

## การพัฒนาระบบการทดสอบแบบปรับเหมาะหลายขั้นตอนที่มีการให้ข้อมูลป้อนกลับใน การทดสอบระดับชาติ (O-NET) วิชาภาษาอังกฤษสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6 System Development for Multistage Adaptive Testing with Feedback in O-NET for The Subject of Grade 12 English Language

◆ ฌภัทร ชัยมงคล

อาจารย์ ดร. ประจำภาควิชาวิจัยและจิตวิทยาการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

Nhabhat Chaimongkol

Lecturer, Ph.D., Department of Educational Research and Psychology,

Faculty of Education, Chulalongkorn University, E-mail: nhabhat.c@chula.ac.th

*Received: July 23, 2020; Revised: September 14, 2020; Accepted: October 12, 2020*

---

### Abstract

Computerized adaptive testing (CAT) is a system that uses a computer process to select items that match the ability of examinees. The classical CAT still has some limitations in ability estimation. Therefore, researchers have developed multistage adaptive testing (MST) to overcome the limitations of the classical CAT. This research develops MST for the preparation of Ordinary National Educational Test (O-NET) in English. This is because most students are unable to pass the O-NET exam in English. The results found that the system consists of 5 components 1) initial stage 2) ability estimation 3) item selection 4) stopping criteria and 5) feedback report, contained both the raw score and ability level of each examinee. Finally, the quality of the system is evaluated by experts using standard evaluation and by students using user satisfaction. The results show that the system is of high quality in every dimension and the overall user satisfaction is quite good.

**Keywords :** Multistage Adaptive Testing, Ordinary National Educational Test, English subject

### บทคัดย่อ

การทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์เป็นการทดสอบที่นำกระบวนการทางคอมพิวเตอร์เข้ามาช่วยในการคัดเลือกข้อสอบที่เหมาะสมกับระดับความสามารถของผู้สอบแต่ละคน ซึ่งการทดสอบแบบปรับเหมาะแบบดั้งเดิมมีข้อจำกัดในประเด็นการประมาณค่าความสามารถที่ต่ำหรือสูงกว่าความเป็นจริงจากข้อสอบข้อเดียว จึงได้มีการพัฒนาระบบการทดสอบแบบปรับเหมาะหลายขั้นตอนขึ้นเพื่อช่วยลดข้อจำกัดดังกล่าว สำหรับงานวิจัยนี้เป็นพัฒนาระบบการทดสอบแบบปรับเหมาะหลายขั้นตอนที่นำมาประยุกต์ใช้กับ

การเตรียมความพร้อมในการสอบระดับชาติวิชาภาษาอังกฤษ เนื่องจากผลการทดสอบระดับชาติวิชาภาษาอังกฤษที่ผ่านมาของประเทศไทย ผู้เรียนส่วนใหญ่ยังมีคะแนนยังไม่ผ่านเกณฑ์มาตรฐาน ผลการวิจัยพบว่า องค์ประกอบของการทดสอบแบบปรับเหมาะหลายขั้นตอนมีองค์ประกอบที่สำคัญ 5 องค์ประกอบ ได้แก่ 1) จุดเริ่มต้นในการทดสอบ 2) การประมาณค่าความสามารถของผู้สอบ 3) การคัดเลือกข้อสอบ 4) จุดยุติการทดสอบ และ 5) รายงานผลข้อมูลป้อนกลับซึ่งประกอบด้วยรายงานคะแนนสังเกตได้และรายงานระดับความสามารถของผู้สอบ สำหรับผลการตรวจสอบคุณภาพระบบโดยผู้เชี่ยวชาญด้วยแบบประเมินระบบแบบอิงมาตรฐาน พบว่า ระบบมีคุณภาพอยู่ในระดับมากที่สุดทุกด้าน และผลการประเมินความพึงพอใจของผู้สอบต่อการใช้งานระบบ พบว่า ความเห็นต่อการใช้งานระบบในภาพรวมอยู่ในระดับดี

**คำสำคัญ :** ระบบการทดสอบแบบปรับเหมาะหลายขั้นตอน การทดสอบระดับชาติ วิชาภาษาอังกฤษ

## บทนำ

การทดสอบแบบดั้งเดิมจะใช้แบบทดสอบฉบับเดียวกับผู้สอบทุกคน ซึ่งอาจส่งผลต่อความถูกต้องแม่นยำในการวัด เนื่องจากแบบทดสอบฉบับหนึ่งประกอบด้วยข้อสอบหลายข้อที่มีระดับความยากแตกต่างกัน ส่งผลให้ผู้สอบบางคนได้ทำข้อสอบที่ยากหรือง่ายเกินไป ดังนั้น การทดสอบในอุดมคติจึงมีแนวคิดที่ว่าข้อสอบแต่ละข้อควรมีความยากที่เหมาะสมกับความสามารถของแต่ละบุคคลและเรียกการทดสอบดังกล่าวว่า การทดสอบแบบปรับเหมาะ (adaptive testing or tailored testing) ต่อมาเมื่อเทคโนโลยีคอมพิวเตอร์เริ่มเข้ามามีบทบาทเพิ่มมากขึ้นจึงเรียกว่าการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ (computerized adaptive testing: CAT) ซึ่งหลักการสำคัญของการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ คือ กระบวนการคัดเลือกข้อสอบ (item selection) ที่แตกต่างกันไปตามการออกแบบการทดสอบ โดยอาศัยคุณลักษณะของข้อสอบตามทฤษฎีการตอบสนองข้อสอบ (item response theory: IRT) ได้แก่ ค่าความยาก ค่าอำนาจจำแนก และค่าความน่าจะเป็นในการเดา (Chang, 2004; Kanjanawasee, 2012)

แม้ว่า CAT จะมีจุดเด่นที่เหนือกว่าการทดสอบแบบดั้งเดิมอยู่หลายประการแต่ยังพบข้อจำกัด คือ การประมาณค่าคะแนนความสามารถของผู้สอบต่ำกว่าและสูงกว่าความเป็นจริง ด้วยเหตุนี้ จึงได้มีการพัฒนากระบวนการทดสอบซึ่งเรียกว่า การทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์แบบหลายขั้นตอน (multistage adaptive testing: MST) ซึ่งจะช่วยให้การประมาณค่าความสามารถของผู้สอบในขั้นตอนแรกเกิดความแม่นยำมากขึ้นเพราะประมาณค่าความสามารถของผู้สอบจากการตอบชุดข้อสอบ จึงช่วยลดปัญหาในการประมาณค่าความสามารถต่ำกว่าหรือสูงกว่าความเป็นจริง (Zheng & Chang, 2015)

นอกจากนี้ MST ยังสามารถรายงานผลการทดสอบตามระดับความสามารถของผู้สอบได้ทันทีหลังจากการทดสอบเสร็จสิ้น ปัจจุบันพบการนำข้อมูลป้อนกลับ (feedback) มาประยุกต์ใช้กับการรายงานผลจาก MST ไม่มากนัก ซึ่งข้อมูลป้อนกลับถือเป็นสารสนเทศที่มีประโยชน์ต่อการพัฒนาตัวผู้รับการทดสอบเป็นอย่างมาก และนอกจากการพัฒนาาระบบเพื่อการทดสอบแล้ว ประเด็นสำคัญที่ต้องคำนึงถึงในการพัฒนาระบบการ

