	1.04	86
A16	-0.58	0.02	-0.15	0.34	1.04	92
A32	1.35	0.03	1.76	0.38	1.07	43
A41	1.10	0.03	1.26	0.40	1.07	51
A7	-0.70	0.03	0.68	0.43	1.06	92
A94	0.47	0.03	1.95	0.44	1.07	72
A59	0.37	0.03	0.93	0.47	1.07	75
A13	0.76	0.04	0.30	0.48	1.08	62
A42	1.10	0.04	1.03	0.49	1.09	51
A11	-0.66	0.03	0.18	0.49	1.07	92
A48	1.10	0.04	2.32	0.51	1.09	51
A71	-0.02	0.04	0.11	0.51	1.07	86
A52	-1.37	0.04	1.67	0.52	1.07	81
A15	0.96	0.04	1.70	0.52	1.09	56
A57	0.22	0.04	1.16	0.54	1.08	80
A49	1.44	0.05	2.88	0.55	1.11	40
A75	-0.33	0.04	1.50	0.57	1.08	91
A19	0.37	0.04	1.74	0.60	1.09	75
A23	0.76	0.05	0.54	0.60	1.10	62
A26	0.58	0.05	2.65	0.63	1.10	68



ตารางที่ 11 (ต่อ)

ITEM NAME	ITEM DIFF	ERR IMPAC	FIT T-TEST		WEIGHTED	
			BETWN	TOTAL	MNSQ	DF
A56	-0.02	0.05	1.75	0.68	1.10	86
A9	0.70	0.06	2.47	0.72	1.12	64
A25	0.47	0.06	3.39	0.72	1.11	72
A46	0.89	0.08	4.25	0.92	1.17	58
A27	-0.62	0.07	1.97	0.94	1.14	92

จากตารางที่ 11 เมื่อพิจารณาจากความยากง่าย (β) ปรากฏว่ามีข้อสอบที่มีค่าความยากง่าย (β) เกิน ± 3 อยู่จำนวน 5 ข้อ ได้แก่ A35 A51 A87 A45 A38 ซึ่งมีค่า Item Diff เท่ากับ 3.35 3.35 -3.24 4.05 และ 4.05 ตามลำดับ

ความสอดคล้องของข้อทดสอบกับโด่ง ICC เมื่อพิจารณาจากค่า Weighted mean square สำหรับแบบทดสอบฉบับนี้ ปรากฏว่ามีข้อสอบที่มีค่า Weighted mean square เกิน $1 + 0.07$ จำนวน 15 ข้อ ได้แก่ข้อ A13 A42 A48 A15 A57 A49 A75 A19 A23 A26 A56 A9 A25 A46 และ A27 เป็นข้อสอบการฟัง 8 ข้อ คือ A9 A13 A15 A19 A23 A25 A26 A27 ข้อสอบการอ่าน 7 ข้อ คือ A42 A46 A48 A49 A56 A57 และ A75

โดยสรุปการดำเนินการสร้างแบบทดสอบสมรรถภาพทั่วไปทางภาษาอังกฤษฉบับที่ 1 มีข้อสอบที่มีความสอดคล้องกับเกณฑ์ของข้อทดสอบที่ดี และสามารถนำมาจัดเก็บไว้ในคลังข้อสอบได้จำนวน 105 ข้อ หรือคิดเป็นร้อยละ 87.50

ตารางที่ 12 ผลการวิเคราะห์ข้อกระทบโดยใช้ Rasch Model ของแบบทดสอบสมิทธิภาพทั่วไป
ทางภาษาอังกฤษ ฉบับที่ 2

ITEM NAME	ITEM DIFF	ERR IMPAC	FIT T-TEST		WEIGHTED	
			BETWN	TOTAL	MNSQ	DF
B09	-0.46	0.00	2.71	-1.17	0.83	89
A35	0.57	0.00	2.01	-1.13	0.83	84
A55	1.63	0.00	1.30	-1.05	0.80	54
A74	-0.21	0.00	2.84	-0.92	0.86	92
A70	0.18	0.00	1.58	-0.90	0.87	90
A57	1.02	0.00	1.42	-0.88	0.85	72
A20	-0.21	0.00	1.11	-0.84	0.98	92
A71	-0.73	0.00	1.62	-0.78	0.88	85
A6	0.76	0.00	0.96	-0.73	0.88	79
B20	-0.82	0.00	1.19	-0.72	0.89	83
A90	0.53	0.00	0.53	-0.72	0.89	85
A9	1.18	0.00	0.20	-0.68	0.88	67
A42	0.97	0.00	-0.75	-0.64	0.88	74
A78	-1.45	0.00	1.79	-0.55	0.90	64
B04	0.30	0.00	-0.07	-0.54	0.91	89
A17	1.78	0.00	-0.97	-0.52	0.89	49
B13	-0.38	0.00	0.81	-0.50	0.92	90
A69	-1.01	0.00	0.87	-0.49	0.91	78
A5	-0.33	0.00	0.22	-0.49	0.92	91
A13	2.13	0.00	0.07	-0.49	0.88	39
A7	-1.01	0.00	-0.04	-0.48	0	78
A68	1.70	0.00	-0.74	-0.46	0.90	51
A19	-1.17	0.00	1.62	-0.42	0.92	73

ตารางที่ 12 (ต่อ)

ITEM NAME	ITEM DIFF	ERR IMPAC	FIT T-TEST		WEIGHTED	
			BETWN	TOTAL	MNSQ	DF
A52	-1.57	0.00	1.19	-0.41	0.92	60
A64	-1.45	0.00	-0.18	-0.40	0.92	64
A89	0.18	0.00	0.32	-0.40	0.94	90
B05	1.36	0.00	0.79	-0.38	0.92	62
A58	0.35	0.00	0.78	-0.38	0.92	88
A45	-0.77	0.00	-0.16	-0.38	0.94	84
A98	-1.39	0.00	0.45	-0.38	0.94	66
A18	0.18	0.00	0.23	-0.37	0.93	90
A54	-0.04	0.00	1.46	-0.37	0.94	92
A75	0.97	0.00	0.22	-0.35	0.94	74
A87	0.53	0.00	1.29	-0.34	0.94	85
A66	-1.51	0.00	0.66	-0.33	0.93	62
A78	-0.87	0.00	0.88	-0.32	0.94	82
A91	-0.92	0.00	-0.14	-0.31	0.94	80
A88	0.62	0.00	-1.74	-0.30	0.95	83
A59	0.53	0.00	0.05	-0.29	0.95	85
A12	0.97	0.00	-0.30	-0.28	0.95	74
A53	0.35	0.00	-1.71	-0.28	0.95	88
A63	-1.78	0.00	-0.40	-0.27	0.94	53
A4	-0.59	0.00	-0.43	-0.27	0.95	87
A34	-1.45	0.00	0.70	-0.27	0.94	64
A26	0.26	0.00	-0.59	-0.26	0.95	89
A1	-2.01	0.00	0.84	-0.26	0.93	45

ตารางที่ 12 (ต่อ)

ITEM NAME	ITEM DIFF	ERR IMPAC	FIT T-TEST		WEIGHTED	
			BETWN	TOTAL	MNSQ	DF
A72	-1.78	0.00	0.63	-0.25	0.94	53
A24	1.43	0.00	0.27	-0.25	0.94	60
A27	-0.12	0.00	-0.62	-0.23	0.96	92
A97	-0.42	0.00	0.18	-0.23	0.96	90
A84	0.00	0.00	0.06	-0.22	0.96	92
A81	0.62	0.00	-0.14	-0.19	0.96	83
B08	-0.29	0.00	0.04	-0.19	0.96	91
A98	1.43	0.00	-0.87	-0.18	0.96	60
A94	0.71	0.00	-1.83	-0.17	0.96	81
B18	0.67	0.00	-0.04	-0.17	0.97	82
A3	-0.77	0.00	-0.81	-0.16	0.97	84
A49	-0.16	0.00	0.08	-0.14	0.97	92
A60	-1.71	0.00	-0.41	-0.14	0.96	55
A38	-0.77	0.00	-2.11	-0.14	0.97	84
A43	-0.42	0.00	-1.24	-0.12	0.97	90
A67	1.43	0.00	1.68	-0.11	0.97	60
A95	-0.92	0.00	-0.54	-0.11	0.97	80
A99	-0.08	0.00	1.75	-0.10	0.98	92
A76	-0.08	0.00	1.62	-0.08	0.98	92
A62	-2.20	0.00	-0.42	-0.17	0.97	40
A96	-2.01	0.00	0.44	-0.05	0.98	45
A36	-0.87	0.00	-0.45	-0.05	0.98	82
B12	0.00	0.00	-1.51	-0.04	0.99	92

ตารางที่ 12 (ต่อ)

ITEM NAME	ITEM DIFF	ERR IMPAC	FIT T-TEST		WEIGHTED	
			BETWN	TOTAL	MNSQ	DF
A16	0.91	0.00	-0.32	-0.02	0.99	75
A29	-0.64	0.00	-0.42	-0.01	0.99	87
A46	0.48	0.00	0.96	-0.01	0.99	86
A85	-2.40	0.00	-0.59	0.00	0.98	34
A65	-1.85	0.00	0.50	0.02	0.99	51
B11	0.76	0.00	-1.03	0.03	1.00	79
A10	0.48	0.00	-1.34	0.05	1.00	86
A13	-0.46	0.00	0.50	0.05	1.00	89
A21	-0.64	0.00	-0.42	0.07	1.00	87
A83	2.03	0.00	0.85	0.12	1.01	42
A80	-2.01	0.01	-0.32	0.13	1.01	45
A44	-2.01	0.01	-0.79	0.13	1.01	45
A2	0.00	0.01	-0.34	0.14	1.01	92
B15	1.13	0.01	0.12	0.14	1.02	69
A50	-0.96	0.01	-0.20	0.18	1.02	79
A23	2.58	0.02	-0.21	0.21	1.03	29
A38	-0.38	0.01	0.49	0.21	1.02	90
A47	0.30	0.02	-0.21	0.25	1.03	89
B07	0.86	0.02	-0.15	0.25	1.03	77
B03	0.13	0.02	0.77	0.28	1.03	91
A3	0.67	0.02	0.15	0.28	1.04	82
A92	-0.33	0.02	0.89	0.29	1.04	91
A22	-0.42	0.02	0.21	0.31	1.04	90

