

# คุณภาพและประสิทธิภาพของแบบทดสอบปรับเปลี่ยน ที่เน้นเนื้อหาโดยใช้เครื่องคอมพิวเตอร์ เพื่อการทดสอบแบบอิงบริเขต

สุวัฒน์ สุกมลสันต์

## ความสำคัญและที่มาของปัญหาการวิจัย

ในการทดสอบตามแบบประเพณีนิยม (conventional testing) นั้น แบบทดสอบจะประกอบด้วยข้อทดสอบจำนวนมาก เพื่อให้การทดสอบมีความตรง (validity) และความเที่ยง (reliability) สูง ข้อทดสอบเหล่านี้จะถูกเรียงลำดับไว้ก่อนแล้วตามเนื้อหาที่ต้องการทดสอบ และมักจะประกอบด้วยข้อทดสอบที่มีความยากง่าย (difficulty index) ในระดับปานกลางเพื่อให้เหมาะสมกับระดับความสามารถทั่วไปของผู้สอบ ซึ่งการกระทำดังกล่าวนี้ ทำให้เกิดความคลาดเคลื่อนในการวัดสำหรับผู้สอบที่มีความรู้ในระดับเก่งหรืออ่อน เพราะข้อทดสอบที่ง่ายเกินไปจะไม่ท้าทายความสามารถในการสอบของผู้สอบที่เก่ง ทำให้ผู้สอบเกิดความเบื่อหน่ายในการตอบอย่างไม่เต็มความสามารถ ส่วนข้อสอบที่ยากเกินไปสำหรับผู้สอบที่อ่อน ก็จะทำให้ผู้สอบรู้สึกท้อถอย และใช้วิธีเดา สิ่งต่างๆ เหล่านี้ทำให้ความถูกต้องและแม่นยำในการทดสอบลดลง และทำให้การประมาณค่าความสามารถของผู้สอบคลาดเคลื่อน (Weiss, 1980: 1) ดังนั้นการทดสอบแบบประเพณีนิยมจึงใช้ได้ผลดีเฉพาะกรณีที่ผู้สอบมีความสามารถในระดับปานกลางเท่านั้น (Lord อ้างโดย Vale and Weiss, 1979: 62) และความคลาดเคลื่อนจะมีมากขึ้น เมื่อความสามารถของผู้สอบเบี่ยงเบนจากความสามารถปานกลางมากขึ้น (Weiss, 1974: 105)

จากข้อจำกัดของการทดสอบแบบประเพณีนิยม และผลทางจิตวิทยาที่เกิดขึ้นกับผู้สอบดังกล่าวแล้ว ทำให้นักทดสอบค้นหาวิธีที่จะทำให้ความคลาดเคลื่อนในการทดสอบลดน้อยลง ด้วยวิธีการจัดข้อทดสอบให้มีความยากง่ายเหมาะกับระดับของผู้สอบแต่ละคน เพื่อลดผลกระทบทางจิตวิทยา วิธีการทดสอบดังกล่าวเรียกว่า การทดสอบเฉพาะบุคคล (tailored testing) หรือการทดสอบปรับเปลี่ยน (Adaptive testing) และหากว่าการจัดระดับความยากง่ายของข้อทดสอบให้อยู่ในกรอบเนื้อหาเดียวกันก็เรียกว่าการทดสอบปรับเปลี่ยนโยงเนื้อหา (Content-based Adaptive Testing) และในระยะต่อมาการทดสอบดังกล่าวนี้อาศัยเครื่องคอมพิวเตอร์ช่วยในการเลือกข้อทดสอบมาเพื่อให้ผู้สอบได้ตอบตามเกณฑ์ที่ผู้สร้างกำหนดขึ้น การทดสอบดังกล่าวก็เรียกว่าการทดสอบปรับเปลี่ยนเนื้อหาโดยใช้คอมพิวเตอร์ (Computerized Content-based Adaptive Testing: CCAT) หรือเรียกสั้นๆ ว่า การทดสอบปรับเปลี่ยนโดยใช้คอมพิวเตอร์ (Computerized Adaptive Testing: CAT) โดยจะเป็นที่เข้าใจกันว่า การปรับเปลี่ยนระดับความยากง่ายของข้อทดสอบนั้นคำนึงถึงเนื้อหาในการสอบด้วยว่าอยู่ในกรอบหรือปริเขต (domain) เดียวกันกับสิ่งที่ต้องการทดสอบ

หลักการทั่วไปของการทดสอบปรับเปลี่ยนมีดังนี้ (Weiss, 1980 : 102)