ทดสอบ คือ การประเมินคุณภาพของระบบซึ่งมีแนวทางการประเมินแตกต่างกันออกไป สำหรับการประเมินระบบ MST สามารถประยุกต์ใช้แนวคิดการประเมินคุณลักษณะของระบบจากผู้ใช้ (user evaluation) คือ การประเมินความพึงพอใจระหว่างอินเตอร์เฟซกับผู้ใช้ (user satisfaction of the human-computer interface) (Lilley, Barker, & Britton, 2004) นอกจากนี้ ยังสามารถประยุกต์ใช้แนวคิดการประเมินตามมาตรฐานการประเมินระบบร่วมด้วย คือ ประเมินคุณสมบัติของระบบในมิติต่าง ๆ ได้แก่ การใช้ประโยชน์ (utility) ความเป็นไปได้ (feasibility) ความเหมาะสม (propriety) และ ความถูกต้อง (accuracy) (Kanjanawasee, 2011) ดังนั้น หากนำ MST มาใช้ร่วมกับการให้ข้อมูลป้อนกลับจะสามารถพัฒนาผู้เรียนได้มากยิ่งขึ้น โดยเฉพาะอย่างยิ่งในการทดสอบระดับชาติ (O-NET) ซึ่งจะเป็นประโยชน์ต่อผู้เรียนและวงการศึกษามากขึ้น

O-NET (Ordinary National Educational Test) คือ การทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน เป็นการทดสอบวัดความรู้และความคิดของนักเรียนชั้นประถมศึกษาปีที่ 6 ชั้นมัธยมศึกษาปีที่ 3 และ ชั้นมัธยมศึกษาปีที่ 6 เป็นการประเมินตามมาตรฐานการเรียนรู้ในหลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พุทธศักราช 2551 การทดสอบประกอบด้วย 5 วิชา ได้แก่ คณิตศาสตร์ วิทยาศาสตร์ ภาษาอังกฤษ ภาษาไทย และ สังคมศึกษา ศาสนาและวัฒนธรรม ซึ่งการทดสอบในระดับชั้นมัธยมศึกษาปีที่ 6 ถือว่าเป็นการทดสอบที่สำคัญต่อนักเรียนเป็นอย่างมาก เนื่องจากนักเรียนต้องนำความรู้พื้นฐานไปใช้ศึกษาต่อในระดับมหาวิทยาลัย โดยวิชาที่มีความจำเป็นต่อการศึกษาในระดับที่สูงขึ้น คือ วิชาภาษาอังกฤษ ซึ่งจากสถิติผลการทดสอบย้อนหลังปี 2558-2560 พบว่า คะแนนเฉลี่ยวิชาภาษาอังกฤษอยู่ในเกณฑ์ต่ำ (National Institute of Educational Testing Service, 2019) ดังนั้น จึงควรเร่งผลักดันให้มีการพัฒนาการเรียนรู้นักเรียนในรายวิชาภาษาอังกฤษ เพื่อให้ให้นักเรียนมีองค์ความรู้ภาษาอังกฤษทัดเทียมกับนานาชาติ และรองรับการเติบโตขององค์ความรู้ใหม่ในอนาคต

ด้วยเหตุนี้ ผู้วิจัยจึงมีความสนใจพัฒนาระบบ MST ที่มีการให้ข้อมูลป้อนกลับในการทดสอบระดับชาติ วิชาภาษาอังกฤษสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6 เพื่อให้นักเรียนสามารถนำข้อมูลป้อนกลับที่ได้รับไปใช้ในการเตรียมความพร้อมและวางแผนก่อนการทดสอบจริง ซึ่งจะช่วยยกระดับมาตรฐานคะแนนการทดสอบระดับชาติให้สูงขึ้นและเป็นไปตามเกณฑ์มาตรฐาน และสามารถใช้เป็นแนวทางในอนาคตสำหรับพัฒนาระบบ MST ในการทดสอบระดับชาติรายวิชาอื่น

## วัตถุประสงค์

1. เพื่อพัฒนาระบบ MST ที่มีการให้ข้อมูลป้อนกลับในการทดสอบระดับชาติวิชาภาษาอังกฤษสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6
2. เพื่อประเมินคุณภาพระบบ MST ที่มีการให้ข้อมูลป้อนกลับในการทดสอบระดับชาติวิชาภาษาอังกฤษสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6

## ขอบเขตการวิจัย

1. MST คือ การทดสอบที่ช่วยลดความคลาดเคลื่อนของการประมาณค่าความสามารถในการทดสอบที่ต่ำกว่าหรือสูงกว่าความเป็นจริงจาก CAT ในขั้นตอนแรก โดย MST ใช้วิธีการแบ่งข้อสอบออกเป็นชั้น (stage) ซึ่งในแต่ละชั้นประกอบด้วยโมดูล (module) ที่มีข้อสอบแบ่งไว้ตามระดับความยาก และอนุญาตให้ผู้สอบกลับไปแก้ไขข้อสอบได้ โดยการแก้ไขจะสามารถทำได้เฉพาะชั้นที่กำลังดำเนินการทดสอบเท่านั้น

2. ข้อสอบที่นำมาใช้ในการพัฒนาระบบการทดสอบ คือ ข้อสอบระดับชาติวิชาภาษาอังกฤษสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6 ระหว่างปี 2558 – 2561 โดยผลการตอบข้อสอบเพื่อนำมาวิเคราะห์คุณภาพข้อสอบในการจัดทำคลังได้รับความอนุเคราะห์จากสถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) ซึ่งโดยทั่วไปคลังข้อสอบที่เหมาะสมควรมีอย่างน้อย 100-200 ข้อ (Weiss อ้างถึงใน Kanjanawasee (2012)) สำหรับคลังข้อสอบในงานวิจัยนี้ได้ดำเนินการวิเคราะห์หามาจากข้อสอบจำนวน 320 ข้อ ซึ่งเป็นข้อสอบระดับชาติวิชาภาษาอังกฤษสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6 ระหว่างปี 2558 – 2561 โดยมีเกณฑ์ในการคัดเลือกข้อสอบเข้าคลัง คือ ค่าความยาก (b) อยู่ในช่วง -4 ถึง 4 ค่าอำนาจจำแนก (a) อยู่ในช่วง 0 ถึง  $\alpha$  และความน่าจะเป็นในการเดา (c) มีค่าไม่เกิน 0.50 ซึ่งในการวิเคราะห์เพื่อประมาณค่าพารามิเตอร์ของข้อสอบ ดำเนินการวิเคราะห์ด้วยซอฟต์แวร์ R เนื่องจากเป็นซอฟต์แวร์ที่สามารถใช้งานโดยไม่เสียค่าใช้จ่าย (freeware) ด้วยแพ็คเกจ (package) ltm ซึ่งเป็นแพ็คเกจที่ใช้ในการวิเคราะห์ข้อสอบตามทฤษฎีการตอบสนองข้อสอบ

