

การพัฒนาแบบทดสอบภาษาอังกฤษ อิงวัตถุประสงค์แบบอุดม สำหรับการศึกษา 3 ระดับ

อัจฉรา วงศ์โสธร

1. ความสำคัญของปัญหาและวัตถุประสงค์ของการวิจัย

การวัดผลการเรียนรู้ภาษาซึ่งเป็นทั้งทักษะความรู้ และมีจิตวิสัยประกอบอยู่ด้วยไม่สามารถใช้วัตถุประสงค์แบบพฤติกรรม และการทดสอบอิงกลุ่มและอิงเกณฑ์เท่านั้น อีกทั้งเป็นที่ยอมรับกันโดยทั่วไปว่า การใช้ภาษาเกิดขึ้นในบริบททางสังคมและวัฒนธรรมมีตัวแปรซึ่งเกิดจากผู้ส่งสาร ผู้รับสาร และสภาพแวดล้อมของการสื่อสารมากมาย จึงควรที่จะใช้วิธีการทดสอบอิงวัตถุประสงค์แบบอุดมที่ใช้ในการทดสอบแบบอิงปริเจตซึ่งรวมเอามิติดังกล่าวเข้าไว้

การพัฒนาข้อทดสอบอิงวัตถุประสงค์แบบอุดมโดยใช้ตารางกำหนดลักษณะข้อสอบตามแนวการสื่อสารทางภาษาสามารถใช้เป็นกรอบอ้างอิงเดียวกันหรือตารางเดียวกันในการพัฒนาข้อสอบภาษาได้เป็นจำนวนมากโดยไม่จำกัดจำนวน ทั้งนี้เพราะธรรมชาติของภาษามีการสร้างตัวเองโดยสามารถเพิ่มคำศัพท์ได้เรื่อย ๆ และถึงแม้ว่าภาษาจะมีโครงสร้างตายตัว ก็สามารถผูกข้อความที่มีความหมายต่าง ๆ กันได้อย่างไม่จำกัดจำนวน (Chomsky, 1971: 13) ดังนั้น ข้อทดสอบภาษาที่สร้างขึ้นสามารถนำมาบรรจุเข้าธนาคารข้อสอบไว้ เพื่อการนำมาใช้ต่อ ๆ ไป ธนาคารข้อสอบมีประโยชน์อย่างมากในกระบวนการประเมินทักษะและความรู้ภาษาของผู้เรียนให้ได้มาตรฐาน ใช้ผลมาปรับปรุงคุณภาพการ

เรียนการสอน และช่วยแบ่งเบาภาระในการออกข้อสอบ ซึ่งเป็นกิจวัตรประจำภาคลงไปได้มาก

การวิจัยเรื่อง“การพัฒนาแบบทดสอบภาษาอังกฤษอิงวัตถุประสงค์แบบอุดมสำหรับการศึกษา 3 ระดับ” มีวัตถุประสงค์เพื่อพัฒนาข้อทดสอบอิงวัตถุประสงค์แบบอุดมให้เป็นมาตรฐาน มีคุณสมบัติตามเกณฑ์ โดยกำหนดประเภทของการใช้ภาษาในทักษะมาลาทางเสียง ได้แก่ การฟัง การพูด และในทักษะมาลาทางตัวอักษร ได้แก่ การอ่าน การเขียน ทั้งที่เป็นทักษะเดี่ยว และทักษะสัมพันธ์ตามหลักสูตรของแต่ละระดับ สำหรับมัธยมศึกษาปีที่ 3 ปีที่ 6 และอุดมศึกษา ระดับละ 3 ฉบับจากวัตถุประสงค์แบบอุดมและตารางกำหนดลักษณะข้อสอบเดียวกัน

2. วิธีการวิจัย

การวิจัยมี 2 ระยะดังรายละเอียดต่อไปนี้

1. ขั้นตอนการดำเนินงาน

การวิจัยระยะที่หนึ่ง (ปีการศึกษา 2529)

1. พัฒนาแบบทดสอบอิงวัตถุประสงค์แบบอุดม
การพัฒนาข้อทดสอบอิงวัตถุประสงค์แบบอุดมมีขั้นตอนดังนี้:

(1) กำหนดปริเจต

ในการกำหนดปริเจตของแบบทดสอบระดับมัธยมศึกษาต้นและมัธยมปลายนั้นผู้วิจัยใช้เอกสารดังต่อไปนี้:

- หลักสูตรภาษาอังกฤษระดับมัธยมศึกษาตอนต้น

พุทธศักราช 2521 และหลักสูตรมัธยมศึกษาตอนปลาย พุทธศักราช 2524 ของกระทรวงศึกษาธิการ

- จุดประสงค์ปลายทางการเรียนรู้วิชาภาษาอังกฤษ ระดับมัธยมศึกษาตอนต้นและจุดประสงค์ปลายทางการเรียนรู้วิชาภาษาอังกฤษระดับมัธยมศึกษาตอนปลาย สร้างขึ้นโดยหน่วยศึกษานิเทศก์ กรมสามัญศึกษา

- ผลการวิจัยของรายงานการวิจัยในระดับบัณฑิตศึกษา มหาวิทยาลัยเกษตรศาสตร์ เรื่อง “การสำรวจความต้องการในการเรียนภาษาอังกฤษของนักเรียน ผู้ปกครอง และครู ชั้นมัธยมศึกษาตอนต้น” ของสุนันทา พากเพียร และเรื่อง “การสำรวจความต้องการในการเรียนภาษาอังกฤษของนักเรียน ผู้ปกครอง และครูชั้นมัธยมศึกษาตอนปลาย” ของ วิริยา สิทธิสาร

- บทเรียนภาษาอังกฤษชุด “The Learning Kit” สำหรับระดับมัธยมศึกษาตอนต้น ของกรมวิชาการ กระทรวงศึกษาธิการ

- บทเรียนภาษาอังกฤษระดับมัธยมศึกษาตอนต้น และตอนปลายของสำนักพิมพ์ต่าง ๆ ที่ได้รับการพิจารณา โดยคณะกรรมการพิจารณาบทเรียนของกระทรวงศึกษาธิการอนุญาตให้นำมาใช้กับนักเรียนมัธยมฯ ได้ ประกอบด้วย

New Horizons in English

Odyssey

Active Context English

Lado English Series

In Touch

Practical English

Tempo

ในการกำหนดปริเฉทสำหรับแบบทดสอบอิงวัตถุประสงค์แบบอุดมในระดัอุดมศึกษานั้นผู้วิจัยได้ใช้หลักสูตรและบทเรียนวิชาภาษาอังกฤษของสถาบันภาษา จุฬาลงกรณ์มหาวิทยาลัย สำหรับรายวิชาภาษาอังกฤษ พื้นฐาน I และ II (รหัส 092-115 และ 092-116) รายวิชาภาษาอังกฤษสำหรับสาขาวิชา I สำหรับคณะ นิเทศศาสตร์ และสำหรับคณะรัฐศาสตร์ (092-223)

สำหรับคณะพาณิชยศาสตร์และการบัญชี (092-213) รายวิชาภาษาอังกฤษระดับสูงอังกฤษสังคม (092-311) อังกฤษธุรกิจ (092-312) ภาษาอังกฤษสื่อสาร (092-341) และรายวิชาภาษาอังกฤษวิชาเลือก การพูดภาษาอังกฤษ (092-192)

2. สร้างวัตถุประสงค์แบบอุดมจากหลักสูตรและบทเรียนของการศึกษาสามระดับที่กล่าวถึงใน ข้อ 1

3. สร้างตารางกำหนดลักษณะของข้อสอบจากวัตถุประสงค์แบบอุดมในข้อ 2 สำหรับเป็นกรอบของการพัฒนาข้อทดสอบอิงวัตถุประสงค์แบบอุดม มีพารามิเตอร์ของหน้าที่ของเนื้อความ (function) สถานการณ์ (situation/setting) ทักษะมาลา (modality) ประเภทของเนื้อความ (genre) ช่องทางของการใช้ภาษา (channel) บทบาทของผู้สื่อสาร (role) สถานภาพของผู้สื่อสาร (status) และหัวข้อของเนื้อความ (topic)

4. ทดลองแบบทดสอบที่สร้างขึ้น 2 ครั้งดังนี้:

ครั้งที่ 1 ทดลองกับกลุ่มตัวอย่างประชากรกลุ่มที่ 1 และนำผลการสอบมาวิเคราะห์หาค่าความยากง่าย และค่าอำนาจจำแนก ตามวิธีการวิเคราะห์แบบประเพณีนิยม เพื่อปรับปรุงเนื้อหาของข้อสอบ การพิจารณาคุณภาพของข้อสอบได้ใช้เกณฑ์ดังนี้:

ค่าความยากง่าย	คุณภาพ
.40-.60	ดีมาก
.31-.39, .61-.69	ดี
.20-.30, .70-.80	พอใช้ได้
.81 ขึ้นไป	ง่ายไป
ต่ำกว่า .20	ยากไป

ค่าอำนาจจำแนก

ค่าอำนาจจำแนกที่ใช้คือค่า r_{pb} เพราะไม่ได้ยึดการกระจายคะแนนแบบปกติ (normal distribution) การพิจารณาปรับปรุงข้อสอบใช้เกณฑ์ดังนี้:

ค่าอำนาจจำแนก	คุณภาพ
.60>	ดีมาก
.35-.59	ดี
.20-.34	พอใช้ได้
ต่ำกว่า .20	ต่ำ
ติดลบ	ต้องตัดทิ้งหรือเปลี่ยนแปลง

แบบทดสอบอิงวัตถุประสงค์แบบอุดมไม่ได้มุ่งที่จะจำแนกผู้เข้าสอบออกเป็นกลุ่มเก่งและกลุ่มอ่อนเหมือนกับแบบทดสอบอิงกลุ่ม ในการปรับปรุงจึงให้ความสำคัญแก่ค่าความยากง่ายมากกว่าค่าอำนาจจำแนก และปรับปรุง/เปลี่ยนแปลงเฉพาะข้อที่มีค่าอำนาจจำแนกติดลบ โดยดูค่าเฉลี่ยความยากง่ายของทั้งฉบับประกอบคุณภาพรายข้อพิจารณาค่าความยากง่ายและค่าอำนาจจำแนก แบ่งเป็น

(1) ข้อที่ใช้ได้คือข้อที่มีค่าความยากง่ายระหว่าง .20-.80 และค่าอำนาจจำแนกเป็นบวก

(2) ข้อที่ต้องปรับปรุงคือข้อที่มีค่าความยากง่ายระหว่าง .20-.80 และมีค่าอำนาจจำแนกติดลบ และข้อที่มีค่าความยากง่ายต่ำกว่า .20 หรือ สูงกว่า .80 และมีค่าอำนาจจำแนกเป็นบวก

(3) ข้อที่ต้องตัดทิ้งหรือเปลี่ยนแปลงคือข้อที่มีค่าความยากง่ายต่ำกว่า .20 หรือ สูงกว่า .80 และมีค่าอำนาจจำแนกติดลบ

สำหรับคุณภาพของตัวลวง (distractor) ต้องปรับปรุงตัวเลือกที่เป็นตัวลวงเมื่อตัวเลือกนั้นมีผู้เลือกตอบไม่ถึงร้อยละ 5 ($p < .05$) และตัวเลือกนั้นมีผู้ตอบในกลุ่มที่ได้คะแนนสูงเลือกตอบมากกว่ากลุ่มที่ได้คะแนนต่ำ ($N_H > N_L$)

ครั้งที่ 2 นำข้อสอบที่ปรับปรุงแล้วไปทดสอบกับกลุ่มตัวอย่างประชากรชุดเดิม ห่างจากการทดสอบครั้งที่ 1 เป็นเวลา 1 เดือนแล้วหาคุณภาพของข้อสอบโดยการหาค่าดัชนี ความง่าย และดัชนีความไว ใช้เกณฑ์การยอมรับข้อทดสอบที่มีดัชนีความง่ายระหว่าง .20 ขึ้นไป เชื่อถือได้ (reliability) แบบ Cronbach Alpha แบบ inter-rater และหาค่าความเที่ยงตรง (validity) โดยวิธีการของ Swaminathan, Hambleton & Algina (สมหวัง พิธิยานุวัฒน์, 2529)

5. นำแบบทดสอบที่ได้ผ่านการวิเคราะห์ 2 ครั้งไปทดสอบกลุ่มตัวอย่างประชากรของการวิจัย

6. วิเคราะห์หาคุณภาพของข้อสอบเป็นรายข้อ โดยใช้วิธี Rasch Model ซึ่งมีขั้นตอนดังนี้: (1) ดูค่า

ความยากง่าย (B) ว่าอยู่ในพิสัยของการยอมรับได้ คือ ± 3.00 หรือไม่

(2) ดูค่าความสอดคล้องของโค้งคุณลักษณะข้อสอบ/ ICC (item characteristic curve) ว่าเป็นไปตามความคาดหวังหรือว่าสอดคล้องมากน้อยเพียงใด โดยดูจากค่า weighted mean square (weighted MNSQ) ว่ามีค่าใกล้เคียงกับ 1.00 หรือไม่ และนำค่าความคลาดเคลื่อนมาตรฐานของการวัด (Se_m) มาบวกและลบในการพิจารณาค่าความสอดคล้องด้วย

7. หาคุณภาพของแบบทดสอบทั้งฉบับโดยวิธีการดังนี้:

ค่าความเชื่อถือได้ สำหรับแบบทดสอบปรนัย ใช้วิธี Cronbach Alpha

แบบทดสอบอัตนัยด้านการเขียน การฟัง-เขียน การอ่าน-เขียน และการฟัง-อ่าน-เขียน ใช้วิธีที่ให้มีผู้ตรวจ 2 คน และหาค่าสหสัมพันธ์ในการหาค่าความเชื่อถือได้แบบ inter-rater

แบบทดสอบอัตนัยด้านการพูดที่ให้ผู้เรียนประเมินความสามารถของตนเองในการพูดตามวัตถุประสงค์ของหลักสูตรนั้น ใช้วิธีหาค่าความคงที่ของการประเมินความสามารถของตนเองโดยให้ผู้เรียนประเมิน 2 ครั้ง ในระยะเวลาห่างกัน ประมาณ 4 สัปดาห์

ค่าความเที่ยงตรง ใช้วิธีวิเคราะห์ความเที่ยงตรงเชิงเนื้อหา (content validity) แบบความสอดคล้องระหว่าง ข้อสอบกับวัตถุประสงค์ (test-item objective congruence) โดยให้ครูที่ได้รับการอบรมเกี่ยวกับหลักสูตรและการทดสอบ จำนวน 10 คนประเมิน สำหรับระดับมัธยมศึกษา ส่วนในระดับอุดมศึกษาให้ผู้สอนรายวิชาที่เป็นกรอบอ้างอิงของการพัฒนาแบบทดสอบ จำนวน 10 คนเป็นผู้ประเมิน

8. สร้างแบบทดสอบคู่ขนานกับแบบทดสอบชุดแรก จำนวน 2 ฉบับ ต่อระดับการศึกษาแต่ละระดับรวมเป็นแบบทดสอบอิงวัตถุประสงค์แบบอุดม 3 ฉบับต่อระดับคือ Form A Form B และ Form C รวมทั้ง 3 ระดับเป็น 9 ฉบับ โดยแต่ละชุดสร้างขึ้นจากวัตถุประสงค์แบบอุดม

และตารางกำหนดลักษณะ ข้อสอบของฉบับแรก (Form A)

9. หาค่าความเที่ยงตรงเชิงเนื้อหาของแบบทดสอบ ทั้ง 3 ฉบับโดยให้ครูผู้สอนในแต่ละระดับการศึกษาจำนวน 10 คน เป็นผู้ประเมินค่าความสอดคล้องของข้อสอบกับ วัตถุประสงค์แบบอุดมของระดับการศึกษาเป็นรายข้อ

การวิจัยระยะที่สอง (ปีการศึกษา 2530)

การดำเนินงานมีขั้นตอนดังนี้:

1. นำแบบทดสอบทั้ง 3 ฉบับที่สร้างจากวัตถุประสงค์แบบอุดมและตารางกำหนดลักษณะข้อสอบเดียวกันไปทดลองใช้กับนักเรียนชั้นมัธยมศึกษาปีที่ 3 และปีที่ 6 ในโรงเรียนขนาดต่าง ๆ ในภาคต่าง ๆ โดยความอนุเคราะห์ของสำนักงานทดสอบ กระทรวงศึกษาธิการ ส่วนแบบทดสอบระดับอุดมศึกษาทั้ง 3 ฉบับได้ให้นิสิตจุฬาลงกรณ์มหาวิทยาลัยในชั้นปีที่ 1 2 และ 3 ที่ลงทะเบียนเรียน รายวิชาภาษาอังกฤษตามหลักสูตรของสถาบันภาษาทำ

2. นำผลการสอบที่ได้มาวิเคราะห์หาค่าความยากง่ายของข้อสอบแต่ละข้อในแต่ละฟอร์มโดยวิธี Rasch Model และหาค่าความสอดคล้องของข้อสอบกับโครงสร้างคุณลักษณะข้อสอบ

3. หาค่าความคงที่ของการให้คะแนนเพื่อเป็นการยืนยันว่าแบบทดสอบฟอร์ม A B และ C ของแต่ละระดับอยู่ในโครงสร้างคุณลักษณะข้อสอบมีคุณสมบัติคู่ขนานกัน โดยที่ค่าเฉลี่ยความยากง่าย ไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติโดยการวิเคราะห์ที่ใช้ F-test

4. หาค่าความเชื่อถือได้แบบ Cronbach Alpha ของแบบทดสอบแต่ละฉบับ

5. หาค่าความคู่ขนานกันของแบบทดสอบทั้ง 3 ฉบับของแต่ละระดับการศึกษา โดยใช้ F-test ค่าเฉลี่ยความยากง่ายของแบบทดสอบแต่ละฉบับที่ประกอบด้วยข้อทดสอบที่สอดคล้องกับโครงสร้างคุณลักษณะของข้อสอบว่าไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ $p.05$ หรือไม่

6. กำหนดคะแนนมาตรฐานตามปริเฉทของเนื้อหา เพื่อกำหนดว่าผู้ที่มีความสามารถทำกลุ่มข้อสอบใดตามมวลของเนื้อหาเป็นผู้ที่มีความสามารถเป็นระดับเปอร์เซ็นต์ใดที่เท่าใด

2. กลุ่มตัวอย่างประชากร

กลุ่มตัวอย่างประชากรในระดับมัธยมศึกษาได้มาจากการสุ่มแบบหลายขั้นตอนอย่างมีชั้นภูมิ โดยขั้นแรกสุ่มจากขนาดของโรงเรียนตามการกำหนดของกรมสามัญศึกษา กระทรวงศึกษาธิการประกอบด้วยโรงเรียนขนาดเล็ก ขนาดกลาง ขนาดใหญ่ และขนาดใหญ่พิเศษ ชั้นที่ 2 สุ่มโดยใช้เกณฑ์มาตรฐานทางวิชาการคือเป็นโรงเรียนดี และปานกลาง และชั้นที่ 3 สุ่มจากชั้นเรียนของโรงเรียนขนาดต่าง ๆ

กลุ่มตัวอย่างประชากรในระดับอุดมศึกษาสุ่มจากนิสิตจุฬาลงกรณ์มหาวิทยาลัยปีที่ 1 2 และ 3 ที่เรียนในรายวิชาภาษาอังกฤษของสถาบันภาษา ได้มาโดยการสุ่มอย่างมีชั้นภูมิ (stratified random sampling)

กลุ่มตัวอย่างในการวิจัยระยะที่หนึ่ง

การเก็บข้อมูลครั้งแรกซึ่งนับเป็นการศึกษานำเป็นพื้นฐานของการคำนวณ เป็นกลุ่มตัวอย่างของการศึกษานำครั้งที่ 1 ในชั้นมัธยมศึกษาปีที่ 3 จำนวน 176 คน มัธยมศึกษาปีที่ 6 จำนวน 142 คน และนิสิตจุฬาลงกรณ์มหาวิทยาลัย 89 คน สำหรับการศึกษานำครั้งที่ 2 เป็นกลุ่มตัวอย่างชั้นมัธยมศึกษาปีที่ 3 จำนวน 42 คน มัธยมศึกษาปีที่ 6 จำนวน 41 คน และนิสิตจุฬาลงกรณ์มหาวิทยาลัย 47 คน

ขนาดของกลุ่มตัวอย่างประชากรในการวิจัยได้มาจากการใช้ค่าเบี่ยงเบนมาตรฐานของการศึกษานำ มีค่าความเชื่อถือได้ร้อยละ 99 เป็นกลุ่มตัวอย่างในระดับมัธยมศึกษาปีที่ 3 จำนวน 697 คน มัธยมศึกษาปีที่ 6 จำนวน 525 คน และเป็นนิสิตจุฬาลงกรณ์มหาวิทยาลัย 496 คน

(ดูรายละเอียดการเก็บข้อมูลใน อัจฉรา วงศ์โสธร. 2536 ระดับทักษะภาษาอังกฤษของนักศึกษาไทย)

กลุ่มตัวอย่างในการวิจัยระยะที่สอง

การเก็บข้อมูลเพื่อหาความคู่ขนานของแบบทดสอบ Form A B และ C ของแต่ละระดับการศึกษา การหาคุณภาพของแบบทดสอบ และการกำหนดคะแนนตามปริเฉทของเนื้อหาใช้กลุ่มตัวอย่างประชากรในระดับมัธยมศึกษาปีที่ 3 จำนวน 563 คน ทำ Form A 183 คน Form B 156 คน Form C 224 คน กลุ่มตัวอย่างในชั้นมัธยมศึกษาปีที่ 6 มี 469 คน ทำ Form A 152 คน Form B 150 คน Form C 167 คน ในระดับอุดมศึกษาใช้กลุ่มตัวอย่างประชากร 257 คน ทำ Form A 87 คน Form B 94 คน Form C 78 คน (ดูภาคผนวกที่ 1)

3. เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการวิจัยเป็นแบบทดสอบอิงวัตถุ-ประสงค์แบบอ้อมที่สร้างโดยใช้ขั้นตอนที่กล่าวไว้ในตอนที่ 2

วิทยวิธีของการทดสอบ (testing method) มี 2 แบบคือแบบทักษะเดียว ได้แก่การทดสอบการฟัง การพูด การอ่าน การเขียน และแบบทักษะสัมพันธ์ ได้แก่การทดสอบการ ฟัง-เขียน การฟัง-อ่าน-เขียน การอ่าน-เขียน และการพูดวัดแบบอ้อมโดยการให้อ่านบทสนทนา การตรวจให้คะแนนมี 2 วิธีคือ

(1) แบบปรนัย มีการให้คะแนน 1 สำหรับข้อที่ถูก และ 0 สำหรับข้อที่ผิดหรือละไว้ไม่ตอบ

(2) แบบอัตนัย ซึ่งมีการให้คะแนนแบบพหุคูณ ได้แก่การทดสอบการฟัง-เขียน การฟัง-อ่าน-เขียน การอ่าน-เขียน การเขียนเรียงความ และการประเมินความสามารถของตนเองในการพูดใช้เกณฑ์การให้คะแนนดังนี้

- 1-5 คือ 1 = อ่อนมากยังต้องแก้ไขปรับปรุงอีกมาก
- 2 = ค่อนข้างอ่อนยังต้องแก้ไขปรับปรุง
- 3 = ปานกลางยังต้องแก้ไขปรับปรุง
- 4 = ค่อนข้างดียังมีที่ต้องแก้ไขปรับปรุงบ้าง
- 5 = ดี-ดีมากไม่ต้องแก้ไขปรับปรุง

และได้กำหนดให้คะแนน 1 และ 2 = 0 (ตก) และ 3-5 = 1 (ได้) ในการวิเคราะห์ข้อทดสอบแบบอัตนัย

4. การเก็บข้อมูล

การเก็บข้อมูลทำระหว่างเดือนสิงหาคม 2529-กุมภาพันธ์ 2531 โดยเก็บเป็น 4 ระยะดังนี้ :

(1) เก็บข้อมูลในการทดลองแบบทดสอบอิงวัตถุ-ประสงค์แบบอ้อมครั้งที่ 1 กับกลุ่มตัวอย่างในชั้นมัธยมศึกษาปีที่ 3 จำนวน 176 คน ชั้นมัธยมศึกษาปีที่ 6 จำนวน 42 คน และนิสิตจุฬาย 89 คน

(2) เก็บข้อมูลในการทดลองแบบทดสอบอิงวัตถุ-ประสงค์แบบอ้อมครั้งที่ 2 กับกลุ่มตัวอย่างในชั้นมัธยมศึกษาปีที่ 3 จำนวน 42 คน ชั้นมัธยมศึกษาปีที่ 6 จำนวน 41 คน และนิสิตจุฬาย 47 คน

(3) เก็บข้อมูลในการวิจัยเพื่อหาคุณภาพของแบบทดสอบอิงวัตถุ-ประสงค์แบบอ้อมเป็นรายข้อ และของทั้งฉบับ โดยใช้กลุ่มตัวอย่างประชากรที่เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 จำนวน 679 คน ชั้นมัธยมศึกษาปีที่ 6 จำนวน 525 คน และนิสิตจุฬาย 493 คน

(4) เก็บข้อมูลเพื่อหาความคู่ขนานกันของแบบทดสอบฟอร์ม A B และ C เพื่อหาคุณภาพของแบบทดสอบและเพื่อกำหนดคะแนนมาตรฐานตามปริเฉทของเนื้อหาของแบบทดสอบในเดือนมกราคม-กุมภาพันธ์ 2531 กลุ่มตัวอย่างประชากรเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 จำนวน 563 คน มัธยมศึกษาปีที่ 6 จำนวน 469 คน และนิสิตจุฬาลงกรณ์มหาวิทยาลัย 257 คน

(5) สอบถามความคิดเห็นของผู้สอนภาษาอังกฤษในชั้นมัธยมศึกษาปีที่ 3 ปีที่ 6 และระดับอุดมศึกษาปีที่ 1 ปีที่ 2 และปีที่ 3 เกี่ยวกับความสอดคล้องระหว่างข้อทดสอบใน Form A Form B และ Form C ของแต่ละระดับการศึกษากับวัตถุ-ประสงค์แบบอ้อมของแต่ละระดับเพื่อหาค่าความเที่ยงตรงเชิงเนื้อหาของแบบทดสอบ

5. การวิเคราะห์ข้อมูล

การวิเคราะห์ข้อมูลเป็นการวิเคราะห์ข้อทดสอบและแบบทดสอบ

การวิเคราะห์ข้อทดสอบ

ใช้วิธีวิเคราะห์ดังนี้:

1. ใช้ 27% ของกลุ่มที่ได้คะแนนสูง และกลุ่มที่ได้คะแนนต่ำตามวิธีของ Chung Teh Fan เพื่อหาค่าความยากง่าย FV และ ค่า delta และค่าอำนาจจำแนก r biserial การวิเคราะห์ที่ใช้โปรแกรมคอมพิวเตอร์ของสถาบันภาษาและเครื่องคอมพิวเตอร์ของสถาบันบริการคอมพิวเตอร์ของจุฬาลงกรณ์มหาวิทยาลัย

2. ใช้วิธีวิเคราะห์หาค่าดัชนีความง่ายและดัชนีความไว (สมหวัง พิธิยานุวัฒน์, 2529)

3. ใช้วิธีวิเคราะห์แบบ Rasch Model ด้วยโปรแกรมคอมพิวเตอร์ BICAL Sub-program UCON ของสถาบันบริการคอมพิวเตอร์จุฬาลงกรณ์มหาวิทยาลัย เพื่อหาความสอดคล้องของค่าความยากง่ายของข้อสอบกับโค้ง ICC และคัดเฉพาะข้อที่สอดคล้องไว้สำหรับนำเข้าไปในธนาคารข้อทดสอบ

4. ใช้วิธีวิเคราะห์หาค่าความสอดคล้องของข้อทดสอบกับวัตถุประสงค์แบบอุดมโดยใช้วิธี item-objective congruence ของแบบทดสอบ Form A Form B และ Form C ของแต่ละระดับการศึกษา และเพื่อยืนยันความเที่ยงตรงเชิงเนื้อหาของแบบทดสอบ

การวิเคราะห์แบบทดสอบ

ใช้วิธีวิเคราะห์ดังนี้:

1. หาค่าความเชื่อถือได้ของแบบทดสอบโดยหาค่า Cronbach Alpha

2. หาค่าความเชื่อถือได้แบบความเท่ากันของการให้คะแนน (inter-rater reliability) และแบบความคงที่ของการประเมินความสามารถในการพูดตามวัตถุประสงค์ของตนเองโดยใช้ค่า สหสัมพันธ์ Pearson-product moment

3. หาค่าความเที่ยงตรงของแบบทดสอบใช้วิธีหาค่าความสอดคล้องของข้อสอบกับวัตถุประสงค์แบบอุดมซึ่งเป็นค่าความเที่ยงตรงเชิงเนื้อหาของแบบทดสอบ Form A Form B และ Form C ของแต่ละระดับการศึกษา

4. พิสูจน์ความคู่ขนานกันของแบบทดสอบทั้ง 3 ฉบับของแต่ละระดับการศึกษาที่ประกอบด้วยข้อสอบที่สอดคล้องกับโค้งลักษณะข้อสอบจากการ วิเคราะห์แบบ

Rasch Model โดยการใช้ F-test หาค่าความแตกต่างกันของค่าเฉลี่ยความยากง่ายของแบบทดสอบว่ามีนัยสำคัญทางสถิติที่ $p.05$ หรือไม่

3. ผลของการวิจัย

การวิจัยได้พัฒนาวัตถุประสงค์แบบอุดมจากหลักสูตรภาษาอังกฤษของระดับมัธยมศึกษาปีที่ 3 ปีที่ 6 และระดับอุดมศึกษาปีที่ 1-3 ของสถาบันภาษา จุฬาลงกรณ์มหาวิทยาลัย พร้อมทั้งตารางกำหนดลักษณะของข้อสอบของแต่ละระดับการศึกษา

แบบทดสอบภาษาอังกฤษอิงวัตถุประสงค์แบบอุดมที่พัฒนาขึ้นในระดับมัธยมศึกษาปีที่ 3 และปีที่ 6 ประกอบด้วยแบบทดสอบวัดทักษะเดี่ยวด้านการฟัง 10 ข้อ การอ่าน 10 ข้อ การเขียน 1 ข้อ การพูดแบบประเมินความสามารถของตนเอง 1 ชุด และแบบทดสอบทักษะสัมพันธ์ วัดการฟัง-เขียน การฟัง-อ่าน-เขียน การพูด-อ่าน (การพูดวัดแบบอ้อม) และการอ่าน-เขียน อย่างละ 1 ชุด สำหรับแบบทดสอบทักษะเดี่ยวระดับอุดมศึกษาประกอบด้วย การอ่าน 15 ข้อ การพูด (แบบประเมินความสามารถของตนเอง) 1 ชุด การเขียน 1 ข้อ และแบบทดสอบทักษะสัมพันธ์ วัดการพูด-อ่าน (การพูดวัดแบบอ้อม) การฟัง-เขียน การฟัง-อ่าน-เขียน และการอ่าน-เขียน อย่างละ 1 ชุด

แบบทดสอบที่นำมาทดลองปรับปรุงให้มีค่าสถิติเข้าตามเกณฑ์มาตรฐานคือแบบทดสอบ Form A ของแต่ละระดับการศึกษา

ผลการวิเคราะห์ตามแบบประเพณีนิยม (Chung Teh Fan) พบว่าแบบทดสอบในระดับมัธยมศึกษาปีที่ 3 มีข้อทดสอบที่มีคุณภาพอยู่ในเกณฑ์มาตรฐาน 24 ข้อที่ต้องแก้ไขปรับปรุงเพราะง่ายเกินไป 2 ข้อ และยากเกินไป 9 ข้อ คิดเป็นข้อสอบที่อยู่ในเกณฑ์ร้อยละ 69 ในระดับมัธยมศึกษาปีที่ 6 มีข้อทดสอบที่อยู่ในเกณฑ์มาตรฐาน 32 ข้อ ต้องแก้ไขปรับปรุงเพราะง่ายเกินไป 2 ข้อ และยากเกินไป 1 ข้อ คิดเป็นข้อสอบที่อยู่ในเกณฑ์ร้อยละ 91 ในระดับอุดมศึกษา มีข้อทดสอบที่มีคุณภาพอยู่ในเกณฑ์ 27 ข้อ ที่ต้องแก้ไขปรับปรุงเพราะ

ยากเกินไป 3 ข้อ และไม่มีข้อที่ง่ายเกินไป คิดเป็นข้อทดสอบที่อยู่ในเกณฑ์ร้อยละ 90

ผลการวิเคราะห์ข้อทดสอบครั้งที่ 2 ที่ปรับปรุงแล้ว ด้วยการหาดัชนีความง่ายและดัชนีความไว พบว่าในระดับมัธยมศึกษาปีที่ 3 มีพิสัยดัชนีความง่ายระหว่าง .357-.571 ซึ่งอยู่ในเกณฑ์มาตรฐาน ส่วนดัชนีความไวมีค่าติดลบเพียงข้อเดียวคือข้อทดสอบการอ่านข้อที่ 5 ซึ่งต้องแก้ไขปรับปรุง ในระดับมัธยมศึกษาปีที่ 6 พบว่าข้อทดสอบมีพิสัยดัชนีความง่ายระหว่าง .340-.634 ซึ่งอยู่ในเกณฑ์มาตรฐาน และมีดัชนีความไวมีค่าบวกทุกข้อเข้าเกณฑ์เช่นเดียวกัน ส่วนแบบทดสอบระดับอุดมศึกษา พบว่าข้อทดสอบมีพิสัยดัชนีความง่าย .319-.590 ซึ่งอยู่ในเกณฑ์มาตรฐาน และมีข้อที่มีดัชนีความไวติดลบ 2 ข้อ ซึ่งต้องปรับปรุงแก้ไขคือการอ่านข้อที่ 2 และ 14

ผลการวิเคราะห์แบบทดสอบทั้งฉบับเพื่อหาค่ามัชฌิมเลขคณิต ค่าส่วนเบี่ยงเบนมาตรฐาน คะแนนสูงสุด คะแนนต่ำสุด และค่ามัชฌิมฐานของคะแนน ค่าความแปรปรวน ค่าเฉลี่ยความยากง่าย และอำนาจจำแนก ตลอดจนค่าความเชื่อถือได้ และค่าความเที่ยงตรงของแบบทดสอบ พบว่า แบบทดสอบในทั้ง 3 ระดับมีค่าความยากง่ายอยู่ในเกณฑ์ดีมากใกล้เคียงกัน คือ ค่าเฉลี่ย p .371, .421 และ .427 ตามลำดับโดยที่ค่าความยากแบบ delta มีค่าเฉลี่ย 14.588, 13.895 และ 13.903 ตามลำดับ ส่วนค่าอำนาจจำแนกแบบ r_{pb} อยู่ในเกณฑ์มาตรฐาน .466, .451 และ .448 ตามลำดับ ค่าความเชื่อถือได้แบบ Cronbach Alpha อยู่ในระดับปานกลาง คือ .568 .641 และ .590 ตามลำดับ การวิเคราะห์เฉพาะแบบทดสอบอันนี้ที่มีการให้คะแนนแบบพหุคูณ พบว่ามีค่าความเชื่อถือได้แบบความเท่ากันของการให้คะแนน r_{xy} .91, .89 และ .95 ตามลำดับ ส่วนค่าความเที่ยงตรงเชิงเนื้อหาแบบความสอดคล้องของข้อทดสอบกับวัตถุประสงค์แบบอุดม พบว่า มีค่า I_{10} .90, .92 และ .86 ตามลำดับ

สรุปได้ว่าแบบทดสอบในทั้ง 3 ระดับมีค่าเฉลี่ยความยากง่ายและค่าอำนาจจำแนกอยู่ในเกณฑ์มาตรฐาน

มีค่าความเชื่อถือได้แบบ Cronbach Alpha ในระดับปานกลาง ค่าความเที่ยงตรงแบบความคงที่ของการให้คะแนนและแบบความสอดคล้องกับวัตถุประสงค์อยู่ในระดับสูง

ผลการพัฒนาแบบทดสอบคู่ขนานเพิ่มขึ้นจาก Form A อีก 2 ฉบับ คือ Form B และ Form C ในระดับมัธยมศึกษาปีที่ 3 และปีที่ 6 ได้แบบทดสอบที่มีค่าความสอดคล้องกับวัตถุประสงค์แบบอุดม หรือ ค่า I_{10} อยู่ในระดับสูง-สูงมาก ในระดับมัธยมศึกษาปีที่ 3 มีค่า I_{10} ของ Form A = .92 Form B = .91 และ Form C = .93 ในระดับมัธยมศึกษาปีที่ 6 ค่า I_{10} ของ Form A = .90 Form B = .90 และ Form C = .95 ในระดับอุดมศึกษา ค่า I_{10} ของ Form A = .94 Form B = .95 และ Form C = .90 สรุปได้ว่าค่าความเที่ยงตรงเชิงเนื้อหาของแบบทดสอบทั้ง 3 ฉบับ ในแต่ละระดับการศึกษามีค่าสูง

ผลการวิเคราะห์ข้อทดสอบตามวิธี Rasch Model เพื่อคัดเลือกเฉพาะข้อทดสอบที่สอดคล้องกับโครงสร้างของข้อสอบ (item characteristic curve/ICC) ได้ข้อที่สอดคล้อง 32 ข้อจาก 35 ข้อในระดับมัธยมศึกษาปีที่ 3 สอดคล้องทุกข้อ ในระดับมัธยมศึกษาปีที่ 6 และสอดคล้อง 28 จาก 30 ข้อในระดับอุดมศึกษา สรุปได้ว่ามีข้อทดสอบที่สามารถนำเข้าธนาคารข้อสอบ ร้อยละ 91 ในระดับมัธยมศึกษาปีที่ 3 ทุกข้อหรือ 100 เปอร์เซ็นต์ในระดับมัธยมศึกษาปีที่ 6 และร้อยละ 93 ในระดับอุดมศึกษา ซึ่งนับว่าเป็นที่น่าพอใจ

การวิเคราะห์ F-test ค่าเฉลี่ยความยากง่ายของแบบทดสอบ Form A B และ C ของแต่ละระดับการศึกษาที่ประกอบด้วยข้อสอบที่สอดคล้องกับโครงสร้างข้อสอบ พบว่าไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ p .05 สรุปได้ว่าแบบทดสอบที่คัดเลือกแล้วมีความคู่ขนานกัน

เกณฑ์ระดับความสามารถในการสื่อสารทางภาษา ทั้งด้านทักษะเดียวและทักษะสัมพันธ์พัฒนาขึ้นจากคะแนนที่ได้จากการสอบแบบทดสอบอิงวัตถุประสงค์แบบอุดม

โดยใช้เกณฑ์ 7 ระดับที่ดัดแปลงจากเกณฑ์ของ The Council of Europe ซึ่งกำหนดให้ระดับ 3 เป็นระดับ Threshold หรือระดับของการเริ่มสื่อสารในภาษาต่างประเทศได้ (ดูเกณฑ์ที่กำหนดในภาคผนวก)

3. อภิปรายผล

แบบทดสอบภาษาอังกฤษอิงวัดจุดประสงค์แบบอุดม ที่สร้างขึ้นและพัฒนาให้เป็นมาตรฐานมีคุณภาพตามเกณฑ์

งานวิจัยเรื่องนี้เป็นการศึกษาวิจัยทางการทดสอบภาษา ตามนัยที่หนึ่งในนัยสามนัยดังนี้ (อัจฉรา วงศ์โสธร, 2535: 4-5):

นัยที่หนึ่ง หมายถึงการพัฒนาข้อทดสอบทางภาษาขึ้นมาให้เป็นมาตรฐาน โดยมีขั้นตอนดำเนินงาน ดังนี้ :

- (1) ศึกษาเกี่ยวกับคุณลักษณะหรือภาวสันนิษฐาน (construct) ที่จะวัดและกำหนดว่าประกอบด้วยอะไร
- (2) สร้างเครื่องมือหรือข้อทดสอบที่มีความเชื่อถือได้ (reliable) และความเที่ยงตรง (valid) ขึ้นเพื่อวัดคุณลักษณะที่ประสงค์
- (3) นำข้อทดสอบไปทดลองใช้และวิเคราะห์คุณภาพเป็นรายข้อและทั้งฉบับ
- (4) ปรับปรุงข้อทดสอบจากสถิติวิเคราะห์ที่ได้จากการทดลองข้อทดสอบ
- (5) นำข้อทดสอบไปทดลองใช้อีกครั้งหนึ่งแล้วตรวจสอบค่าความยากง่ายค่าความเที่ยงและค่าความตรง

นัยที่สอง หมายถึงการศึกษาเกี่ยวกับความสามารถ สมรรถิผล สมรรถวิสัยทางภาษาโดยการนำเครื่องมือที่เป็นแบบทดสอบทางภาษาไปใช้ในการเก็บข้อมูล

นัยที่สาม หมายถึงการศึกษาเกี่ยวกับองค์ประกอบหรือปัจจัย (factor) ของทักษะและความสามารถในการใช้ภาษาว่ามีลักษณะเช่นใด เช่นงานวิจัยของ Oller & Hinofotis (1980) Bachman & Palmer (1981) Carrol (1987) และ อัจฉรา วงศ์โสธร (1990)

งานวิจัยเรื่อง“การพัฒนาแบบทดสอบภาษาอังกฤษ

อิงวัดจุดประสงค์แบบอุดมสำหรับการศึกษา 3 ระดับ” เป็น การวิจัยทางการทดสอบโดยนัยที่หนึ่งตามที่ได้กล่าวไว้ข้างต้นซึ่งสามารถนำไปสู่การสร้างธนาคารข้อทดสอบภาษาขึ้นใช้ได้อย่างต่อเนื่องและมีการเพิ่มพูนจำนวนของข้อสอบในธนาคารได้อย่างต่อเนื่อง

การวิเคราะห์ข้อทดสอบโดยวิธีประเพณีนิยม (classical) และวิธี Rasch Model เพื่อการตัดทิ้ง ปรับปรุง และเปลี่ยนแปลงข้อทดสอบเพื่อพัฒนาให้มีคุณสมบัติตามเกณฑ์มาตรฐาน และเพื่อการนำเข้าไว้ในธนาคารข้อทดสอบนั้นเป็นวิธีที่นักทดสอบ นักวัดและประเมินผลทางภาษานิยมใช้กัน เช่น Lim Tock Keng (1994) ได้ใช้วิธีวิเคราะห์แบบประเพณีนิยมสำหรับหาค่าสถิติของตัวเลือก และใช้ Rasch Model ในการหาความสอดคล้องของข้อสอบวัดความเข้าใจในการอ่านในระดับประถมศึกษากับโคลงลักษณะข้อสอบและการพิจารณานำข้อสอบบรรจุไว้ในแบบทดสอบฉบับต่อไป

การใช้ Rasch Model ในการวิเคราะห์และนำเข้าข้อทดสอบไว้ในธนาคาร มีประโยชน์ในการใช้ข้อสอบที่ผู้มีความสามารถสูงทำได้เป็นข้อสอบจัดระดับหรือยกเว้นการเรียนและจัดข้อสอบได้เหมาะสมกับระดับความสามารถของผู้เข้าสอบ โดยใช้ข้อสอบที่ผู้มีความสามารถปานกลางและอ่อนมีแนวโน้มทำได้ในการใช้ผลวินิจฉัยปัญหาเพื่อการเรียนการสอนซ่อมเสริม

Bachman (1990: 206) ได้กล่าวถึงการ ใช้ Rasch Model วิเคราะห์ข้อทดสอบอย่างแพร่หลาย และมีประสิทธิภาพยิ่งดังที่กล่าวถึงในงานของ Davidson & Henning (1975) Griffin (1985) Henning, et al (1985) Madsen & Larson (1986) Pollitt & Hutchinson (1987) Larson (1987) นอกจากนี้ McNamara (1990) ได้กล่าวถึงการ ใช้ Rasch Model วิเคราะห์ข้อสอบภาษาเพื่อการสื่อสารได้อย่างมีประสิทธิภาพ

เกณฑ์ในการตัดสินระดับความสามารถเป็นเกณฑ์ที่ใช้วัดจุดประสงค์แบบอุดมซึ่งกำหนดขอบเขตของเนื้อหา และทักษะภาษาและกิจกรรมการใช้ภาษาในการวัดผลไว้เกณฑ์ทั้ง 7 ระดับนี้ให้คำอธิบายความสามารถในการ

รับสารที่เป็นทักษะเดียวได้แก่การฟัง และการอ่านไว้ พร้อมกับทักษะส่งสารด้านการพูดและเขียน รวมทั้งเกณฑ์ของการใช้ทักษะสัมพันธ์ ด้านการฟัง-เขียน การฟัง-อ่าน-เขียน และการอ่าน-เขียน การใช้เกณฑ์เป็นการให้ความหมายแก่คะแนนซึ่งโดยตัวเองไม่มีความหมาย ถ้าปราศจากคำอธิบายและการแปลผลคะแนน การใช้เกณฑ์นี้เป็นแบบมาตราส่วนประเมินค่าที่นิยมใช้ช่วงคะแนน (band scores) ในการให้ความหมายคะแนน มาตราประเมินค่าสำหรับการทดสอบทางภาษาที่นิยมใช้กันทั่วไปมี 4 แบบด้วยกัน คือ

1. แบบของ British Council ซึ่งมี 9 ระดับ ใช้สำหรับทักษะการฟัง-พูด (L-S) การฟัง (L) การอ่าน-เขียน (R-W) การอ่าน (R) และคำเฉลียว (AV) ความหมายของแต่ละระดับมีดังนี้ ระดับ 1/0 Non-user, ระดับ 2 Intermittent, ระดับ 3 Extremely limited, ระดับ 4 Marginal, ระดับ 5 Modest, ระดับ 6 Competent, ระดับ 7 Good, ระดับ 8 Very good, ระดับ 9 Expert ระดับฟังประสงค์คือระดับ 7 สำหรับผู้ใช้ภาษาได้ในชีวิตประจำวันและอาชีพ สามารถเข้าใจสถานการณ์ส่วนใหญ่ในสิ่งแวดล้อมที่ใช้ภาษาอังกฤษประจำ มีความไม่เข้าใจหรือข้อจำกัดทางภาษาเป็นครั้งคราวแต่ไม่เป็นอุปสรรคในการสื่อสาร

2. แบบของ Foreign Service Institute (FSI) ใช้สำหรับการวัดความสามารถทางการพูด มีเกณฑ์ 5 ระดับ ระดับ 1 เป็นความสามารถเบื้องต้นของผู้ที่ยังมีประสบการณ์ด้านการใช้ภาษาจำกัด มีความสามารถสื่อสารในเรื่องทั่ว ๆ ไป เช่น เรื่องถามและบอกทิศทาง บอกเวลาและชื่อของได้ ระดับ 2 สนทนาเกี่ยวกับอาชีพ ครอบครัวและเรื่องของตนเองได้ สามารถบอกวิธีทำสิ่งต่าง ๆ ได้ ใช้ไวยากรณ์ผิด ๆ อยู่บ้าง ระดับ 3 ใช้ภาษาที่เป็นพิธีการได้ ใช้ศัพท์และไวยากรณ์ได้อย่างถูกต้อง สำเนียงไม่เหมือนเจ้าของภาษาแต่ไม่ก่อความรำคาญให้เจ้าของภาษา ระดับ 4 สามารถสื่อสารได้ดี มีที่ผิดน้อย ระดับ 5 พูดได้ดีเหมือนเจ้าของภาษาที่มีการศึกษา ใช้ภาษาได้ดีจนเป็นที่ยอมรับของเจ้าของภาษาที่มีความรู้

ในทุกลักษณะ เช่น รัศมีทั้งหมดทั้งสำนวน ภาษาถิ่น และวัฒนธรรม เพื่อการกำหนดระดับคะแนนที่ละเอียดยิ่งขึ้น FSI เพิ่มเกณฑ์ + เข้าไปในคำอธิบายการตีความคะแนนที่ได้ว่าอยู่ในระดับใดจึงมีทั้งหมด 9 ระดับ เช่น คะแนน 16-25 คือระดับ 0+ 63-72 คือระดับ 3 และ 73-82 คือระดับ 3+ เป็นต้น

3. แบบของ Clark ประเมินความสามารถในการใช้ภาษาใน 4 ด้าน คือ ด้านการออกเสียง การใช้ศัพท์ การใช้โครงสร้าง และความคล่องในการใช้ภาษา

4. แบบของ The Council of Europe ซึ่ง Wilkins (1977) พัฒนาขึ้น มี 7 ระดับ จากระดับเริ่มเรียนหรือระดับความสามารถแบบ “งู ๆ ปลา ๆ” ของระดับ 1 จนถึงระดับ 7 ซึ่งเป็นระดับของผู้ที่มีความสามารถเทียบเท่าเจ้าของภาษา เกณฑ์นี้ใช้ประเมินความสามารถในการสื่อสาร ฟัง พูด อ่านและเขียน

เกณฑ์การตีความเกี่ยวกับระดับความสามารถสื่อสารทางภาษาอังกฤษของผู้เรียนที่ใช้ในงานวิจัยนี้เป็นเกณฑ์ 7 ระดับที่ดัดแปลงให้เหมาะสมกับแบบทดสอบที่ใช้และเหมาะสมกับกลุ่มตัวอย่างประชากรในการวิจัยที่เป็นนักศึกษาไทยเกณฑ์นี้ ระบบปฏิสัมพันธ์ระหว่างผู้ส่งสารและผู้รับสารรวมทั้งสื่อที่ใช้ว่าเป็นสื่อสิ่งพิมพ์แบบใด ตลอดจนช่องทางของการสื่อสารว่าเป็นแบบตัวต่อตัวหรือแบบผ่านเทป โทรศัพท์ ฯลฯ ซึ่งเหมาะสมในการประเมินการใช้ภาษาเพื่อการสื่อสารของนักศึกษาไทย (ดูเกณฑ์ที่กำหนดในภาคผนวก)

ข้อเสนอแนะ

ข้อเสนอแนะทางการวิจัย

ควรมีการวิจัยทางการทดสอบภาษาอังกฤษที่ใช้วัดอุปประสงค์แบบอุดมในการพัฒนาข้อทดสอบขึ้นใช้ใน ระดับการศึกษาอื่น ๆ และควรนำแบบทดสอบอิงวัดอุปประสงค์แบบอุดมที่ได้รับการพัฒนาให้มีคุณสมบัติตามเกณฑ์มาตรฐานไปใช้ในการวิจัยทางการทดสอบภาษา ตามนัยที่สอง คือศึกษาเกี่ยวกับระดับทักษะและความรู้ความสามารถทางภาษาอังกฤษของนักศึกษาในระดับ

ต่าง ๆ นอกจากนี้ควรได้นำแบบทดสอบมาตรฐานภาษาอังกฤษอิงวัตถุประสงค์แบบอุดมไปวิจัยตามนัยที่สาม คือ การศึกษาเกี่ยวกับปัจจัยทางภาษาในระดับการศึกษาต่าง ๆ ว่ามีรูปแบบและองค์ประกอบเช่นใด

ข้อเสนอแนะทางการเรียนการสอน

ควรมีการพัฒนาธนาคารข้อทดสอบภาษาอังกฤษขึ้นใช้ในการศึกษาทุกระดับเพราะจะช่วยให้การวัดและประเมินผลมีมาตรฐาน และเชื่อถือได้อย่างปราศจากข้อ

สงสัยเกี่ยวกับคุณภาพของเครื่องมือที่ใช้ตัดสินผล การเรียนซึ่งหมายถึงอนาคตของผู้เรียน นอกจากนี้ควรนำผลการทดสอบโดยใช้วัตถุประสงค์แบบอุดมไปใช้ในการจัดการเรียนการสอนซ่อมเสริม เพื่อปรับปรุงคุณภาพของการเรียนภาษาอังกฤษของผู้เรียนเป็นรายบุคคล และเพื่อการช่วยครูผู้สอนในการจัดทำหรือจัดเตรียมแบบฝึกทางภาษาซ่อมเสริมให้แก่ผู้เรียนที่อ่อนตลอดจนการจัดทำบทเรียนหรือแบบฝึกหัดเสริมทักษะให้แก่ผู้เรียนที่เรียนดี

เอกสารอ้างอิง

- กระทรวงศึกษาธิการ. 2533. *หลักสูตร มัธยมศึกษาตอนปลาย พ.ศ. 2524 ฉบับปรับปรุง พ.ศ. 2533*. กรุงเทพฯ :
คุรุสภา.
- วิริยา สิทธิสาร. 2530. “การสำรวจความต้องการในการเรียนภาษาอังกฤษของ นักเรียน ผู้ปกครองและครูชั้นมัธยม
ศึกษาตอนปลายในโรงเรียนรัฐบาลเขตกรุงเทพมหานครโดยใช้วัตถุประสงค์แบบอุดม.”วิทยานิพนธ์มหาบัณฑิต
สาขาศึกษาศาสตร์การสอน, มหาวิทยาลัยเกษตรศาสตร์.
- สมหวัง พิธิยานุวัฒน์. 2529. “การทดสอบอิงปริเขต.” 67-84. บทคัดสรรทางวิชาการทดสอบ. ศูนย์พัฒนาศึกษา
แห่งชาติของประเทศ. ทบวงมหาวิทยาลัย กรุงเทพฯ: โรงพิมพ์มหาวิทยาลัยสุโขทัยธรรมมาธิราช
- สุนันทา พากเพียร. 2530. “การสำรวจความต้องการในการเรียนภาษาอังกฤษของนักเรียน ผู้ปกครองและครูชั้น
มัธยมศึกษาตอนต้นในโรงเรียนรัฐบาลเขตกรุงเทพมหานคร โดยใช้วัตถุประสงค์แบบอุดม” วิทยานิพนธ์
มหาบัณฑิต สาขาศึกษาศาสตร์การสอน, มหาวิทยาลัยเกษตรศาสตร์.
- อัจฉรา วงศ์โสธร. 2529. *เทคนิควิธีการสร้างข้อสอบภาษาอังกฤษ สำหรับวัดและประเมินผลการใช้ภาษาเพื่อการ
สื่อสาร*. กรุงเทพฯ: สำนักพิมพ์อักษรเจริญทัศน์.
- อัจฉรา วงศ์โสธร และคณะ. 2535. *การสอบสมิทธิภาพทั่วไปทางภาษาอังกฤษ ครั้งที่ 2 ปี พ.ศ. 2535 ในโครงการ
“การสร้างและพัฒนาคุณภาพแบบทดสอบสมิทธิภาพทั่วไปทางภาษาอังกฤษและระบบการทดสอบที่เกี่ยวข้อง.”*
(รายงานการวิจัย). กรุงเทพฯ: สถาบันภาษาและฝ่ายวิชาการจุฬาลงกรณ์มหาวิทยาลัย
- อัจฉรา วงศ์โสธร. 2536. *ทักษะภาษาอังกฤษของนักศึกษาไทย*. (รายงานการวิจัย). สถาบันภาษา, จุฬาลงกรณ์
มหาวิทยาลัย
- Bachman, L. F. 1990. *Fundamental Considerations in Language Testing*. Oxford: Oxford University Press.
- Carroll. J.B. 1987, “Psychometric Theory and Language Testing.” In Grotjahn, R., Klein-Braley, C. & Stevenson, D.K. (eds.) *Taking Their Measure. The Validity and Validation of Language Tests. Quantitative Linguistics*, Vol. 34. Bochum: Studienver. Brockmeyer.
- Clark, J. L. D. 1980. “Toward a common measure of speaking proficiency” in J. R. Frith (ed.) *Measuring Spoken Language Proficiency*. Washington, DC: Georgetown University Press.
- Cronbach, L.J. 1951 “Coefficient alpha and the internal structure of tests.” *Psychometrika* 16:297-34.
- Cronbach, L. J., N. Rajaratnam, and G.C. Gleser. 1963. “Theory of generalizability: A liberalization of reliability theory.” *British Journal of Statistical Psychology* 16:137-163.
- Cronbach, L. J. 1971. “Validity” in R. L.Thorndike (ed.) *Educational Measurement*.(2nd.ed.) Washington, DC: American Council on Education.
- Davidson, F., and G. Henning. 1985. “A self-rating scale of English difficulty: Rasch scalar analysis of items and rating categories.” *Language Testing* 2, 2: 164-179.
- Fan, C.T. 1952. *Item Analysis Table*. Princeton, N.J: Educational Testing Service.
- Lim T.K. 1994 “Analyzing Primary Level Reading Tests Using the Rasch Model and Classical Model” *Pasaa Journal*. Vol.24. December 1994.

- Madsen, H. S., and J. W. Larson, 1986. "Computerized Rasch analysis of item bias in ESL tests" in C.W. Stansfield (ed.) *Technology and Language Testing*. Washington, DC: TESOL.
- 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: Danish Institute for Educational Research.
- The British Council. 1978. *An English Language Testing Service Specifications*. London: The British Council.
- The British Council and University of Cambridge Local Examinations Syndicate. 1987. *English Language Testing Service: An Introduction to ELTS*. London: The British Council.
- University of Cambridge Local Examinations Syndicate. International Examinations. 1987. *English as a Foreign Language: General Handbook*. Cambridge: University of Cambridge Local Examinations Syndicate.
- Wilkins, D. A. 1977. *Language Performance Scale*. Strasbourg: The Council of Europe.
- Wright, B. D., and M. H. Stone. 1979. *Best Test Design*. Chicago: MESA Press.

ภาคผนวก

การกำหนดระดับความสามารถใช้ภาษาอังกฤษ

เพื่อการใช้ผลคะแนนสอบในการตีความเกี่ยวกับความสามารถในการใช้ภาษาอังกฤษตามวัตถุประสงค์แบบอุดมของแต่ละระดับการศึกษาได้ใช้เกณฑ์จัดระดับความสามารถ 7 ระดับที่พัฒนาจากเกณฑ์ของ The Council of Europe (1977) (อัจฉรา วงศ์โสธร, 2529) และผลที่คาดหวังตามหลักสูตรของแต่ละระดับการศึกษาในการอธิบายความสามารถใช้ภาษาอังกฤษเพื่อการสื่อสารจากรากฐานคะแนนที่ได้จากแบบทดสอบอิงวัตถุประสงค์แบบอุดมในแต่ละระดับการศึกษา

ระดับมัธยมศึกษาตอนต้นและตอนปลาย

1. ทักษะการฟัง (คะแนน 0-10)

0-1.5: ระดับ 1 ทราบแต่เพียงว่าบทสนทนาที่ได้ฟังเป็นภาษาอังกฤษโดยไม่เข้าใจว่าสิ่งที่ได้ยินมีความหมายอย่างไร ทราบแต่เพียงคำหรือข้อความสั้น ๆ เดี่ยว ๆ ที่ได้ฟังและที่ผู้พูดต้องกล่าวซ้ำ ๆ

1.6-2.5: ระดับ 2 เข้าใจข้อความบางข้อความที่เคยชินและง่าย ๆ เช่น การทักทาย การกล่าวคำอำลา

2.6-4.5: ระดับ 3 เข้าใจบทสนทนาได้เพียงเลา ๆ และเข้าใจผิดบ่อยครั้งถ้าไม่ได้ยินข้อความซ้ำ ๆ

4.6-6.5: ระดับ 4 สามารถเข้าใจข้อความที่ใช้ภาษาธรรมดาไม่มีศัพท์สำนวนยาก ๆ ได้แต่ยังไม่เข้าใจข้อความทั้งหมดหรือรายละเอียดในบางตอน

6.6-7.5: ระดับ 5 สามารถเข้าใจเนื้อความที่สำคัญ ๆ ได้ แต่ยังไม่มีความลำบากในการฟังข้อความที่ซับซ้อน ต้องฟังซ้ำ ๆ

7.6-8.5: ระดับ 6 สามารถเข้าใจเนื้อความได้เกือบทั้งหมดและที่พูดด้วยความเร็วปกติ

8.6-10: ระดับ 7 สามารถเข้าใจข้อความที่ได้ยินทั้งหมดและสามารถตีความเกี่ยวกับเรื่องที่ได้ยิน

ระดับมัธยมศึกษาและอุดมศึกษา

2. **ทักษะการพูด** (คะแนน 0-30 ในระดับมัธยมศึกษา, 0-60 ในระดับอุดมศึกษา)

0-2.5 (ระดับมัธยมศึกษา) หรือ 0-5.5 (ระดับอุดมศึกษา): ระดับ 1 พูดไม่ได้เลย-ได้น้อยมากจนไม่สามารถสื่อสารได้

2.6-5.5 (ระดับมัธยมศึกษา) หรือ 5.6-11.5 (ระดับอุดมศึกษา): ระดับ 2 พูดติดตะกุกตะกัก มีที่ผิดพลาดไวยากรณ์และการออกเสียงบ่อยมาก

5.6-9.5 (ระดับมัธยมศึกษา) หรือ 11.6-18.5 (ระดับอุดมศึกษา): ระดับ 3 สามารถพูดประโยคง่าย ๆ ได้ ยังมีที่ผิดพลาดไวยากรณ์และการใช้เสียงแปร่งอยู่ ผู้ที่สนทนาด้วยยังมีความลำบากที่จะทำความเข้าใจ

9.6-14.5 (ระดับมัธยมศึกษา) หรือ 18.6-29.5 (ระดับอุดมศึกษา): ระดับ 4 สามารถพูดให้ผู้สนทนาด้วยฟังเข้าใจได้ ยังไม่สามารถพูดข้อความที่ต่อเนื่องยาว ๆ ได้

14.6-19.5 (ระดับมัธยมศึกษา) หรือ 29.6-39.5 (ระดับอุดมศึกษา): ระดับ 5 สามารถพูดสื่อความหมายได้ตามวัตถุประสงค์ต่าง ๆ ที่ระบุและสามารถแสดงความคิดเห็นได้แต่อาจจะยังคลุมเครืออยู่ ผู้ที่สนทนาด้วยต้องซักถามบ่อยครั้งเพื่อให้เข้าใจชัดเจนขึ้น

19.6-24.5 (ระดับมัธยมศึกษา) หรือ (39.6-49.5) (ระดับอุดมศึกษา): ระดับ 6 สามารถพูดสื่อความหมายได้ทุกวัตถุประสงค์อย่างคล่องแคล่วชัดเจนสามารถแสดงความคิดเห็นได้ มีการใช้เสียง ศัพท์และไวยากรณ์ผิดน้อยมาก

24.6-30.0 (ระดับมัธยมศึกษา) หรือ 49.6-60.0 (ระดับอุดมศึกษา): ระดับ 7 สามารถพูดได้ชัดเจนและอย่างคล่องแคล่ว ตามวัตถุประสงค์ได้ในทุกสถานการณ์

3. ทักษะการอ่าน (คะแนน 0-10 ในระดับมัธยมศึกษา, คะแนน 0-15 ในระดับอุดมศึกษา)

0-1.5 (ระดับมัธยมศึกษา) หรือ 0-2.5 (ระดับอุดมศึกษา) : ระดับ 1 สามารถอ่านข้อความง่าย ๆ หรือประกาศสั้น ๆ ได้อย่างเลา ๆ

1.6-2.5 (ระดับมัธยมศึกษา) หรือ 2.6-4.5 (ระดับอุดมศึกษา): ระดับ 2 สามารถเข้าใจหัวเรื่องที่อ่านได้อย่างคร่าว ๆ

2.6-4.5 (ระดับมัธยมศึกษา) หรือ 4.6-6.5 (ระดับอุดมศึกษา): ระดับ 3 สามารถเข้าใจหัวเรื่องและเนื้อหาได้อย่างคร่าว ๆ โดยไม่ต้องใช้พจนานุกรม อีกทั้งจะเข้าใจรายละเอียดได้เมื่อใช้พจนานุกรมช่วย

4.6-6.5 (ระดับมัธยมศึกษา) หรือ 6.6-8.5 (ระดับอุดมศึกษา): ระดับ 4 สามารถเข้าใจเนื้อเรื่องที่อ่านได้อย่างพอเพียง แต่ต้องอ่านข้อความที่ซับซ้อนซ้ำ ๆ ยังต้องใช้พจนานุกรมช่วยให้เข้าใจมากขึ้น ยังอ่านอย่างช้า ๆ

6.6-7.5 (ระดับมัธยมศึกษา) หรือ 8.6-10.5 (ระดับอุดมศึกษา): ระดับ 5 สามารถอ่านข้อความทุกอย่างได้ตามวัตถุประสงค์ด้วยความเข้าใจดีและไม่ลำบาก ยังต้องใช้พจนานุกรมและอ่านอย่างช้า ๆ ในบางตอนในเรื่องที่ตนไม่คุ้นเคย

7.6-8.5 (ระดับมัธยมศึกษา) หรือ 10.6-12.5 (ระดับอุดมศึกษา): ระดับ 6 สามารถอ่านเพื่อความได้ตามวัตถุประสงค์ต่าง ๆ ด้วยความเข้าใจดีมาก ถึงแม้ว่าจะต้องอ่านซ้ำ ๆ ในบางตอน

8.6-10.0 (ระดับมัธยมศึกษา) หรือ 12.6-15.0 (ระดับอุดมศึกษา): ระดับ 7 เข้าใจทุกเรื่องที่อ่านอย่างไม่บกพร่อง อ่านอย่างคล่องแคล่วและไม่ต้องใช้พจนานุกรมช่วย

4. ทักษะการเขียน

เกณฑ์นี้อธิบายคะแนนระดับมัธยมศึกษาทั้งตอนต้นและตอนปลายและระดับอุดมศึกษาต่างกันตามวัตถุประสงค์แบบอุดมที่ระบุไว้สำหรับแต่ละระดับ(คะแนน 0-5)

0-1.0: ระดับ 1 สามารถเขียนคัดลอกข้อความหรือประโยคง่าย ๆ จากความจำได้

1.1-2.0: ระดับ 2 สามารถเขียนเปลี่ยนแปลงประโยคได้แต่ยังมีที่ผิดทั้งทางด้านศัพท์และไวยากรณ์มากทำให้ผู้อ่านไม่สามารถเข้าใจได้นัก

2.1-2.5: ระดับ 3 สามารถเขียนประโยคเดี่ยว ๆ ได้บ้างแต่ถ้าเป็นข้อความที่ซับซ้อนและต่อเนื่องจะสามารถเขียนได้น้อยมาก จำเป็นต้องเปิดพจนานุกรมบ่อยครั้ง

2.6-3.0: ระดับ 4 สามารถเขียนเนื้อความต่อเนื่องได้บ้าง แต่ยังใช้รูปประโยคและสำนวนผิด ๆ อยู่ ยังต้องใช้พจนานุกรมช่วยอยู่มาก

3.1-3.5: ระดับ 5 สามารถเขียนเนื้อความต่าง ๆ ตามวัตถุประสงค์ได้แต่ยังมีอิทธิพลของภาษาไทยในการใช้ภาษาอยู่ ยังต้องใช้พจนานุกรมช่วยอยู่บ้าง

3.6-4.0: ระดับ 6 สามารถเขียนเนื้อความได้ทุกวัตถุประสงค์โดยไม่ผิดไวยากรณ์ตัวสะกดและเครื่องหมายวรรคตอน มีการใช้รูปประโยคและศัพท์สำนวนต่าง ๆ ได้ ไม่จำเป็นต้องใช้พจนานุกรมช่วยในการเขียน ยังเขียนช้าและอาจผิดที่ลีลาภาษา

4.1-5.0: ระดับ 7 สามารถเขียนได้ทุกวัตถุประสงค์อย่างคล่องแคล่ว และแทบจะไม่มีที่ผิดเลย ผู้อ่านสามารถเข้าใจข้อเขียนได้ทุกตอนอย่างไม่ลำบาก

5. ทักษะสัมพันธ์ด้านการฟัง-เขียน (คะแนน 0-5)

0-1.0: ระดับ 1 ไม่เข้าใจข้อความที่ได้ฟังและไม่สามารถเขียนตอบคำถามที่เกี่ยวข้องได้อย่างสัมพันธ์กับข้อความที่ฟัง อาจเขียนเป็นคำหรือข้อความภาษาอังกฤษสั้น ๆ ที่จำไว้แต่ไม่ถูกต้องด้านภาษาและความหมาย และไม่สามารถเขียนตามคำบอกได้อย่างต่อเนื่อง อาจเขียนได้คำเดียวและไม่ถูกที่ตามคำบอก

1.1-2.0: ระดับ 2 เขียนตามคำบอกได้เป็นคำ ๆ การตอบคำถามยังมีที่ผิดมากทั้งด้านศัพท์และไวยากรณ์ทำให้ไม่สามารถสื่อความหมายได้นัก

2.1-2.5: ระดับ 3 เขียนตามคำบอกได้เป็นช่วง ๆ มีความยาวมากขึ้น พอสื่อความหมายได้ถึงแม้จะมีที่ผิดพลาดการใช้ภาษาและเครื่องหมายวรรคตอน การตอบคำถามสื่อความได้พอควรยังมีการใช้ศัพท์และไวยากรณ์ที่ไม่ถูกต้องอยู่

2.6-3.0: ระดับ 4 เขียนตามคำบอกและสื่อความหมายได้ ยังมีที่บกพร่องทางการใช้ภาษาอยู่บ้าง การเขียนตอบคำถามข้อความที่ได้ฟังได้ความหมายพอควร

3.1-3.5 ระดับ 5 สามารถเขียนตามคำบอกได้ดี มีความเข้าใจข้อความที่ฟังและมีความสามารถทางการเขียนดีพอที่จะถ่ายทอดความเข้าใจออกมาเป็นภาษาเขียนที่สื่อความหมายได้

3.6-4.0: ระดับ 6 ไม่มีปัญหาที่สำคัญในการเขียนตามคำบอก มีที่ผิดเล็กน้อยเท่านั้น ความสามารถด้านการฟังและการเขียนดี สามารถรับสารและส่งสารได้อย่างมีความหมายถูกต้อง อาจจะมีที่ผิดพลาดการใช้ศัพท์และไวยากรณ์เล็กน้อย

4.1-5.0: ระดับ 7 สามารถเขียนตามคำบอกได้ดีมาก ความสามารถด้านการฟังและการเขียนเพื่อสื่อความเข้าใจในการฟังถูกต้องได้ใจความสมบูรณ์

6. ทักษะสัมพันธ์ด้านการฟัง-อ่าน-เขียน

0-1.0: ระดับ 1 ความสามารถด้านการรับสารทางการฟังและอ่านอ่อนมากไม่เข้าใจสารพอที่จะเขียนแสดงความเข้าใจได้ อาจเขียนได้เป็นคำ ๆ แต่ไม่สัมพันธ์กับสารที่ฟังและอ่าน

1.1-2.0: ระดับ 2 สามารถฟังและอ่านได้บ้าง แต่ไม่สามารถเขียนสื่อความหมายได้เป็นข้อความที่ต่อเนื่อง อาจเขียนได้บ้างเป็นคำ ๆ หรือข้อความสั้น ๆ แต่ไม่ได้ความหมายนัก

2.1-2.5: ระดับ 3 สามารถจับหัวข้อของสารที่ฟังและอ่านได้แต่ไม่เข้าใจใจความที่สำคัญและรายละเอียดบางตอน สามารถเขียนตอบคำถามเกี่ยวกับสารได้บ้าง ยังมีการใช้ศัพท์และไวยากรณ์ที่ผิด ๆ บ้าง

2.6-3.0: ระดับ 4 สามารถเข้าใจเนื้อความที่ฟังและอ่านได้เป็นส่วนใหญ่ และเขียนสื่อความหมายให้ผู้อ่านเข้าใจได้ มีที่ผิดพลาดการใช้ศัพท์และไวยากรณ์บ้าง

3.1-3.5: ระดับ 5 สามารถเข้าใจเนื้อความที่ฟังและอ่านได้ทั้งที่เป็นใจความสำคัญและรายละเอียดและถ่ายทอดความเข้าใจเป็นภาษาเขียนได้ มีที่ผิดพลาดการใช้ศัพท์และไวยากรณ์เล็กน้อย แต่การใช้สำนวนและลีลาภาษายังเป็นไทย ๆ อยู่

3.6-4.0: ระดับ 6 สามารถเข้าใจเนื้อความที่ฟังและอ่านได้ทั้งหมด และสามารถตีความได้ สามารถเขียนแสดงความเข้าใจเกี่ยวกับสิ่งที่อ่านได้เกือบทั้งหมด แต่สำนวนและลีลาภาษายังไม่สละสลวย

4.1-5.0: ระดับ 7 สามารถเข้าใจเนื้อความที่ฟังและอ่านได้ทั้งหมด ตีความได้และเขียนถ่ายทอดสิ่งที่เข้าใจและตีความได้โดยใช้ภาษาถูกต้อง สามารถใช้สำนวนและลีลาภาษาได้อย่างสละสลวย

7. ทักษะสัมพันธ์ด้านการอ่าน-เขียน

การใช้เกณฑ์และการตีความคะแนนเป็นระดับความสามารถ พิจารณาได้จากเกณฑ์ของแบบทดสอบทักษะสัมพันธ์ด้านการฟัง-อ่าน-เขียน ในข้อ 6