ตารางที่ 12 (ต่อ)

ITEM NAME	ITEM DIFF	ERR IMPAC	FIT T-TEST		WEIGHTED	
			BETWN	TOTAL	MNSQ	DF
A11	1.02	0.02	0.29	0.31	1.04	72
A37	-0.25	0.02	0.13	0.33	1.04	91
A22	1.30	0.03	0.01	0.35	1.05	64
A56	0.26	0.03	-0.32	0.43	1.06	89
B14	0.97	0.03	1.74	0.43	1.06	74
B10	0.18	0.03	0.11	0.45	1.06	90
A33	-0.08	0.03	0.25	0.45	1.06	92
A31	0.35	0.03	0.50	0.51	1.07	88
B18	1.18	0.04	0.63	0.54	1.00	67
A40	-1.12	0.04	0.70	0.55	1.08	75
A30	-0.12	0.04	0.08	0.60	1.08	92
A61	-0.46	0.05	0.93	0.64	1.09	89
A86	0.44	0.05	0.84	0.65	1.09	87
B06	-0.12	0.05	1.23	0.70	1.10	92
A82	0.35	0.06	-0.53	0.74	1.11	88
A21	0.86	0.06	2.80	0.75	1.12	77
A14	1.24	0.06	1.73	0.75	1.13	66
A77	-1.17	0.06	2.68	0.7	1.12	73
A15	0.26	0.06	1.40	0.85	1.12	89
B16	1.36	0.09	3.32	1.10	1.20	62
A41	-0.87	0.09	2.98	1.10	1.10	82
A51	2.03	0.13	3.71	1.21	1.27	42
B17	0.26	0.09	2.80	1.22	1.18	89

ตารางที่ 12 (ต่อ)

ITEM NAME	ITEM DIFF	ERR IMPAC	FIT T-TEST		WEIGHTED	
			BETWN	TOTAL	MNSQ	DF
A25	1.07	0.10	2.50	1.24	1.21	71
B00	0.44	0.09	2.52	1.26	1.19	87
A32	-0.42	0.11	3.65	1.52	1.24	90
A28	1.07	0.14	2.50	1.65	1.29	71
A48	1.13	0.18	3.54	2.12	1.39	69

จากตารางที่ 12 เมื่อพิจารณาค่าความยากง่าย (β) ปรากฏว่าไม่มีข้อสอบที่มีค่าความยากง่าย (β) เกิน ± 3 เลย

ความสอดคล้องของข้อทดสอบกับโค้ง ICC เมื่อพิจารณาจากค่า Weighted mean square สำหรับแบบทดสอบฉบับนี้ ปรากฏว่ามีข้อสอบที่มีค่า Weighted mean square เกิน $1 + 0.10$ จำนวน 14 ข้อ ได้แก่ข้อ A82 A21 A14 A77 A15 B16 A41 A51 B17 A25 B00 A32 A28 และ A48 เป็นข้อสอบการฟัง 6 ข้อ คือ A14 A15 A21 A25 A28 A32 ข้อสอบการอ่าน 5 ข้อ คือ A41 A48 A51 A77 A82 เป็นข้อสอบการเขียน 3 ข้อ คือ B00 B16 B17

โดยสรุปการดำเนินการสร้างแบบทดสอบสมรรถภาพทั่วไปทางภาษาอังกฤษฉบับที่ 2 มีข้อสอบที่มีความสอดคล้องกับเกณฑ์ของข้อทดสอบที่ดี และสามารถนำมาจัดเก็บไว้ในคลังข้อสอบได้ จำนวน 106 ข้อ หรือคิดเป็นร้อยละ 88.33

ตารางที่ 13 ผลการวิเคราะห์ข้อกระทบโดยใช้ Rasch Model ของแบบทดสอบสมรรถภาพทั่วไป
ทางภาษาอังกฤษ ฉบับที่ 3

ITEM NAME	ITEM DIFF	ERR IMPAC	FIT T-TEST		WEIGHTED	
			BETWN	TOTAL	MNSQ	DF
B17	0.23	0.00	2.85	-1.27	0.83	102
A76	1.20	0.00	2.08	-1.12	0.82	77
A70	0.19	0.00	1.84	-1.95	0.87	103
A22	-0.15	0.00	2.01	-0.80	0.89	103
A35	0.09	0.00	0.12	-0.71	0.90	103
B11	0.23	0.00	0.97	-0.69	0.90	102
A2	-0.15	0.00	1.39	-0.58	0.91	103
A41	0.78	0.00	0.99	-0.59	0.91	81
A71	1.20	0.00	0.72	-0.58	0.90	77
A82	0.04	0.00	0.92	-0.55	0.92	102
A15	-0.11	0.00	0.65	-0.52	0.92	103
B09	1.53	0.00	1.49	-0.52	0.90	65
A82	0.46	0.00	-0.04	-0.52	0.92	99
A83	1.01	0.00	0.13	-0.51	0.92	84
A94	-0.61	0.00	1.55	-0.49	0.92	96
A43	-0.11	0.00	1.62	-0.49	0.93	103
A69	-0.78	0.00	0.92	-0.48	0.92	91
A80	-0.70	0.00	-0.70	-0.48	0.93	94
A73	-0.82	0.00	2.10	-0.48	0.92	90
A87	-1.10	0.00	0.76	-0.47	0.92	81
A52	-0.34	0.00	2.42	-0.47	0.93	101
A59	-0.34	0.00	0.67	-0.46	0.93	101
A79	0.34	0.00	1.50	-0.44	0.93	101

ตารางที่ 13 (ต่อ)

ITEM NAME	ITEM DIFF	ERR IMPAC	FIT T-TEST		WEIGHTED	
			BETWN	TOTAL	MNSQ	DF
B14	0.23	0.00	0.64	-0.42	0.94	102
A93	0.30	0.00	0.31	-0.42	0.94	101
A27	-1.19	0.00	0.31	-0.42	0.93	77
A25	1.65	0.00	1.06	-0.41	0.92	61
A42	-0.65	0.00	2.01	-0.40	0.94	95
A72	-1.19	0.00	-0.10	-0.37	0.93	77
A60	-1.65	0.00	-0.03	-0.35	0.93	60
A28	-0.38	0.00	0.78	-0.35	0.95	100
A95	-0.49	0.00	0.56	-0.34	0.95	98
A54	-0.70	0.00	0.97	-0.34	0.94	94
A65	0.27	0.00	0.91	-0.32	0.95	102
A10	0.23	0.00	-0.14	-0.31	0.95	102
B12	-0.96	0.00	0.44	-0.31	0.95	86
A67	0.38	0.00	-0.47	-0.31	0.95	100
B03	0.58	0.00	0.86	-0.26	0.96	96
A97	-2.19	0.00	0.40	-0.25	0.93	41
B10	2.47	0.00	0.57	-0.25	0.92	34
B18	0.70	0.00	-0.12	-0.25	0.96	93
A26	-0.34	0.00	-0.75	-0.24	0.96	101
A86	1.20	0.00	0.17	0.22	0.96	77
A46	0.54	0.00	-0.34	-0.22	0.96	97
B05	-0.61	0.00	-0.40	-0.21	0.96	96
A12	0.00	0.00	-1.23	-0.20	0.97	103

ตารางที่ 13 (ต่อ)

ITEM NAME	ITEM DIFF	ERR IMPAC	FIT T-TEST		WEIGHTED	
			BETWN	TOTAL	MNSQ	DF
A49	-1.00	0.00	-0.47	-0.19	0.96	84
A45	-1.46	0.00	1.13	-0.19	0.96	67
A51	-2.29	0.00	0.37	-0.17	0.95	38
A98	0.00	0.00	-1.47	-0.16	0.97	103
A58	0.00	0.00	-1.02	-0.15	0.97	103
A7	0.30	0.00	0.57	-0.13	0.98	102
A77	-0.19	0.00	-0.28	-0.13	0.98	103
B74	-0.53	0.00	-0.24	-0.12	0.98	98
A85	0.62	0.00	-1.18	-0.10	0.98	95
A30	-2.29	0.00	-0.75	-0.09	0.96	38
A90	-0.57	0.00	-1.82	-0.09	0.98	97
A63	-1.19	0.00	-1.12	-0.08	0.98	77
A84	-0.15	0.00	-1.77	-0.08	0.98	103
A38	0.04	0.00	-0.58	-0.06	0.96	103
A37	-1.19	0.00	-0.86	-0.06	0.98	77
A36	-0.65	0.00	-0.20	-0.06	0.98	95
A33	1.15	0.00	1.40	-0.05	0.98	79
A48	0.11	0.00	0.27	-0.04	0.99	103
A14	-0.82	0.00	1.24	-0.01	0.99	90
A57	1.15	0.00	-0.28	-0.00	0.99	79
A89	0.54	0.0	-0.82	-0.00	0.99	97
A1	-2.51	0.00	0.78	-0.00	0.98	31
A66	-0.53	0.00	0.87	0.02	1.00	98

ตารางที่ 13 (ต่อ)