1. เป็นการทดสอบที่พยายามคัดเลือกข้อทดสอบที่มีความยากง่ายเหมาะสมกับระดับความสามารถของผู้สอบมาให้ผู้สอบตอบ
2. ข้อทดสอบที่คัดเลือกมาอยู่ในปริเขตของเนื้อหาเดียวกัน
3. ถ้าผู้สอบตอบข้อสอบได้ถูกต้อง ข้อสอบข้อต่อไปจะเป็นข้อสอบที่มีความยากเพิ่มขึ้นแต่หากว่าตอบผิดข้อทดสอบข้อต่อไปจะเป็นข้อที่ง่ายกว่าทำเช่นนั้นจนกว่าการสอบจะสิ้นสุด
4. การทดสอบแต่ละครั้งผู้สอบไม่จำเป็นต้องตอบข้อสอบเหมือนกันทุกข้อทั้งฉบับ และไม่จำเป็นต้องตอบจำนวนในข้อเท่ากัน

ดังนั้น จะเห็นได้ว่าการปรับเปลี่ยนข้อทดสอบจะเป็นสิ่งที่ยุ่งยาก จึงเหมาะที่จะใช้เครื่องคอมพิวเตอร์ส่วนบุคคล (personal computer) มาช่วยในการคัดเลือกข้อทดสอบที่เหมาะสมกับระดับความสามารถของผู้สอบ (Weiss and Betz, 1973 : 110)

เนื่องจากการทดสอบปรับเปลี่ยนดังกล่าวเป็นวิธีการทดสอบที่ค่อนข้างใหม่มากในวงการทดสอบในประเทศไทย โดยเฉพาะการทดสอบปรับเปลี่ยนเนื้อหาโดยอาศัยเครื่องคอมพิวเตอร์ (CCAT) ผู้วิจัยจึงมีความสนใจที่จะศึกษาเกี่ยวกับคุณภาพและประสิทธิภาพของการทดสอบดังกล่าวเพื่อการทดสอบภาษาอังกฤษแบบอิงปริเขต (domain-referenced testing) โดยเฉพาะเมื่อเป็นการทดสอบโครงสร้างทางภาษาอังกฤษในระดับจุลภาค (micro-level)

### วัตถุประสงค์ของการวิจัย

1. เพื่อจัดสร้างโปรแกรมคอมพิวเตอร์ที่สามารถทำงานเพื่อใช้ทดสอบปรับเปลี่ยนเนื้อหา (Content-based Adaptive Testing) ได้
2. เพื่อศึกษาคุณภาพและประสิทธิภาพของการทดสอบปรับเปลี่ยนเนื้อหาเพื่อทดสอบอิงปริเขต โดยการเปรียบเทียบกับคุณภาพและประสิทธิภาพของการทดสอบตามประเพณีนิยม

### สมมุติฐานการวิจัย

1. โปรแกรมคอมพิวเตอร์ที่สร้างขึ้นน่าจะสามารถทำงานเพื่อใช้ทดสอบปรับเปลี่ยนเนื้อหาตามกรอบแนวคิดที่กำหนดขึ้นได้

จากการศึกษาของ Weiss (Weiss, 1984 : 362) พบว่าการทดสอบปรับเปลี่ยนทำให้แบบทดสอบมีความเที่ยงและความตรงไม่แตกต่างหรืออาจมีสูงกว่าแบบทดสอบแบบประเพณีนิยม และสามารถลดความยาวของข้อสอบได้ถึงร้อยละ 50 เพื่อวัดตัวแปรเดียวกัน และ Green และคณะ (Green and Others, 1984 : 70) ยังพบว่าการทดสอบแบบประเพณีนิยมเหมาะสำหรับผู้สอบที่มีความสามารถในระดับปานกลาง แต่ทำให้เกิดมีความคลาดเคลื่อนในการวัดสูงสำหรับผู้สอบที่มีความสามารถสูงหรือต่ำมาก ดังนั้นผู้วิจัยจึงมีสมมุติฐานในการวิจัยเพิ่มดังนี้

2. ในการทดสอบความสามารถทางโครงสร้างทางภาษาอังกฤษของนิสิตชั้นปีที่ 1 ด้วยการทดสอบแบบปรับเปลี่ยนเนื้อหา แบบทดสอบน่าจะมีความตรงและความเที่ยงไม่แตกต่างจากการทดสอบแบบประเพณีนิยม
3. ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบจากการสอบแบบปรับเปลี่ยนเนื้อหา น่าจะต่ำกว่าค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถจากการทดสอบแบบประเพณีนิยม

### ขอบเขตของการวิจัย

#### 1. ประชากร

ประชากรของการวิจัยครั้งนี้ได้แก่นิสิตชั้นปีที่ 1 ของจุฬาลงกรณ์มหาวิทยาลัยที่ลงทะเบียนเรียนรายวิชาภาษาอังกฤษพื้นฐาน I ในภาคปลาย ปีการศึกษา 2535