3. ระบบ MST ที่มีการให้ข้อมูลป้อนกลับในการทดสอบระดับชาติวิชาภาษาอังกฤษสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6 พัฒนาด้วยภาษา PHP ซึ่งสามารถใช้งานผ่านระบบออนไลน์ โดยเป้าหมายของการพัฒนาระบบ คือ เพื่อเตรียมความพร้อมผู้เรียนระดับชั้นมัธยมศึกษาปีที่ 6 ในการพัฒนาตนเองก่อนการทดสอบระดับชาติวิชาภาษาอังกฤษจริง

4. การประเมินระบบการทดสอบประกอบด้วย การประเมิน 2 ส่วน ได้แก่ 1) การประเมินแบบอิงมาตรฐาน ได้แก่ การใช้ประโยชน์ ความเป็นไปได้ ความเหมาะสม และความถูกต้อง (Kanjanawasee, 2011) และ 2) การประเมินความพึงพอใจของผู้สอบต่อการใช้งานระบบ ซึ่งมีหลักเกณฑ์ในการประเมิน 5 ด้าน ได้แก่ หน้าจอระบบ คำศัพท์และสารสนเทศของระบบ การเรียนรู้ และสมรรถนะของระบบ (Chin, Diehl, & Norman, 1988)

**กรอบแนวคิด**



## นิยามศัพท์

1. ระบบ MST ที่มีการสะท้อนข้อมูลป้อนกลับ หมายถึง ชุดขององค์ประกอบที่มีความสัมพันธ์กันใน MST ที่สามารถสะท้อนข้อมูลป้อนกลับให้ผู้เข้ารับการทดสอบเพื่อนำไปพัฒนาและปรับปรุงตนเอง โดยระบบประกอบด้วย 5 องค์ประกอบ ได้แก่ 1) จุดเริ่มต้นในการทดสอบ (initial stage) 2) การประมาณค่าความสามารถของผู้สอบ (ability estimation) 3) การคัดเลือกข้อสอบ (item selection) 4) จุดยุติการทดสอบ (stopping criteria) และ 5) รายงานผลข้อมูลป้อนกลับซึ่งประกอบด้วยรายงาน 2 รูปแบบ ได้แก่ 1) รายงานคะแนนสังเกตได้ (observe score) และ 2) รายงานระดับความสามารถของผู้สอบ (ability)

2. MST หมายถึง การทดสอบที่มีการคัดเลือกข้อสอบให้เหมาะสมกับระดับความสามารถของผู้สอบ ซึ่งการประมาณค่าความสามารถของผู้สอบมาจากโมดูลของข้อสอบในแต่ละชั้น โดยแต่ละชั้นจะประกอบไปด้วยโมดูลของข้อสอบที่มีการจัดเตรียมไว้โดยผู้บริหารการทดสอบ และแต่ละโมดูลจะประกอบไปด้วยข้อสอบที่มีความยากใกล้เคียงกัน

3. พารามิเตอร์ของข้อสอบ หมายถึง คุณลักษณะของข้อสอบแต่ละข้อซึ่งประกอบด้วย ค่าพารามิเตอร์ความยาก (difficulty:  $b$ ) ค่าพารามิเตอร์อำนาจจำแนก (discrimination:  $a$ ) และค่าพารามิเตอร์ความน่าจะเป็นในการเดา (guessing:  $c$ )

4. แบบทดสอบระดับชาติวิชาภาษาอังกฤษสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6 หมายถึง แบบทดสอบทางการศึกษาระดับชาติด้านพื้นฐานซึ่งใช้เพื่อวัดความรู้และความคิดของนักเรียนชั้นมัธยมศึกษาปีที่ 6 ประเมินตามมาตรฐานการเรียนรู้ในหลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พุทธศักราช 2551 โดยวิชาภาษาอังกฤษเป็นหนึ่งในกลุ่มสาระการเรียนรู้จากทั้งหมด 5 กลุ่มสาระการเรียนรู้

5. ข้อมูลป้อนกลับ หมายถึง ชุดข้อมูลที่สะท้อนความสามารถของผู้เรียนโดยระบุถึงจุดเด่น จุดที่ควรปรับปรุง เพื่อนำไปใช้เป็นแนวทางในการพัฒนาหรือปรับปรุงตนเองของผู้เรียนให้ดียิ่งขึ้น

6. การประเมินระบบการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ หมายถึง การประเมินระบบจากผู้เชี่ยวชาญด้วยการประเมินแบบอิงมาตรฐาน ได้แก่ ความมีประโยชน์ ความเป็นไปได้ ความเหมาะสม และความถูกต้อง และการประเมินจากผู้สอบด้วยการประเมินความพึงพอใจของผู้สอบต่อการใช้งานระบบ

## การทบทวนวรรณกรรม

การวิจัยในครั้งนี้ ผู้วิจัยแบ่งเนื้อหาในการทบทวนวรรณกรรมออกเป็น 5 ตอน ได้แก่ ตอนที่ 1 มโนทัศน์เกี่ยวกับ CAT ตอนที่ 2 มโนทัศน์เกี่ยวกับ MST ตอนที่ 3 มโนทัศน์เกี่ยวกับการให้ข้อมูลป้อนกลับ ตอนที่ 4 มโนทัศน์เกี่ยวกับการทดสอบระดับชาติวิชาภาษาอังกฤษ และตอนที่ 5 มโนทัศน์เกี่ยวกับการประเมินผลระบบการทดสอบ ซึ่งแต่ละตอนมีรายละเอียด ดังนี้

*ตอนที่ 1 มโนทัศน์เกี่ยวกับการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์*

CAT เป็นเทคโนโลยีทางการทดสอบที่มีจุดแข็งในเชิงทฤษฎี คือ เป็นการทดสอบที่ผู้สอบจะได้ทำข้อสอบที่มีความยากตามระดับความสามารถของตน และจำนวนข้อสอบที่ใช้ในการทดสอบจะลดลงเมื่อเทียบ

กับการทดสอบแบบดั้งเดิมโดยไม่เสียความถูกต้องแม่นยำในการประมาณค่าความสามารถของผู้สอบ (Zheng & Chang, 2015) นอกจากนี้ CAT ยังเป็นวิธีที่สามารถนำมาใช้ในการประเมินความก้าวหน้าของนักเรียนทั้งแบบรายบุคคลและแบบกลุ่ม (Weiss, 2004) แม้ว่า CAT จะมีจุดแข็งที่เหนือกว่าการทดสอบแบบดั้งเดิมแต่ยังมีจุดอ่อนในเชิงทฤษฎี คือ อาจเกิดการประมาณค่าความสามารถที่ต่ำหรือสูงกว่าความเป็นจริง (Zheng & Chang, 2015) ด้วยเหตุนี้ จึงมีการพัฒนา MST ขึ้น ซึ่งจากงานวิจัยที่ศึกษาเปรียบเทียบระหว่าง CAT กับ MST พบว่า การเพิ่มจำนวนขั้นในการทดสอบสามารถช่วยลดความคลาดเคลื่อนในการประมาณค่าความสามารถของผู้สอบได้ (Patsula, 1999)