ITEM NAME	ITEM DIFF	ERR IMPAC	FIT T-TEST		WEIGHTED	
			BETWN	TOTAL	MNSQ	DF
A64	0.78	0.00	-1.02	0.02	1.00	91
A40	-0.49	0.00	-0.73	0.03	1.00	98
A21	-1.14	0.00	-0.52	0.05	1.00	79
A8	1.10	0.00	1.16	0.07	1.00	81
A96	-0.87	0.00	0.33	0.09	1.01	88
B06	-0.38	0.01	-2.22	0.13	1.01	100
A62	-0.91	0.01	0.13	0.17	1.02	87
A32	0.42	0.01	-0.41	0.18	1.02	100
B16	2.59	0.01	0.52	0.19	1.03	81
A68	2.00	0.01	0.45	0.21	1.03	48
B13	1.41	0.01	0.21	0.22	1.03	69
A24	-0.78	0.01	-1.50	0.22	1.03	91
A23	0.62	0.01	-1.37	0.22	1.03	95
A75	1.20	0.01	0.58	0.23	1.03	77
A9	-0.19	0.01	1.39	0.24	1.03	103
A38	0.15	0.02	-1.74	0.26	1.03	103
A34	0.83	0.02	0.57	0.26	1.03	90
B04	-0.07	0.02	-0.40	0.27	1.03	103
A13	0.66	0.02	-0.86	0.27	1.03	94
A88	-0.53	0.02	-0.61	0.20	1.04	98
A61	0.23	0.02	0.48	0.30	1.04	102
A5	-0.49	0.02	0.80	0.31	1.04	98
A28	-0.11	0.02	1.92	0.35	1.04	103

ตารางที่ 13 (ต่อ)

ITEM NAME	ITEM DIFF	ERR IMPAC	FIT T-TEST		WEIGHTED	
			BETWN	TOTAL	MNSQ	DF
B08	0.47	0.02	-1.13	0.36	1.04	101
A21	-0.22	0.02	1.12	0.28	1.05	102
A4	0.83	0.03	-0.36	0.41	1.05	90
B00	-0.38	0.03	1.50	0.42	1.05	100
A6	1.85	0.04	1.37	0.44	1.07	53
A50	-0.87	0.03	1.90	0.48	1.07	88
A47	-0.91	0.03	0.77	0.50	1.07	87
A91	1.25	0.04	0.60	0.51	1.08	75
A99	-1.05	0.04	0.74	0.53	1.08	82
A16	0.15	0.03	0.73	0.54	1.07	103
A22	-0.96	0.04	1.78	0.54	1.08	86
A11	0.38	0.04	0.41	0.60	1.08	100
A55	0.70	0.04	2.02	0.60	1.08	93
B18	1.14	0.05	0.02	0.62	1.10	69
A44	1.47	0.05	1.27	0.63	1.10	67
B20	-0.45	0.04	0.75	0.63	1.09	99
B07	-0.15	0.04	0.32	0.63	1.08	108
A53	-0.53	0.04	1.73	0.65	1.09	98
A19	1.41	0.05	3.03	0.67	1.11	69
A81	-0.22	0.05	2.28	0.75	1.10	102
A78	1.53	0.06	3.85	0.79	1.13	65
A3	-0.30	0.06	1.49	0.84	1.11	102
B15	0.27	0.06	1.17	0.89	1.12	102

ตารางที่ 13 (ต่อ)

ITEM NAME	ITEM DIFF	ERR IMPAC	FIT T-TEST		WEIGHTED	
			BETWN	TOTAL	MNSQ	DF
A17	-1.05	0.07	3.00	0.90	1.14	82
A31	0.78	0.07	2.79	0.98	1.14	91
A18	0.62	0.09	2.01	1.22	1.18	95
A20	-0.38	0.09	3.03	1.26	1.18	100
A56	-0.27	0.11	3.69	1.62	1.24	102

จากตารางที่ 13 เมื่อพิจารณาค่าความยากง่าย (β) ปรากฏว่าไม่มีข้อสอบที่มีค่าความยากง่าย (β) เกิน ± 3 เลย

ความสอดคล้องของข้อทดสอบกับโค้ง ICC เมื่อพิจารณาจากค่า Weighted mean square สำหรับแบบทดสอบฉบับนี้ ปรากฏว่ามีข้อสอบที่มีค่า Weighted mean square เกิน $1 + 0.07$ จำนวน 20 ข้อ ได้แก่ข้อ A91 A99 A22 A11 A55 B18 A44 B20 B07 A53 A19 A81 A78 A3 B15 A17 A31 A18 A20 และ A56 เป็นข้อสอบการฟัง 8 ข้อ คือ A3 A11 A17 A28 A19 A20 A22 A31 ข้อสอบการอ่าน 8 ข้อ A44 A53 A55 A56 A78 A81 A91 A99 ข้อสอบการเขียน 4 ข้อ คือ B07 B15 B18 B20

โดยสรุปการดำเนินการสร้างแบบทดสอบสมรรถภาพทั่วไปทางภาษาอังกฤษฉบับที่ 3 มีข้อสอบที่มีความสอดคล้องเกณฑ์ของข้อทดสอบที่ดี และสามารถนำมาจัดเก็บไว้ในคลังข้อสอบได้ จำนวน 100 ข้อ หรือคิดเป็นร้อยละ 83.33

ตารางที่ 14 ผลการวิเคราะห์ข้อกระทบงโดยใช้ Rasch Model ของแบบทดสอบสมรรถภาพทั่วไป
ทางภาษาอังกฤษ ฉบับที่ 4

ITEM NAME	ITEM DIFF	ERR IMPAC	FIT T-TEST		WEIGHTED	
			BETWN	TOTAL	MNSQ	DF
A44	-0.49	0.00	3.15	-1.45	0.82	119
A91	-0.39	0.00	2.84	-1.26	0.84	121
A74	-0.62	0.00	2.43	-1.07	0.86	115
A32	-0.23	0.00	1.73	-1.01	0.87	124
A72	-1.01	0.00	2.04	-0.98	0.86	101
A52	-0.86	0.00	1.82	-0.87	0.98	107
A93	-0.07	0.00	1.55	-0.80	0.90	126
A43	-0.42	0.00	-0.44	-0.66	0.91	121
B12	0.02	0.00	0.29	-0.62	0.92	126
A20	0.62	0.00	1.21	-0.62	0.92	118
B05	1.07	0.00	-0.42	-0.54	0.92	102
A61	-1.64	0.00	1.42	-0.54	0.90	72
A85	0.62	0.00	-0.48	-0.52	0.33	118
A92	-1.17	0.00	0.70	-0.50	0.92	94
A49	-1.26	0.00	1.10	-0.49	0.92	90
B04	-1.05	0.00	0.66	-0.48	0.93	99
B03	-1.59	0.00	0.60	-0.46	0.92	75
B07	-1.26	0.00	0.35	-0.43	0.93	90
A54	-1.64	0.00	0.64	-0.40	0.93	72
A68	-1.76	0.00	0.75	-0.40	0.92	67
A46	-0.11	0.00	0.39	-0.39	0.95	126
A87	-1.64	0.00	-0.05	-0.37	0.93	72
A8	0.93	0.00	1.10	-0.37	0.94	108

ตารางที่ 14 (ต่อ)

ITEM NAME	ITEM DIFF	ERR IMPAC	FIT T-TEST		WEIGHTED	
			BETWN	TOTAL	MNSQ	DF
B14	0.08	0.00	0.11	-0.36	0.95	126
A79	-1.59	0.00	-0.09	-0.36	0.93	75
A90	0.14	0.00	0.53	-0.35	0.95	126
A10	0.89	0.00	-0.21	-0.33	0.95	109
B00	-0.76	0.00	-0.64	-0.32	0.95	111
A63	-1.88	0.00	1.30	-0.31	0.93	62
A57	0.08	0.00	1.62	-0.31	0.96	126
A40	-0.65	0.00	0.89	-0.31	0.95	114
A90	-1.49	0.00	0.67	-0.30	0.94	79
A5	1.00	0.00	1.51	-0.30	0.95	105
B15	-1.22	0.00	-0.72	-0.29	0.95	92
A84	-2.40	0.00	1.36	-0.29	0.92	42
A56	-2.40	0.00	0.27	-0.28	0.93	42
A35	0.14	0.00	0.27	-0.26	0.96	126
B08	-1.39	0.00	-0.28	-0.26	0.95	84
A98	0.96	0.00	-1.79	-0.26	0.96	107
A48	0.49	0.00	-0.58	-0.26	0.96	121
A23	0.89	0.00	-0.08	-0.24	0.96	109
A19	1.00	0.00	0.11	-0.23	0.96	105
A76	-0.14	0.00	0.18	-0.22	0.97	125
A55	-1.76	0.00	-0.67	-0.21	0.96	67
A38	-1.05	0.00	-0.37	-0.20	0.96	99
A51	-0.36	0.00	0.49	-0.20	0.97	122

ตารางที่ 14 (ต่อ)

ITEM NAME	ITEM DIFF	ERR IMPAC	FIT T-TEST		WEIGHTED	
			BETWN	TOTAL	MNSQ	DF
A36	4.21	0.00	0.54	-0.18	0.96	10
A77	0.65	0.00	-1.15	-0.18	0.97	117
A89	0.24	0.00	-1.04	-0.18	0.97	125
A70	1.85	0.00	-0.20	-0.18	0.96	60
A50	2.98	0.00	1.39	-0.17	0.94	30
A73	-0.49	0.00	0.76	-0.17	0.97	119
A64	-1.35	0.00	0.78	-0.15	0.97	86
A8	0.24	0.00	-0.25	-0.15	0.98	125
A99	-2.08	0.00	-0.33	-0.14	0.96	54
B18	-0.04	0.00	-0.00	-0.14	0.98	90
A38	1.36	0.00	0.49	-0.11	0.98	90
A59	-1.13	0.00	0.83	-0.11	0.98	95
A67	-0.94	0.00	-1.53	-0.11	0.98	104
A58	-0.98	0.00	-0.79	-0.08	0.98	102
A88	0.14	0.00	1.11	-0.08	0.98	126
A21	-1.94	0.00	1.47	-0.06	0.98	59
A4	1.15	0.00	-0.58	-0.05	0.99	90
A9	0.86	0.00	0.70	0.04	0.99	111
B09	-0.11	0.00	0.56	-0.04	0.99	126
A34	-0.11	0.00	-0.88	-0.03	0.99	126
A24	0.36	0.0	1.59	-0.01	0.99	124
A25	1.15	0.00	-0.11	-0.00	0.99	99
A78	-0.14	0.00	0.62	-0.00	0.99	125