#### 2. ตัวแปรที่ต้องการศึกษา

ก. ตัวแปรต้น (independent variables) ได้แก่ วิธีการทดสอบ 2 วิธี คือ

1. การทดสอบแบบประเพณีนิยม (Conventional Testing: CT)

2. การทดสอบแบบปรับเปลี่ยนโยงเนื้อหา โดยใช้คอมพิวเตอร์ (Computerized Content-based Adaptive Testing: CCAT)

ข. ตัวแปรตาม (dependent variables) ได้แก่

1. ค่าความจริง
2. ค่าความเที่ยง
3. ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่า
4. ความยาวของแบบทดสอบ
5. เวลาที่ใช้ในการทดสอบ
6. เจตคติต่อการทดสอบ

## มโนทัศน์เกี่ยวกับการทดสอบแบบปรับเปลี่ยน (Adaptive Testing)

1. ความหมายและหลักการในการทดสอบปรับเปลี่ยน

แบบทดสอบปรับเปลี่ยน (Adaptive Test) หมายถึงแบบทดสอบที่สามารถปรับเปลี่ยนความยากง่าย (difficulty index) ของข้อทดสอบแต่ละข้อให้เหมาะสมกับระดับความสามารถของผู้สอบแต่ละบุคคล (Lord, 1971: 80, Hambleton and Swaminathan, 1975 : 296) ดังนั้นจึงทำให้นักทดสอบบางท่านเรียกแบบทดสอบชนิดนี้ว่าแบบทดสอบเฉพาะบุคคล (Tailored Test) หรือแบบทดสอบทรงพีระมิด (Pyramidal Test) ตามลักษณะของการเรียงข้อทดสอบตามระดับความยากง่ายเป็นรูปพีระมิด และหากการปรับเปลี่ยนข้อทดสอบให้เหมาะสมกับระดับความสามารถของผู้สอบแต่ละบุคคลโดยอาศัยเครื่องคอมพิวเตอร์ แบบทดสอบชนิดนี้มักเรียกกันว่าแบบทดสอบปรับเปลี่ยนโดยอาศัยเครื่องคอมพิวเตอร์ (Computerized Adaptive Test) หรือที่รู้จักกันดีทั่วไปว่า CAT และหากการปรับเปลี่ยนข้อทดสอบไม่ได้ใช้เกณฑ์ความยากง่ายของข้อทดสอบเท่านั้น หากแต่อาศัยการโยงเนื้อหาของข้อทดสอบด้วย และใช้เครื่องคอมพิวเตอร์ช่วยในการสุ่มข้อทดสอบ ก็เรียกว่าแบบทดสอบปรับเปลี่ยนโยงเนื้อหาโดยใช้คอมพิวเตอร์ (Computerized Content-based Adaptive Test หรือ CCAT) โดยหลักการแล้วการทดสอบโดยใช้วิธีแบบทดสอบปรับเปลี่ยน ผู้สอบแต่ละคนไม่จำเป็นต้องตอบข้อทดสอบเหมือนกันทุกข้อ จำนวน

ข้อทดสอบไม่จำเป็นต้องเท่ากัน ทั้งนี้ขึ้นอยู่กับผลของการตอบข้อทดสอบแต่ละข้อของผู้สอบแต่ละคน (Weiss and Kingsbury 1984 : 68) ปกติแล้ว ผู้สอบจะเริ่มตอบข้อทดสอบที่มีระดับความยากง่ายปานกลางก่อน ถ้าตอบถูกก็จะตอบข้อทดสอบที่มียากขึ้นกว่าเดิม แต่หากตอบผิดก็จะตอบข้อทดสอบที่ง่ายกว่าเดิม การตอบข้อทดสอบข้อต่อไปจะเป็นเช่นนั้นจนกว่าการทดสอบจะสิ้นสุด (Green and Others, 1984 : 72)

ดังนั้นจะเห็นได้ว่า ความพยายามของนักทดสอบที่ต้องการการทดสอบแบบปรับเปลี่ยนก็เพื่อต้องการให้การทดสอบความสามารถของผู้สอบที่มีความสามารถมาก (เก่ง) กับผู้ที่มีความสามารถน้อย (อ่อน) ได้ถูกต้องมากยิ่งขึ้น รวมทั้งเพื่อลดอิทธิพลจากผลทางจิตวิทยาต่อผู้สอบที่มีความสามารถมากและสามารถน้อยที่ต้องตอบข้อทดสอบที่ไม่เหมาะสมกับระดับความสามารถของตนเอง

## 2. ส่วนประกอบของการทดสอบปรับเปลี่ยน

การทดสอบปรับเปลี่ยน เป็นวิธีการทดสอบที่มีรูปแบบแตกต่างไปจากวิธีการทดสอบแบบประเพณีนิยม เนื่องจากการทดสอบวิธีนี้จะต้องกำหนดส่วนประกอบต่างๆ ที่ใช้ในการดำเนินการสอบไว้ล่วงหน้าทั้งสิ้น ซึ่งส่วนประกอบที่กล่าวมีดังนี้