#### *ตอนที่ 2 มโนทัศน์เกี่ยวกับการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์แบบหลายขั้นตอน*

ใน MST ผู้สอบสามารถข้ามคำถาม ทบทวน หรือแก้ไขคำตอบในขั้นที่กำลังทดสอบได้ (Zheng, Nozawa, Gao, & Chang, 2012) ซึ่ง MST จะมีการจัดชุดข้อสอบไว้ล่วงหน้า (pre-assembled) โดยแบ่งเป็นขั้นซึ่งแต่ละขั้นจะประกอบด้วยชุดข้อสอบหลายชุด แต่ละชุดจะประกอบไปด้วยจำนวนโมดูลของข้อสอบ ในการสอบผู้สอบจะได้ชุดแบบทดสอบที่มีความยากสอดคล้องกับระดับความสามารถของตนเอง ดังนั้น MST จึงเป็นการปรับชุดของข้อสอบให้เหมาะสมกับความสามารถของผู้สอบแทนการปรับข้อสอบเพียงครั้งละหนึ่งข้อ (Wang, 2017)

#### *ตอนที่ 3 มโนทัศน์เกี่ยวกับการให้ข้อมูลป้อนกลับ*

ข้อมูลป้อนกลับเป็นมโนทัศน์ที่ได้รับการกล่าวถึงเป็นจำนวนมากทั้งในบริบททางธุรกิจและบริบททางการศึกษา เนื่องจากข้อมูลป้อนกลับมีอิทธิพลต่อการเรียนรู้ของมนุษย์ โดยในบริบททางการศึกษาการให้ข้อมูลป้อนกลับในชั้นเรียนเป็นการรับข้อมูลด้านความสามารถหรือความเข้าใจของผู้เรียนจากผู้มีส่วนได้ส่วนเสีย เช่น ครู เพื่อน หนังสือ ประสบการณ์ หรือตนเอง และการให้ข้อมูลป้อนกลับยังเป็นองค์ประกอบหลักของการประเมินความก้าวหน้าระหว่างเรียน (formative assessment) ซึ่งเป็นปัจจัยที่มีความสำคัญในการเรียนรู้สำหรับผู้เรียน วัตถุประสงค์หลักของการให้ข้อมูลป้อนกลับ คือ การปิดช่องว่างระหว่างความสามารถในปัจจุบันของผู้เรียนกับความสามารถที่ควรจะเป็น (Hattie & Timperley, 2007)

#### *ตอนที่ 4 มโนทัศน์เกี่ยวกับการทดสอบระดับชาติวิชาภาษาอังกฤษ*

การทดสอบระดับชาติ (O-NET) คือ การทดสอบทางการศึกษาระดับชาติขั้นพื้นฐานเป็นการทดสอบเพื่อวัดความรู้และความคิดของนักเรียนชั้นประถมศึกษาปีที่ 6 ชั้นมัธยมศึกษาปีที่ 3 และชั้นมัธยมศึกษาปีที่ 6 ประเมินตามมาตรฐานการเรียนรู้ในหลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พุทธศักราช 2551 โดยการทดสอบประกอบด้วย 5 วิชา ได้แก่ คณิตศาสตร์ วิทยาศาสตร์ ภาษาอังกฤษ ภาษาไทย และสังคมศึกษา ศาสนาและวัฒนธรรม ซึ่งการทดสอบในระดับชั้นมัธยมศึกษาปีที่ 6 เป็นสิ่งที่มีความสำคัญอย่างยิ่ง เนื่องจากนักเรียนจำเป็นต้องมีความรู้พื้นฐานเพื่อนำไปใช้ศึกษาต่อในระดับที่สูงขึ้น โดยวิชาที่มีความจำเป็นต่อการศึกษาค้นคว้าเพิ่มเติมจากบทความหรือหนังสือต่างประเทศ คือ วิชาภาษาอังกฤษ จากสถิติผลการทดสอบย้อนหลังปี 2558-2560 พบว่า คะแนนเฉลี่ยวิชาภาษาอังกฤษอยู่ในเกณฑ์ต่ำ (National Institute of

Educational Testing Service, 2019) ดังนั้น ควรเร่งผลักดันการเรียนรู้อาษาภาษาอังกฤษ เพื่อให้ นักเรียนมีองค์ความรู้ด้านภาษาอังกฤษที่สูงขึ้น

#### *ตอนที่ 5 มโนทัศน์เกี่ยวกับการประเมินผลระบบการทดสอบ*

สิ่งสำคัญในการพัฒนาระบบการทดสอบ คือ การประเมินคุณภาพระบบ ซึ่งมีหลากหลายวิธีการ สำหรับการประเมินคุณภาพระบบ CAT นั้นมีการนำแนวคิดการประเมินมาใช้เรียกว่าการประเมินตาม มาตรฐานการประเมินระบบของ Stufflebeam ประกอบด้วยเนื้อหาหลักในการประเมิน 4 ด้าน ได้แก่ การใช้ ประโยชน์ ความเป็นไปได้ ความเหมาะสม และความถูกต้อง (Kanjanawasee, 2011) และการประเมินจาก ผู้ใช้งานซึ่งการประเมินจากผู้ใช้งานที่นิยมใช้ คือ การประเมินความพึงพอใจระหว่างอินเตอร์เฟสกับผู้ใช้งาน ประกอบด้วยเนื้อหาหลัก 5 ด้าน ได้แก่ ภาพรวมปฏิสัมพันธ์ของระบบ (overall reaction to the software) หน้าจอ (screen) การใช้คำศัพท์และสารสนเทศระบบ (terminology and system information) การเรียนรู้ (learning) และความสามารถของระบบ (system capabilities) โดยในการวิจัยนี้เรียกว่าการประเมินความพึงพอใจของผู้สอบต่อการใช้งานระบบ

### **วิธีดำเนินการวิจัย**

การพัฒนาระบบ MST ที่มีการให้ข้อมูลป้อนกลับในการทดสอบระดับชาติวิชาภาษาอังกฤษสำหรับ นักเรียนระดับชั้นมัธยมศึกษาปีที่ 6 ใช้ระเบียบวิธีวิจัยและพัฒนา (research and development) ใน การศึกษาโดยมีขั้นตอนการวิจัยทั้งสิ้น 3 ระยะ ได้แก่ ระยะที่ 1 การพัฒนาระบบ MST ที่มีการให้ข้อมูล ป้อนกลับในการทดสอบระดับชาติวิชาภาษาอังกฤษสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6 ระยะที่ 2 การ ทดลองใช้ระบบ MST ที่มีการให้ข้อมูลป้อนกลับในการทดสอบระดับชาติวิชาภาษาอังกฤษสำหรับนักเรียน ระดับชั้นมัธยมศึกษาปีที่ 6 และระยะที่ 3 การประเมินผลระบบ MST ที่มีการให้ข้อมูลป้อนกลับในการทดสอบ ระดับชาติวิชาภาษาอังกฤษสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6 โดยมีรายละเอียดแต่ละระยะดังนี้