ตารางที่ 14 (ต่อ)

ITEM NAME	ITEM DIFF	ERR IMPAC	FIT T-TEST		WEIGHTED	
			BETWN	TOTAL	MNSQ	DF
B10	-0.36	0.00	-0.51	0.00	0.99	122
B18	-0.11	0.00	-0.56	0.02	1.00	126
A11	1.63	0.00	-0.28	0.04	1.00	78
A42	-0.58	0.00	0.69	0.05	1.00	116
B17	0.46	0.00	0.36	0.05	1.00	122
A13	0.05	0.00	0.89	0.06	1.00	126
A28	1.15	0.00	0.56	0.06	1.00	99
A75	1.11	0.00	-1.31	0.06	1.00	101
A81	-0.20	0.00	-0.05	0.07	1.00	125
A26	-1.44	0.00	1.34	0.09	1.01	81
A21	-0.20	0.01	1.49	0.14	1.01	125
A7	0.89	0.01	-0.28	0.14	1.01	109
A29	-0.07	0.01	-1.45	0.15	1.01	126
B13	0.62	0.01	-0.72	0.19	1.02	118
A62	0.75	0.01	0.40	0.20	1.03	114
B06	-0.65	0.01	-1.81	0.20	1.02	114
A15	-0.26	0.01	0.39	0.21	1.02	124
A80	-0.26	0.01	-1.55	0.22	1.02	124
A96	-0.86	0.01	-0.90	0.22	1.02	107
A47	-0.17	0.01	-1.78	0.24	1.02	125
A82	1.23	0.01	-0.08	0.24	1.03	96
A37	-0.42	0.01	-0.75	0.24	1.03	121
A87	1.04	0.02	0.68	0.30	1.04	104

ตารางที่ 14 (ต่อ)

ITEM NAME	ITEM DIFF	ERR IMPAC	FIT T-TEST		WEIGHTED	
			BETWN	TOTAL	MNSQ	DF
A94	0.05	0.02	1.03	0.34	1.04	126
A22	1.31	0.02	1.36	0.35	1.04	92
A33	0.27	0.02	1.63	0.35	1.04	125
A1	0.46	0.02	-0.00	0.39	1.04	122
A14	-0.17	0.02	1.99	0.40	1.05	125
A27	1.85	0.04	2.09	0.47	1.07	62
A71	0.20	0.03	-0.41	0.48	1.06	126
A66	3.89	0.07	3.84	0.49	1.15	13
A45	0.33	0.03	0.14	0.50	1.06	124
A31	0.72	0.03	0.50	0.55	1.07	115
B11	0.36	0.04	1.15	0.59	1.07	124
A41	1.85	0.05	3.14	0.60	1.10	69
B20	0.93	0.04	0.37	0.64	1.08	108
A16	0.72	0.04	1.53	0.67	1.08	115
A18	1.04	0.05	1.86	0.70	1.09	104
A2	-1.22	0.05	1.65	0.71	1.10	92
A22	0.08	0.04	1.77	0.72	1.09	126
A88	-0.01	0.04	0.77	0.78	1.09	126
A65	0.49	0.05	0.40	0.75	1.09	121
A12	-0.11	0.05	2.57	0.84	1.10	126
A53	0.62	0.06	2.17	1.01	1.13	118
B16	0.33	0.07	1.50	1.10	1.14	124
A6	0.42	0.07	1.56	1.11	1.14	123

ตารางที่ 14 (ต่อ)

ITEM NAME	ITEM DIFF	ERR IMPAC	FIT T-TEST		WEIGHTED	
			BETWN	TOTAL	MNSQ	DF
A69	-0.17	0.07	3.93	1.13	1.14	125
A86	0.96	0.09	2.69	1.29	1.18	107
A17	-0.20	0.08	3.02	1.30	1.17	125
A60	0.68	0.10	3.20	1.52	1.21	116
A30	0.46	0.11	3.83	1.76	1.24	122

จากตารางที่ 14 เมื่อพิจารณาค่าความยากง่าย (β) ปรากฏว่ามีข้อสอบที่มีค่าความยากง่าย (β) เกิน ± 3 อยู่จำนวน 2 ข้อ ได้แก่ ข้อสอบวัดทักษะการอ่านข้อ A36 A66 ซึ่งมีค่า Item Diff เท่ากับ 4.21 และ 3.89

ความสอดคล้องของข้อทดสอบกับโค้ง ICC เมื่อพิจารณาจากค่า Weighted mean square สำหรับแบบทดสอบฉบับนี้ ปรากฏว่ามีข้อสอบที่มีค่า Weighted mean square เกิน $1 + 0.08$ จำนวน 16 ข้อ ได้แก่ข้อ A66 A41 A18 A2 A22 A83 A65 A12 A53 B16 A6 A69 A86 A17 A60 A30 เป็นข้อสอบการฟัง 6 ข้อ คือ A2 A6 A12 A17 A18 และ A22 ข้อสอบการอ่าน 9 ข้อ คือ A30 A61 A53 A60 A65 A68 A69 และ A83 A86 วัดทักษะการเขียน 1 ข้อ คือ B16

โดยสรุปการดำเนินการสร้างแบบทดสอบสมิทริภาพทั่วไปทางภาษาอังกฤษฉบับที่ 4 มีข้อสอบที่มีความสอดคล้องกับเกณฑ์ของข้อทดสอบที่ดี และสามารถนำมาจัดเก็บไว้ในคลังข้อสอบได้จำนวน 103 ข้อ หรือคิดเป็นร้อยละ 85.83

ตอนที่ 4 สรุปผลการวิเคราะห์

4.1 การวิเคราะห์ค่าความเที่ยง

ตารางที่ 15 สรุปผลการวิเคราะห์ค่าความเที่ยง

แบบทดสอบฉบับที่	ค่าความเที่ยง			
	KR ₂₀	Phi	inter-rater	α
1				
ปรนัย	.878	.859		
อัตนัย 1,2,3			.741, .829, .834	
2				
ปรนัย	.910	.893		
อัตนัย 1,2,3			.594, .730, .800	
3				
ปรนัย	.854	.822		
อัตนัย 1,2,3			.727, .808, .684	
4				
ปรนัย	.877	.854		
อัตนัย 1,2,3			.778, .685, .840	
ปรับปรุง				
ปรนัย	.727	.676		
อัตนัย			.846	.826

จากตารางที่ 15 จะเห็นได้ว่าแบบทดสอบมีค่าความเที่ยงอยู่ในเกณฑ์ปานกลาง - สูง มีค่าตั้งแต่ .594 - .910 โดยข้อทดสอบแบบอัตนัยข้อที่ 1 ของฉบับที่ 2 มีค่าความเที่ยงต่ำสุด และข้อทดสอบปรนัยของฉบับที่ 2 มีค่าความเที่ยงสูงสุด นอกจากนี้อยู่ในระดับสูง .727 - .899

4.2 การวิเคราะห์ข้อกระทง

4.2.1 ผลการวิเคราะห์ข้อสอบปรนัย

ค่าความยากง่ายและค่าอำนาจจำแนกของข้อสอบปรนัยในแบบทดสอบสมิทธิภาพทั่วไปทางภาษาอังกฤษของแบบทดสอบฉบับที่ 1, 2, 3, 4 และแบบทดสอบฉบับปรับปรุง (ป) นำเสนอในตารางที่ 16



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 16 แสดงค่าความยากง่ายและค่าอำนาจจำแนกของข้อสอบปลายปี ของแบบทดสอบสมิทธิภาศัพท์ทั่วไปทางภาษาอังกฤษ

ฉบับที่	ค่าความยากง่าย					ค่าอำนาจจำแนก															รวม
	ดีมาก					ดี					พอใช้					ต่ำ					
	.60					.35 - .59					.20 - .34					> .20					
	1	2	3	4	๕	1	2	3	4	๕	1	2	3	4	๕	1	2	3	4	๕	
ดีมาก					6	5	17	14	6	2	33	18	14	27	2	12	7	18	11		192
.40- .60																					
ดี					6	2	8	7	4	5	8	13	15	13	2	3	2	10	6	1	105
.31- .39, .61- .69																					
พอใช้					14	4	14	3	8	8	27	14	14	17	1	12	9	12	7	5	169
.20- .29, .70- .80																					
ง่ายไป									2	1	6	8	4	11		1	3	2	1	3	42
< .81																					
ยากไป							2	2		4	6	1		1	2	1	4	5	6	1	35
> .19																					
รวม					26	11	41	26	20	20	80	54	47	69	7	29	25	47	31	10	543

จุฬาลงกรณ์มหาวิทยาลัย

จากตารางที่ 16 จะเห็นได้ว่าแบบทดสอบสอบสมิทธิภาพทั่วไปทางภาษาอังกฤษฉบับปรับปรุงมีคุณภาพดีที่สุดในข้อที่มีคุณภาพดีมากที่สุดทั้งในด้านความยากง่ายและอำนาจจำแนก ดีในด้านค่าความยากง่าย และดีมากในด้านค่าอำนาจจำแนก และพอใช้ในด้านค่าความยากง่าย แต่ยังคงดีมากในด้านค่าอำนาจจำแนกในขณะที่ข้อสอบสมิทธิภาพทั่วไปทางภาษาอังกฤษ ฉบับที่ 1-4 ไม่มีข้อใดที่เข้าเกณฑ์นี้เลย



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตารางต่อไปที่นำเสนอปริมาณร้อยละของแบบทดสอบที่เข้าเกณฑ์ตามการวิเคราะห์แบบ Classical Model และ Rasch Model

ตารางที่ 17 จำนวนและร้อยละของข้อกระทงของแบบทดสอบสมรรถภาพทั่วไปทางภาษาอังกฤษที่เข้าเกณฑ์ตามการวิเคราะห์แบบ Classical Model และ แบบ Rasch Model