1. ข้อทดสอบ (Test Item) ที่ใช้กับการทดสอบวิธีนี้เป็นข้อทดสอบที่ผ่านการวิเคราะห์ข้อทดสอบตามทฤษฎีการตอบสนองข้อทดสอบมา (Item Response Theory) แล้วตามทฤษฎีการตอบสนองข้อสอบ มีข้อตกลงเบื้องต้นว่าคุณลักษณะที่ทดสอบต้องเป็นมิติเดียวกัน ค่าพารามิเตอร์ของข้อทดสอบมีความคงที่ข้ามกลุ่มและมีความเป็นอิสระในการตอบข้อสอบ (Lord and Novick, cited in Koch, and Reckase, 1978 : 162) ดังนั้นการทดสอบวิธีนี้จึงสามารถประยุกต์ใช้ได้กับรูปแบบการวิเคราะห์ข้อทดสอบที่มี 1, 2 หรือ 3 พารามิเตอร์การเลือกใช้รูปแบบใดก็ควรพิจารณาธรรมชาติของข้อสอบและต้องเหมาะสมกับข้อมูลการตอบข้อทดสอบด้วย แต่จากงานวิจัยของ Kinley and Reckase (Kinley, and Reckase, 1980 : 75) เปรียบเทียบการทดสอบปรับเปลี่ยนโดยใช้รูปแบบการตอบข้อทดสอบที่มี 1 พารามิเตอร์และ 3 พารามิเตอร์ พบว่ารูปแบบที่มี 3 พารามิเตอร์ มีความเหมาะสมมากกว่ารูปแบบที่มี 1 พารามิเตอร์

นอกจากนี้ จากการศึกษาของ Thorndike (Thorndike, 1983 : 112) เกี่ยวกับแบบสอบปรับเปลี่ยนชี้ให้เห็นว่า ความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของการทดสอบปรับเปลี่ยนขึ้นอยู่กับจำนวนข้อและค่าอำนาจจำแนก กล่าวคือ ข้อทดสอบที่มีค่าอำนาจจำแนกสูงขึ้นค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าก็จะลดลง โดยปกติแล้วถ้าข้อทดสอบที่มีค่าอำนาจจำแนกโดยเฉลี่ย 0.7 ควรจะใช้ข้อสอบประมาณ 10 ถึง 15 ข้อ

2. กลุ่มข้อทดสอบ (Item Pool) จะต้องประกอบด้วยข้อทดสอบที่ประมาณค่าลักษณะเฉพาะของข้อทดสอบตามทฤษฎีการตอบสนองข้อทดสอบ พารามิเตอร์ของข้อทดสอบทุกข้อต้องได้รับการจัดให้อยู่ในมาตรวัดเดียวกันโดยใช้กระบวนการที่เหมาะสม ข้อทดสอบในกลุ่มข้อทดสอบจะต้องครอบคลุมช่วงของระดับความสามารถของประชากรและวัดความสามารถมิติเดียวกัน Urry (Urry, 1984 : 79) กล่าวว่า กลุ่มข้อทดสอบควรประกอบด้วยข้อทดสอบอย่างน้อย 100 ข้อ ค่าอำนาจจำแนกของข้อทดสอบเกิน 0.8 ความยากกระจายอยู่ในช่วง -3.00 ถึง +3.00 และค่าการเดาต้องน้อยกว่า 0.3

3. ระดับในการเริ่มต้น (Entry Level) ในการทดสอบปรับเปลี่ยน ผู้สอบแต่ละคนจะตอบข้อทดสอบต่างกัน ดังนั้นจึงสามารถที่จะให้ผู้สอบเริ่มต้นทำข้อทดสอบที่มีระดับความยากที่ต่างกันได้ การเริ่มทำข้อสอบที่คลาดเคลื่อนไปจากความสามารถที่เหมาะสมไม่มีผลกระทบต่อการประมาณความสามารถเท่าใดนัก แต่ถ้าให้ผู้สอบเริ่มทำข้อทดสอบที่ตรงกับความสามารถของเขาก็จะทำให้ลดจำนวนข้อที่จะใช้ในการทดสอบลงได้

4. กฎการเลือกข้อทดสอบ (Item Selection Rule) จะเลือกใช้ข้อทดสอบที่สามารถให้สารสนเทศสูงสุด ณ ระดับความสามารถของผู้สอบ การเลือกข้อทดสอบให้เหมาะสมกับระดับความสามารถของผู้สอบ นอกจากจะพิจารณาจากข้อที่ให้สารสนเทศสูงสุดแล้วผลการศึกษาเชิงประจักษ์ของ Lord (Lord, 1971 : 116) ชี้ให้เห็นว่าคะแนนจากแบบทดสอบจะสะท้อนให้เห็นความสามารถของแต่ละบุคคลได้แม่นยำเมื่อความน่าจะเป็นในการตอบข้อสอบได้ถูกต้องของแต่ละบุคคลเป็น .50 ดังนั้นแบบทดสอบจะมีความ