*ระยะที่ 1 การพัฒนาระบบการทดสอบแบบปรับเหมาะหลายขั้นตอนที่มีการให้ข้อมูลป้อนกลับในการ ทดสอบระดับชาติวิชาภาษาอังกฤษสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6*

สำหรับการพัฒนาระบบ MST ที่มีการให้ข้อมูลป้อนกลับในการทดสอบระดับชาติวิชาภาษาอังกฤษ สำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6 แบ่งการดำเนินงานออกเป็น 3 ตอนได้แก่ 1) ศึกษาองค์ประกอบ ของ MST และรูปแบบการรายงานผลข้อมูลป้อนกลับ 2) ออกแบบและพัฒนาระบบการทดสอบและรูปแบบ การรายงานผลข้อมูลป้อนกลับ และ 3) ตรวจสอบคุณภาพของระบบการทดสอบที่มีการสะท้อนข้อมูล ป้อนกลับ



ระยะที่ 2 การทดลองใช้ระบบการทดสอบแบบปรับเหมาะหลายขั้นตอนที่มีการให้ข้อมูลป้อนกลับในการทดสอบระดับชาติวิชาภาษาอังกฤษสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6

ในการทดลองใช้ระบบ MST ที่มีการให้ข้อมูลป้อนกลับในการทดสอบระดับชาติวิชาภาษาอังกฤษสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6 ใช้ตัวอย่างวิจัยในการทดลองใช้ระบบ จำนวน 53 คน โดยใช้วิธีการเลือกแบบเฉพาะเจาะจง (purposive sampling) จากนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6 ในโรงเรียนที่มีความพร้อมสำหรับการทดสอบด้วยระบบคอมพิวเตอร์แบบออนไลน์และสะดวกในการทดลองใช้ระบบภายในเขตกรุงเทพมหานคร

ระยะที่ 3 การประเมินผลระบบการทดสอบแบบปรับเหมาะหลายขั้นตอนที่มีการให้ข้อมูลป้อนกลับในการทดสอบระดับชาติวิชาภาษาอังกฤษสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6

การประเมินผลระบบการทดสอบแบ่งออกเป็น 3 ตอน ได้แก่ ตอนที่ 1 การตรวจสอบคุณภาพแบบประเมินระบบ ใช้การวิเคราะห์ด้วยค่าดัชนีความสอดคล้องระหว่างข้อคำถามหรือรายการตรวจสอบกับวัตถุประสงค์ (IOC) ซึ่งเป็นการตรวจสอบคุณภาพแบบประเมินโดยอาศัยดุลพินิจจากผู้เชี่ยวชาญด้านการวัดและประเมินผลการศึกษา จำนวน 5 ท่าน ตอนที่ 2 การประเมินระบบแบบอิงมาตรฐาน ใช้การประเมินระบบโดยผู้เชี่ยวชาญด้านการทดสอบ จำนวน 5 ท่าน ก่อนนำระบบไปใช้งานด้วยแบบประเมินแบบอิงมาตรฐานตามแนวคิดของ Stufflebeam และตอนที่ 3 การประเมินความพึงพอใจของผู้สอบต่อการใช้งานระบบ ใช้การประเมินความพึงพอใจของผู้สอบต่อการใช้งานระบบหลังจากจบการทดสอบ

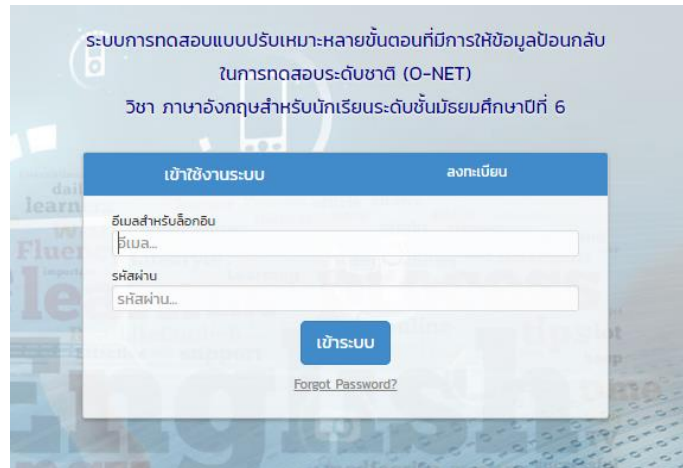
## สรุปผลการวิจัย

สรุปผลการวิจัยมีดังนี้

1. ผลการพัฒนา ระบบ MST ที่มีการให้ข้อมูลป้อนกลับในการทดสอบระดับชาติวิชาภาษาอังกฤษสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6 พบว่า องค์ประกอบของ MST ที่มีการให้ข้อมูลป้อนกลับประกอบด้วย 5 องค์ประกอบ ได้แก่ 1) จุดเริ่มต้นในการทดสอบ 2) การประมาณค่าความสามารถของผู้สอบ 3) การคัดเลือกข้อสอบ 4) จุดยุติการทดสอบ และ 5) รายงานผลข้อมูลป้อนกลับซึ่งประกอบด้วยรายงาน 2 รูปแบบ ได้แก่ 1) รายงานคะแนนสังเกตได้ และ 2) รายงานระดับความสามารถของผู้สอบ สำหรับผลการวิเคราะห์ข้อสอบวิชาภาษาอังกฤษของการทดสอบระดับชาติสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6 พบว่า ข้อสอบ จำนวน 320 ข้อ ซึ่งใช้ในการทดสอบระหว่างปี 2558-2561 เมื่อนำมาวิเคราะห์คุณภาพข้อสอบตามเกณฑ์ คือ ค่าความยาก (b) อยู่ในช่วง -4 ถึง 4 ค่าอำนาจจำแนก (a) อยู่ในช่วง 0 ถึง  $\alpha$  และความน่าจะเป็นในการเดา (c) ไม่เกิน 0.50 ตามทฤษฎีการตอบสนองข้อสอบ พบว่า มีข้อสอบที่ผ่านเกณฑ์ จำนวน 179 ข้อ หลังจากนั้นดำเนินการบรรจุลงในคลังข้อสอบในระบบ MST

ผลการตรวจสอบคุณภาพระบบก่อนนำระบบไปใช้งานจริงโดยผู้เชี่ยวชาญด้วยแบบประเมินระบบแบบอิงมาตรฐาน แบ่งประเด็นในการตรวจสอบออกเป็น 4 ด้าน พบว่า ระบบมีความเหมาะสมทั้ง 4 ด้าน โดย

อันดับหนึ่งได้แก่ ด้านความเป็นไปได้ในการนำไปใช้ ( $M=4.96, SD=0.15$ ) รองลงมาคือ ด้านความถูกต้อง ( $M=4.88, SD=0.18$ ) ด้านความมีประโยชน์ ( $M=4.84, SD=0.22$ ) และด้านความเหมาะสม ( $M=4.83, SD=0.14$ )