เกณฑ์วิเคราะห์	ฉบับที่ (120 ข้อ)				ฉบับปรับปรุง (63 ข้อ)
	1	2	3	4	
Classical	109 (90.83)	102 (85.00)	104 (84.17)	102 (85.00)	54 (85.71)
Rasch	105 (87.50)	106 (88.33)	100 (83.33)	103 (85.83)	

สรุปได้ว่าเมื่อนำข้อสอบที่ต้องปรับปรุงไปแก้ไข และนำมาทดลองใช้อีกครั้งหนึ่งเป็นข้อสอบที่ใช้ได้ร้อยละ 85.71 จำนวน 54 ข้อ มี 9 ข้อ หรือ ร้อยละ 14.29 ที่ยังใช้ไม่ได้ และควรตัดทิ้ง

ข้อสอบที่เข้าเกณฑ์ Classical Model และ Rasch Model มีจำนวนที่แตกต่างกัน ทั้งนี้เป็นเพราะการวิเคราะห์มีพื้นฐานของทฤษฎีที่แตกต่างกัน

4.2.2 ผลการวิเคราะห์ข้อสอบอัตนัย

ตารางที่ 18 แสดงค่าความยากง่ายและค่าอำนาจจำแนกของข้อสอบอัตนัย ของแบบทดสอบ
สมิทธิภาพทั่วไปทางภาษาอังกฤษ

ลำดับที่/ข้อที่	ค่าความยากง่าย		ค่าอำนาจจำแนก		
	ดีมาก	ดี	พอใช้	ต่ำ	รวม
	.60	.35 - .59	.20 - .34	> .20	
ดีมาก		1/2, 1/3	2/1, 2/2		7 ข้อ
ดี		3/2	4/1		3 ข้อ
พอใช้		2/3	1/1		2 ข้อ
ยากไป			2/3		1 ข้อ
> .19			นำไปปรับปรุง		
รวม		5 ข้อ	7 ข้อ		12 ข้อ

สรุปได้ว่าข้อสอบอัตนัยฉบับปรับปรุงแล้วอยู่ในเกณฑ์ดี

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

สรุป อภิปรายผล และข้อเสนอแนะ

สรุป

งานวิจัยเรื่อง "การสร้างและพัฒนาคุณภาพแบบทดสอบสมิทริททั่วไประบบภาษาอังกฤษและระบบการทดสอบที่เกี่ยวข้อง" ในระยะที่ 2 ประจำปีการศึกษา 2534 มีวัตถุประสงค์เฉพาะดังนี้

1. เพื่อพัฒนาข้อทดสอบสมิทริททั่วไประบบภาษาอังกฤษด้านการฟัง การอ่าน และเขียน เพิ่มมากขึ้นจากข้อทดสอบที่ได้จากการสร้างขึ้นในการวิจัยระยะปีที่ 1 และให้ได้แบบทดสอบ 4 ฉบับคู่ขนานกันในด้านเนื้อหาและสถิติโดยใช้ตารางวิเคราะห์เดียวกัน
2. เพื่อวิเคราะห์และปรับปรุงแบบทดสอบทั้งฉบับและเป็นรายข้อ
3. เพื่อพัฒนาแบบทดสอบให้เป็นมาตรฐานและเก็บไว้ใช้ในคลังข้อทดสอบโดยใช้คอมพิวเตอร์

กลุ่มผลวิจัยมี 2 กลุ่ม ประกอบด้วย กลุ่มที่ 1 นิสิตจุฬาลงกรณ์มหาวิทยาลัยตั้งแต่ระดับชั้นปีที่ 1 ถึง บัณฑิตศึกษา รวมทั้งบุคคลภายนอกที่สมัครสอบตามประกาศของสถาบันภาษา และเข้าสอบในภาคปลายปีการศึกษา 2533 เมื่อวันที่ 16-31 มกราคม 2534 จำนวน 481 คน กลุ่มที่ 2 คือบุคคลภายนอกที่มีพื้นฐานความรู้ระดับปริญญาตรีและปริญญาโท ซึ่งเข้าทดสอบในภาคปลายปีการศึกษา 2534 จำนวน 87 คน ซึ่งในจำนวนนี้ 75 คน ทำข้อสอบปรนัยของแบบทดสอบสมิทริททั่วไประบบภาษาอังกฤษฉบับปรับปรุง และ 12 คนทำข้อสอบอัตนัยวัดความสามารถในการเขียน

ประโยชน์ที่คาดว่าจะได้รับคือ

1. สถาบันภาษา จุฬาลงกรณ์มหาวิทยาลัยได้แบบทดสอบสมิทริททั่วไประบบภาษาอังกฤษมีมาตรฐานและเป็นที่ยอมรับทั่วไป มีจำนวน 4 ชุด และจำนวนข้อเพิ่มขึ้น
2. สถาบันภาษาได้ข้อทดสอบที่มีมาตรฐานที่สามารถเก็บไว้ใช้ในคลังข้อสอบทางภาษาซึ่งใช้คอมพิวเตอร์ช่วยในการจัดเก็บและสามารถนำมาใช้และพัฒนาข้อสอบต่อไปได้สะดวก

ผลของการวิจัยสรุปได้ดังนี้

1. ได้แบบทดสอบปรนัยวัดทักษะการฟัง อ่าน และเขียน และแบบอัตนัยวัดทักษะการเขียน จำนวน 4 ฉบับ ที่มีค่าความเที่ยงอยู่ในเกณฑ์ ดี - ดีมาก ($KR_{20} = .854 - .910$ แบบ Phi = $.822 - .896$) และแบบอัตนัยวัดทักษะการเขียนที่มีค่าความเที่ยงอยู่ในเกณฑ์ ดี - ดีมาก ($R_{xy} = .594 - .840$) โดยมีค่าอำนาจในการทำนาย $48.00 - 58.50$
2. ได้ข้อทดสอบแบบปรนัยวัดสมรรถภาพทั่วไปทางภาษาอังกฤษที่มีคุณภาพดีมาก - พอใช้ได้ จำนวน 471 ข้อ
3. ได้ข้อทดสอบแบบอัตนัยวัดสมรรถภาพทางการเขียนที่มีคุณภาพ ดี - พอใช้ได้ 12 ข้อ
4. ได้ข้อทดสอบแบบปรนัยวัดสมรรถภาพทั่วไปทางภาษาอังกฤษซึ่งได้มาตรฐานเข้าเกณฑ์ของ Rasch Model จำนวน 414 ข้อ

อภิปรายผล

การวิเคราะห์แบบทดสอบสมรรถภาพทั่วไปทางภาษาอังกฤษแบบปรนัย ซึ่งมีการให้คะแนน 0 และ 1 ใช้วิธีการต่างกันแบบทดสอบอัตนัย ซึ่งมีการให้คะแนนถึง 2 หลัก ทั้งการหาค่าความเที่ยง ค่าความยากง่าย และค่าอำนาจจำแนก และเนื่องจากว่าแบบทดสอบวัดสมรรถภาพทั่วไปทางภาษาอังกฤษนี้เป็นแบบทดสอบอิงเกณฑ์ การวิเคราะห์จึงให้ความสำคัญต่อค่าความยากง่ายมากกว่าค่าอำนาจจำแนก

การวิเคราะห์ข้อทดสอบโดยใช้ Classical Model เหมาะสมกับการปรับปรุงข้อกระตงเพราะให้ข้อมูลเกี่ยวกับตัวเลือกที่ถูกและตัวลวง ส่วนการวิเคราะห์โดยใช้ Rasch Model เหมาะสมสำหรับการคัดเลือกข้อกระตงเพื่อนำไปเก็บไว้ในคลังข้อทดสอบเพราะเป็นการวิเคราะห์ที่ไม่อิงกลุ่ม (Sample-free)

การปรับปรุงข้อกระตงโดยยึดค่าความยากง่ายเป็นเกณฑ์และอำนาจจำแนกที่ไม่ดีลบบจะทำให้ค่าอำนาจจำแนกของข้อสอบดีขึ้นตามไปด้วย

ข้อเสนอแนะ

1. ควรจัดเก็บข้อกระทงที่มีคุณภาพมาตรฐานตามเกณฑ์ไว้ใช้ในคลังข้อทดสอบโดยใช้คอมพิวเตอร์
2. ควรพัฒนาข้อทดสอบทักษะภาษาแบบอิงเกณฑ์ที่วัดสมรรถภาพทั่วไปทางภาษาอังกฤษที่ใช้ตารางวิเคราะห์อื่น ๆ ซึ่งสร้างขึ้นจากสถานการณ์ของการใช้ภาษาเพื่อการสื่อสารที่แตกต่างหลากหลายออกไป
3. ควรทำการวิจัยเปรียบเทียบระดับสมรรถภาพทางการใช้ภาษาอังกฤษของนิสิตนักศึกษาและบุคคลทั่วไปที่วัดโดยแบบทดสอบของสถาบันภาษากับแบบทดสอบมาตรฐานที่มีการใช้อย่างแพร่หลาย เช่น แบบทดสอบ TOFEL (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

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

บรรณานุกรม

โกวิท ประวาลพุกษ์ และสมศักดิ์ สินธุรเวชกุล. 2527. การประเมินในชั้นเรียน.

กรุงเทพฯ: วัฒนาพานิช.

สุพันธ์ สุกมลสันต์ และคณะ. 2533. รายงานการวิจัยฉบับที่ 1 ในโครงการ การสร้าง และพัฒนาคุณภาพของแบบทดสอบวัดสมรรถภาพทั่วไปทางภาษาอังกฤษและระบบการทดสอบ ที่เกี่ยวข้องกับเรื่อง "แนวโน้มระดับสมรรถภาพทั่วไปทางภาษาอังกฤษของนิสิตจุฬาลงกรณ์ มหาวิทยาลัยและบุคคลภายนอกที่สนใจ ปี พ.ศ. 2533". สถาบันภาษา และฝ่ายวิชาการ จุฬาลงกรณ์มหาวิทยาลัย.