แม่นยำของการวัดค่า กล่าวคือ มีความคลาดเคลื่อนมาตรฐานสูงในการประมาณค่าความสามารถ เมื่อความน่าจะเป็นของการตอบข้อสอบถูกต้องของแต่ละบุคคลมีค่ามากหรือน้อยกว่า .05 (Weiss, 1974 : 118)

5. เกณฑ์การยุติการทดสอบ (Terminal Criterion) ขึ้นอยู่กับรูปแบบของการทดสอบแต่ละชนิดซึ่งสามารถวัดระดับความสามารถของผู้สอบได้อย่างแม่นยำมากน้อยต่างกัน เช่น รูปแบบปิรามิดขนาดชั้นคงที่ 10 ชั้น จะยุติการสอบเมื่อผู้สอบทำข้อสอบครบ 10 ข้อ เป็นต้น

### 3. การดำเนินการสอบ

ผู้สอบจะสอบทีละข้อ ผลการตอบข้อทดสอบแต่ละข้อจะถูกประมาณความสามารถเพื่อคัดเลือกข้อทดสอบให้เหมาะสมกับระดับความสามารถและมีเนื้อหาในปริเขตเดียวกันกับข้อเดิม เพื่อใช้เป็นข้อทดสอบที่จะตอบต่อไป กระบวนการจะดำเนินเช่นนี้จนสิ้นสุดการทดสอบ วิธีการนี้เหมาะใช้กับเครื่องคอมพิวเตอร์ (Urry, 1977 : 81; Weiss and Kingsbury, 1984 : 96)

การดำเนินการสอบ เริ่มต้นโดยการให้ผู้สอบนั่งหน้าจอคอมพิวเตอร์ ที่มีจานขับแม่เหล็ก (disk drive) ที่สามารถนำแผ่นบันทึกข้อมูล (diskette) ใส่หรือถอดออกมาได้ในแผ่นบันทึกข้อมูลจะมีโปรแกรมการทดสอบที่สร้างขึ้นตามองค์ประกอบต่างๆ ของการทดสอบปรับเปลี่ยนดังกล่าวข้างต้น ผู้ดำเนินการสอบจัดเตรียมเครื่องให้อยู่ในสภาพที่พร้อมจะทดสอบเครื่องคอมพิวเตอร์จะให้คำแนะนำในการตอบและรับข้อมูลส่วนตัวของนักเรียน เช่น ชื่อ-ชื่อสกุล เลขประจำตัว ชั้น/ห้อง เกรดเฉลี่ย เป็นต้น หลังจากนั้นเครื่องคอมพิวเตอร์จะแสดงข้อสอบข้อแรกบนจอภาพเพื่อให้ผู้สอบได้ตอบ ข้อสอบข้อแรกนี้เป็นข้อทดสอบที่มีค่าความยากปานกลางถ้าผู้สอบตอบข้อทดสอบข้อแรกถูกข้อทดสอบข้อถัดไปก็จะเป็นข้อทดสอบที่มีความยากเพิ่มขึ้น จากนั้นก็จะประมาณค่าความสามารถที่แท้จริงของผู้สอบ การทดสอบจะดำเนินต่อไปตามแนวทางต่อไปนี้คือ

1. หลังจากประมาณค่าความสามารถของผู้สอบแล้วเครื่องคอมพิวเตอร์ก็จะคัดเลือกข้อสอบข้อใหม่ที่มีประโยชน์ในการกำหนดความสามารถจริง

2. แสดงข้อทดสอบข้อใหม่บนจอภาพและผู้สอบ

ตอบข้อทดสอบโดยการเลือกตัวเลือกที่ต้องการ แล้วพิมพ์ตัวเลือกที่ต้องการ คอมพิวเตอร์จะรับคำตอบแล้วตรวจสอบข้อสอบ

3. ประมาณค่าความสามารถของผู้สอบอีกครั้งหนึ่งตามผลของการตอบข้อทดสอบข้อที่ผ่านมา

4. พิจารณาว่าผู้สอบควรจะทำข้อทดสอบข้อต่อไปนี้หรือไม่ตามเกณฑ์การยุติการสอบที่กำหนดไว้

5. ถ้าผู้สอบจะต้องทำข้อทดสอบข้อต่อไปก็ย้อนกลับไปขั้นที่ 1 อีกครั้ง

เมื่อสิ้นสุดการทดสอบ สารสนเทศเกี่ยวกับการตอบข้อทดสอบของผู้สอบก็จะถูกบันทึกไว้ในเครื่องคอมพิวเตอร์ และแสดงผลการทดสอบให้ผู้สอบทราบ