ภาพที่ 1 หน้าจอระบบการทดสอบ

2. ผลวิเคราะห์การทดลองใช้ระบบ MST ที่มีการให้ข้อมูลป้อนกลับในการทดสอบระดับชาติวิชาภาษาอังกฤษสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6 ดำเนินการเก็บรวบรวมข้อมูลจากตัวอย่างวิจัยจำนวน 53 คน เป็นนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6 ในเขตกรุงเทพมหานคร จำนวน 53 คน โดยเป็นเพศหญิงและเรียนสายวิทยาศาสตร์ทั้งหมด สำหรับอายุของตัวอย่างในการวิจัย พบว่า มีอายุ 17 ปี จำนวน 14 คน คิดเป็นร้อยละ 26.41 และอายุ 18 ปี จำนวน 39 คิดเป็นร้อยละ 73.53

ข้อมูลที่ได้จากการนำระบบไปทดลองใช้ พบว่า กลุ่มอายุที่ผ่านการทดสอบมากที่สุด คือ กลุ่มอายุ 18 ปี ซึ่งผ่านการทดสอบ จำนวน 2 คน คิดเป็นร้อยละ 3.78 ในส่วนของข้อมูลคะแนนรวมและคะแนนการให้ข้อมูลป้อนกลับรายด้าน พบว่า คะแนนรวมมีค่าเฉลี่ยเท่ากับ 10.58 ( $SD=2.53$ ) ในส่วนคะแนนรายด้านสำหรับการให้ข้อมูลย้อนกลับ พบว่า ความสามารถในการใช้งานมีค่าเฉลี่ยเท่ากับ 5.20 ( $SD=1.92$ ) ความสามารถในการเขียนมีค่าเฉลี่ยเท่ากับ 2.18 ( $SD=1.57$ ) และความสามารถในการอ่านมีค่าเฉลี่ยเท่ากับ 3.18 ( $SD=1.66$ ) และเมื่อพิจารณาค่าเฉลี่ยความสามารถและความคลาดเคลื่อนมาตรฐานที่ได้จากการนำระบบไปทดลองใช้ พบว่า มีค่าความสามารถ ( $\theta$ ) เฉลี่ย เท่ากับ -1.28 และค่าความคลาดเคลื่อนมาตรฐาน (standard error) เฉลี่ย เท่ากับ 0.88

3. ผลการประเมินระบบ MST ที่มีการให้ข้อมูลป้อนกลับในการทดสอบระดับชาติวิชาภาษาอังกฤษสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6 พบว่า แบบประเมินระบบแบบอิงมาตรฐานที่พัฒนาขึ้นจำนวน 20 ข้อ มีค่าดัชนีความสอดคล้องมากกว่า 0.5 มีจำนวน 14 ข้อ ผู้วิจัยจึงได้ดำเนินการปรับแก้ก่อนนำไปใช้งาน สำหรับแบบประเมินความพึงพอใจของผู้สอบกับการใช้งานระบบที่พัฒนาขึ้นมีจำนวน 22 ข้อ มีค่าดัชนีความสอดคล้องมากกว่า 0.5 ทุกข้อ

ผลการประเมินระบบก่อนนำระบบไปใช้จริงด้วยแบบประเมินระบบแบบอิงมาตรฐาน พบว่า ผู้เชี่ยวชาญมีความเห็นว่าระบบมีความเหมาะสมอยู่ในระดับดีมากทั้ง 4 ด้าน โดยอันดับหนึ่งได้แก่ ด้านความเป็นไปได้ในการนำไปใช้ ( $M=4.96, SD=0.15$ ) รองลงมา คือ ด้านความถูกต้อง ( $M=4.88, SD=0.18$ ) ด้านความมีประโยชน์ ( $M=4.84, SD=0.22$ ) และด้านความเหมาะสม ( $M=4.83, SD=0.14$ )

ผลการประเมินความพึงพอใจของผู้สอบกับการใช้งานระบบ พบว่า ความเห็นต่อการใช้งานระบบในภาพรวม ค่อนข้างดี ( $M=5.88, SD=1.97$ ) ด้านหน้าจอกการแสดงผลในประเด็นการใช้งานของเครื่องมือ ต่าง ๆ บนหน้าจอมีค่าสูงสุด หมายความว่า การใช้งานค่อนข้างง่าย ( $M=6.10, SD=2.03$ ) ด้านคำศัพท์และสารสนเทศของระบบในประเด็นข้อความบนหน้าจอที่แสดงแก่ผู้ใช้งานเพื่อป้อนข้อมูลมีค่าสูงสุด หมายความว่า ระบบข้อความบนหน้าจอมีความชัดเจน ( $M=6.17, SD=1.67$ ) ด้านการเรียนรู้ในประเด็นความสามารถในการจดจำข้อมูลและคำสั่งที่เคยป้อนมีค่าสูงสุด หมายความว่า ระบบสามารถจดจำข้อมูลและคำสั่งที่เคยป้อนทำให้ง่ายต่อการใช้งาน ( $M=6.43, SD=2.13$ ) และด้านความสามารถของระบบในประเด็นความเร็วในการตอบสนองต่อการใช้งานระบบมีค่าสูงสุด หมายความว่า การทำงานของระบบสามารถตอบสนองการใช้งานของผู้ใช้ได้อย่างรวดเร็ว ( $M=6.33, SD=2.20$ )

ระบบการทดสอบแบบปรับเหมาะหลายขั้นตอนที่มีการให้ข้อมูลป้อนกลับ ในการทดสอบระดับชาติ (O-NET) วิชา ภาษาอังกฤษสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6			
ชื่อ : ตัวอย่าง โรงเรียน : ตัวอย่าง	นามสกุล : ทดสอบ	สายการเรียน : ศิลป์ จังหวัด : กรุงเทพมหานคร	ระดับชั้น : ม.6
	ข้อสอบ	ตอบถูก	ร้อยละ
<b>Part I: Language Use</b>	<b>25</b>	<b>17</b>	<b>68.00%</b>
Dialog Completion	5	3	60.00%
Situational Dialog	5	3	60.00%
Sentence Completion	5	4	80.00%
Error Correction	10	7	70.00%
<b>Part II: Writing Ability</b>	<b>10</b>	<b>7</b>	<b>70.00%</b>
<b>Part III: Reading Ability</b>	<b>15</b>	<b>11</b>	<b>73.33%</b>
Reading Ability -Vocabulary	5	4	80.00%
Reading Ability -Reading Comprehension	10	7	70.00%
<b>รวม</b>	<b>50</b>	<b>35</b>	<b>70.00%</b>

ค่า Theta ( $\Theta$ )	ค่าคะแนนมาตรฐาน (Scale Score)	ผลการทดสอบ
1.995	69.95	ผ่านการทดสอบ

**ข้อมูลป้อนกลับ**  
นักเรียนมีความสามารถภาษาอังกฤษผ่านเกณฑ์มาตรฐานทุกด้าน (มากกว่า 50%) ได้แก่ การใช้งานภาษา (Language use) ความสามารถในการเขียน (Writing ability) และความสามารถในการอ่านภาษาอังกฤษ (Reading ability)

ภาพที่ 3 รายงานผลการทดสอบและข้อมูลป้อนกลับ



ภาพที่ 2 แผนภาพแสดงผลการทดสอบและข้อมูลป้อนกลับ

## อภิปรายผล

จากผลการวิจัยสามารถนำมาอภิปรายได้ 4 ประเด็น ได้แก่ 1) การพัฒนาระบบ MST 2) การทดลองใช้ระบบ MST 3) รายงานการสะท้อนข้อมูลป้อนกลับ และ 4) การประเมินระบบ MST รายละเอียดมีดังนี้

1. CAT มีแนวคิดที่สำคัญ คือ ผู้สอบจะเริ่มทำข้อสอบจากข้อที่มีระดับความยากปานกลาง หากผู้สอบตอบข้อสอบผิดในข้อแรกข้อถัดไปจะง่ายลง ในทางตรงข้ามหากผู้สอบตอบข้อสอบถูกข้อถัดไปจะยากขึ้น ผู้สอบจะดำเนินการสอบไปเรื่อย ๆ กระทั่งระดับความสามารถของผู้สอบคงที่หรืออยู่ในเกณฑ์ความคลาดเคลื่อนที่ยอมรับได้ การทดสอบจะยุติลง (Kanjanawasee, 2012) ต่อมา พบว่า CAT มีการประมาณค่าความสามารถที่สูงและต่ำกว่าจริง จึงได้มีการพัฒนา MST ขึ้น โดยผลจากการศึกษาของ Lord และ Novick พบว่า MST สามารถประมาณค่าความสามารถได้แม่นยำกว่า CAT (Zenisky, Hambleton, & Luecht, 2010) และสำหรับผลจากการวิจัยครั้งนี้ พบว่า ระบบ MST ที่มีการให้ข้อมูลป้อนกลับในการทดสอบระดับชาติวิชาภาษาอังกฤษสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6 มีองค์ประกอบที่สำคัญ ได้แก่ 1) จุดเริ่มต้นในการทดสอบ 2) การประมาณค่าความสามารถของผู้สอบ 3) การคัดเลือกข้อสอบ 4) จุดยุติการทดสอบ และ 5) รายงานผลข้อมูลป้อนกลับซึ่งสอดคล้องกับ Chaimongkol, Pasiphol, and Kanjanawasee (2016)

2. การนำระบบ MST ที่มีการให้ข้อมูลป้อนกลับในการทดสอบระดับชาติวิชาภาษาอังกฤษสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6 ไปทดลองใช้กับตัวอย่างวิจัย จำนวน 53 คน พบว่า มีผู้สอบผ่านเพียง 2 คน อาจเนื่องมาจากตัวอย่างวิจัยไม่ได้เตรียมความพร้อมก่อนการสอบ เนื่องจากผลการสอบไม่มีผลได้ผลเสีย (high stake) กับตัวอย่างวิจัย ทำให้ตัวอย่างวิจัยอาจไม่ได้ใช้ความพยายามเต็มที่ในการทำข้อสอบ ซึ่งสอดคล้องกับแนวคิดของ Beardsley and Berliner (2003) และ Mesler (2008) ว่าข้อสอบที่มีผลได้ผลเสียกับผู้สอบจะทำให้ผู้สอบมีความพยายามในการทำข้อสอบมากกว่าปกติ

3. การให้ข้อมูลป้อนกลับเป็นองค์ประกอบที่สำคัญในการบวนการจัดการเรียนการสอนและเป็นส่วนหนึ่งในการประเมินความก้าวหน้าระหว่างเรียน (formative assessment) เพื่อระบุให้ผู้เรียนทราบ

จุดเด่นและข้อจำกัด และนำข้อจำกัดดังกล่าวไปพัฒนาตนเองจนนำไปสู่การบรรลุผลสัมฤทธิ์ตามศักยภาพของผู้เรียน ในการวิจัยครั้งนี้ใช้การให้ข้อมูลป้อนกลับแบบทั่วไป โดยข้อมูลในส่วนของรายงานผลการทดสอบ จะแสดงจำนวนข้อสอบทั้งหมดที่ตัวอย่างวิจัยทำการสอบ พร้อมแสดงคะแนนผลการตอบข้อสอบทั้งหมด หลังจากนั้นปรับข้อมูลให้เป็นร้อยละเพื่อนำไปใช้แสดงผลการผ่านการทดสอบรายด้าน โดยหากมีคะแนนมากกว่าร้อยละ 50 ถือว่าผ่านการทดสอบรายด้าน สำหรับเงื่อนไขการผ่านการทดสอบจะพิจารณาจากจากคะแนนในมาตร (scale score) ซึ่งมีการปรับเทียบให้เป็นคะแนนมาตรฐานที่ (standardized t-score) โดยอิงจากค่าความสามารถผู้สอบ ( $\Theta$ ) หากค่าคะแนนในมาตรมากกว่าร้อยละ 50 ถือว่าผู้สอบผ่านการทดสอบ จากข้อมูลข้างต้นเห็นได้ว่าการให้ข้อมูลป้อนกลับเป็นสิ่งจำเป็นต่อการพัฒนาผู้เรียน ดังนั้น การพัฒนาระบบการทดสอบที่มีการให้ข้อมูลป้อนกลับในครั้งนี้จึงมีความสอดคล้องและเป็นไปตามแนวโน้มการเปลี่ยนแปลงของโลก

4. ประเด็นสำคัญในการพัฒนาระบบ คือ การตรวจสอบคุณภาพ โดยการวิจัยครั้งนี้ได้นำวิธีการประเมิน จำนวน 2 วิธี มาใช้ในการตรวจสอบคุณภาพระบบ วิธีแรกใช้การประเมินคุณภาพระบบก่อนนำระบบไปใช้งานจริง ซึ่งใช้การประเมินระบบแบบอิงมาตรฐานตามแนวคิดของ Stufflebeam โดยครอบคลุม 4 ประเด็น ได้แก่ ด้านความมีประโยชน์ ด้านความเป็นไปได้ในการนำไปใช้ ด้านความเหมาะสม และด้านความถูกต้อง (Kanjanawasee, 2011) ซึ่งจะช่วยประเมินเบื้องต้นได้ว่าระบบที่พัฒนาขึ้นสามารถนำไปใช้งานได้จริงหรือไม่ และในการนำระบบไปทดลองจะประเมินคุณภาพด้วยวิธีที่สอง คือ การประเมินความพึงพอใจของผู้สอบต่อการใช้งานระบบซึ่งพัฒนามาจากแบบประเมินความพึงพอใจผู้ใช้งานของอินเตอร์เฟซระหว่างผู้ใช้งานกับคอมพิวเตอร์ (user satisfaction of the human-computer interface) โดยเป้าหมายหลักของการประเมิน คือ ต้องการทราบว่าระบบทดสอบที่พัฒนาขึ้นสามารถตอบสนองการใช้งานของผู้ใช้งานได้มากน้อยเพียงใด เนื่องจากผลการประเมินที่ได้จะนำไปใช้เป็นแนวทางในการพัฒนาระบบที่ตอบสนองการใช้งานของผู้สอบในอนาคตต่อไป ซึ่งจากข้อมูลข้างต้นจึงเชื่อมั่นได้ว่าระบบ MST ที่มีการให้ข้อมูลป้อนกลับในการทดสอบระดับชาติวิชาภาษาอังกฤษสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6 มีคุณภาพ ได้มาตรฐาน และสามารถนำไปใช้งานได้จริง

## ข้อเสนอแนะ

### ข้อเสนอแนะเพื่อการวิจัยครั้งต่อไป

1. การวิจัยครั้งนี้มีข้อสอบที่ผ่านเกณฑ์เพียง 179 ข้อ ซึ่งข้อสอบส่วนใหญ่เป็นข้อสอบที่มีการจัดเป็นชุด (testlets) จึงทำให้โอกาสที่ผู้สอบจะพบข้อสอบซ้ำมีอัตราสูง ด้วยเหตุนี้ งานวิจัยในอนาคตจึงควรเพิ่มจำนวนข้อสอบบรรจุลงในคลังข้อสอบ เพื่อให้เนื้อหาในการทดสอบแบบปรับเหมาะมีระดับความยากที่หลากหลายและครอบคลุมเนื้อหาการทดสอบระดับชาติวิชาภาษาอังกฤษระดับชั้นมัธยมศึกษาปีที่ 6 มากยิ่งขึ้น

2. การวิจัยครั้งนี้ยังไม่ได้ดำเนินการเปรียบเทียบคะแนนสอบของแบบทดสอบแต่ละฉบับเนื่องจากข้อจำกัดของจำนวนข้อสอบที่ผ่านเกณฑ์ ดังนั้น งานวิจัยครั้งต่อไปควรมีการปรับเทียบคะแนนของแบบทดสอบแต่ละปีเพื่อให้คลังข้อสอบของระบบการทดสอบแบบปรับเหมาะมีความเท่าเทียมกันมากยิ่งขึ้น

3. การวิจัยครั้งนี้มีการจัดชุดข้อสอบซึ่งอาจมีสัดส่วนที่แตกต่างจากการทดสอบระดับชาติเนื่องมาจากข้อจำกัดของจำนวนข้อสอบที่ผ่านเกณฑ์ ดังนั้น ในการวิจัยครั้งต่อไปควรมีการจัดชุดข้อสอบที่มีสัดส่วนใกล้เคียงกับข้อสอบระดับชาติวิชาภาษาอังกฤษระดับชั้นมัธยมศึกษาปีที่ 6 ให้มากยิ่งขึ้น

4. ปัจจุบันการวินิจฉัยผู้เรียน (diagnosis) เป็นประเด็นที่เข้ามามีบทบาทสำคัญในการพัฒนาผู้เรียน ดังนั้น ในการวิจัยครั้งต่อไปควรพัฒนาระบบการทดสอบแบบปรับเหมาะร่วมกับการประเมินเพื่อวินิจฉัยทางพุทธิปัญญา (cognitive diagnosis) เพื่อให้ผู้เรียนได้รับข้อสอบสำหรับการวินิจฉัยที่ตรงกับระดับความสามารถของผู้เรียน และหลังจากการทดสอบเสร็จสิ้นสามารถให้ข้อมูลป้อนกลับไปยังผู้เรียนเพื่อนำไปใช้ในการพัฒนาตนเองได้ทันที

### กิตติกรรมประกาศ

งานวิจัยนี้ได้รับเงินอุดหนุนการวิจัยจากสถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) ประจำปีงบประมาณ พ.ศ. 2562

### References

- Beardsley, A., & Berliner, D. (2003). The effects of high-stakes testing on student motivation and learning. *Educational Leadership*, 60(5), 32-38.
- Chang, H.-H. (2004). Understanding Computerized Adaptive Testing: From Robbins-Monro to Lord and Beyond. In D. Kaplan (Ed.), *The SAGE Handbook of Quantitative Methodology for the Social Sciences* (pp. 117). California: Sage Publication, Inc.
- Chin, J. P., Diehl, V. A., & Norman, K. L. (1988). *Development of An Instrument Measuring User Satisfaction of the Human-Computer Interface*. Paper presented at the SIGCHI Conference on Human Factors in Computing Systems.
- Chaimongkol, N., Pasiphol, S., & Kanjanawasee, S. (2016). Computerized Adaptive Testing with Reflective Feedback: A Conceptual Framework. *Procedia - Social and Behavioral Sciences*, 217(2016), 806 – 812.
- Han, K. T., & Guo, F. (2013). *An Approach to Assembling Optimal Multistage Testing Modules on the Fly*. Graduate Management Admission Council.
- Hattie, J., & Timperley, H. (2007). The Power of Feedback. *Review of Educational Research*, 77(1), 81-112.

- Kanjanawasee, S. (2011). *Evaluation Theory* (7<sup>th</sup> ed.). Bangkok: Chulalongkorn University Press.
- Kanjanawasee, S. (2012). *Modern Test Theory* (4<sup>th</sup> ed.). Bangkok: Chulalongkorn University Press.
- Lilley, M., Barker, T., & Britton, C. (2004). The development and evaluation of a software prototype for computer-adaptive testing. *Computers & Education*, 43(2004), 109-123.
- Mesler, A. (2008). *High Stakes and Higher Impacts: The Effects of High-Stakes Testing on School and Classroom Operations*. Duke University, Durham, NC. Retrieved from [http://uwp.duke.edu/uploads/assets/Mesler\\_09.pdf](http://uwp.duke.edu/uploads/assets/Mesler_09.pdf).
- National Institute of Educational Testing Service (Public Organization). (2019). *O-NET score statistics in Grade 12 (2014-2017)*. Retrieved from <https://www.niets.or.th/uploads/editor/files/Download/ช่วงคะแนน%20ม.6.pdf>
- Patsula, L. N. (1999). *A comparison of computerized adaptive testing and multi-stage testing*. (Doctor of Philosophy), University of Massachusetts Amherst. Retrieved from [https://scholarworks.umass.edu/dissertations\\_1/3282](https://scholarworks.umass.edu/dissertations_1/3282)
- Wang, K. (2017). *A fair comparison of the performance of computerized adaptive testing and multistage adaptive testing*. (Doctor of Philosophy), Michigan State University.
- Zenisky, A., Hambleton, R. K., & Luecht, R. M. (2010). Multistage Testing: Issues, Designs, and Research. In W. J. v. d. Linden & C. A. W. Glas (Eds.), *Elements of Adaptive Testing*: Springer.
- Zheng, Y., & Chang, H.-H. (2015). On-the-Fly Assembled Multistage Adaptive Testing. *Applied Psychological Measurement*, 39(2), 1-15.
- Zheng, Y., Nozawa, Y., Gao, X., & Chang, H.-H. (2012). *Multistage Adaptive Testing for a Large-Scale Classification Test: Design, Heuristic Assembly, and Comparison with Other Testing Modes*: ACT.



Nhabhat Chaimongkol, Ph.D., Lecturer, Department of Educational Research and Psychology, Faculty of Education, Chulalongkorn University.