อัจฉรา วงศ์โสธร. 2529. เทคนิควิธีการสร้างข้อสอบภาษาอังกฤษสำหรับวัดและประเมิน ผลการใช้ภาษาเพื่อการสื่อสาร. กรุงเทพฯ: สำนักพิมพ์อักษรเจริญทัศน์.

_____. 2532. การวิจัยเพื่อพัฒนาแบบทดสอบอิงปริเจตภาษาอังกฤษสำหรับใช้กับ นักศึกษาไทยในระดับต่าง ๆ. งานวิจัยจุฬาลงกรณ์มหาวิทยาลัย ลำดับที่ 21, ฝ่ายวิจัย จุฬาลงกรณ์มหาวิทยาลัย. กรุงเทพฯ : โรงพิมพ์จุฬาลงกรณ์มหาวิทยาลัย.

Adams, R.J., R.E. Griffin, and L. Martin. 1987. A latent trait method for measuring dimension in second language proficiency. Language Testing 4, 1: 9-27.

Anastasi, A. 1976. Psychological testing (4th ed.). New York: Collier Macmillan.

Bachman, L.F. 1985. "Critical review of current practices in language testing and evaluation." Lecture presented at CULI National Seminar on Problems and Issues in ELT and ELTE at the Tertiary Level, Bangkok 15 May-4 June 1985.

_____. 1986. "The development and use of criterion-referenced tests of language proficiency in language program evaluation." in K. Prapphal et al. (eds.): Trends in Language Programme Evaluation (Papers presented at CULI's First International Conference, 9-11 December, 1986, Bangkok). Bangkok: Chulalongkorn University Language Institute.

_____. and A.S. Palmer. 1982. "The construct validation of some components of communicative proficiency." TESOL Quarterly 16, 4: 449-65.

- Bachman, L.F. 1990. Fundamental considerations in language testing. Oxford: Oxford University Press.
- Brown, J.D. 1989. "Short-cut estimates of criterion-referenced test reliability." Paper presented at the 11th Annual Language Testing Research Colloquium, San Antonio, March. 1989.
- Carroll, J.B. 1983. "The difficulty of a test and its factor composition revisited" in H. Wainer and S. Messick (eds.): Principles of Modern Psychological Measurement: A Festschrift for Frederic M. Lord. Hillsdale, NJ: Lawrence Erlbaum Associates: 257-82.
- Cronbach, L.J. 1951. "Coefficient Alpha and the internal structure of tests". Psychometrika, 16, 297-334.
- . 1971. "Validity". In R.L. Thorndike (ed.) Educational Measurement. Second edition. Washington, D.C.: American Council on Education.
- Davidson, F. and G. Henning. 1985. "A self-rating scale of English proficiency: Rasch scalar analysis of items and rating categories." Language Testing 2, 2: 164-79.
- De Saussure, F. 1916. Cours de linguistique generale. Paris: Payot. [Course in general linguistics (translated by Wade Baskin). New York: Philosophical Library, 1985.]
- Fan, C.T. 1952. Item Analysis Table. Princeton, N.J.: Educational Testing Service.
- Griffin, P.E. 1985. "The use of latent trait models in the calibration of tests of spoken language in large-scale selection-placement programs." in Y.P. Lee, A.C.Y. Fok, R. Lord and G. Low. (eds.) New directions in language testing. Oxford: Pergamon Press.
- Henning, G., T. Hudson, and J. Turner. 1985. "Item response theory and the assumption of unidimensionality." Language Testing 2, 2: 141-54.
- Larson, J.W. 1987. "Computerized adaptive language testing: A Spanish placement exam." in K.M. Bailey, T.L. Dale, and R.T. Clifford. (eds.) 1987. Language Testing Research: Selected Papers from the 1986 Colloquium. Monterey, CA: Defense Language Institute, 1-19.

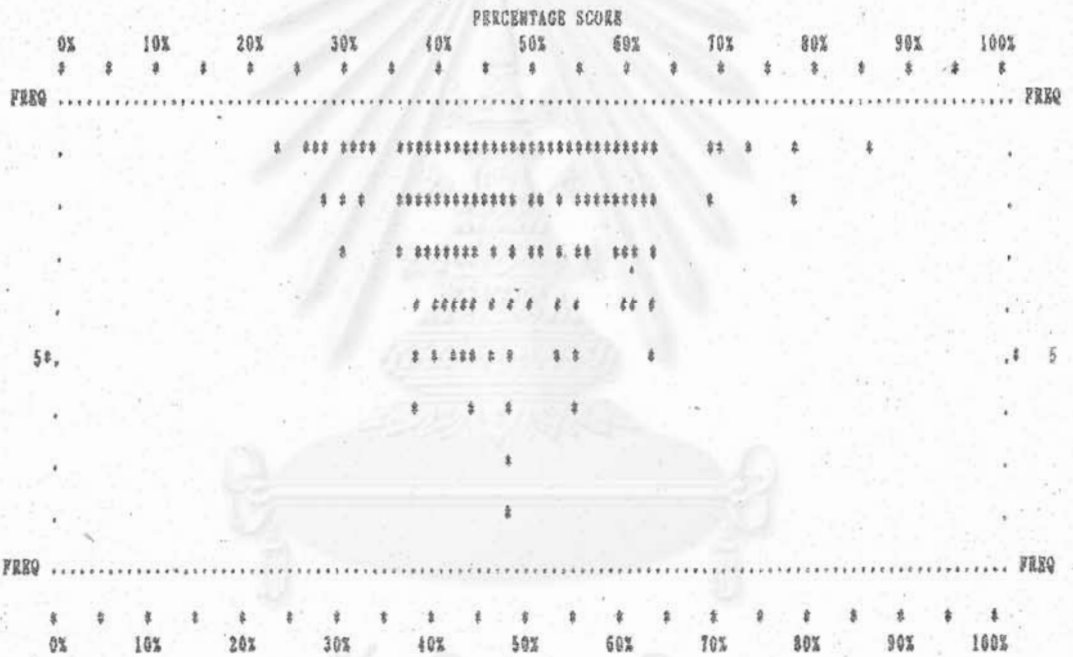
- Laurier, M. 1990. "What we can do with computerized adaptive testing and what we cannot do." Paper presented at SEAMEO RELC 25th Regional Seminar on Language Testing and Language Programme Evaluation, 9-12 April 1990, Singapore.
- Madsen, H.S. 1987. "Utilizing Rasch analysis to detect cheating on language examinations" in K.M. Bailey, T.L. Dale, and R.T. Clifford (eds.) 1987. Language Testing Research: Selected Papers from the 1986 Colloquium. Monterey, CA: Defense Language Institute.
- _____ and J.W. Larson. 1986. "Computerized Rasch analysis of item bias in ESL tests" in C.W. Stansfield. (ed.) 1986. Technology and Language Testing. (Collected papers from the 1985 Colloquium.) Washington, DC: TESOL, 11-23.
- McNamara, T.F. 1990. "The role of item response theory in language test validation." Paper presented at SEAMEO RELC 25th Regional Seminar on Language Testing and Language Programme Evaluation, 9-12 April 1990, Singapore.
- Morrow, K. 1981. "Communicative language testing: Revolutions or evolution" in C.J. Alderson and A. Hughes. (eds.) ELT Documents 111-- Issues in Language Testing. London: The British Council.
- Oller, J.W. Jr. and F.B. Hinofotis. 1980. "Two mutually exclusive hypotheses about second language ability: Indivisible and partly divisible competence. In J.W. Oller, Jr. and K. Perkins (eds.) 1980. Research in language testing. Rowley, MA: Newbury House, 13-23.
- _____. (Ed.) 1983. Issues in language testing research. Rowley, MA: Newbury House.
- Pilliner, A.G. 1983. "Norm-referenced and criterion-referenced tests -- and evaluation." (Paper presented at CULI National Seminar on Evaluation Procedures for Communicative Syllabuses, 25 October - 4 November 1983, Bangkok: Chulalongkorn University Language Institute.
- Pollitt, A. 1979. "A simple example to illustrate the sample-free principle in Rasch measurement." Paper presented at the British Council Course on Evaluation in ESP, December 1979.

- Pollitt, A. and C. Hutchinson. 1987. "Calibrating graded assessments: Rasch partial credit analysis of performance in writing." Language Testing 4, 1: 72-92.
- Rasch, G. 1960. Probabilistic models for some intelligence and attainment tests. Copenhagen: Danmarks Pædagogiske Institut.
- Scannell, D.P. and D.B. Tracy. 1975. Testing and Measurement in the Classroom. Boston: Houghton Mifflin Co.
- Stansfield, C.W. (ed.) 1986. Technology and Language Testing. (A collection of papers from the Seventh Annual Language Testing Research Colloquium, held at Educational Testing Service, Princeton, New Jersey, April 6-9, 1985). Washington, DC: Teachers of English to Speakers of Other Languages.
- Whitney, D.R. and D.L. Sabers. 1970. "Improving essay examinations III. Use of item analysis". Technical Bulletin. 11. Iowa City: University Evaluation and Examinations Service (Mimeo.)
- Wright, B.D. and M.H. Stone. 1979. Best test design. Chicago: Mesa Press.