#### 4. ประเภทของการทดสอบแบบปรับเปลี่ยน

การทดสอบแบบปรับเปลี่ยน แบ่งออกเป็น 2 ประเภทใหญ่ๆ (Weiss, 1974 : 98 ; Hambleton and Swaminathan, 1985 : 86) และในแต่ละประเภทยังแบ่งย่อยได้อีกดังนี้

1. กลวิธีสองขั้นตอน (Two-stage Strategies)
2. กลวิธีหลายขั้นตอน (Multi-stage Strategies) ซึ่งแบ่งเป็น

2.1 รูปแบบแยกทางคงที่ (Fixed Branching Model) เช่น รูปแบบปิรามิด (Pyramidal Model) ซึ่งมีหลายรูปแบบ

2.2 รูปแบบแยกทางแปรผัน (Variable Branching Model) ซึ่งมีอีกหลายรูปแบบ

### วิธีดำเนินการวิจัย

#### ประชากร

ประชากรของการวิจัยครั้งนี้ได้แก่นิสิตชั้นปีที่ 1 ของจุฬาลงกรณ์มหาวิทยาลัยที่ลงทะเบียนเรียนรายวิชาภาษาอังกฤษพื้นฐาน II ในภาคปลายปีการศึกษา 2535 จำนวนประมาณ 2,404 คน (ยกเว้นนิสิตชั้นปีที่ 1 ของคณะวิศวกรรมศาสตร์ คณะวิทยาศาสตร์ และคณะอักษรศาสตร์)

### พลวิจัย

กำหนดเกณฑ์นิสิตที่จะนำมาใช้ในการวิจัยครั้งนี้ว่าจะต้องเป็นนิสิตของคณะที่มีพื้นฐานความรู้ทางภาษาอังกฤษทั้งในระดับเก่ง ปานกลาง และอ่อน อย่างละประมาณ 30 คน เป็นอย่างน้อย และคณะต้องมีเครื่องไมโครคอมพิวเตอร์เพื่อการเรียนการสอนอย่างน้อย 15 เครื่อง ทั้งนี้เพื่อความสะดวกในการบริหารงานวิจัยครั้งนี้ ปรากฏว่าคัดเลือกได้ คณะแพทยศาสตร์ นิเทศศาสตร์ เกษศาสตร์ รัฐศาสตร์ สถาปัตยกรรมศาสตร์ ครุศาสตร์ และเศรษฐศาสตร์ จึงทำการสุ่มอย่างง่าย

ผู้วิจัยจึงดำเนินการแบ่งนิสิต โดยอาศัยคะแนนสอบภาษาอังกฤษพื้นฐาน I FE I เป็นเกณฑ์และได้พลวิจัยจำนวน 360 คน แล้วแบ่งเป็น 2 กลุ่มๆ ละ 180 คน

#### เครื่องมือการวิจัย

1. แบบทดสอบวัดสมรรถภาพทั่วไปทางภาษาอังกฤษ (general English proficiency test)

แบบอิงเกณฑ์เพื่อวัดความสามารถด้านโครงสร้างทางภาษาอังกฤษ แบบทดสอบนี้เป็นแบบทดสอบปรนัย ชนิดเลือกตอบ 4 ตัวเลือก จำนวน 105 ข้อ มีความตรง (validity) ซึ่งวัดโดยดัชนีความสอดคล้องของข้อทดสอบกับวัตถุประสงค์ (objective-item congruence index) = 0.92 และมีความเชื่อถือได้ (dependability) เมื่อจุดตัดเท่ากับ 80% = 0.836 จึงนับว่าเป็นแบบทดสอบอิงเกณฑ์ที่มีคุณสมบัติเพียงพอที่จะใช้เพื่อการวิจัยครั้งนี้

2. แบบทดสอบปรับเปลี่ยนที่โยงเนื้อหาโดยใช้คอมพิวเตอร์ (Computerized Content-based Adaptive Test : CCAT) เป็นแบบทดสอบที่พัฒนามาจากแบบทดสอบวัดสมรรถภาพทั่วไปทางภาษาอังกฤษที่กล่าวแล้วในข้อที่ 1 โดยกระบวนการดังนี้

2.1 นำแบบทดสอบวัดสมรรถภาพทั่วไปทางภาษาอังกฤษไปทดสอบกับนิสิตชั้นปีที่ 1 คณะพาณิชยศาสตร์และการบัญชี แพทยศาสตร์ และเศรษฐศาสตร์ จำนวน 256 คน

- 2.2 นำผลการสอบมาวิเคราะห์ข้อทดสอบรายข้อ

โดยวิธีประเพณีนิยม (Classical Model) เพื่อปรับปรุงคุณภาพของข้อทดสอบ และคัดเลือกข้อทดสอบที่ดีที่สุดไว้ใช้ 144 ข้อ

2.3 นำแบบทดสอบที่ปรับปรุงแล้วไปสอบกับนิสิตชั้นปีที่ 1 ของคณะแพทยศาสตร์ คณะพาณิชยศาสตร์และการบัญชี คณะนิเทศศาสตร์และคณะเศรษฐศาสตร์ที่ไม่ได้สอบแบบทดสอบครั้งแรก จำนวน 1,123 คน

2.4 นำผลการสอบมาวิเคราะห์ข้อทดสอบรายข้อโดยวิธีโลจิสติกแบบ 3 พารามิเตอร์ (3-parameter Logistic Model) ได้ข้อทดสอบที่ดีที่สุดไว้ใช้ต่อไป 125 ข้อแต่เลือกมาใช้เพียง 105 ข้อ

2.5 จัดข้อทดสอบที่ผ่านการวิเคราะห์ในขั้นที่ 4 เป็นแบบปิรามิด 15 ชั้น โดยแต่ละชั้นมีข้อทดสอบเท่ากับจำนวนชั้น กล่าวคือชั้นที่ 1 มีข้อสอบ 1 ข้อ ชั้นที่ 2 มี 2 ข้อ และชั้นที่ 3 มีข้อสอบ 3 ข้อ เรื่อยๆ ไปจนถึงชั้นที่ 15 รวมมีข้อทดสอบทั้งหมด 105 ข้อ

แบบทดสอบชุดนี้มีค่าอำนาจจำแนก (a) โดยเฉลี่ยเท่ากับ 1.512 (ระหว่าง 0.802-2.00) ค่าความยาก (b) โดยเฉลี่ยเท่ากับ 0.187 (ระหว่าง -0.904-3.298) ค่าการเดา (c) โดยเฉลี่ยเท่ากับ 0.201 (ระหว่าง 0.035-0.298) ซึ่งอยู่ในเกณฑ์ของแบบทดสอบที่ดีทั่วไป

3. แบบสำรวจเจตคติในการใช้แบบทดสอบวัดสมรรถภาพทั่วไปทางภาษาอังกฤษ และแบบทดสอบปรับเปลี่ยนที่เน้นเนื้อหาโดยใช้คอมพิวเตอร์ ซึ่งเป็นแบบสำรวจที่ใช้มาตราวัดความแตกต่างของความหมายของคำ (semantic differential scale) ชนิด 7 ระดับ แบบสำรวจนี้มีความเที่ยง (reliability) ชนิด Cronbach alpha เท่ากับ 0.93 ซึ่งนับว่าค่อนข้างสูงมาก

## การเก็บรวบรวมข้อมูล

1. จัดให้พลวิจัยกลุ่มที่ 1 สอบแบบทดสอบวัดสมรรถภาพทั่วไปทางภาษาอังกฤษมาแล้วตามแบบประเพณีนิยม (conventional method) และพลวิจัยอีกกลุ่มหนึ่งสอบแบบทดสอบแบบปรับเปลี่ยนที่เน้นเนื้อหาโดยใช้คอมพิวเตอร์ แล้วบันทึกผลการสอบของผู้สอบแต่ละคนและเวลาที่ใช้ในการสอบไว้โดยอาศัยเครื่องคอมพิวเตอร์

2. ใช้แบบสำรวจเจตคติเพื่อสำรวจเจตคติของผู้สอบเกี่ยวกับการสอบแบบทดสอบดังกล่าวแบบประเพณีนิยมและแบบปรับเปลี่ยนโดยใช้คอมพิวเตอร์ รวมทั้งปัญหาและความคิดเห็นต่างๆ ในการสอบดังกล่าวแล้วด้วย

## การวิเคราะห์ข้อมูล

1. ทำการวิเคราะห์ข้อทดสอบโดยวิธีประเพณีนิยม (Classical Model) โดยใช้โปรแกรม ITEMX ของผู้วิจัยที่เขียนขึ้นไว้

2. ทำการวิเคราะห์ข้อทดสอบโดยวิธีโลจิสติกแบบ 3 พารามิเตอร์ (3-parameter Logistic Model) โดยโปรแกรม LOGIST5 ซึ่งสร้างโดย ETS (Educational Testing Service) ของอเมริกา

3. ใช้โปรแกรม SPSSX ชื่อ Factor Analysis เพื่อวิเคราะห์หาความมีมิติเดียวของแบบทดสอบ และใช้โปรแกรม SPSS/PC เพื่อวิเคราะห์หาสถิติบรรยายต่างๆ ที่เกี่ยวข้อง

4. ผู้วิจัยเขียนโปรแกรมคอมพิวเตอร์ขึ้นใช้เองเพื่อคำนวณหาค่าสถิติอื่นๆ ที่เกี่ยวข้อง เช่น ค่าฟังก์ชันสารสนเทศของแบบทดสอบ ค่าความคลาดเคลื่อนมาตรฐานในการประมาณหาความสามารถ ค่าฟังก์ชันสารสนเทศของข้อทดสอบ ค่าความเชื่อถือได้ของแบบทดสอบอิงปริเขต และค่าดัชนีความสอดคล้องของข้อทดสอบกับวัตถุประสงค์ เป็นต้น

## ผลการวิจัย

### 1. การจัดสร้างโปรแกรมคอมพิวเตอร์

สามารถจัดสร้างโปรแกรมคอมพิวเตอร์ที่สามารถทำงานเพื่อใช้ทดสอบปรับเปลี่ยนเนื้อหาโดยใช้คอมพิวเตอร์ (Computerized Content-Based Adaptive Testing) ได้โปรแกรมนี้เขียนโดยใช้โปรแกรม Foxbase Version 2.0 สามารถใช้เป็นโปรแกรมบรรณาธิการ (authoring program) ได้ กล่าวคือ ผู้ใช้สามารถเปลี่ยนแปลงข้อมูลของข้อทดสอบที่จะนำมาใช้เพื่อเก็บไว้โปรแกรมได้นอกจากนี้โปรแกรมยังสามารถบันทึกข้อมูลต่างๆ เกี่ยวกับผู้สอบแต่ละครั้งได้ เช่น เวลาในการสอบ ข้อที่ทำถูกต้อง กระสวน

(Pattern) ของการตอบและความสามารถโดยเฉลี่ยของผู้สอบ  
เป็นต้น

## 2. คุณภาพและประสิทธิภาพของการสอบ ปรนัย

ตารางที่ 1: ค่าความตรงเชิงพยากรณ์ของคะแนนความสามารถของการทดสอบแบบประเพณีนิยมและแบบปรับเปลี่ยนเนื้อหา (CCAT)

คะแนนสอบ	วิธีการทดสอบ		z
	ประเพณีนิยม	CCAT	
CCAT	0.876*	-	
ผลสัมฤทธิ์เฉลี่ย	0.785*	0.762*	0.665

\*  $p < 0.05$

ตารางที่ 2: ค่าเฉลี่ยและค่าเบี่ยงเบนมาตรฐานของความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของแบบทดสอบทั้ง 2 แบบ

วิธีการทดสอบ	n	SEE	S.D. <sub>see</sub>	t
CCAT	31	0.3045	0.024	-8.108*
ประเพณีนิยม	31	0.5332	0.172	

\*  $p < 0.05$

จากตารางที่ 1 และ 2 แสดงว่า คุณภาพของการทดสอบแบบปรับเปลี่ยนมีมากกว่าคุณภาพของการทดสอบแบบประเพณีนิยม กล่าวคือการทดสอบทั้ง 2 แบบ มีความตรงเชิงพยากรณ์กับคะแนนผลสัมฤทธิ์เฉลี่ยเท่ากับ 0.762 และ 0.785 ตามลำดับ และแตกต่างกันอย่างไม่มีนัยสำคัญ และผลการสอบของการทดสอบทั้ง 2 แบบ มีความสัมพันธ์ซึ่งกันและกันเท่ากับ 0.876 ซึ่งนับว่า ค่อนข้างสูงมากและการทดสอบแบบปรับเปลี่ยนมีความคลาด-มาตรฐานในการประมาณค่าความสามารถโดยเฉลี่ย (SEE) น้อยกว่าการทดสอบแบบประเพณีนิยมอย่างมีนัยสำคัญ

นอกจากนี้การทดสอบแบบปรับเปลี่ยนยังมีประสิทธิภาพสูงกว่าการทดสอบแบบประเพณีนิยม เนื่องจาก

ใช้เวลาในการสอบน้อยกว่าประมาณ 3 เท่า ผู้สอบมีความเครียดในการสอบน้อยกว่า มีเจตคติในการสอบดีกว่าและใช้ข้อทดสอบน้อยกว่ากันมาก คือใช้ 50 ข้อ แทนที่จะเป็น 120 ข้อ แต่ได้ผลการสอบเท่าเทียมกัน ซึ่งรายละเอียดบางอย่างได้แสดงไว้แล้วในตารางที่ 1-2 ข้างต้น

### แหล่งที่มา

สุพัฒน์ สุขมลสันต์, รายงานการวิจัย เรื่อง "คุณภาพและประสิทธิภาพของแบบทดสอบปรับเปลี่ยนที่เน้นเนื้อหา โดยใช้เครื่องคอมพิวเตอร์ เพื่อการทดสอบแบบอิงปริเซต", สถาบันภาษา จุฬาลงกรณ์มหาวิทยาลัย, 2536