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

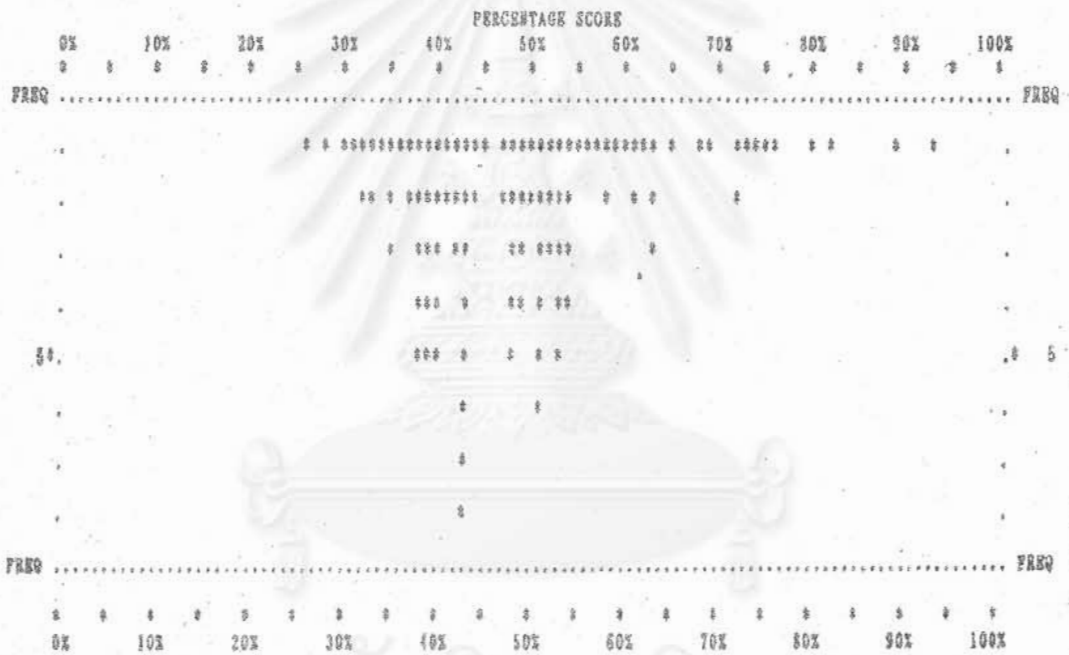
ภาคผนวก ก.

แผนภูมิที่ 3 : การกระจายของคะแนนจากแบบทดสอบวัดสมิทึภาพทั่วไปทางภาษาอังกฤษฉบับที่ 1



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

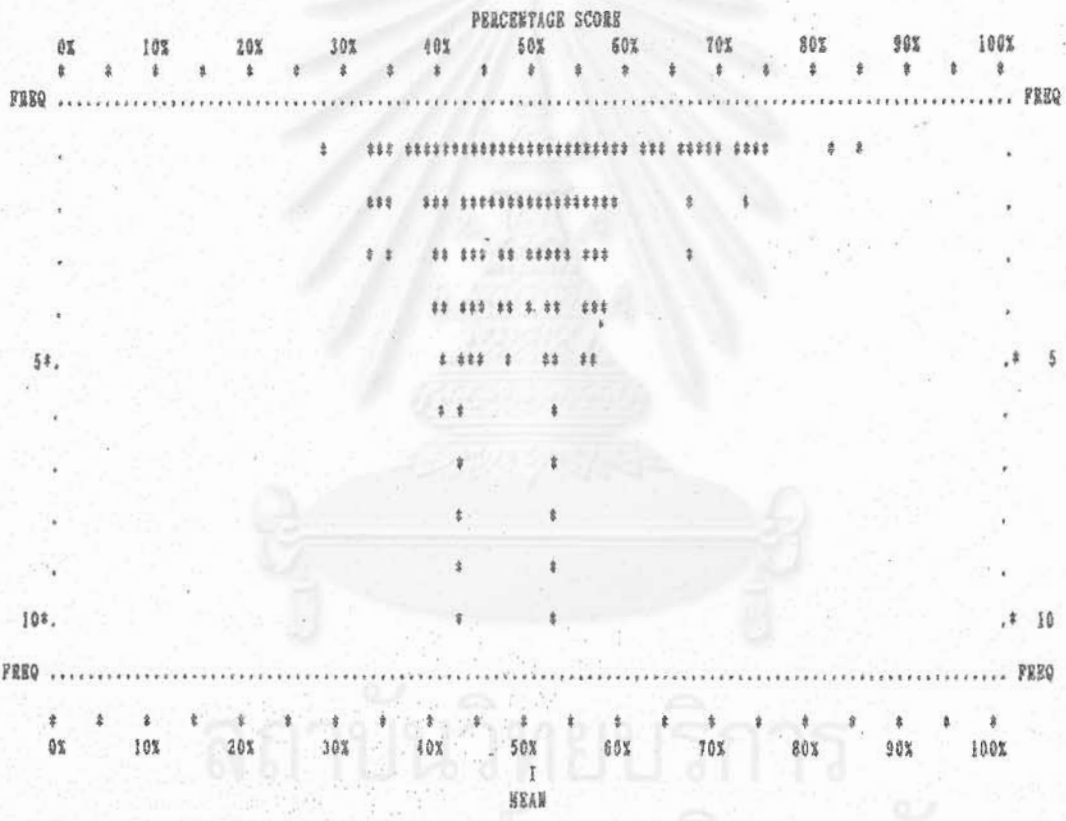
แผนภูมิที่ 4 การกระจายของคะแนนจากแบบทดสอบวัดสมรรถภาพทั่วไปทางภาษาอังกฤษฉบับที่ 2



สถาบันวิทยาลัยบริการ

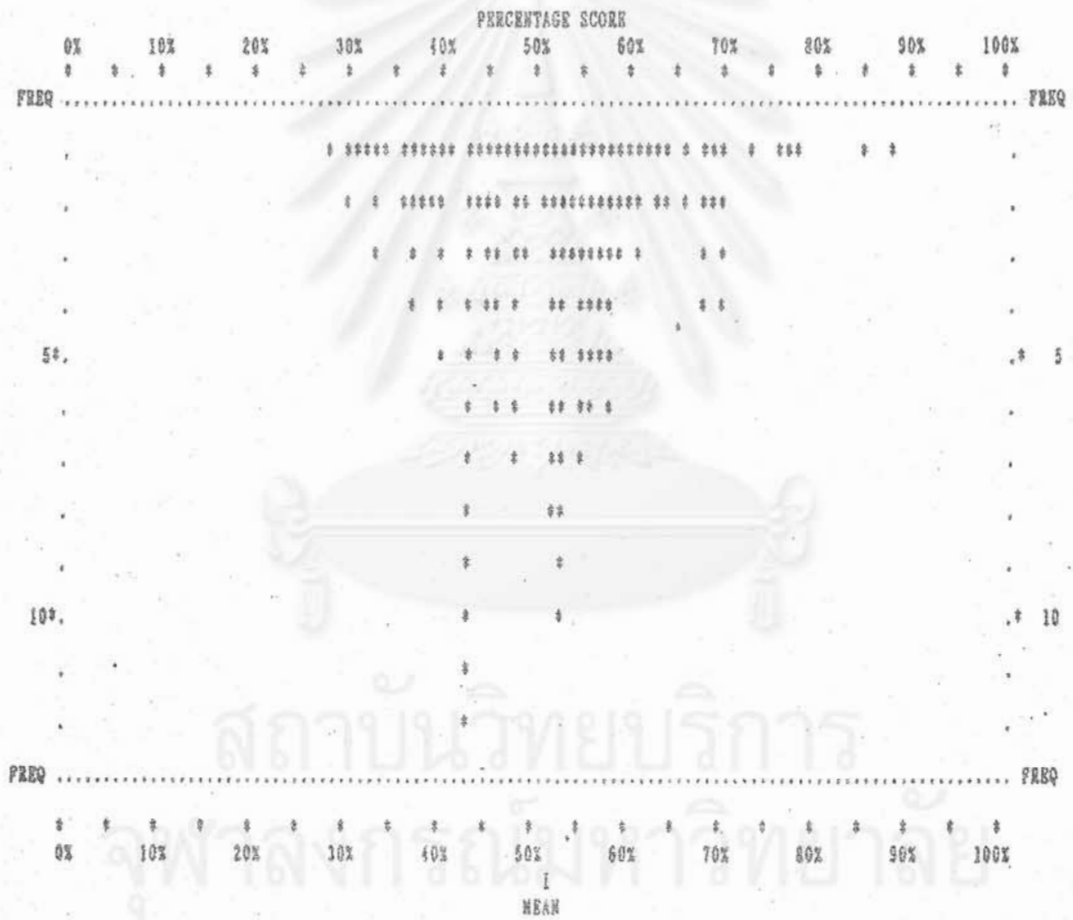
จุฬาลงกรณ์มหาวิทยาลัย

แผนภูมิที่ 5 การกระจายของคะแนนจากแบบทดสอบวัดสมิทริภาพทั่วไปทางภาษาอังกฤษฉบับที่ ๓

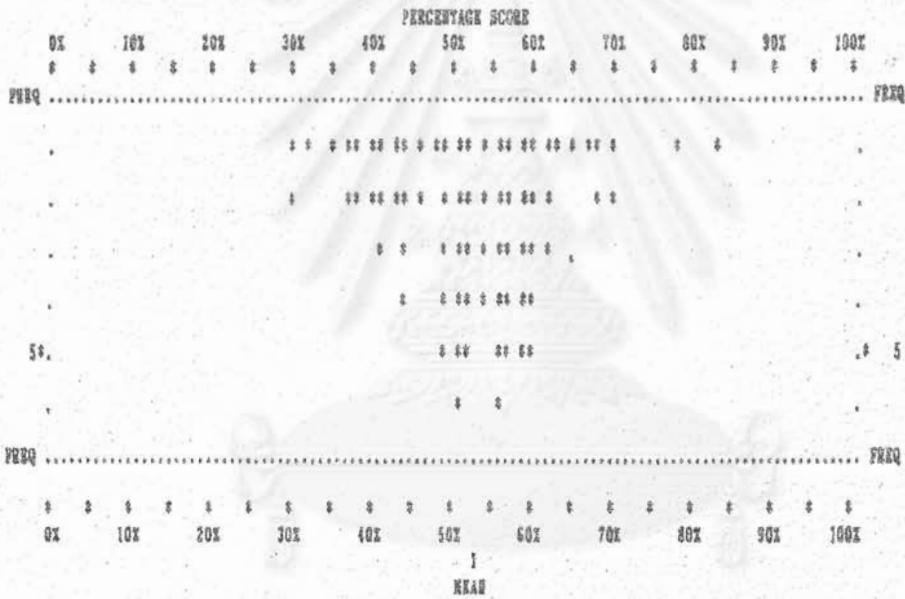


สถาบันวิจัยประชากร
จุฬาลงกรณ์มหาวิทยาลัย

แผนภูมิที่ 6 การกระจายของคะแนนจากแบบทดสอบวัดสมรรถภาพทั่วไปทางภาษาอังกฤษฉบับที่ 4



แผนภูมิที่ 7 การกระจายของคะแนนจากแบบทดสอบวัดสมรรถภาพทั่วไปทางภาษาอังกฤษฉบับปรับปรุง



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ภาคผนวก ข.

ตารางที่ 19 สรุปค่าสถิติพื้นฐานของแบบทดสอบวัดสมรรถภาพทั่วไปทางภาษาอังกฤษ ฉบับที่ 1

	mean	Min	Max	SD	Variance
Test Scores	57.909	27.000	103.000	13.974	195.293
Difficulty	0.494	0.074	0.917	0.502	0.252
Delta	13.220	7.410	18.810	2.392	5.721
Point-Biserial (RPB)	0.250	0.000	0.547	0.134	0.018
Sum of Item PQ	25.194				
P	57.909				
Reliability Statistics					
KR ₂₀ =	0.878	SE 20 =	4.875		
KR ₂₁ =	0.854	SE 21 =	5.345		
Split - Half					
RTT =	0.870	SETT =	5.029		

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 20 สรุปค่าสถิติพื้นฐานของแบบทดสอบวัดสมิทธิภาพทั่วไปทางภาษาอังกฤษ ฉบับที่ 2

	mean	Min	Max	SD	Variance
Test Scores	59.320	31.000	111.000	15.996	255.867
Difficulty	0.496	0.117	0.903	0.506	0.256
Delta	13.093	7.769	17.809	2.357	5.554
Point-Biserial (RPB)	0.371	-0.223	0.702	0.184	0.034
Sum of Item PQ	24.982				
P	59.320				
Reliability Statistics					
$KR_{20} = 0.910$		$SE_{20} = 4.800$			
$KR_{21} = 0.890$		$SE_{21} = 5.301$			
Split - Half					
$RTT = 0.899$		$SETT = 5.086$			

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 21 สรุปค่าสถิติพื้นฐานของแบบทดสอบวัดสมิทริภาพทั่วไปทางภาษาอังกฤษ ฉบับที่ 3

	mean	Min	Max	SD	Variance
Test Scores	60.737	34.000	102.000	12.945	167.563
Difficulty	0.506	0.076	0.924	0.505	0.255
Delta	12.941	7.238	18.762	2.227	4.962
Point-Biserial (RPB)	0.230	-0.175	0.566	0.141	0.020
Sum of Item PQ	25.616				
P	60.737				
Reliability Statistics					
KR ₂₀ = 0.854	SE 20 = 4.942				
KR ₂₁ = 0.828	SE 21 = 5.370				
Split - Half					
RTT = 0.864	SETT = 4.776				

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 22 สรุปค่าสถิติพื้นฐานของแบบทดสอบวัดสมรรถภาพทั่วไปทางภาษาอังกฤษ ชั้นปีที่ 4

	mean	Min	Max	SD	Variance
Test Scores	62.029	33.000	106.000	13.788	190.115
Difficulty	0.517	0.014	0.914	0.503	0.253
Delta	12.831	7.506	21.783	2.553	6.516
Point-Biserial (RPB)	0.250	-0.144	0.563	0.131	0.017
Sum of Item PQ	24.756				
P	62.029				
Reliability Statistics					
$KR_{20} = 0.877$		$SE_{20} = 4.834$			
$KR_{21} = 0.849$		$SE_{21} = 5.350$			
Split - Half					
$RTT = 0.879$		$SETT = 4.798$			

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 23 สรุปค่าสถิติพื้นฐานของแบบทดสอบวัดสมิทริภาพทั่วไปทางภาษาอังกฤษ ฉบับปรับปรุง

	mean	Min	Max	SD	Variance
Test Scores	33.120	19.000	52.000	6.572	43.188
Difficulty	0.487	0.040	0.867	0.510	0.260
Delta	13.235	8.517	20.042	2.802	7.850
Point-Biserial (RPB)	0.677	0.027	1.213	0.285	0.081
Reliability Statistics					
KR ₂₀ = 0.727	SE 20 = 3.437				
KR ₂₁ = 0.647	SE 21 = 3.907				
Cronbach Alpha					
alpha = 0.727	SEM-ALP = 3.437				
Split - Half					
RTT = 0.745	SETT = 3.321				

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ภาคผนวก ค.

ตารางที่ 24 แสดงค่าสหสัมพันธ์ระหว่างทักษะ การฟัง การอ่าน และการเขียน ของข้อสอบ
วัดสมรรถภาพทั่วไปทางภาษาอังกฤษ

แบบทดสอบ	ทักษะ	การฟัง	การอ่าน	การเขียน
ฉบับที่ 1	การฟัง	1.000	.365 ^{**}	.292 [*]
	การอ่าน		1.000	.573 ^{**}
	การเขียน			1.000
ฉบับที่ 2	การฟัง	1.000	.634 ^{**}	.526 ^{**}
	การอ่าน		1.000	.627 ^{**}
	การเขียน			1.000
ฉบับที่ 3	การฟัง	1.000	.474 ^{**}	.468 ^{**}
	การอ่าน		1.000	.596 ^{**}
	การเขียน			1.000
ฉบับที่ 4	การฟัง	1.000	.550 ^{**}	.494 ^{**}
	การอ่าน		1.000	.623 ^{**}
	การเขียน			1.000

จากตารางที่ 24 แสดงค่าสหสัมพันธ์ระหว่างทักษะการฟัง การอ่าน และการเขียนของคะแนนที่ผู้เข้าสอบได้จากแบบทดสอบทั้ง 4 ฉบับ จะเห็นได้ว่าค่าสหสัมพันธ์ที่มีนัยสำคัญทางสถิติที่ .01 จำนวน 1 คู่ ระหว่างการฟัง และการเขียน ของแบบทดสอบฉบับที่ 1 ค่าสหสัมพันธ์คู่อื่น ๆ อีก 11 คู่ ล้วนแต่มีนัยสำคัญทางสถิติที่ .001 ทั้งสิ้น ค่าสหสัมพันธ์เรียงลำดับจากมากไปหาน้อย เป็นดังนี้

ฉบับที่ 1	อ่าน - เขียน	= .579
	ฟัง - อ่าน	= .365
	ฟัง - เขียน	= .292
ฉบับที่ 2	ฟัง - อ่าน	= .634
	อ่าน - เขียน	= .627
	ฟัง - เขียน	= .526
ฉบับที่ 3	อ่าน - เขียน	= .596
	ฟัง - อ่าน	= .474
	ฟัง - เขียน	= .468
ฉบับที่ 4	อ่าน - เขียน	= .623
	ฟัง - อ่าน	= .550
	ฟัง - เขียน	= .494

สรุปได้ว่า ทักษะทางตัวอักษร คือ อ่าน - เขียน มีค่าสหสัมพันธ์กันมากที่สุด รองลงมาเป็นลำดับสอง คือ ทักษะ ฟัง - อ่าน ซึ่งเป็นทักษะรับสารเหมือนกัน ลำดับสามคือ ทักษะ ฟัง - เขียน ซึ่งเป็นทักษะทางเสียงและทักษะทางตัวอักษร

จุฬาลงกรณ์มหาวิทยาลัย

คณะกรรมการ

คณะกรรมการสร้างและพัฒนาแบบทดสอบสมรรถภาพทั่วไปทางภาษาอังกฤษ

คณะกรรมการ

รองศาสตราจารย์ ดร. อัจฉรา วงศ์โสธร	ประธาน
รองศาสตราจารย์ ภัคคินี นิยมเหตุ	กรรมการ
ผู้ช่วยศาสตราจารย์ ศิริพร พงษ์สุรพันธ์	กรรมการ
Mr. John Condon	กรรมการ
Mr. Melvyn Sharman	กรรมการ
Mr. Peter McQuay	กรรมการ
Mr. William Ferrie	กรรมการ
ผู้ช่วยศาสตราจารย์ อัมพรชาติ อ๋ปสร	กรรมการและเลขานุการ

ผู้สร้างและพัฒนาแบบทดสอบ

ผู้ช่วยศาสตราจารย์ ศิริพร พงษ์สุรพันธ์	ประธาน
รองศาสตราจารย์ กลุยา คัมภีรานนท์	กรรมการ
ผู้ช่วยศาสตราจารย์ ปิยนารถ นีกทองพรรณ	กรรมการ
ผู้ช่วยศาสตราจารย์ อัมพรชาติ อ๋ปสร	กรรมการ
ผู้ช่วยศาสตราจารย์ วัลยาพร นาวิการ	กรรมการ
ผู้ช่วยศาสตราจารย์ สุมลมาลย์ วรเสียงสุข	กรรมการ
ผู้ช่วยศาสตราจารย์ เสียงทิพย์ สุขศรี	กรรมการ
อาจารย์ อังทิพย์ รัตตะรังสี	กรรมการ
อาจารย์ ชลาธิป ชาญชัยฤกษ์	กรรมการ
Mr. Melvyn Sharman	กรรมการ
Mr. Peter McQuay	กรรมการ



คณะกรรมการดำเนินการวิจัย

ผู้วิจัย

รองศาสตราจารย์ ดร.อัจฉรา วงศ์โสธร

ผู้ช่วยศาสตราจารย์ศิริพร พงษ์สุรนิพนธ์

ผู้ช่วยวิจัย

นางสาวกิ่งกมน ทวีชาติวิทยากุล

นางจรัสศรี โภทวิฑูรธรรม

นางสาวพรทิพย์ อันทิวโรทัย



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย