



รายงานวิจัย

การพัฒนาระบบการทดสอบแบบปรับเหมาะหลายขั้นตอนที่มีการให้ข้อมูล
ป้อนกลับในการทดสอบระดับชาติ (O-NET) วิชาภาษาอังกฤษสำหรับนักเรียน
ระดับชั้นมัธยมศึกษาปีที่ 6

System Development for Multistage Adaptive Testing with
Feedback in O-NET for The Subject of Grade 12
English Language

โดย

อาจารย์ ดร.ณภัทร ชัยมงคล

งานวิจัยนี้ได้รับเงินอุดหนุนการวิจัยจากสถาบันทดสอบทางการศึกษาแห่งชาติ
(องค์การมหาชน)
ประจำปีงบประมาณ พ.ศ. 2562

สารบัญ

หน้า

สารบัญ	๒
สารบัญภาพ	๓
สารบัญตาราง	๓
บทคัดย่อ	๓
Abstract	๒
กิตติกรรมประกาศ	๒
บทที่ 1	1
บทนำ.....	1
ความเป็นมาและความสำคัญของปัญหา.....	1
วัตถุประสงค์ของการวิจัย	5
ขอบเขตของการวิจัย.....	5
คำจำกัดความที่ใช้ในการวิจัย	6
ประโยชน์ที่ได้รับ	7
บทที่ 2.....	8
เอกสารและงานวิจัยที่เกี่ยวข้อง	8
ตอนที่ 1 มโนทัศน์เกี่ยวกับการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์.....	8
ตอนที่ 2 มโนทัศน์เกี่ยวกับการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์แบบหลายขั้นตอน ...	17
ตอนที่ 3 มโนทัศน์เกี่ยวกับการให้ข้อมูลป้อนกลับ	22
ตอนที่ 4 มโนทัศน์เกี่ยวกับการทดสอบระดับชาติ (O-NET) วิชาภาษาอังกฤษ.....	27
ตอนที่ 5 มโนทัศน์เกี่ยวกับการประเมินผลระบบการทดสอบ	28
ตอนที่ 6 กรอบแนวคิดการวิจัย.....	30
บทที่ 3	32
วิธีดำเนินการวิจัย.....	32
ระยะที่ 1 การพัฒนาระบบการทดสอบแบบปรับเหมาะหลายขั้นตอนที่มีการให้ข้อมูลป้อนกลับใน การทดสอบระดับชาติ (O-NET) วิชาภาษาอังกฤษสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6....	32
ระยะที่ 2 การทดลองใช้ระบบการทดสอบแบบปรับเหมาะหลายขั้นตอนที่มีการให้ข้อมูลป้อนกลับ ในการทดสอบระดับชาติ (O-NET) วิชาภาษาอังกฤษสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6	34
ระยะที่ 3 การประเมินผลระบบการทดสอบแบบปรับเหมาะหลายขั้นตอนที่มีการให้ข้อมูล ป้อนกลับในการทดสอบระดับชาติ (O-NET) วิชาภาษาอังกฤษสำหรับนักเรียนระดับชั้น มัธยมศึกษาปีที่ 6	35

บทที่ 4	37
ผลการวิเคราะห์ข้อมูล	37
ระยะที่ 1 ผลการพัฒนากระบวนการทดสอบแบบปรับเหมาะหลายขั้นตอนที่มีการให้ข้อมูลป้อนกลับ ในการทดสอบระดับชาติ (O-NET) วิชาภาษาอังกฤษสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6	37
ระยะที่ 2 ผลการทดลองใช้ระบบการทดสอบแบบปรับเหมาะหลายขั้นตอนที่มีการให้ข้อมูล ป้อนกลับในการทดสอบระดับชาติ (O-NET) วิชาภาษาอังกฤษสำหรับนักเรียนระดับชั้น มัธยมศึกษาปีที่ 6	48
ระยะที่ 3 ผลการประเมินระบบการทดสอบแบบปรับเหมาะหลายขั้นตอนที่มีการให้ข้อมูล ป้อนกลับในการทดสอบระดับชาติ (O-NET) วิชาภาษาอังกฤษสำหรับนักเรียนระดับชั้น มัธยมศึกษาปีที่ 6	50
บทที่ 5	57
สรุปผลการวิจัย อภิปรายผล และข้อเสนอแนะ	57
สรุปผลการวิจัย	58
อภิปรายผลการวิจัย	61
ข้อเสนอแนะในการวิจัยครั้งต่อไป.....	63
บรรณานุกรม.....	64
ภาคผนวก	68
ประวัติผู้วิจัย	107

สารบัญญภาพ

	หน้า
ภาพที่ 1 ขั้นตอนการดำเนินการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์	12
ภาพที่ 2 แผนภาพองค์ประกอบและเส้นทางการทำงานของ	19
ภาพที่ 3 รูปแบบโมเดล 1-2-3-4 ของการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์แบบหลาย ขั้นตอน	20
ภาพที่ 4 กรอบแนวคิดการให้ข้อมูลป้อนกลับ (Hattie & Timperley, 2007)	23
ภาพที่ 5 กรอบแนวคิดในการวิจัย	32
ภาพที่ 6 หน้าจอหลักระบบการทดสอบ	41
ภาพที่ 7 หน้าจอการลงทะเบียน	41
ภาพที่ 8 การแสดงสถานะเมื่อกดปุ่มลงทะเบียน	42
ภาพที่ 9 การเข้าสู่ระบบการทดสอบ	42
ภาพที่ 10 การแก้ไขข้อมูลส่วนตัว	43
ภาพที่ 11 หน้าจอคำชี้แจง	43
ภาพที่ 12 หน้าจอการทดสอบ	44
ภาพที่ 13 หน้าจอการกดเลือกคำตอบ	44
ภาพที่ 14 หน้าจอเมื่อตอบคำถามครบในขั้นนั้น ๆ	45
ภาพที่ 15 หน้าจอเมื่อตอบคำถามครบทุกขั้น	45
ภาพที่ 16 หน้าจอการประเมินความพึงพอใจต่อการใช้งานระบบ	46
ภาพที่ 17 หน้าจอเมื่อประเมินความพึงพอใจต่อการใช้งานระบบครบทุกข้อ	46
ภาพที่ 18 รายงานผลการทดสอบและข้อมูลป้อนกลับ	48

สารบัญตาราง

	หน้า
ตารางที่ 1 แนวทางในการให้ข้อมูลป้อนกลับที่พึงปฏิบัติและพึงหลีกเลี่ยง	25
ตารางที่ 2 รายงานผลการทดสอบระดับชาติระหว่างปี พ.ศ. 2558-2560.....	27
ตารางที่ 3 รูปแบบรายงานคะแนนจากการสอบ	39
ตารางที่ 4 รูปแบบรายงานระดับความสามารถของผู้สอบ	40
ตารางที่ 5 ข้อมูลอายุของตัวอย่างในการวิจัย	48
ตารางที่ 6 ผลการทดสอบจากระบบจำแนกตามอายุ.....	49
ตารางที่ 7 คะแนนรวมและคะแนนการให้ข้อมูลป้อนกลับรายด้าน	50
ตารางที่ 8 ค่าเฉลี่ยความสามารถและความคลาดเคลื่อนมาตรฐาน.....	50
ตารางที่ 9 ผลการประเมินระบบด้วยแบบประเมินระบบแบบอิงมาตรฐาน	52
ตารางที่ 10 ผลการประเมินความพึงพอใจของผู้สอบกับการใช้งานระบบ	54

บทคัดย่อ

การทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์เป็นการทดสอบที่นำกระบวนการทางคอมพิวเตอร์เข้ามาช่วยในการคัดเลือกข้อสอบที่เหมาะสมกับระดับความสามารถของผู้สอบแต่ละคน ซึ่งการทดสอบแบบปรับเหมาะแบบดั้งเดิมมีข้อจำกัดในประเด็นการประมาณค่าความสามารถที่ต่ำหรือสูงกว่าความเป็นจริงจากข้อสอบข้อเดียว จึงได้มีการพัฒนาระบบการทดสอบแบบปรับเหมาะหลายขั้นตอนขึ้นเพื่อช่วยลดข้อจำกัดดังกล่าว สำหรับงานวิจัยนี้เป็นพัฒนาระบบการทดสอบแบบปรับเหมาะหลายขั้นตอนที่นำมาประยุกต์ใช้กับการเตรียมความพร้อมในการสอบระดับชาติวิชาภาษาอังกฤษ เนื่องจากผลการทดสอบระดับชาติวิชาภาษาอังกฤษที่ผ่านมาของประเทศไทย ผู้เรียนส่วนใหญ่ยังมีคะแนนยังไม่ผ่านเกณฑ์มาตรฐาน ผลการวิจัย พบว่า องค์ประกอบของการทดสอบแบบปรับเหมาะหลายขั้นตอนมีองค์ประกอบที่สำคัญ 5 องค์ประกอบ ได้แก่ 1) จุดเริ่มต้นในการทดสอบ 2) การประมาณค่าความสามารถของผู้สอบ 3) การคัดเลือกข้อสอบ 4) การยุติการทดสอบ และ 5) รายงานการให้ข้อมูลป้อนกลับซึ่งประกอบด้วยรายงานคะแนนสังเกตได้และรายงานระดับความสามารถของผู้สอบ สำหรับผลการตรวจสอบคุณภาพระบบโดยผู้เชี่ยวชาญด้วยแบบประเมินระบบแบบอิงมาตรฐาน พบว่า ระบบมีคุณภาพอยู่ในระดับมากที่สุดทุกด้าน และผลการประเมินความพึงพอใจของผู้สอบต่อการใช้งานระบบ พบว่า ความเห็นต่อการใช้งานระบบในภาพรวมค่อนข้างดี

คำสำคัญ: ระบบการทดสอบแบบปรับเหมาะหลายขั้นตอน การทดสอบระดับชาติ วิชาภาษาอังกฤษ

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

กิตติกรรมประกาศ

ผู้วิจัยขอขอบคุณสถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) ที่มอบทุนวิจัยประจำปี พ.ศ.2562 ในครั้งนี้ ขอขอบคุณผู้ทรงคุณวุฒิทุกท่านที่เสียสละเวลาในการประเมินคุณภาพแบบ ประเมินและประเมินคุณภาพระบบพร้อมทั้งให้คำแนะนำเพื่อนำไปพัฒนางานวิจัยให้มีคุณภาพมากยิ่งขึ้น ขอขอบคุณโรงเรียนสายปัญญา ในพระบรมราชินูปถัมภ์ ที่ให้ความอนุเคราะห์ในการเก็บข้อมูล และขอขอบคุณตัวอย่างวิจัยทุกท่านที่เสียสละเวลาในการทดลองใช้งานระบบสำหรับการวิจัยครั้งนี้

ณภัทร ชัยมงคล

ผู้วิจัย

บทที่ 1

บทนำ

ความเป็นมาและความสำคัญของปัญหา

การทดสอบแบบดั้งเดิมจะใช้แบบทดสอบฉบับเดียวกับผู้สอบทุกคน ซึ่งอาจส่งผลกระทบต่อความถูกต้องแม่นยำในการวัด เนื่องจากแบบทดสอบฉบับหนึ่ง ประกอบด้วยข้อสอบหลายข้อที่มีระดับความยากแตกต่างกัน ส่งผลให้ผู้สอบบางคนได้ทำข้อสอบที่ยากหรือง่ายเกินไป ดังนั้น การทดสอบในอุดมคติจึงมีแนวคิดที่ว่า ข้อสอบแต่ละข้อควรมีความยากที่เหมาะสมกับความสามารถของแต่ละบุคคล และเรียกการทดสอบนั้นว่า การทดสอบแบบปรับเหมาะ (adaptive testing or tailored testing)

แนวคิดเรื่องการทดสอบแบบปรับเหมาะ เริ่มนำมาใช้ในการทดสอบเขาวัวปัญญาด้วยแบบสอบของบินเน็ต (Binet) โดยผู้สอบจะได้ทำข้อสอบครั้งละหนึ่งข้อ โดยเริ่มจากข้อแรก หากข้อสอบข้อแรกง่ายเกินไป ข้อถัดไปก็จะง่ายลงเพื่อให้เหมาะสมกับระดับเขาวัวปัญญาของผู้สอบ และเมื่อทำข้อสอบไปเรื่อย ๆ จนถึงข้อที่ยากมากขึ้นจนผู้สอบไม่สามารถทำข้อสอบได้ถูกเป็นจำนวนหนึ่ง จึงยุติการทดสอบ (ศิริชัย กาญจนวาสี, 2555)

ต่อมา เมื่อเทคโนโลยีคอมพิวเตอร์เริ่มเข้ามามีบทบาทเพิ่มมากขึ้น David Weiss จึงได้นำคอมพิวเตอร์เมนเฟรม (mainframe computer) มาใช้กับการทดสอบแบบปรับเหมาะ และเรียกว่า การทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ (computerized adaptive testing: CAT) ซึ่งเป็นที่นิยมใช้ในการประเมินผลทางการศึกษาในประเทศสหรัฐอเมริกา เช่น การทดสอบเพื่อศึกษาต่อระดับบัณฑิตศึกษา (Graduate Record Examination: GRE) (Chang, 2004; Linden, 2008)

การทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ (CAT) เป็นวิธีการทดสอบที่ผสมผสานทฤษฎีการวัดและเทคโนโลยีเข้าด้วยกันเพื่อให้การวัดมีความแม่นยำมากยิ่งขึ้น และมีองค์ประกอบที่สำคัญ 3 องค์ประกอบ ได้แก่ (1) จุดเริ่มต้นในการทดสอบ (2) กระบวนการคัดเลือกข้อสอบและการประมาณค่าความสามารถ และ (3) เกณฑ์ที่ใช้ในการยุติการทดสอบ ซึ่งหลักการสำคัญของการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์คือกระบวนการคัดเลือกข้อสอบ (item selection) ที่แตกต่างกันไปตามวิธีการออกแบบการทดสอบ โดยอาศัยคุณลักษณะของข้อสอบตามทฤษฎีการตอบสนองข้อสอบ (item response theory) ได้แก่ ค่าความยาก ค่าอำนาจจำแนก และค่าความน่าจะเป็นในการเดา (Chang, 2004, 2015; Phankokkrud, 2012; ศิริชัย กาญจนวาสี, 2555)

แม้ว่าการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์จะมีข้อดีว่าการทดสอบแบบดั้งเดิมอยู่หลายประการ แต่ยังคงพบประเด็นปัญหา คือ ในการทดสอบ GRE และ GMAT ปี 2000 และ 2003 พบว่า คะแนนของผู้สอบกว่าพันคนขาดความน่าเชื่อถือ เพราะมีการประมาณค่าคะแนนความสามารถของผู้สอบต่ำกว่าและสูงกว่าความเป็นจริง เนื่องจากกระบวนการของการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์อยู่บนพื้นฐานของวิธีการให้สารสนเทศสูงสุด (maximum information) ที่ค่อนข้าง

ซับซ้อน และหากผู้สอบตอบข้อสอบได้ถูกโดยบังเอิญ หรือหากผู้สอบประมาณคำตอบข้อสอบผิดในข้อสอบข้อแรก อาจทำให้การประมาณค่าความสามารถของผู้สอบเกิดความคลาดเคลื่อน ส่งผลให้การเลือกข้อสอบในขั้นถัดไปคลาดเคลื่อนไปด้วย และผู้สอบอาจไม่สามารถกลับไปสู่ข้อสอบที่อยู่ในระดับที่ตรงกับความสามารถของตนเองได้ ดังนั้น จึงได้มีการพัฒนากระบวนการทดสอบที่เรียกว่า การทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์แบบหลายขั้นตอน (multistage adaptive testing: MST) ซึ่งจะช่วยทำให้การประมาณค่าความสามารถของผู้สอบในขั้นตอนแรกเกิดความแม่นยำมากขึ้น เพราะประมาณค่าความสามารถของผู้สอบจากการตอบชุดข้อสอบ ไม่ใช่เพียงข้อสอบเพียงข้อเดียว จึงช่วยลดปัญหาในการประมาณค่าความสามารถต่ำกว่าหรือสูงกว่าความเป็นจริง (Chang, 2015; Zheng & Chang, 2015)

การทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์แบบหลายขั้นตอน (MST) ถูกพัฒนาขึ้นเพื่อลดความคลาดเคลื่อนจากการประมาณค่าคะแนนที่ต่ำกว่าหรือสูงกว่าความเป็นจริงจากการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์แบบดั้งเดิม (CAT) โดยการทดสอบแบบ MST ใช้วิธีการแบ่งข้อสอบออกเป็นขั้น (stage) แต่ละขั้นประกอบด้วยโมดูล (module) แต่ละโมดูลมีข้อสอบที่จัดแบ่งไว้ตามระดับความยาก ในการทดสอบผู้สอบจะเริ่มทำการทดสอบในขั้นแรกที่มีระดับความยากปานกลาง หลังจากทำการทดสอบขั้นแรกเสร็จสิ้น ระบบจะทำการประมาณค่าความสามารถของผู้สอบจากผลการตอบข้อสอบ แล้วจึงทำการเลือกข้อสอบในขั้นที่สองตามความสามารถของผู้สอบจากผลการตอบข้อสอบในขั้นแรก เมื่อผู้สอบทำขั้นที่สองเสร็จสิ้น ระบบจะทำการประมาณค่าความสามารถของผู้สอบและทำการเลือกข้อสอบในขั้นถัดไปตามความสามารถของผู้สอบ และทำกระบวนการเช่นนี้จนกระทั่งผลการประมาณค่าความสามารถของผู้สอบมีความคงเส้นคงวาหรือเป็นไปตามเกณฑ์ในการยุติการทดสอบ จึงหยุดการทดสอบ วิธีการทดสอบแบบ MST นี้ นอกจากจะลดความคลาดเคลื่อนในการประมาณค่าความสามารถของผู้สอบในขั้นตอนแรกแล้ว การทดสอบแบบ MST ยังช่วยลดความเครียดของผู้สอบ เนื่องจากผู้สอบสามารถกลับไปแก้ไขคำตอบในขั้นที่กำลังทดสอบอยู่ได้ ซึ่งแตกต่างจากการทดสอบแบบ CAT ที่ผู้สอบไม่สามารถกลับไปแก้ไขคำตอบได้ อย่างไรก็ตาม การแก้ไขคำตอบในการทดสอบแบบ MST นี้ สามารถทำได้เฉพาะในขั้นที่กำลังดำเนินการทดสอบเท่านั้น ไม่สามารถกลับไปแก้ไขคำตอบในขั้นอื่นได้ (Han & Guo, 2013; Zheng & Chang, 2015; Zheng, Nozawa, Gao, & Chang, 2012)

การทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์แบบหลายขั้นตอน (MST) ยังมีข้อดีอีกประการหนึ่งคือ สามารถรายงานผลการทดสอบตามระดับความสามารถของผู้สอบได้ทันทีหลังจากการทดสอบเสร็จสิ้น แต่ในปัจจุบัน ยังไม่ค่อยพบการนำข้อมูลป้อนกลับ (feedback) มาประยุกต์ใช้กับการรายงานผลจากการทดสอบแบบ MST ทั้ง ๆ ที่ข้อมูลป้อนกลับถือเป็นสารสนเทศที่มีประโยชน์ต่อการพัฒนาตัวผู้รับการทดสอบเป็นอย่างมาก

ในบริบททางการศึกษา โดยทั่วไปข้อมูลป้อนกลับจะเกี่ยวข้องกับการพัฒนาความรู้และทักษะของผู้เรียน แต่หากกล่าวถึงผลสัมฤทธิ์ (achievement) ข้อมูลป้อนกลับจะหมายถึงองค์ประกอบที่มีนัยสำคัญในการสร้างแรงจูงใจในการเรียนรู้ (Shute, 2007) ซึ่งจากการศึกษาผลการพัฒนารูปแบบแฟ้มสะสมงานอิเล็กทรอนิกส์ที่สะท้อนข้อมูลป้อนกลับในการประเมินพัฒนาการการเรียนรู้ของนักเรียนในสังกัดสำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน พบว่าการใช้แฟ้มสะสมงานอิเล็กทรอนิกส์ที่ให้การสะท้อนข้อมูลป้อนกลับเพื่อการปรับปรุง ทำให้มีจำนวนนักเรียนที่มีพัฒนาการมากกว่านักเรียนที่ไม่มีพัฒนาการ (โชติกา ภาชีผล, ประกอบ กรณีกิจ, & พิทักษ์ โสตถยาคม, 2558) สะท้อนให้เห็นว่าการสะท้อนข้อมูลป้อนกลับที่มีสารสนเทศที่ทำให้เกิดการปรับปรุงหรือพัฒนา มีประโยชน์มากกว่าการให้ข้อมูลป้อนกลับเป็นคะแนนการทดสอบเพียงอย่างเดียว ยิ่งไปกว่านั้น หากนำแนวคิดดังกล่าวมาใช้ร่วมกับการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์แบบหลายขั้นตอน (MST) จะช่วยทำให้การรายงานผลและการให้ข้อมูลป้อนกลับมีความสะดวก และรวดเร็ว อีกทั้งผู้สอบยังสามารถนำสารสนเทศที่ได้จากข้อมูลป้อนกลับไปใช้ในการพัฒนาตนเองได้ทันที

ดังที่ได้กล่าวไปข้างต้นว่า การทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์แบบหลายขั้นตอน (MST) มีข้อได้เปรียบที่เหนือกว่าการทดสอบโดยใช้กระดาษ (paper-and-pencil test: P&P) และการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ (CAT) อยู่หลายประการ อีกทั้งผู้พัฒนาระบบการทดสอบแบบ MST ยังอาจสามารถต่อยอดพัฒนาให้ระบบรายงานผลแบบสะท้อนข้อมูลป้อนกลับให้แก่ผู้สอบทันที ซึ่งเป็นประโยชน์ต่อการพัฒนาผู้สอบอย่างมาก ด้วยเหตุนี้ การพัฒนาระบบคอมพิวเตอร์เพื่อการทดสอบแบบ MST นี้ จึงเป็นสิ่งสำคัญที่จะต้องดำเนินการอย่างรอบคอบ และจำเป็นจะต้องมีการประเมินการใช้งานของระบบ ซึ่งมีวิธีการและแนวทางในการประเมินแตกต่างกันออกไป

สำหรับการประเมินระบบหรือโปรแกรมคอมพิวเตอร์การทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์แบบหลายขั้นตอน (MST) นี้ สามารถทำได้โดยประยุกต์ใช้แนวคิดการประเมินคุณลักษณะของระบบจากผู้ใช้ (user evaluation) คือ การประเมินความพึงพอใจระหว่างอินเตอร์เฟซกับผู้ใช้ (user satisfaction of the human-computer interface) ซึ่งในการประเมินคุณลักษณะของระบบด้วยแนวคิดนี้ หากเป็นบริบทที่เกี่ยวข้องกับการศึกษา อาจจะใช้ผู้เชี่ยวชาญในสาขาวิชาคอมพิวเตอร์และการศึกษาประเมินร่วมกัน เนื่องจากการพัฒนาระบบต้องใช้เทคโนโลยีทางคอมพิวเตอร์เข้ามาช่วยในการพัฒนา ส่วนเนื้อหานั้นมีความเกี่ยวข้องกับด้านการศึกษา (M. Lilley, Barker, & Britton, 2004; Nielsen, 1992) นอกจากนี้ อาจประยุกต์ใช้แนวคิดการประเมินตามมาตรฐานการประเมินระบบมาร่วมด้วย คือ ประเมินคุณสมบัติของระบบในมิติต่าง ๆ ได้แก่ การใช้ประโยชน์ (utility) ความเป็นไปได้ (feasibility) ความเหมาะสม (propriety) และ ความถูกต้อง (accuracy) (ศิริชัย กาญจนวาสี, 2554) ซึ่งผลจากการประเมินทั้งหมดจะช่วยให้ผู้พัฒนาระบบมี

ข้อมูลในการพัฒนาและปรับปรุงระบบให้สามารถทำงานได้อย่างราบรื่น และเป็นประโยชน์ต่อผู้ใช้งานมากยิ่งขึ้น

ในปัจจุบัน มีนำการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ (CAT) มาใช้อย่างแพร่หลายมากขึ้น แต่ดังที่ได้กล่าวไปแล้วว่าการทดสอบแบบ CAT ยังมีข้อจำกัดบางประการ ดังนั้น หากนำการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์แบบหลายขั้นตอน (MST) มาพัฒนาเพื่อใช้งาน จะก่อให้เกิดความมั่นใจมากขึ้นในการประมาณค่าความสามารถของผู้สอบ อีกทั้งยังสามารถให้ข้อมูลป้อนกลับเพื่อการพัฒนาผู้เรียนได้มากยิ่งขึ้น โดยเฉพาะอย่างยิ่งในการทดสอบระดับชาติ (O-NET) หากมีการพัฒนาการทดสอบให้เป็นแบบ MST ก็จะเป็นประโยชน์ต่อผู้เรียน และวงการศึกษามากขึ้น

O-NET (Ordinary National Educational Test) คือ การทดสอบทางการศึกษาระดับชาติขั้นพื้นฐานเป็นการทดสอบวัดความรู้และความคิดของนักเรียนชั้นประถมศึกษาปีที่ 6 ชั้นมัธยมศึกษาปีที่ 3 และ ชั้นมัธยมศึกษาปีที่ 6 เป็นการประเมินตามมาตรฐานการเรียนรู้ในหลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พุทธศักราช 2551 การทดสอบประกอบด้วย 5 วิชา ได้แก่ คณิตศาสตร์ วิทยาศาสตร์ ภาษาอังกฤษ ภาษาไทย และ สังคมศึกษา ศาสนาและวัฒนธรรม ซึ่งการทดสอบในระดับชั้นมัธยมศึกษาปีที่ 6 ถือว่าเป็นการทดสอบที่สำคัญยิ่งต่อนักเรียน เนื่องจากนักเรียนจำเป็นต้องใช้ความรู้พื้นฐานเพื่อนำไปใช้ศึกษาต่อในระดับมหาวิทยาลัย โดยวิชาที่มีความจำเป็นต่อการศึกษาในระดับสูง คือ วิชาภาษาอังกฤษ เพราะเมื่อศึกษาต่อในระดับสูงการค้นคว้าหาความรู้ใหม่ ๆ จำเป็นต้องใช้ภาษาอังกฤษ วิชาภาษาอังกฤษจึงถือเป็นวิชาที่มีความสำคัญในการพัฒนาองค์ความรู้ในระดับสูงของผู้เรียน และมีผลต่อการพัฒนาประเทศให้ทัดเทียมกับนานาชาติอารยประเทศ จากสถิติผลการทดสอบย้อนหลังปี 2558-2560 พบว่า คะแนนเฉลี่ยวิชาภาษาอังกฤษอยู่ในเกณฑ์ต่ำ (สถาบันทดสอบทางการศึกษาแห่งชาติ, 2562) ดังนั้น จึงควรเร่งผลักดันให้มีการพัฒนาสมรรถนะการเรียนรู้ของนักเรียนในรายวิชาภาษาอังกฤษ เพื่อให้ นักเรียนมีศักยภาพทางภาษาที่สูงขึ้นทัดเทียมกับนานาชาติอารยประเทศ และรองรับการเติบโตขององค์ความรู้ใหม่ ๆ ในอนาคต

ด้วยเหตุนี้ ผู้วิจัยจึงมีความสนใจพัฒนาระบบการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์แบบหลายขั้นตอน (MST) ที่มีการให้ข้อมูลป้อนกลับในการทดสอบระดับชาติ (O-NET) วิชาภาษาอังกฤษสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6 เพื่อลดความคลาดเคลื่อนจากการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์แบบดั้งเดิม (CAT) โดยระบบการทดสอบที่พัฒนาขึ้นนี้ สามารถสะท้อนข้อมูลป้อนกลับให้กับผู้สอบได้ทันที ซึ่งมีสารสนเทศเพื่อพัฒนาการเรียนรู้ของผู้สอบในรายวิชาภาษาอังกฤษ นอกจากนี้ ยังใช้เป็นแนวทางในอนาคตสำหรับพัฒนาระบบการทดสอบแบบ MST ในการทดสอบระดับชาติรายวิชาอื่น ๆ เพื่อให้ นักเรียนนำข้อมูลป้อนกลับที่ได้รับไปใช้ในการเตรียมตัวและวางแผนการทดสอบก่อนการทดสอบจริง ซึ่งจะช่วยยกระดับมาตรฐานคะแนนการทดสอบระดับชาติให้สูงขึ้น และเป็นไปตามเกณฑ์มาตรฐาน

วัตถุประสงค์ของการวิจัย

1. เพื่อพัฒนาระบบการทดสอบแบบปรับเหมาะหลายขั้นตอนที่มีการให้ข้อมูลป้อนกลับในการทดสอบระดับชาติ (O-NET) วิชาภาษาอังกฤษสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6
2. เพื่อประเมินคุณภาพระบบการทดสอบแบบปรับเหมาะหลายขั้นตอนที่มีการให้ข้อมูลป้อนกลับในการทดสอบระดับชาติ (O-NET) วิชาภาษาอังกฤษสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6

ขอบเขตของการวิจัย

1. การทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์แบบหลายขั้นตอน หรือการทดสอบแบบปรับเหมาะหลายขั้นตอน (MST) คือ การทดสอบที่ลดความคลาดเคลื่อนการประมาณค่าความสามารถในการทดสอบที่ต่ำกว่าหรือสูงกว่าความเป็นจริงจากการทดสอบแบบปรับเหมาะแบบดั้งเดิมในขั้นตอนแรก โดยการทดสอบแบบปรับเหมาะหลายขั้นตอนใช้วิธีการแบ่งข้อสอบออกเป็นขั้น (stage) ซึ่งในแต่ละขั้นประกอบด้วยโมดูล (module) ที่มีข้อสอบแบ่งไว้ตามระดับความยาก และอนุญาตให้ผู้สอบกลับไปแก้ไขข้อสอบได้ โดยการแก้ไขจะสามารถทำได้เฉพาะขั้นที่กำลังดำเนินการทดสอบเท่านั้น

2. ข้อสอบที่นำมาใช้ในการพัฒนาระบบการทดสอบ คือ ข้อสอบระดับชาติ (O-NET) วิชาภาษาอังกฤษสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6 ระหว่างปี 2558 – 2561 โดยผลการตอบข้อสอบเพื่อนำมาวิเคราะห์คุณภาพข้อสอบในการจัดทำคลังได้รับความอนุเคราะห์จากสถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) ซึ่งโดยทั่วไปคลังข้อสอบที่เหมาะสมควรมีอย่างน้อย 100-200 ข้อ (Weiss อ้างถึงใน ศิริชัย กาญจนวาสิ (2555) สำหรับคลังข้อสอบในงานวิจัยนี้ได้ดำเนินการวิเคราะห์มาจากรายการข้อสอบจำนวน 320 ข้อ ซึ่งเป็นข้อสอบระดับชาติ (O-NET) วิชาภาษาอังกฤษสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6 ระหว่างปี 2558 – 2561 โดยมีเกณฑ์ในการคัดเลือกข้อสอบเข้าคลัง คือ ค่าความยาก (b) อยู่ในช่วง -4 ถึง 4 ค่าอำนาจจำแนก (a) อยู่ในช่วง 0 ถึง α และความน่าจะเป็นในการเดา (c) ไม่เกิน 0.50 ซึ่งในการวิเคราะห์เพื่อประมาณค่าพารามิเตอร์ของข้อสอบดำเนินการวิเคราะห์ด้วยโปรแกรม R เนื่องจากเป็นโปรแกรมประเภทไม่มีค่าใช้จ่าย (open source) ด้วยชุดคำสั่งสำเร็จรูป (package) irtos ซึ่งเป็นชุดคำสั่งสำเร็จรูปที่ใช้ในการวิเคราะห์ข้อสอบตามทฤษฎีการตอบสนองข้อสอบ (item response theory: IRT)

3. ระบบการทดสอบแบบปรับเหมาะหลายขั้นตอนที่มีการให้ข้อมูลป้อนกลับในการทดสอบระดับชาติ (O-NET) วิชาภาษาอังกฤษสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6 พัฒนาโดยภาษา PHP ซึ่งสามารถใช้งานผ่านระบบออนไลน์ โดยเป้าหมายของการพัฒนาระบบ คือ เพื่อเตรียมความพร้อมผู้เรียนระดับชั้นมัธยมศึกษาปีที่ 6 ในการพัฒนาตนเองก่อนการทดสอบระดับชาติ (O-NET) วิชาภาษาอังกฤษจริง

4. การประเมินระบบการทดสอบประกอบด้วย การประเมิน 2 ส่วน ได้แก่ 1) การประเมินแบบอิงมาตรฐาน ได้แก่ การใช้ประโยชน์ (utility) ความเป็นไปได้ (feasibility) ความเหมาะสม (propriety) และความถูกต้อง (accuracy) (ศิริชัย กาญจนวาสี, 2554) และ 2) การประเมินความพึงพอใจจากผู้ใช้งานระบบด้วยการประเมินความพึงพอใจระหว่างอินเทอร์เฟซกับผู้ใช้ (user satisfaction of the human-computer interface) ซึ่งมีหลักเกณฑ์ในการประเมิน 5 ด้าน ได้แก่ หน้าจอระบบ คำศัพท์และสารสนเทศของระบบ การเรียนรู้ และสมรรถนะของระบบ (Chin, Diehl, & Norman, 1988; Navas et al., 2007)

คำจำกัดความที่ใช้ในการวิจัย

1. ระบบการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ที่มีการสะท้อนข้อมูลป้อนกลับ หมายถึง ชุดขององค์ประกอบที่มีความสัมพันธ์กันในการทดสอบแบบปรับเหมาะหลายขั้นตอนด้วยวิธีออนไลน์ที่สามารสะท้อนข้อมูลป้อนกลับให้ผู้เข้ารับการทดสอบเพื่อนำไปพัฒนาและปรับปรุงตนเอง โดยระบบประกอบด้วยองค์ประกอบหลัก 6 องค์ประกอบ ได้แก่ 1) จุดเริ่มต้นในการทดสอบ (initial stage) 2) การประมาณค่าความสามารถของผู้สอบ (ability estimation) 3) การคัดเลือกข้อสอบ (item selection) 4) การควบคุมอัตราการเปิดเผยของข้อสอบ (exposure rate) และ 5) จุดยุติการทดสอบ (stopping criteria) และรูปแบบรายงานผลที่สะท้อนข้อมูลป้อนกลับประกอบด้วย รายงาน 2 รูปแบบ ได้แก่ 1) รายงานคะแนนสังเกตได้ (observe score) และ 2) รายงานระดับความสามารถของผู้สอบ (ability)

2. การทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์แบบหลายขั้นตอน (multi stage testing: MST) หมายถึง การทดสอบที่มีการคัดเลือกข้อสอบให้เหมาะสมกับระดับความสามารถของผู้สอบ ซึ่งการประมาณค่าความสามารถของผู้สอบมาจากโมดูล (module) ของข้อสอบในแต่ละขั้น (stage) โดยแต่ละขั้นจะประกอบไปด้วยโมดูลของข้อสอบที่มีการจัดเตรียมไว้โดยผู้บริหารการทดสอบ และแต่ละโมดูลจะประกอบไปด้วยข้อสอบที่มีความยากใกล้เคียงกัน

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

4. แบบทดสอบระดับชาติ (O-NET) วิชาภาษาอังกฤษสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6 หมายถึง แบบทดสอบทางการศึกษาระดับชาติขั้นพื้นฐานซึ่งใช้เพื่อวัดความรู้และความคิดของนักเรียนชั้นมัธยมศึกษาปีที่ 6 ประเมินตามมาตรฐานการเรียนรู้ในหลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พุทธศักราช 2551 โดยวิชาภาษาอังกฤษเป็นหนึ่งในกลุ่มสาระการเรียนรู้จากทั้งหมด 5 กลุ่มสาระการเรียนรู้

5. ข้อมูลป้อนกลับ (feedback) หมายถึง ชุดข้อมูลที่สะท้อนความสามารถของผู้เรียนโดยระบุถึงจุดเด่น จุดที่ควรปรับปรุง เพื่อนำไปใช้เป็นแนวทางในการพัฒนาหรือปรับปรุงตนเองของผู้เรียนให้ดียิ่งขึ้น

6. การประเมินระบบการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ หมายถึง การประเมินระบบจากผู้เชี่ยวชาญด้วยการประเมินแบบอิงมาตรฐาน ได้แก่ ความมีประโยชน์ ความเป็นไปได้ ความเหมาะสม และความถูกต้อง และการประเมินจากผู้สอบด้วยการประเมินความพึงพอใจของผู้สอบต่อการใช้งานระบบ

ประโยชน์ที่ได้รับ

1. ได้ระบบการทดสอบแบบปรับเหมาะหลายขั้นตอนที่มีการให้ข้อมูลป้อนกลับในการทดสอบระดับชาติ (O-NET) วิชาภาษาอังกฤษสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6
2. นักเรียนระดับชั้นมัธยมศึกษาปีที่ 6 ได้รับข้อมูลป้อนกลับเพื่อเตรียมความพร้อมสำหรับการทดสอบระดับชาติ (O-NET) วิชาภาษาอังกฤษ

บทที่ 2

เอกสารและงานวิจัยที่เกี่ยวข้อง

การวิจัยในครั้งนี้ ผู้วิจัยแบ่งเนื้อหาออกเป็น 5 ตอน ได้แก่ ตอนที่ 1 มโนทัศน์เกี่ยวกับการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ ตอนที่ 2 มโนทัศน์เกี่ยวกับการทดสอบแบบปรับเหมาะแบบหลายขั้นตอน ตอนที่ 3 มโนทัศน์เกี่ยวกับการให้ข้อมูลป้อนกลับ ตอนที่ 4 มโนทัศน์เกี่ยวกับการทดสอบระดับชาติ (O-NET) วิชาภาษาอังกฤษ ตอนที่ 5 มโนทัศน์เกี่ยวกับการประเมินผลระบบการทดสอบ และตอนที่ 6 กรอบแนวคิดการวิจัย ซึ่งแต่ละตอนมีรายละเอียด ดังนี้

ตอนที่ 1 มโนทัศน์เกี่ยวกับการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์

รูปแบบของการทดสอบมีวิวัฒนาการมาอย่างต่อเนื่อง เริ่มจากการทดสอบโดยใช้กระดาษ (paper-and-pencil test: P&P) ต่อมาเมื่อเทคโนโลยีด้านคอมพิวเตอร์และอินเทอร์เน็ตมีความเจริญก้าวหน้าขึ้น จึงเกิดรูปแบบการทดสอบที่จัดการโดยใช้คอมพิวเตอร์ (computer-administered test) ซึ่งแบ่งได้เป็น 3 รูปแบบหลัก คือ การทดสอบโดยใช้คอมพิวเตอร์ (computer-based testing: CBT) การทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ (computerized adaptive testing: CAT) และ การทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์แบบหลายขั้นตอน (multistage adaptive tests: MST) โดยการทดสอบแบบ P&P และ CBT ผู้สอบทุกคนที่สอบในครั้งเดียวกันจะได้แบบทดสอบฉบับเดียวกัน ในขณะที่ CAT และ MST ที่เป็นการทดสอบแบบปรับเหมาะ (adaptive) ผู้สอบแต่ละคนจะได้รับข้อสอบที่มีความยากง่ายสอดคล้องกับความสามารถของตนเอง ทำให้แบบทดสอบที่ใช้สำหรับผู้สอบแต่ละคนสามารถประมาณค่าความสามารถของผู้สอบได้แม่นยำขึ้น และความยาวของแบบทดสอบลดลง (Patsula, 1999; Zheng & Chang, 2015) ซึ่งจากการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องกับการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ มีแนวคิดที่เกี่ยวข้อง โดยนำเสนอเป็นหัวข้อต่าง ๆ ดังนี้ 1.1) ทฤษฎีการตอบสนองข้อสอบ 1.2) ความหมายและหลักการของการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ 1.3) องค์ประกอบของการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ และ 1.4) จุดแข็งและจุดอ่อนของการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์

1.1 ทฤษฎีการตอบสนองข้อสอบ

ทฤษฎีการตอบสนองข้อสอบ (item response theory: IRT) หรือทฤษฎีคุณลักษณะแฝง (Latent Trait Theory) มีแนวคิดว่าการประมาณค่าความสามารถ (ability) ของผู้สอบนั้นขึ้นกับลักษณะการตอบของผู้สอบและคุณสมบัติของข้อคำถามที่ถูกจัดให้แก่ผู้สอบ (Embretson & Reise, 2000) และอธิบายความสัมพันธ์ระหว่างความสามารถของผู้สอบ กับผลการตอบข้อสอบหรือข้อคำถามโดยใช้โค้งคุณลักษณะข้อสอบ (item characteristic curve: ICC)

คุณสมบัติของข้อสอบหรือข้อคำถาม ถูกกำหนดเป็นค่าพารามิเตอร์ของข้อสอบ 3 ค่า ได้แก่ ค่าความยาก (difficulty: b) ค่าอำนาจจำแนก (discrimination power: a) และค่าความน่าจะเป็นในการเดา (guessing probability: c) ทฤษฎีการตอบสนองข้อสอบสามารถจำแนกออกเป็นสองประเภทตามลักษณะการให้คะแนน สามารถแบ่งได้เป็น 2 ประเภท ตามลักษณะการให้คะแนน คือ แบบตรวจให้คะแนนสองค่า (dichotomous IRT) ใช้สำหรับการประมาณค่าพารามิเตอร์ของเครื่องมือที่มีการตรวจให้คะแนนรายข้อเพียงสองค่า เช่น การตรวจให้คะแนนแบบตอบถูกได้ 1 ตอบผิดได้ 0 เป็นต้น และแบบตรวจให้คะแนนมากกว่าสองค่า (polytomous IRT) ใช้สำหรับการประมาณค่าพารามิเตอร์ของเครื่องมือที่รายการคำตอบแต่ละข้อต่างกัน เช่น เครื่องมือแบบมาตราประมาณค่า ซึ่งเครื่องมือวัดทางการศึกษาและจิตวิทยาจำนวนมากมีลักษณะเช่นนี้ เช่น แบบวัดความสนใจทางอาชีพ แบบวัดบุคลิกภาพ เป็นต้น (ศิริชัย กาญจนวาสี, 2555)

สำหรับการทดสอบแบบหลายตัวเลือก (multiple choice) เป็นการทดสอบที่ตรวจให้คะแนนแบบสองค่า โมเดลการตอบสนองข้อสอบแบบให้คะแนนสองค่าจึงเป็นความสัมพันธ์ระหว่างความสามารถของผู้สอบ (ability: θ) กับโอกาสในการตอบข้อสอบถูก (probability: P_i) และแสดงเป็นโค้งคุณลักษณะของข้อสอบในรูปแบบฟังก์ชันโลจิสติก (logistic function) หรือฟังก์ชันปกติสะสม (normal ogive function) โดยมีโมเดลหลัก 3 โมเดล ได้แก่ (ทัศน์ศิริินทร์ สว่างบุญ, ศิริเดช สุชีวะ, ศิริชัย กาญจนวาสี, & Muraki, 2555)

(1) โมเดลโลจิสติกแบบ 3 พารามิเตอร์ (three parameter logistic model: 3PL) ซึ่งมีรูปแบบฟังก์ชันดังนี้

$$P_i(\theta) = c_i + \frac{(1 - c_i)}{1 + e^{-Da_i(\theta - b_i)}}$$

โดยที่ θ คือ ค่าความสามารถของผู้สอบซึ่งได้จากการประมาณค่าตามทฤษฎีการตอบสนองข้อสอบ ซึ่งนิยมปรับให้เป็นมาตรฐาน (standardize) ที่มีค่าเฉลี่ยเป็น 0 และความแปรปรวนเป็น 1 และค่าที่ใช้ส่วนใหญ่อยู่ระหว่าง -3 ถึง 3

$P_i(\theta)$ คือ ค่าความน่าจะเป็นของผู้ที่มีความสามารถระดับ θ จะสามารถตอบข้อสอบข้อที่ i ได้ถูกต้อง

b_i คือ ค่าพารามิเตอร์ความยากของข้อสอบ (difficulty parameter) ค่าที่นิยมใช้อยู่ระหว่าง -2.5 ถึง +2.5

a_i คือ ค่าพารามิเตอร์อำนาจจำแนกของข้อสอบ (discrimination parameter) ค่าที่นิยมใช้อยู่ระหว่าง +0.50 ถึง +2.50

c_i คือ ค่าพารามิเตอร์โอกาสในการเดาของข้อสอบ (guessing parameter) ค่าที่นิยมใช้คือค่าไม่เกิน 0.30

e คือ ค่าลอการิทึมฐานธรรมชาติ (natural log) มีค่าประมาณ 2.7182

D คือ ค่าองค์ประกอบของการปรับสเกลเพื่อให้ฟังก์ชัน (logistic) กับนอมอลอจไฟฟ์ (normal ogive) มีค่าใกล้เคียงกัน หรือค่าประมาณความสามารถมีค่าต่างกันไม่เกิน 0.01

(2) โมเดลโลจิสติกแบบ 2 พารามิเตอร์ (two parameter logistic model: 2PL) ซึ่งมีรูปแบบฟังก์ชันดังนี้

$$P_i(\theta) = \frac{1}{1 + e^{-Da_i(\theta - b_i)}}$$

ซึ่งค่าความยาก (b) และค่าอำนาจจำแนก (a) มีค่าแปรเปลี่ยนไปตามลักษณะของข้อสอบแต่ละข้อ และค่าโอกาสในการเดามีค่าเป็น 0

(3) โมเดลโลจิสติกแบบ 1 พารามิเตอร์ (one parameter logistic model: 1PL) หรือเรียกอีกอย่างว่าโมเดลแบบราซ (rasch model) ซึ่งมีรูปแบบฟังก์ชันดังนี้

$$P_i(\theta) = \frac{1}{1 + e^{-(\theta - b_i)}}$$

ซึ่งค่าความยาก (b) มีค่าแปรเปลี่ยนไปตามลักษณะของข้อสอบแต่ละข้อ ค่าอำนาจจำแนก (a) มีค่าคงที่และค่าโอกาสในการเดามีค่าเป็น 0

นอกจากนี้ ตามทฤษฎีการตอบสนองข้อสอบยังสามารถวิเคราะห์ฟังก์ชันสารสนเทศของข้อสอบ (item information) และของแบบสอบ (test information) ซึ่งเป็นดัชนีที่แสดงถึงคุณภาพของข้อสอบและแบบสอบ สำหรับค่าฟังก์ชันสารสนเทศของข้อสอบเป็นดัชนีที่ได้มาจากค่าพารามิเตอร์ของข้อสอบ ได้แก่ ค่าความยาก (b parameter) ค่าอำนาจจำแนก (a parameter) และค่าความแปรปรวนของคะแนนรายข้อ แสดงถึงคุณภาพของข้อสอบรายข้อ ซึ่งค่าฟังก์ชันสารสนเทศของข้อสอบจะสูงขึ้นเมื่อผู้สอบมีระดับความสามารถใกล้เคียงกับค่าความยากของข้อสอบ (b) และข้อสอบจะมีค่าฟังก์ชันสารสนเทศของข้อสอบสูงขึ้นเมื่อค่าอำนาจจำแนกของข้อสอบ (a) สูงขึ้น และมีค่าโอกาสในการเดา (c) มีค่าเข้าใกล้ 0 ส่วนค่าฟังก์ชันสารสนเทศของแบบสอบเป็นผลรวมเชิงพีชคณิตของค่าฟังก์ชันสารสนเทศของข้อสอบแต่ละข้อ แสดงคุณภาพของแบบสอบทั้งฉบับเป็นดัชนีที่แสดงถึงความสามารถของแบบสอบในการประมาณค่าความสามารถของผู้สอบได้อย่างถูกต้องและแม่นยำ ซึ่งค่าฟังก์ชันสารสนเทศของแบบสอบนี้สัมพันธ์ผกผันกับค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่า หากสารสนเทศของแบบสอบมีค่าสูงในช่วง θ ใด แสดงว่าแบบสอบสามารถประมาณค่าความสามารถของผู้สอบในช่วง θ นั้นได้อย่างถูกต้องและแม่นยำสูง โดยจะมีความคลาดเคลื่อนมาตรฐานของการประมาณค่าต่ำ (ศิริชัย กาญจนวาสี, 2555)

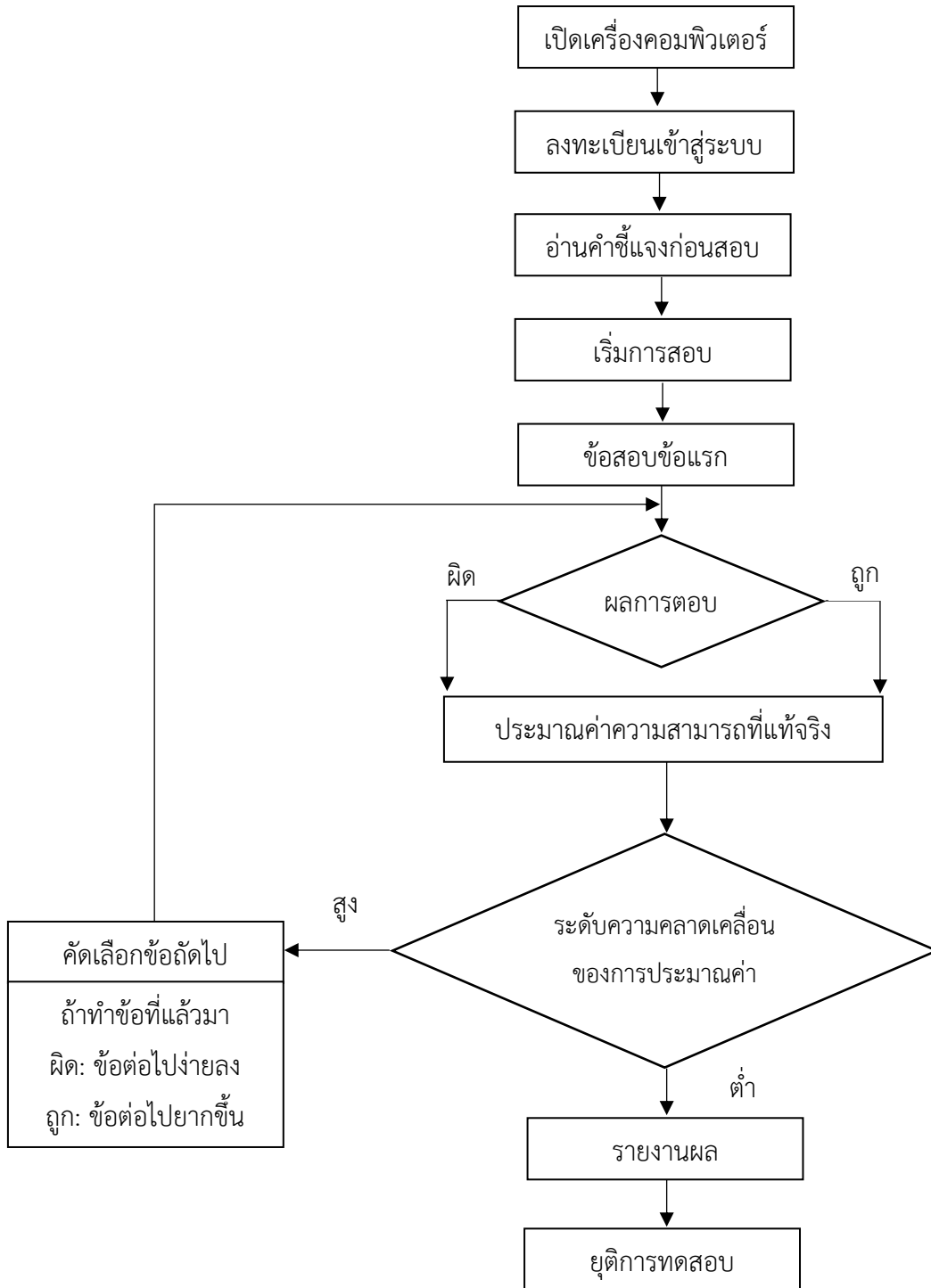
ทฤษฎีการตอบสนองข้อสอบนี้ ถูกนำมาประยุกต์ใช้ในการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ ที่คัดเลือกข้อสอบที่มีความยากง่ายแตกต่างกันให้แก่ผู้สอบแต่ละคนตามระดับความสามารถของผู้สอบ ซึ่งรายละเอียดเกี่ยวกับการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ดังนี้จะนำเสนอในหัวข้อถัดไป

1.2 ความหมายและหลักการของการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์

การทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ (computerized adaptive testing: CAT) หมายถึง การทดสอบที่ใช้วิธีการทางคอมพิวเตอร์เข้ามาช่วยในการคัดเลือกข้อสอบ และประมาณค่าความสามารถของผู้สอบ ผู้สอบแต่ละคนจึงได้รับแบบทดสอบที่มีความยากง่ายแตกต่างกัน ขึ้นกับระดับความสามารถของผู้สอบแต่ละคน (ทัศนศิริรินทร์ สว่างบุญ และคณะ, 2555)

การใช้คอมพิวเตอร์ช่วยทำให้คะแนนการทดสอบมีความถูกต้องในเชิงสถิติเพิ่มขึ้น แนวคิดของการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์นี้ เริ่มจากการนำมาใช้กับแบบวัดเชาว์ปัญญาของ Binet-Simon ในปี 1905 โดยคอมพิวเตอร์จะคัดเลือกข้อคำถามให้เหมาะสมกับอายุสมองของผู้สอบ ซึ่งอาศัยการอ้างอิงจากอายุสมองของผู้สอบที่ได้มาจากผลการตอบข้อคำถามข้อก่อนหน้า และจะเลือกข้อคำถามให้ผู้สอบตอบจนเพียงพอต่อการประมาณค่าอายุสมองที่แท้จริงของผู้สอบได้ ความเป็นจริงแนวคิดการทดสอบแบบปรับเหมาะมีมานานแล้ว ดังเห็นได้จากการทดสอบปากเปล่า (oral test) ที่ผู้จัดสอบสามารถเลือกคำถามให้เหมาะสมกับความสามารถของผู้สอบได้ตามระดับความรู้จากการสอบถามและพูดคุย (Linden, 2000)

สำหรับการคัดเลือกข้อคำถามให้เหมาะสมกับความสามารถของผู้สอบนั้น มีแนวทางในการคัดเลือกข้อสอบสำหรับผู้สอบแต่ละคน โดยอ้างอิงจากผลการตอบข้อสอบข้อที่ผ่านมา ในการทดสอบโดยทั่วไปผู้สอบจะได้เริ่มทำข้อสอบข้อแรกที่มีค่าความยากระดับปานกลาง จากนั้นระบบคอมพิวเตอร์จะทำการประมาณค่าความสามารถของผู้สอบจากผลการตอบข้อสอบ แล้วคัดเลือกข้อสอบข้อถัดไปที่มีความยากและอำนาจจำแนกที่เหมาะสมกับระดับความสามารถของผู้สอบ โดยมีหลักการคือหากผู้สอบตอบข้อสอบข้อที่ผ่านมาถูกต้อง ผู้สอบจะได้ข้อสอบข้อถัดไปที่มีความยากมากขึ้น แต่หากผู้สอบตอบข้อสอบผิด ผู้สอบจะได้ข้อสอบข้อถัดไปที่มีความยากลดลง จนกระทั่งการประมาณค่าความสามารถผู้สอบมีความเชื่อถือได้คืออยู่ในเกณฑ์ของความคลาดเคลื่อนที่กำหนดจึงยุติการทดสอบ โดยภาพขั้นตอนการดำเนินการทดสอบแสดงดังภาพที่ 1 (Ćisar, Radosav, Markoski, Pinter, & Ćisar, 2010; ทัศนศิริรินทร์ สว่างบุญ และคณะ, 2555)



ภาพที่ 1 ขั้นตอนการดำเนินการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์

1.3 องค์ประกอบของการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์

ในการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ มีองค์ประกอบหลักในกระบวนการทดสอบ ได้แก่ จุดเริ่มต้นการทดสอบ (initial item) การคัดเลือกข้อสอบและการประมาณค่าความสามารถ (item selection and ability estimation) เกณฑ์ยุติการทดสอบ (stopping criteria) มีรายละเอียด ดังนี้ (ทศน์ศิริรินทร์ สว่างบุญ และคณะ, 2555; ศิริชัย กาญจนวาสี, 2538)

1.3.1 จุดเริ่มต้นการทดสอบ (initial item)

การทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์มีจุดเริ่มต้นจากการคัดเลือกข้อสอบข้อแรก (initial item) ที่จะแสดงต่อผู้สอบ โดยปกติข้อสอบข้อแรกที่ถูกคัดเลือกขึ้นมาจะ เป็นข้อที่มีความยากปานกลาง ซึ่งสามารถแบ่งวิธีการเลือกข้อสอบข้อแรกออกเป็น 2 วิธี ได้แก่

(1) เมื่อประชากรของผู้สอบมีความสามารถใกล้เคียงกัน (homogenous) หรือไม่ทราบระดับความสามารถหรือข้อมูลผลสัมฤทธิ์ที่ผ่านมาของประชากร เช่น หากเป็นการทดสอบ เฉพาะนักเรียนระดับมัธยมศึกษาปีที่ 6 ควรเริ่มต้นทำข้อสอบที่มีความยากปานกลาง เนื่องจาก นักเรียนระดับชั้นเดียวกันควรมีระดับความสามารถหรือผลสัมฤทธิ์ที่ใกล้เคียงกัน

(2) เมื่อประชากรของผู้สอบมีความสามารถแตกต่างกัน (heterogeneous) หรือมีข้อมูลสารสนเทศเกี่ยวกับระดับความสามารถหรือผลสัมฤทธิ์ที่ผ่านมาของประชากร เช่น นักเรียน ระดับมัธยมศึกษาตอนต้น ควรเริ่มต้นทำข้อสอบที่มีความยากปานกลางในแต่ละระดับชั้น เนื่องจาก นักเรียนแต่ละระดับชั้นมีความสามารถที่แตกต่างกัน

อย่างไรก็ดี จุดเริ่มต้นการทดสอบนี้ ไม่มีผลต่อความแม่นยำในการประมาณค่า ความสามารถของผู้สอบมากนัก ดังที่ Lord (1977 อ้างถึงใน ศิริชัย กาญจนวาสี, 2555) ได้ศึกษา โดยการจำลองข้อมูลการทดสอบที่ข้อสอบข้อแรกที่มีระดับความยากแตกต่างกันพบว่า เกณฑ์การ คัดเลือกข้อสอบข้อแรกไม่ค่อยมีความสำคัญ โดยเฉพาะการทดสอบที่แบบสอบมีจำนวนข้อมากกว่า 25 ข้อ กล่าวคือเมื่อใช้ข้อสอบเกิน 25 ข้อจะไม่ส่งผลต่อความแม่นยำในการประมาณค่าความสามารถ ของผู้สอบ จึงสามารถสรุปได้ว่าระดับความยากของข้อสอบข้อแรกไม่มีผลต่อการทดสอบแบบปรับ เหมาะโดยใช้คอมพิวเตอร์เมื่อแบบสอบมีความยาวเพิ่มขึ้น

1.3.2 การคัดเลือกข้อสอบและการประมาณค่าความสามารถ (item selection and ability estimation)

หลังจากผู้สอบทำข้อสอบข้อแรกเรียบร้อยแล้ว ระบบคอมพิวเตอร์จะทำการ ประมาณค่าความสามารถของผู้สอบจากผลการตอบข้อสอบ จากนั้นจะทำการเลือกข้อสอบข้อถัดไปที่ เหมาะสมกับระดับความสามารถของผู้สอบต่อไป รายละเอียดของการคัดเลือกข้อสอบและการ ประมาณค่าความสามารถของผู้สอบมีดังนี้

(1) การประมาณค่าความสามารถ (ability estimation) : วิธีการประมาณค่า ความสามารถที่แท้จริงของผู้สอบตามทฤษฎีการตอบสนองข้อสอบ (item response theory: IRT) ที่

นิยมใช้มีอยู่ 2 วิธี ได้แก่ วิธีการประมาณค่าด้วยความเป็นไปได้สูงสุด (maximum likelihood estimation) และ วิธีการประมาณค่าแบบเบเซียน (bayesian estimation) ซึ่งโปรแกรมที่ใช้ในการวิเคราะห์เพื่อประมาณค่าความสามารถของผู้สอบตามทฤษฎีการตอบสนองข้อสอบมีอยู่ด้วยกันหลายโปรแกรม เช่น โปรแกรม LOGISTS โปรแกรม BILOG เป็นต้น

(2) การคัดเลือกข้อสอบ (item selection) : การคัดเลือกข้อสอบข้อถัดไปสำหรับการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ สามารถเลือกใช้กับวิธีการประมาณค่าความสามารถของผู้สอบวิธีใดก็ได้ แต่เพื่อความสะดวกในการคำนวณจึงนิยมเลือกใช้วิธีการประมาณค่าความสามารถของผู้สอบและวิธีการคัดเลือกข้อสอบที่สอดคล้องกัน การคัดเลือกข้อสอบซึ่งเป็นที่ยอมรับประกอบด้วยวิธีการประมาณค่าด้วยความเป็นไปได้สูงสุด (maximum likelihood estimation) และวิธีการประมาณค่าแบบเบเซียน (bayesian estimation) โดยมีรายละเอียดดังนี้

(2.1) วิธีการประมาณค่าด้วยความเป็นไปได้สูงสุด (maximum likelihood estimation) มีวิธีการคัดเลือกข้อสอบ 3 วิธี ได้แก่ 1) การคัดเลือกข้อสอบที่มีค่าความยาก (b) สอดคล้องกับระดับความสามารถที่ได้จากการประมาณค่า 2) การคัดเลือกข้อสอบที่มีตำแหน่งสารสนเทศสูงสุดสอดคล้องกับระดับความสามารถที่ได้จากการประมาณค่า และ 3) การคัดเลือกข้อสอบที่ให้สารสนเทศสูงสุดตรงตำแหน่งระดับความสามารถ

- การคัดเลือกข้อสอบที่มีค่าความยาก (b) สอดคล้องกับระดับความสามารถของผู้สอบที่ได้จากการประมาณค่า โดยระบบจะทำการตรวจสอบข้อสอบทุกข้อในคลังเพื่อเลือกข้อสอบที่มีค่าความยากใกล้เคียงกับค่าความสามารถของผู้สอบที่ได้จากการประมาณค่าที่แท้จริงมากที่สุด และจะใช้ข้อสอบข้อดังกล่าวเป็นข้อสอบข้อถัดไปในการทดสอบ ซึ่งวิธีนี้ใช้กันมาเป็นระยะเวลานาน เนื่องจากเป็นวิธีที่มีการคำนวณไม่ซับซ้อนและค่อนข้างประหยัดค่าใช้จ่าย

- การคัดเลือกข้อสอบที่มีตำแหน่งสารสนเทศสูงสุดสอดคล้องกับระดับความสามารถที่ได้จากการประมาณค่า เนื่องจากการทดสอบโดยทั่วไปมักมีปัจจัยจากการเดาเข้ามาเกี่ยวข้อง ดังนั้นวิธีการการคัดเลือกข้อสอบที่มีค่าความยากสอดคล้องกับระดับความสามารถของผู้สอบที่ได้จากการประมาณค่าจึงมีแนวโน้มที่ระบบจะเลือกข้อสอบที่ยากเกินความสามารถของผู้สอบ จึงมีการนำวิธีการเลือกจากตำแหน่งของข้อสอบที่มีสารสนเทศสูงสุดสอดคล้องกับระดับความสามารถของผู้สอบมาใช้ โดยข้อสอบแต่ละข้อจะให้สารสนเทศสูงสุดที่ตำแหน่ง $\theta = m_i$ เมื่อค่า m_i สำหรับโมเดลการตอบสนองข้อสอบแบบ 3 พารามิเตอร์ สามารถคำนวณได้จาก

$$m_i = b_i + \frac{1}{Da_1} \ln \left[\frac{1 + \sqrt{1 + 8c_i}}{2} \right]$$

เมื่อ $D =$ ค่าคงที่

$\ln =$ ค่าลอการิทึมธรรมชาติ (natural logarithm)

$a_i =$ ค่าอำนาจจำแนกของข้อสอบ

$b_i =$ ค่าความยากของข้อสอบ

$c_i =$ ค่าโอกาสในการเดาข้อสอบถูก

ค่า m_i เป็นค่าความสามารถของผู้สอบ (θ) ณ ตำแหน่งที่ข้อสอบข้อนั้นให้สารสนเทศสูงที่สุด ซึ่งหากค่า $m_i > b_i$ และ $c_i > 0$ แสดงว่าข้อสอบมีค่าความยากเหมาะสมที่จะนำมาใช้ประมาณค่าความสามารถของผู้สอบที่มีความสามารถสูงกว่าความยากของข้อสอบเล็กน้อย ดังนั้นการนำค่า m_i มาใช้ในการพิจารณาคัดเลือกข้อสอบจึงทำให้การคัดเลือกข้อสอบดีกว่าวิธีเลือกจากค่าระดับความยากเพียงอย่างเดียว

- การคัดเลือกข้อสอบที่ให้สารสนเทศสูงสุดตรงตำแหน่งระดับความสามารถ โดยระบบคอมพิวเตอร์จะตรวจสอบข้อสอบทุกข้อที่ยังมิได้นำมาใช้สอบซึ่งสามารถให้สารสนเทศได้สูงที่สุดตรงตำแหน่งความสามารถ วิธีการนี้จะต้องคำนวณหาสารสนเทศของข้อสอบ (item information) ณ ตำแหน่งระดับความสามารถของผู้สอบ (θ)

(2.2) วิธีการประมาณค่าแบบเบเซียน (bayesian estimation) คือการเลือกข้อสอบที่ยังไม่เคยมีการนำมาใช้ในการทดสอบ ซึ่งจะให้ค่าความแปรปรวนของค่าความสามารถที่คาดหวังต่ำที่สุด (smallest posterior variance) หมายความว่า เป็นการคัดเลือกข้อสอบข้อที่คาดว่าจะสามารถลดความคลาดเคลื่อนในการประมาณค่าความสามารถของผู้สอบได้มากที่สุด มีรายละเอียดของทฤษฎีเบส์ (bayes' theorem) ดังนี้

$$f(\theta/U) = K L(\theta/U) f(\theta)$$

เมื่อ $f(\theta/U) =$ การแจกแจงภายหลัง (posterior distribution) ของค่าความสามารถ

$L(\theta/U) =$ ฟังก์ชันความเป็นไปได้ (likelihood function) ของเวกเตอร์ U

(ผลการตอบข้อสอบ)

$f(\theta) =$ การแจกแจงภายหลัง (posterior distribution) ของค่าความสามารถ

$K =$ ค่าคงที่ (constant)

1.3.3 เกณฑ์ยุติการทดสอบ (stopping criteria)

การทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ ผู้สอบจะต้องทดสอบจนกว่าจะถึงเกณฑ์ยุติการทดสอบ การทดสอบจึงจะสิ้นสุดลงและทำการประมาณค่าความสามารถของผู้สอบ

ต่อไป เกณฑ์ยุติการทดสอบที่นิยมใช้มีอยู่ 2 ประเภทได้แก่ เกณฑ์กำหนดจำนวนข้อสอบคงที่ และ เกณฑ์กำหนดระดับความคลาดเคลื่อนที่ยอมรับได้

(1) เกณฑ์กำหนดจำนวนข้อสอบคงที่ : เป็นเกณฑ์ที่กำหนดให้ผู้สอบทุกคนทำข้อสอบที่มีความยาวของแบบสอบหรือมีจำนวนข้อเท่ากันทุกคน เช่น หากกำหนดเกณฑ์แบบจำนวนข้อสอบคงที่โดยให้ความยาวแบบสอบเท่ากับ 30 ข้อ ผู้สอบทุกคนจะต้องทำข้อสอบจำนวน 30 ข้อเท่ากันทุกคน เมื่อทำครบข้อที่ 30 การทดสอบจะยุติลง เกณฑ์ลักษณะเช่นนี้เหมาะกับการศึกษากระบวนการทดสอบแบบปรับเหมาะด้วยการจำลองข้อมูล เนื่องจากเมื่อข้อสอบมีจำนวนเท่ากัน จะทำให้สามารถเปรียบเทียบสารสนเทศของแบบสอบได้อย่างชัดเจน

(2) เกณฑ์กำหนดระดับความคลาดเคลื่อนที่ยอมรับได้ : เป็นเกณฑ์ที่ใช้กับการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ในทางปฏิบัติ คือผู้สอบจะต้องทำการการทดสอบไปเรื่อย ๆ จนกระทั่งความคลาดเคลื่อนมาตรฐาน (standard error) ในการประมาณค่าความสามารถของผู้สอบลดต่ำลงจนถึงระดับที่สามารถยอมรับได้หรือถึงเกณฑ์ที่กำหนดไว้ การทดสอบจึงยุติลง ความคลาดเคลื่อนมาตรฐานสามารถคำนวณได้จากสูตร

$$SE(\theta) = \frac{1}{\sqrt{I(\theta)}}$$

เมื่อ $SE(\theta)$ = ความคลาดเคลื่อนมาตรฐานของการประมาณค่าความสามารถ θ

$I(\theta)$ = สารสนเทศของแบบสอบของผู้มีความสามารถ θ

1.4 จุดแข็งและจุดอ่อนของการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์

การทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ (computer adaptive testing: CAT) เป็นเทคโนโลยีทางการทดสอบ ที่มีจุดแข็งในเชิงทฤษฎี คือ เป็นการทดสอบที่ผู้สอบจะได้ทำข้อสอบที่มีความยากตามระดับความสามารถของตน โดยที่ข้อสอบไม่ง่ายหรือยากจนเกินไปเหมือนกับการทดสอบแบบดั้งเดิมที่ใช้กระดาษในการทดสอบ และจำนวนข้อสอบที่ใช้ในการทดสอบก็ลดลงโดยที่ไม่สูญเสียความแม่นยำในการประมาณค่าความสามารถของผู้สอบ โดยจะเลือกข้อสอบแต่ละข้อจากคลังข้อสอบด้วยกระบวนการทางคอมพิวเตอร์ภายหลังการประมาณค่าความสามารถของผู้สอบบนพื้นฐานของผลการตอบข้อสอบจากข้อก่อนหน้า ซึ่งเป็นการควบคุมคุณภาพที่มีประสิทธิภาพ เนื่องจากสามารถลดความคลาดเคลื่อนที่เกิดจากคนในการจัดการข้อสอบได้ อีกทั้งยังควบคุมเงื่อนไขต่าง ๆ เช่น ความครอบคลุมของเนื้อหา (content coverage) และอัตราการเปิดเผยของข้อสอบ (item exposure) เป็นต้น (Zheng & Chang, 2015; ทิศน์ศิริรินทร์ สว่างบุญ และคณะ, 2555) นอกจากนี้ การทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ ยังสามารถนำมาปรับใช้กับการประเมินความรู้ของนักเรียนได้ ซึ่งจากผลการวิจัยพบว่า การทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์เป็นเครื่องมือที่มี

พลังสำหรับการประเมินเป็นอย่างยิ่ง (Andjelic & Cekerevac, 2014) และการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ยังเป็นวิธีการทดสอบที่มีประสิทธิภาพที่สามารถนำมาใช้ในการประเมินความก้าวหน้าของนักเรียนทั้งแบบรายบุคคลและแบบกลุ่มได้อีกด้วย (Weiss, 2004)

แม้ว่าการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์จะมีจุดแข็งที่เหนือกว่าการทดสอบโดยใช้กระดาษ (paper-and-pencil testing) และการทดสอบโดยใช้คอมพิวเตอร์ (computer-based testing) แต่ก็ยังมีจุดอ่อนในเชิงทฤษฎี คือ อาจเกิดการประมาณค่าความสามารถที่สูงกว่าความเป็นจริง (over estimate) หรือต่ำกว่าความเป็นจริง (under estimate) ซึ่งมีการศึกษาที่พบกรณีดังกล่าวในการทดสอบเข้าเรียนต่อในระดับบัณฑิตศึกษา (graduate record examination: GRE) และการทดสอบเพื่อเข้าศึกษาต่อด้านการจัดการในระดับบัณฑิตศึกษา (graduate management admission test: GMAT) ที่พบว่าผลคะแนนการทดสอบไม่มีความคงเส้นคงวาและมีการประมาณค่าความสามารถของผู้สอบกว่าพันคนคลาดเคลื่อนไปจากความเป็นจริง (Zheng & Chang, 2015)

การทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ยังมีข้อจำกัดในเชิงปฏิบัติคือ ไม่อนุญาตให้ผู้สอบย้อนกลับไปเปลี่ยนคำตอบ เนื่องจากข้อสอบแต่ละข้อถูกใช้ในการประมาณค่าความสามารถของผู้สอบ นอกจากนี้ การควบคุมอัตราการเปิดเผยของข้อสอบ การควบคุมความเท่าเทียมของเนื้อหาซึ่งทำได้ยาก และเนื่องจากการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์เป็นการเลือกข้อสอบจากคลังมาจัดเป็นชุดแบบทดสอบ ทำให้มีจำนวนชุดแบบทดสอบแบบไม่จำกัด จึงเป็นไปได้ยากที่ผู้เชี่ยวชาญหรือคณะกรรมการจัดการทดสอบจะสามารถตรวจสอบคุณภาพของแบบทดสอบทุกชุดได้ (Luecht & Nungester, 1998 cited in Patsula, 199)

ด้วยเหตุที่ การทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์แบบดั้งเดิม (CAT) ยังมีข้อจำกัดบางประการ จึงมีการพัฒนาการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์แบบหลายขั้นตอน (multistage adaptive tests: MST) ขึ้น ซึ่งมีงานวิจัยที่ศึกษาเปรียบเทียบระหว่างการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์แบบดั้งเดิมกับการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์แบบหลายขั้นตอน พบว่าการเพิ่มจำนวนขั้นตอน (stage) ในการทดสอบ ช่วยลดความคลาดเคลื่อนในการประมาณค่าความสามารถของผู้สอบได้ (Patsula, 1999)

ตอนที่ 2 มโนทัศน์เกี่ยวกับการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์แบบหลายขั้นตอน

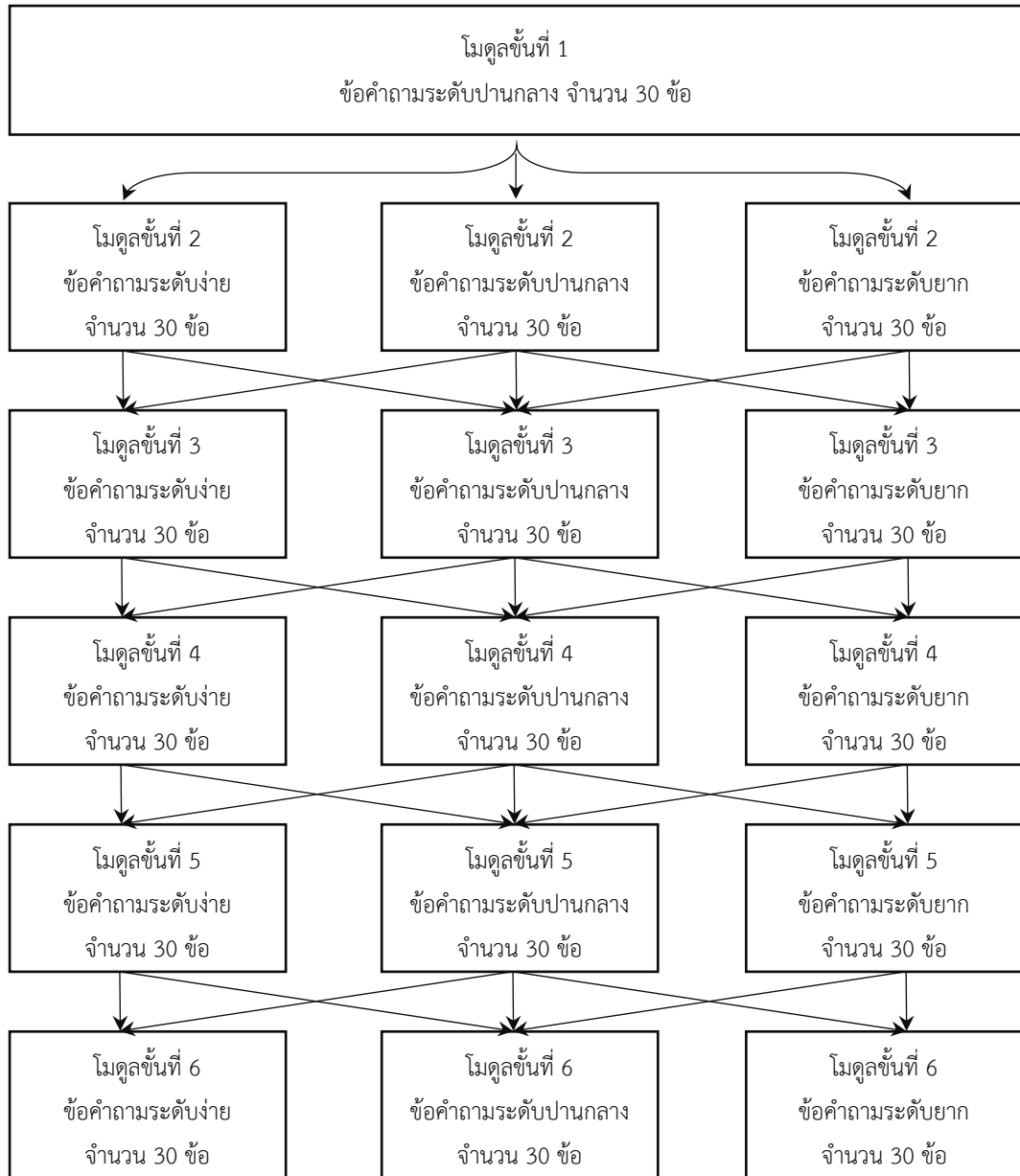
การทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์แบบหลายขั้นตอน (MST) ได้รับความนิยมนมากขึ้นในช่วงหลายปีที่ผ่านมา เนื่องจากเป็นการผสมผสานและรวมข้อดีของการทดสอบโดยใช้กระดาษ และการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์แบบดั้งเดิม กล่าวคือ ในการทดสอบแบบ MST นักพัฒนาแบบทดสอบสามารถทบทวนชุดแบบทดสอบได้ก่อนที่จะนำไปใช้ทดสอบ ส่วนผู้ทดสอบยัง

สามารถข้ามข้อคำถาม ทบทวน หรือแก้ไขคำตอบในชั้นที่กำลังทดสอบได้อีกด้วย (Mead, 2006; Sari, Sari, & Huggins-Manley, 2016; Zheng, Nozawa, Gao, & Chang, 2012)

นักวิจัยที่เริ่มนำแนวคิดการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์แบบหลายขั้นตอน (MST) มาใช้คือ Lord และ Novick ซึ่งเริ่มแรกได้นำมาใช้กับการทดสอบแบบสองขั้น (two-stage testing) และพบว่าผลการทดสอบมีความแม่นยำมากขึ้นในการประมาณค่าความสามารถ (Zenisky, 2010) ซึ่งข้อสอบในการทดสอบแบบ MST จะมีการจัดชุดไว้ล่วงหน้า (pre-assembled) โดยแบ่งเป็นหลายขั้น แต่ละขั้นจะประกอบด้วยชุด (panel) ข้อสอบหลายชุด แต่ละชุดจะประกอบไปด้วยจำนวนโมดูล (modules) ของข้อสอบที่แน่นอน ซึ่งโมดูลจะมีระดับความยากที่หลากหลาย ระหว่างการทดสอบ ผู้สอบแต่ละคนจะได้รับชุดข้อสอบที่มีการเตรียมไว้ล่วงหน้าอย่างสุ่ม (random) และจะเริ่มทำการทดสอบในโมดูลสำหรับขั้นที่หนึ่ง (stage 1) หลังจากทำข้อสอบในขั้นที่หนึ่งเสร็จเรียบร้อยแล้วในขั้นถัดไป ผู้สอบจะได้รับการจัดชุดโมดูลที่มีระดับความยากเหมาะสมกับความสามารถของตน โดยแต่ละโมดูลจะได้รับการเลือกผ่านเส้นทาง (pathway) ของแต่ละโมดูลจากชุดข้อสอบที่ถูกเลือก (Zheng & Chang, 2015) กล่าวคือโมดูลในขั้นที่หนึ่งเป็นการทดสอบเพื่อประมาณค่าความสามารถของผู้สอบ และในโมดูลถัดไป ผู้สอบจะได้รับการทดสอบด้วยชุดแบบทดสอบที่มีความยากสอดคล้องกับความสามารถของตนเอง ดังนั้น การทดสอบแบบ MST จึงเป็นการปรับชุดของข้อคำถามให้เหมาะสมกับความสามารถของผู้สอบ แทนการปรับข้อคำถามที่ละเอียดตามความสามารถของผู้สอบ (Sari et al., 2016; Wang, 2017)

2.1 องค์ประกอบของการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์แบบหลายขั้นตอน

ในการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์แบบหลายขั้นตอน (MST) มีกระบวนการที่แตกต่างจากการทดสอบการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ (CAT) ซึ่งองค์ประกอบของกระบวนการที่สำคัญ ได้แก่ ชุด (panel) โมดูล (module) ขั้น (stage) และเส้นทาง (pathway) ดังภาพที่ 2 (Zheng et al., 2012)



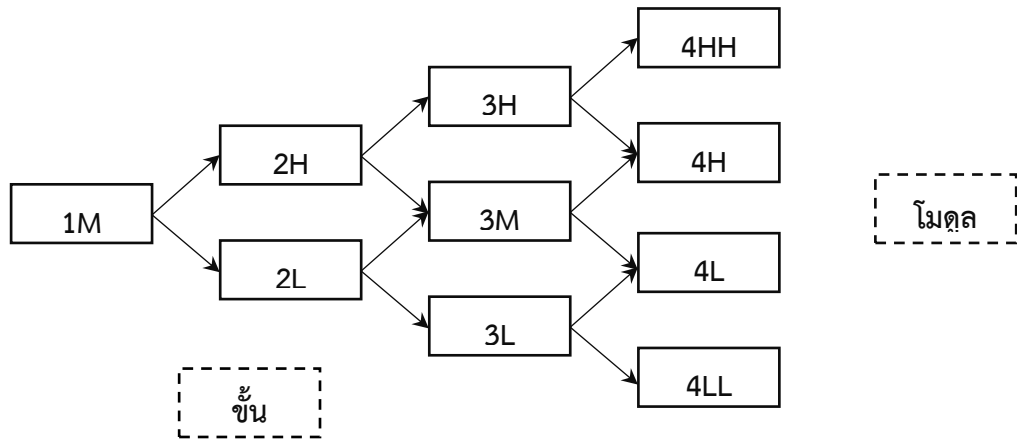
ภาพที่ 2 แผนภาพองค์ประกอบและเส้นทางการทำงานของ
การทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์แบบหลายขั้นตอน (MST)

ในการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์แบบหลายขั้นตอน (MST) ประกอบด้วย 4 องค์ประกอบสำคัญ โดยมีรายละเอียดดังนี้

— ชุด (panel) ในชุดข้อสอบหนึ่งชุดจะประกอบด้วยข้อสอบที่มีโมดูลความยากคละกัน และในแต่ละชุดข้อสอบจะต้องมีจำนวนโมดูลที่เท่ากัน ซึ่งหมายความว่าข้อสอบต้องเป็นข้อสอบคู่ขนานกัน โดยชุดข้อสอบจะมีที่ชุดก็ได้ ขึ้นอยู่กับผู้บริหารการทดสอบจะจัดเตรียมชุดข้อสอบจากคลังข้อสอบไว้ล่วงหน้าก่อนทำการทดสอบจริง

— โมดูล (module or testlets) ในหนึ่งชุดจะสามารถประกอบไปด้วยข้อสอบหลายโมดูลขึ้นอยู่กับการจัดข้อสอบ ในแต่ละโมดูลสามารถกำหนดระดับความยากได้ตามความต้องการขึ้นอยู่กับผู้บริหารการทดสอบ ซึ่งโดยทั่วไปจะประกอบด้วยระดับความยากตั้งแต่ ง่าย (easy) ปานกลาง (medium) และยาก (difficult)

— ชั้น (stage) คือขั้นตอนในทดสอบ สามารถแบ่งได้ตามความต้องการของผู้บริหารการทดสอบ โดยชั้นแรกจะเป็นชั้นที่มีระดับความยากปานกลางเพื่อใช้ในการประมาณค่าความสามารถของผู้สอบเพื่อคัดเลือกโมดูลข้อสอบที่เหมาะสมกับผู้สอบในชั้นถัดไป ดังตัวอย่างในภาพที่ 3 แสดงข้อมูลโมเดล 1-2-3-4 ของการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์แบบหลายขั้นตอน



ภาพที่ 3 รูปแบบโมเดล 1-2-3-4 ของการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์แบบหลายขั้นตอน

— เส้นทาง (pathway or routing) เมื่อผู้สอบทำการทดสอบเสร็จในแต่ละชั้น ระบบจะทำการประมาณค่าความสามารถของผู้สอบ หลังจากนั้นระบบจะค้นหาโมดูลตามเส้นทางที่เหมาะสมกับความสามารถของผู้สอบ เมื่อพบโมดูลที่เหมาะสมจะดำเนินการวิ่งไปตามเส้นทางเพื่อนำผู้สอบไปยังโมดูลที่เหมาะสมตามความสามารถของผู้สอบ

2.2 จุดแข็งและจุดอ่อนของการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์แบบหลายขั้นตอน

แม้ว่าการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์แบบหลายขั้นตอน (MST) จะถูกพัฒนาขึ้นเพื่อแก้ปัญหาบางประการของการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์แบบดั้งเดิม (CAT) ซึ่งกลายเป็นจุดแข็งของการทดสอบแบบ MST นั่นคือ กระบวนการเลือกข้อสอบจะกระทำเฉพาะระหว่างชั้น (stage) เท่านั้น ไม่ได้ใช้กระบวนการเลือกภายในชั้น ดังนั้น ผู้สอบจึงสามารถย้อนหรือข้ามข้อสอบภายในชั้นที่กำลังดำเนินการทดสอบได้ อีกทั้งยังสามารถทบทวนหรือเปลี่ยนคำตอบภายในชั้นที่กำลังทดสอบได้ ซึ่งต่างจากการทดสอบแบบ CAT ซึ่งผู้สอบไม่สามารถกลับไปแก้ไข

ข้อสอบก่อนหน้าได้ ดังนั้นการทดสอบแบบ MST จึงช่วยลดความเครียดและความวิตกกังวลของผู้สอบในระหว่างการทดสอบ นอกจากนี้กระบวนการในการจัดการทดสอบแบบ MST ยังอนุญาตให้ผู้พัฒนาแบบทดสอบสามารถกำกับติดตามและควบคุมคุณภาพของชุดแบบทดสอบ เนื่องจากชุดแบบทดสอบจะถูกเตรียมขึ้นก่อนการทดสอบ ดังนั้นผู้เชี่ยวชาญในการพัฒนาการทดสอบสามารถทบทวนชุดแบบทดสอบจากหลากหลายมุมมอง เช่น การควบคุมความเท่าเทียมของเนื้อหา (content balance) ความเท่าเทียมกันของการเฉลยคำตอบ (answer key) เป็นต้น (Zheng et al., 2012) นอกจากนี้ในการประมาณค่าความสามารถของผู้สอบเพื่อคัดเลือกชุดข้อสอบที่เหมาะสมกับความสามารถของผู้สอบจะทำการประมาณค่าจากชุดของข้อสอบในขั้นแรก ซึ่งต่างจากการทดสอบแบบ CAT จึงช่วยลดความผิดพลาดจากการประมาณค่าความสามารถผู้สอบจากข้อสอบข้อแรกเพียงข้อเดียว ซึ่งหากเกิดความผิดพลาดในการประมาณค่าความสามารถจากข้อสอบข้อแรกจะทำให้ผู้สอบอาจให้เวลานานกว่าปกติในการทดสอบ (Zheng & Chang, 2015) ในปัจจุบัน การทดสอบมาตรฐานหรือการทดสอบระดับสากลของต่างประเทศ เช่น การทดสอบเพื่อการศึกษาต่อในระดับบัณฑิตศึกษา (The Graduate Record Examination: GRE®) จึงมีการพัฒนารูปแบบการทดสอบมาเป็นแบบ MST (Dallas, 2014; Wang, 2017)

อย่างไรก็ดี การทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์แบบหลายขั้นตอน (MST) ยังมีจุดอ่อน คือ ต้องมีการจัดเตรียมชุดข้อสอบ (panel) และโมดูล (module) ของข้อสอบไว้ล่วงหน้า ต้องใช้คนในการดำเนินการ ดังนั้นหากเป็นการทดสอบที่มีขนาดใหญ่อาจเกิดความผิดพลาดในการจัดชุดข้อสอบได้ จึงถือว่าเป็นอีกหนึ่งจุดอ่อนที่สำคัญของการทดสอบด้วยวิธีนี้ (Zheng & Chang, 2015) ปัจจุบันได้มีการพัฒนาการจัดชุดข้อสอบแบบอัตโนมัติขึ้น เพื่อแก้ไขปัญหาความผิดพลาดที่เกิดขึ้นจากคนในการจัดชุดข้อสอบ เรียกว่าการจัดชุดข้อสอบแบบออนเดอะฟลาย (on-the-fly) โดยมีการศึกษากระบวนการจัดชุดข้อสอบแบบอัตโนมัติ (automated test assembly) เพื่อใช้สำหรับการทดสอบแบบ MST และมีการพัฒนาการจัดชุดข้อสอบอัตโนมัติแบบฮิวริสติกขึ้น ซึ่งพบว่าการจัดชุดข้อสอบอัตโนมัติแบบฮิวริสติกมีสมรรถนะเหมาะกับการนำไปใช้มากกว่ากระบวนการจัดชุดข้อสอบด้วยวิธีอื่น (Zheng, Nozawa, Zhu, & Gao, 2014)

การทดสอบแบบปรับเหมาะหลายขั้นตอนถูกพัฒนาขึ้นเพื่อลดการประมาณค่าที่ผิดพลาดจากการประมาณค่าความสามารถของผู้สอบจากข้อสอบเพียงข้อเดียว ซึ่งในการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์แบบหลายขั้นตอน (MST) จะประมาณค่าความสามารถจากชุดข้อสอบในขั้นแรก (initial stage) จากการศึกษาวิจัยได้มีการเปรียบเทียบการทดสอบแบบ MST กับการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์แบบดั้งเดิม (CAT) โดยการจำลองข้อมูลการทดสอบเพื่อรับใบประกอบวิชาชีพแพทย์ และเป็นการทดสอบขนาดกลาง (medium-size) พบว่าการทดสอบแบบ MST มีการประมาณค่าความสามารถที่ถูกต้องและแม่นยำมากกว่าการทดสอบแบบ CAT (Brossman & Guille, 2014) ซึ่งสอดคล้องกับงานวิจัยที่นำวิธีการทดสอบแบบ MST มาใช้ในการ

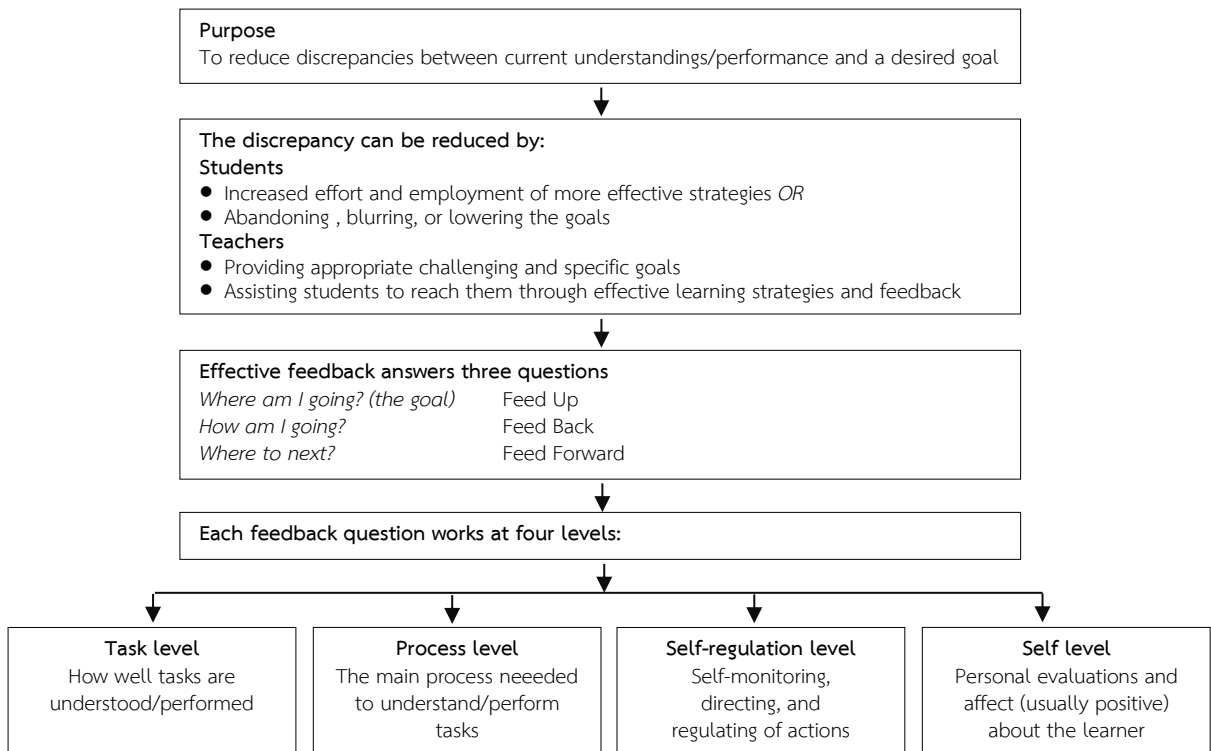
การประเมินนานาชาติด้านสมรรถนะวัยผู้ใหญ่ (The Programme for the International Assessment of Adult Competencies: PIAAC) ซึ่งเป็นการประเมินขนาดใหญ่ระดับนานาชาติ (international large-scale assessments) พบว่าการประเมินด้วยวิธีทดสอบแบบ MST มีประสิทธิภาพสูงกว่าการประเมินด้วยรูปแบบดั้งเดิม (nonadaptive) ทั้งในกลุ่มผู้สอบที่มีสมรรถนะสูงและต่ำ นอกจากนี้ ข้อสอบที่ใช้ยังครอบคลุมเนื้อหาที่ต้องการ และเชื่อว่าการออกแบบวิธีการทดสอบเป็นชุดข้อคำถามยังช่วยลดปัญหาการทำหน้าที่ต่างกันของข้อสอบ (differential item functioning) ซึ่งเป็นผลกระทบที่เกิดจากความแตกต่างระหว่างประเทศอีกด้วย (Yamamoto, Khorramdel, & Shin, 2018) และจากการศึกษาด้วยการจำลองข้อมูลสำหรับการทดสอบขนาดใหญ่ (large-scale) พบว่าความถูกต้องในการจำแนกของการทดสอบแบบ MST ดีกว่าการทดสอบแบบ CAT (Zheng et al., 2012) นอกจากนี้ จากงานวิจัยที่ศึกษาในนักเรียนประถมศึกษาและมัธยมศึกษาตอนต้น พบว่าการทดสอบแบบ MST ช่วยให้การวัดผลสัมฤทธิ์ของนักเรียนมีความแม่นยำขึ้น มีผลทางบวกต่อแรงจูงใจและความเชื่อมั่นผูกพันต่อการสอบของนักเรียน และนักวิจัยยังแนะนำให้ใช้วิธีการทดสอบแบบ MST กับนักเรียนที่มีแรงจูงใจและความเชื่อมั่นผูกพันต่อการสอบต่ำอีกด้วย (Martin & Lazendic, 2018)

ตอนที่ 3 มโนทัศน์เกี่ยวกับการให้ข้อมูลป้อนกลับ

ข้อมูลป้อนกลับ หรือข้อมูลย้อนกลับ (feedback) เป็นมโนทัศน์ที่ถูกกล่าวถึงเป็นอย่างมากทั้งในบริบททางธุรกิจและบริบททางการศึกษา เพราะข้อมูลป้อนกลับมีอิทธิพลอย่างมากต่อการเรียนรู้ของมนุษย์ (Hattie & Timperley, 2007) ในบริบททางการศึกษา การให้ข้อมูลป้อนกลับในชั้นเรียนเป็นมโนทัศน์ที่เกี่ยวกับการได้รับข้อมูลข้อมูลด้านความสามารถหรือความเข้าใจของผู้เรียน จากผู้มีส่วนได้ส่วนเสีย เช่น ครู เพื่อน หนังสือ ประสบการณ์ หรือตนเอง และการให้ข้อมูลป้อนกลับยังเป็นองค์ประกอบหลักของการประเมินความก้าวหน้าระหว่างเรียน (formative assessment) ซึ่งเป็นปัจจัยที่มีความสำคัญในการเรียนรู้สำหรับผู้เรียน วัตถุประสงค์หลักของการให้ข้อมูลป้อนกลับคือการปิดช่องว่างระหว่างความสามารถในปัจจุบันของผู้เรียนกับความสามารถที่ปรารถนา (Hattie & Timperley, 2007; Ludvigsen, Krumsvik, & Furnes, 2015)

Hattie & Timperley (2007) ได้เสนอกรอบแนวคิดของการให้ข้อมูลป้อนกลับโดยเริ่มตั้งแต่การกำหนดวัตถุประสงค์ของการให้ข้อมูลป้อนกลับว่าเพื่อลดความแตกต่างระหว่างความเข้าใจหรือประสิทธิภาพ ณ ปัจจุบัน กับเป้าหมายที่คาดหวัง ซึ่งความแตกต่างดังกล่าวจะสามารถลดลงได้จากการที่นักเรียนมีความพยายามเพิ่มขึ้น มีกลยุทธ์ที่มีประสิทธิภาพมากขึ้น หรือนักเรียนอาจตัดทอนหรือลดเป้าหมายให้ต่ำลง และยังสามารถลดลงจากครูที่จัดหาเป้าหมายที่ท้าทายนักเรียนอย่างเหมาะสม มีเป้าหมายเฉพาะเจาะจงลงไป และให้การช่วยเหลือนักเรียนให้ไปถึงเป้าหมายโดยใช้กลยุทธ์ในการเรียนรู้และข้อมูลป้อนกลับ ซึ่งข้อมูลป้อนกลับที่มีประสิทธิภาพจะต้องตอบคำถามได้ว่าเป้าหมายที่

นักเรียนต้องไปให้ถึงนั่นคืออะไร จะไปสู่เป้าหมายนั้นได้อย่างไร และควรจะทำอย่างไรต่อไป ซึ่งแต่ละคำถามมีระดับการให้ข้อมูลป้อนกลับแตกต่างกันแบ่งได้เป็น 4 ระดับ คือ ได้แก่ 1) ระดับความสามารถตามหน้าที่ 2) ระดับกระบวนการ 3) ระดับการกำกับตนเอง และ 4) ระดับส่วนบุคคล ดังภาพที่ 4



ภาพที่ 4 กรอบแนวคิดการให้ข้อมูลป้อนกลับ (Hattie & Timperley, 2007)

การให้ข้อมูลป้อนกลับนั้น แสดงให้เห็นถึงสถานะทางการเรียน ณ ขณะปัจจุบันของผู้เรียน ช่วยเสริมสร้างแรงจูงใจให้ผู้เรียนมีเป้าหมายผลสัมฤทธิ์ทางการเรียนที่สูงขึ้น และมีศักยภาพที่พัฒนาขึ้น โดยเฉพาะอย่างยิ่งหากได้รับข้อมูลป้อนกลับจากผู้สอนร่วมกับการใช้เครื่องมือในการพัฒนาอื่น ๆ (Baranczyk & Best, 2019)

การสะท้อนข้อมูลป้อนกลับมีอยู่หลากหลายรูปแบบ เช่น การให้ข้อมูลป้อนกลับระหว่างเรียน (formative feedback) การให้ข้อมูลป้อนกลับหลังการเรียน (summative feedback) การให้ข้อมูลป้อนกลับเชิงประเมิน (evaluative feedback) การให้ข้อมูลป้อนกลับแบบอิงกลุ่ม (norm-referenced feedback) เป็นต้น ซึ่งการเลือกใช้ข้อมูลป้อนกลับแบบใดนั้นขึ้นอยู่กับแนวทางที่จะนำไปใช้งาน (Batool & Akhter, 2019) โดยวัตถุประสงค์ของการสะท้อนข้อมูลป้อนกลับนอกจากจะเพื่อการลดช่องว่างระหว่างความสามารถในปัจจุบันของผู้เรียนกับระดับความสามารถที่พึงปรารถนาแล้วนั้น ยังมีวัตถุประสงค์เพื่อประเมินผลสัมฤทธิ์ของผู้เรียน และที่สำคัญคือเพื่อเสริมสร้างแรงจูงใจและความมั่นใจของผู้เรียน และยกระดับความสามารถของผู้เรียนให้สูงขึ้น (Hyland, 2000 cited in Hatziapostolou & Paraskakis, 2010) โดยมีงานวิจัยที่พบว่าการให้ข้อมูลป้อนกลับทำให้คะแนน

เฉลี่ยของผู้เรียนกลุ่มอ่อน (low performers) สูงขึ้นอย่างมีนัยสำคัญทางสถิติ (Baranczyk & Best, 2019) นอกจากนี้ ยังมีงานวิจัยที่ทดลองใช้วิธีการให้ข้อมูลป้อนกลับรูปแบบต่าง ๆ ได้แก่ การให้ข้อมูลป้อนกลับแบบมีแบบแผนหรือเป็นทางการ (planned or formal feedback) และการให้ข้อมูลป้อนกลับแบบปฏิสัมพันธ์หรือไม่เป็นทางการ (interactive or informal feedback) ที่มีผลต่อผลสัมฤทธิ์ในรายวิชาคณิตศาสตร์ พบว่าตัวอย่างในกลุ่มทดลองที่ได้รับการให้ข้อมูลป้อนกลับมีคะแนนทดสอบหลังเรียน (posttest) สูงกว่าคะแนนทดสอบก่อนเรียน (pretest) และแตกต่างกับกลุ่มควบคุมอย่างมีนัยสำคัญทางสถิติ (Batool & Akhter, 2019) จากงานวิจัยแสดงให้เห็นถึงประสิทธิผลของการให้ข้อมูลป้อนกลับในการยกระดับผลสัมฤทธิ์ของผู้เรียน

ส่วนการให้ข้อมูลป้อนกลับแก่ผู้เรียนที่เรียนภาษาต่างประเทศเป็นภาษาที่สอง (second language) พบว่าข้อมูลป้อนกลับมีความสำคัญต่อกระบวนการเรียนรู้ของผู้เรียนอย่างมาก มีงานวิจัยที่พบว่าการให้ข้อมูลป้อนกลับที่มีประสิทธิภาพช่วยให้ผู้เรียนพัฒนาคุณภาพและทักษะด้านการเขียน กระตุ้นการให้เหตุผลเชิงวิพากษ์ และส่งเสริมความเป็นตัวของตัวเองของผู้เรียน (Fithriani, 2019)

นอกจากนี้ ยังพบว่าการให้ข้อมูลป้อนกลับมีความสำคัญต่อการเรียนการสอนภาษาต่างประเทศ ซึ่งไม่ว่าจะเป็นการให้ข้อมูลป้อนกลับโดยครูผู้สอน หรือจากเพื่อน มีผลทำให้ผู้เรียนมีความสามารถทางภาษาต่างประเทศในด้านการเขียนเพิ่มมากขึ้น (Jahin, 2012; Kamberi, 2013)

จากงานวิจัยข้างต้น แสดงให้เห็นว่าการให้ข้อมูลป้อนกลับมีผลทางบวกต่อผลสัมฤทธิ์ และความสามารถในการเรียนรู้หรือสมรรถนะของผู้เรียน และถึงแม้ว่าการให้ข้อมูลป้อนกลับนั้นจะช่วยในการพัฒนาผู้เรียนได้ แต่ในบางครั้งก็อาจลดทอนกำลังใจของผู้เรียนได้เช่นกัน Iron (2008) ได้เสนอแนวทางในการให้ข้อมูลป้อนกลับที่เหมาะสมไว้ว่าครูควรให้ข้อมูลป้อนกลับในเวลาที่เหมาะสม อย่างทันท่วงที สะท้อนถึงสมรรถนะของผู้เรียนในปัจจุบัน มีเป้าหมายที่ชัดเจน และทำให้นักเรียนมีสมรรถนะเพิ่มขึ้น นอกจากนี้ ครูควรให้ข้อมูลป้อนกลับแก่นักเรียนทุกคนอย่างเท่าเทียมกัน ให้สารสนเทศที่ครบถ้วน และสอดคล้องกัน มีจุดเน้นที่เฉพาะเจาะจงเพื่อให้ผู้เรียนเกิดแรงจูงใจในการพัฒนาตนเอง นอกจากนี้ Shute (2007) ยังได้เสนอแนวทางที่ควรปฏิบัติและควรหลีกเลี่ยงในการให้ข้อมูลป้อนกลับระหว่าง สรุปได้ดังตารางที่ 1

ตารางที่ 1 แนวทางในการให้ข้อมูลป้อนกลับที่พึงปฏิบัติและพึงหลีกเลี่ยง

การให้ข้อมูลป้อนกลับที่พึงปฏิบัติ	การให้ข้อมูลป้อนกลับที่พึงหลีกเลี่ยง
1. ควรให้ข้อมูลป้อนกลับที่ระบุแนวทางการเรียนรู้ของผู้เรียนที่เชื่อมโยงกับภารกิจ โดยให้คำแนะนำว่าควรพัฒนาอย่างไร	1. ควรหลีกเลี่ยงการให้ข้อมูลป้อนกลับที่เปรียบเทียบกับผู้เรียนคนอื่นทั้งทางตรงและทางอ้อม
2. ควรให้ข้อมูลป้อนกลับที่อธิบายว่าปัญหาเกิดจากอะไร อย่างไร และ/หรือทำไม	2. ควรหลีกเลี่ยงการให้ข้อมูลป้อนกลับเฉพาะภาพรวม แต่ควรระบุจุดแข็งและวิธีการพัฒนาตนเอง
3. ควรให้ข้อมูลป้อนกลับเป็นลำดับขั้นเพื่อควบคุมความผิดพลาดและให้สารสนเทศที่เพียงพอกับผู้เรียน	3. ควรหลีกเลี่ยงการให้ข้อมูลป้อนกลับที่ทำให้นักเรียนหมดกำลังใจ และบ่อนทำลายการเรียนรู้ของนักเรียน
4. ควรให้ข้อมูลป้อนกลับที่ชัดเจน และสอดคล้องกับเป้าหมาย เพราะข้อมูลป้อนกลับที่ไม่ชัดเจน อาจขัดขวางการเรียนรู้และทำให้ผู้เรียนท้อแท้ได้	4. ควรหลีกเลี่ยงการชมเชยมากเกินไป เพราะอาจทำให้ผู้เรียนเกิดความสับสนต่อผลลัพธ์ที่ได้จากการเรียนรู้
5. ควรให้ข้อมูลป้อนกลับที่เข้าใจง่าย ไม่ซับซ้อน	5. ควรหลีกเลี่ยงการให้ข้อมูลป้อนกลับด้วยการพูด เพราะอาจเกิดการเอนเอียง แต่ควรให้ข้อมูลป้อนกลับในรูปแบบที่มีความเป็นกลาง เช่น การให้ข้อมูลป้อนกลับผ่านคอมพิวเตอร์
6. ควรให้ข้อมูลป้อนกลับที่ทำให้เกิดความชัดเจนระหว่างสมรรถนะในการเรียนรู้ของผู้เรียนกับสิ่งที่ต้องทำเพื่อบรรลุเป้าหมาย	6. ควรหลีกเลี่ยงการให้ข้อมูลป้อนกลับแก่นักเรียนที่มีความใส่ใจในการเรียนอยู่แล้ว เพราะอาจขัดขวางการเรียนรู้ของนักเรียน
7. ข้อมูลป้อนกลับควรมาจากแหล่งที่มีความน่าเชื่อถือ ไม่เอนเอียง นอกจากนี้การให้ข้อมูลป้อนกลับผ่านทางคอมพิวเตอร์จะลดความเอนเอียงได้มากกว่าการให้ข้อมูลผ่านบุคคล	7. ควรหลีกเลี่ยงการบอกคำตอบที่ถูกต้อง เพราะผู้เรียนอาจนำไปใช้ในการที่ผิด จึงควรพิจารณาความพร้อมในการให้ข้อมูลป้อนกลับที่เหมาะสม
8. ควรให้ข้อมูลป้อนกลับที่ส่งเสริมสมรรถนะและทิศทางการเรียนรู้	8. ควรหลีกเลี่ยงการให้ข้อมูลป้อนกลับเป็นข้อความเพียงอย่างเดียว แต่อาจนำเสนอข้อมูลป้อนกลับในรูปแบบอื่นควบคู่ไปด้วย
9. ควรให้ข้อมูลป้อนกลับเมื่อผู้เรียนพยายามปรับปรุง และแก้ไขปัญหา	9. ควรหลีกเลี่ยงการให้ข้อมูลป้อนกลับที่เป็นการวิเคราะห้ความผิดพลาด เพราะอาจไม่ให้ประโยชน์ที่เพียงพอต่อผู้เรียน

นอกจากการให้ข้อมูลป้อนกลับจะมีผลทางบวกต่อผลสัมฤทธิ์ทางการศึกษาแล้ว ยังมีงานวิจัยที่ศึกษาโดยนำการให้ข้อมูลป้อนกลับ (feedback) มาใช้ในการพิจารณาทักษะของความเป็นผู้นำเป็นระยะเวลานาน เพื่อให้องค์กรบรรลุเป้าหมายที่ตั้งไว้ แต่ปัญหาเบื้องต้นของการให้ข้อมูลป้อนกลับคือเน้นกิจกรรมที่เกิดขึ้นในอดีต และมีโอกาสน้อยมากที่จะเกิดขึ้นอีกในอนาคต ดังนั้นการให้ข้อมูลป้อนกลับจึงมีข้อจำกัด ไม่เป็นพลวัต จึงได้มีการศึกษาผู้นำมากกว่าห้าพันคนโดยให้ฝึกปฏิบัติ ซึ่งในการฝึกปฏิบัติกำหนดให้มีบทบาทหนึ่งคือการถามเพื่อให้ข้อมูลป้อนกลับเพื่อปรับปรุง (feedforward) โดยให้ผู้นำให้ข้อมูลสำหรับอนาคตมากเท่าที่จะสามารถทำได้ และอีกบทบาทหนึ่งคือการรับฟังข้อมูลป้อนกลับเพื่อการปรับปรุง โดยใช้เวลาประมาณ 10-15 นาที ในแต่ละบทบาท หลังจากเสร็จสิ้นการฝึกปฏิบัติให้ผู้นำให้คำนิยามหนึ่งคำสำหรับประสบการณ์ที่ได้ ซึ่งส่วนใหญ่เป็นคำตอบในเชิงบวก เช่น เยี่ยม เป็นประโยชน์ และคำตอบในภาพรวมคือทุกคนสนุก เพราะว่าการฝึกปฏิบัติไม่ทำให้รู้สึกแย่ ไม่ทำให้เสียหน้า ไม่ทำให้รู้สึกไม่สบายใจ อย่างไรก็ตาม มีได้หมายความว่าไม่ต้องให้ข้อมูลป้อนกลับแบบปกติ แต่ต้องการแสดงให้เห็นว่าการให้ข้อมูลป้อนกลับเพื่อการปรับปรุงสามารถทำได้บ่อยกว่า รวมถึงประสิทธิภาพและประสิทธิผลที่มีมากกว่า การให้ข้อมูลป้อนกลับเพื่อการปรับปรุงทำให้การดำเนินชีวิตมีความสุขมากขึ้น (Goldsmith, 2003)

มีงานวิจัยที่พัฒนารูปแบบการให้ข้อมูลป้อนกลับให้ครอบคลุม และเป็นประโยชน์ต่อผู้เรียนมากขึ้น โดยมีการศึกษานำข้อมูลป้อนกลับมาประยุกต์ใช้ในรูปแบบแฟ้มสะสมงานซึ่งประกอบด้วยข้อมูลป้อนกลับ 2 ส่วน ได้แก่ ข้อมูลป้อนกลับทั่วไป (feedback) และการให้ข้อมูลป้อนกลับเพื่อการปรับปรุง (feedforward) การให้ข้อมูลป้อนกลับทั่วไป ได้แก่ การให้ข้อความที่สามารถอธิบายผลงานของนักเรียนว่าถูกหรือผิดและดีหรือไม่ดีอย่างไร โดยอิงจากเกณฑ์ในการประเมิน ส่วนการให้ข้อมูลป้อนกลับเพื่อการปรับปรุง ได้แก่ การให้ข้อความซึ่งเป็นข้อมูลที่สามารถนำไปใช้เป็นแนวทางในการพัฒนาและปรับปรุงผลงานในอนาคต กระบวนการให้ข้อมูลป้อนกลับทั้งสองส่วน เมื่อนำมาใช้ร่วมกันเรียกว่าเป็นการสะท้อนข้อมูลป้อนกลับ (reflective feedback) ซึ่งหมายถึงผู้เรียนจะได้รับข้อมูลผลการเรียนตามความสามารถ ซึ่งจะตัดสินตามเกณฑ์ เช่น ผู้เรียนได้ 70 คะแนนจาก คะแนนเต็ม 100 คะแนน และเมื่อทราบคะแนนแล้วก็จะนำข้อมูลมาสะท้อนกลับไปยังผู้เรียนโดยการให้คำแนะนำเพื่อการปรับปรุง เช่น คะแนนส่วนใหญ่ได้มาจากเนื้อหาส่วนใดและควรพัฒนาในส่วนใด เป็นต้น (โชติกาภาชีผล, ประกอบ กรณีกิจ, & พิทักษ์ โสตถยาคม, 2558)

การศึกษาการสะท้อนข้อมูลป้อนกลับโดยศึกษาจากครูฝึกสอนซึ่งสะท้อนข้อมูลป้อนกลับไปยังครูพี่เลี้ยง พบว่าความคาดหวังหรือความต้องการของครูฝึกสอนที่ต้องการจากครูพี่เลี้ยงมีมากกว่าประสบการณ์ที่ได้รับจริงจากการฝึกสอน แสดงให้เห็นว่าครูฝึกสอนได้รับการช่วยเหลือในด้านต่าง ๆ ไม่เพียงพอ อาจเกิดมาจากภาระงานที่มากเกินไปของครูพี่เลี้ยง หรือครูพี่เลี้ยงมีความรู้ในการสอนงานไม่เพียงพอ ดังนั้นจึงต้องจัดสรรภาระงานของครูพี่เลี้ยงหรือทำการพัฒนาความรู้ของครูพี่เลี้ยงในด้านต่าง ๆ ให้เพียงพอต่อการสอนงานครูฝึกสอน (Damar, 2013) และยังมีการนำการสะท้อนข้อมูล

ป้อนกลับมาใช้ในการเพิ่มประสิทธิภาพการพัฒนาทักษะการให้เหตุผลทางคลินิกของผู้เรียนอีกด้วย (Wojcikowski & Brownie, 2013)

การทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์โดยทั่วไป จะรายงานผลโดยให้ข้อมูลป้อนกลับแบบปกติ ซึ่งต่อมา ฌักทรี ชัยมงคล (2558) ได้พัฒนาระบบการทดสอบแบบปรับเหมาะหลายขั้นตอนด้วยวิธีออนเดอะฟลาย (on the fly assembled multistage adaptive testing: OMST) ในการทดสอบมาตรฐานวิชาชีพของบุคลากรสาขาไอที และได้้นำการให้ข้อมูลป้อนกลับแบบปกติ (feedback) และการให้ข้อมูลป้อนกลับเพื่อการปรับปรุง (feedforward) มาใช้ร่วมกัน เรียกว่าการสะท้อนข้อมูลป้อนกลับ (reflective feedback) โดยรายงานคะแนนเป็นข้อมูลป้อนกลับแบบปกติให้แก่ผู้สอบทุกคนทันทีหลังการทดสอบ แต่ระบบจะให้การสะท้อนข้อมูลป้อนกลับแก่ผู้ที่ผ่านเกณฑ์การทดสอบเท่านั้น โดยแสดงเป็นข้อความที่ให้แนวทางในการพัฒนาตนเองและแนวโน้มอาชีพที่เหมาะสม จึงเป็นสารสนเทศที่ผู้สอบนำไปใช้ในการพัฒนาได้อย่างสอดคล้องกับความสามารถของตน

ตอนที่ 4 มโนทัศน์เกี่ยวกับการทดสอบระดับชาติ (O-NET) วิชาภาษาอังกฤษ

การทดสอบระดับชาติ (Ordinary National Educational Test: O-NET) คือ การทดสอบทางการศึกษาระดับชาติขั้นพื้นฐานเป็นการทดสอบเพื่อวัดความรู้และความคิดของนักเรียนชั้นประถมศึกษาปีที่ 6 ชั้นมัธยมศึกษาปีที่ 3 และชั้นมัธยมศึกษาปีที่ 6 ประเมินตามมาตรฐานการเรียนรู้ในหลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พุทธศักราช 2551 โดยการทดสอบประกอบด้วย 5 วิชา ได้แก่ คณิตศาสตร์ วิทยาศาสตร์ ภาษาอังกฤษ ภาษาไทย และ สังคมศึกษา ศาสนาและวัฒนธรรม (สถาบันทดสอบทางการศึกษาแห่งชาติ, 2562) รายละเอียดดังตารางที่ 2

ตารางที่ 2 รายงานผลการทดสอบระดับชาติระหว่างปี พ.ศ. 2558-2560

ปีการศึกษา 2558						
รหัสวิชา	วิชา	ผู้เข้าสอบ (N)	คะแนน เฉลี่ย (Mean)	ส่วนเบี่ยงเบน มาตรฐาน (S.D.)	คะแนน ต่ำสุด	คะแนน สูงสุด
01	ภาษาไทย	422,625	49.36	15.57	0.50	96.50
02	สังคมศึกษา ศาสนาและ วัฒนธรรม	423,519	39.70	8.28	0.00	82.00
03	ภาษาอังกฤษ	423,417	24.98	12.46	0.00	99.00
04	คณิตศาสตร์	423,654	26.59	14.79	0.00	100.00
05	วิทยาศาสตร์	422,718	33.40	8.43	0.00	90.00

ตารางที่ 2 รายงานผลการทดสอบระดับชาติระหว่างปี พ.ศ. 2558-2560 (ต่อ)

ปีการศึกษา 2559

รหัสวิชา	วิชา	ผู้เข้าสอบ (N)	คะแนน เฉลี่ย (Mean)	ส่วนเบี่ยงเบน มาตรฐาน (S.D.)	คะแนน ต่ำสุด	คะแนน สูงสุด
01	ภาษาไทย	378,161	52.29	16.27	0.00	97.00
02	สังคมศึกษา ศาสนาและ วัฒนธรรม	378,924	35.89	8.86	0.00	83.00
03	ภาษาอังกฤษ	378,779	27.76	13.39	0.00	100.00
04	คณิตศาสตร์	379,064	24.88	14.73	0.00	100.00
05	วิทยาศาสตร์	378,268	31.62	8.96	0.00	92.00

ปีการศึกษา 2560

รหัสวิชา	วิชา	ผู้เข้าสอบ (N)	คะแนน เฉลี่ย (Mean)	ส่วนเบี่ยงเบน มาตรฐาน (S.D.)	คะแนน ต่ำสุด	คะแนน สูงสุด
01	ภาษาไทย	372,043	49.25	15.77	0.00	98.00
02	สังคมศึกษา ศาสนาและ วัฒนธรรม	372,567	34.70	9.10	0.00	84.00
03	ภาษาอังกฤษ	372,587	28.31	14.65	1.25	100.00
04	คณิตศาสตร์	372,853	24.53	17.59	0.00	100.00
05	วิทยาศาสตร์	372,232	29.37	12.03	1.60	96.80

การทดสอบในระดับชั้นมัธยมศึกษาปีที่ 6 เป็นสิ่งที่มีความสำคัญอย่างยิ่ง เนื่องจากนักเรียนจำเป็นต้องใช้ความรู้พื้นฐานเพื่อนำไปใช้ศึกษาต่อในระดับสูง โดยวิชาที่มีความจำเป็นต่อการศึกษาหาความรู้เพิ่มเติมจากตำราหรือหนังสือต่างประเทศ คือ วิชาภาษาอังกฤษ จากสถิติผลการทดสอบย้อนหลังปี 2558-2560 พบว่า คะแนนเฉลี่ยวิชาภาษาอังกฤษอยู่ในเกณฑ์ต่ำ (สถาบันทดสอบทางการศึกษาแห่งชาติ, 2562) ดังนั้น ควรเร่งผลักดันการเรียนรู้อายวิชาภาษาอังกฤษ เพื่อให้ให้นักเรียนมีศักยภาพทางภาษาที่สูงขึ้นทัดเทียมกับนานาชาติ

ตอนที่ 5 มโนทัศน์เกี่ยวกับการประเมินผลระบบการทดสอบ

สิ่งสำคัญในการพัฒนาระบบทดสอบ หรือระบบคอมพิวเตอร์ต่าง ๆ นั้น คือ การประเมินผลการใช้งานของระบบ ซึ่งมีหลากหลายวิธีการและแนวทาง สำหรับการประเมินผลระบบการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์นั้นมีการนำแนวคิดการประเมินคุณลักษณะระบบที่นิยม เรียกว่า การประเมินตามมาตรฐานการประเมินระบบ (heuristic evaluation) และการประเมินจากผู้ใช้งาน (user evaluation) มาใช้ร่วมกัน ซึ่งการประเมินจากผู้ใช้งานที่นิยมใช้คือการประเมินความพึงพอใจ

ระหว่างอินเตอร์เฟซกับผู้ใช้งาน (user satisfaction of the human-computer interface) โดยในการประเมินจะอาศัยผู้เชี่ยวชาญทั้งศาสตร์ของคอมพิวเตอร์และการศึกษาเข้าด้วยกัน ซึ่งรายละเอียดในการประเมินผลระบบการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์มีรายละเอียดดังนี้ (Lilley, Barker, & Britton, 2004)

5.1 การประเมินตามมาตรฐานการประเมินระบบ

การประเมินตามมาตรฐานการประเมินระบบของ Stufflebeam ประกอบด้วยเนื้อหาหลักในการประเมินทั้งสิ้น 4 หมวด ได้แก่ การใช้ประโยชน์ (utility) ความเป็นไปได้ (feasibility) ความเหมาะสม (propriety) และความถูกต้อง (accuracy) ซึ่งมีรายละเอียดดังนี้ (ศิริชัย กาญจนวาสี, 2554)

- การใช้ประโยชน์ (utility) เป็นการประกันความเป็นประโยชน์ของระบบในการตอบสนองต่อความต้องการในการใช้สารสนเทศของผู้ที่เกี่ยวข้องให้ครอบคลุม ทันเวลา และมีผลต่อการนำไปใช้
- ความเป็นไปได้ (feasibility) เป็นการประกันความสอดคล้องของระบบกับสภาพความเป็นจริง เหมาะสมกับสถานการณ์ ปฏิบัติได้ ยอมรับได้ ประหยัดและคุ้มค่า
- ความเหมาะสม (propriety) เป็นการประกันว่าระบบมีความเหมาะสมซึ่งเป็นไปตามกฎ ระเบียบ จรรยาบรรณ คำนึงถึงสวัสดิภาพของผู้ที่เกี่ยวข้องและผู้ที่ได้รับผลกระทบ
- ความถูกต้อง (accuracy) เป็นการประกันว่าระบบมีการใช้เทคนิคที่เหมาะสม ให้สารสนเทศที่เพียงพอ

5.2 การประเมินความพึงพอใจระหว่างอินเตอร์เฟซกับผู้ใช้งาน (user satisfaction of the human-computer interface)

นอกจากการประเมินความสามารถของระบบจากผู้เชี่ยวชาญด้วยการประเมินเพื่อช่วยแก้ปัญหาแล้ว การประเมินความพึงพอใจระหว่างอินเตอร์เฟซกับผู้ใช้งานก็มีความสำคัญมาก เพราะผู้ใช้งานเป็นหัวใจสำคัญของการพัฒนาระบบ หากระบบสามารถตอบสนองการใช้งานของผู้ใช้งานได้เป็นอย่างดี ทำให้ผู้ใช้งานเกิดความพึงพอใจ ระบบก็จะได้รับการขับเคลื่อนให้นำไปใช้งานจริงในองค์กรหรือสังคม แต่หากระบบไม่ทำให้ผู้ใช้งานเกิดความพึงพอใจ ก็อาจทำให้ระบบไม่ได้รับการขับเคลื่อนหรืออาจต้องสูญเสียเวลาเปล่าในการพัฒนา ดังนั้นจึงได้มีการพัฒนาเครื่องมือการประเมินความพึงพอใจระหว่างอินเตอร์เฟซกับผู้ใช้งานขึ้น ซึ่งเป็นเครื่องมือเพื่อประเมินความพึงพอใจของอินเตอร์เฟซกับผู้ใช้งานโดยทั่วไป (generic user interface questionnaire: QUIS) โดยรายละเอียดในการประเมินประกอบด้วยเนื้อหาหลัก 5 ด้านดังนี้ (Chin, Diehl, & Norman, 1988; Navas et al., 2007)

- ภาพรวมปฏิสัมพันธ์ของระบบ (overall reaction to the software) เป็นความพึงพอใจโดยรวมของผู้ใช้งานในการใช้งานระบบ

- หน้าจอ (screen) ผู้ใช้งานสามารถอ่านหน้าจอได้อย่างง่าย มีความชัดเจน เน้นหน้าจอในส่วนที่ใช้ประโยชน์ รวมถึงรูปแบบในการจัดวาง
- การใช้คำศัพท์และสารสนเทศระบบ (terminology and system information) ผู้ใช้งานสามารถพบคำศัพท์ในระบบที่มีความสอดคล้องและเหมาะสม ข้อความแจ้งเตือนบนหน้าจอมีความชัดเจน ระบบมีการแจ้งเตือนว่าผู้ใช้งานดำเนินการอะไรอยู่พร้อมทั้งมีข้อความแจ้งเตือนความผิดพลาดที่มีประโยชน์
- การเรียนรู้ (learning) ผู้ใช้งานมีความคิดเสมอว่าระบบง่ายต่อการเรียนรู้ ไม่ยุ่งยาก ซ้ำซ้อน ชื่อและคำสั่งง่ายต่อการจำ
- ความสามารถของระบบ (system capabilities) ผู้ใช้งานสามารถใช้งานระบบได้อย่างรวดเร็ว และแก้ไขข้อผิดพลาดระบบได้ง่าย ซึ่งการใช้งานสัมพันธ์กับประสบการณ์ของผู้ใช้งาน

ตอนที่ 6 กรอบแนวคิดการวิจัย

จากการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้อง ผู้วิจัยได้พัฒนากอบแนวคิดการวิจัยในการพัฒนาระบบการทดสอบแบบปรับเหมาะหลายขั้นตอนที่มีการให้ข้อมูลป้อนกลับในการทดสอบระดับชาติ (O-NET) วิชาภาษาอังกฤษสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6 ซึ่งแบ่งออกเป็น 4 ส่วน ได้แก่ 1) หลักการและแนวคิดในการพัฒนาระบบการทดสอบแบบปรับเหมาะหลายขั้นตอนที่มีการให้ข้อมูลป้อนกลับ 2) การพัฒนาระบบการทดสอบแบบปรับเหมาะหลายขั้นตอนที่มีการให้ข้อมูลป้อนกลับ 3) ระบบการทดสอบแบบปรับเหมาะหลายขั้นตอนที่มีการให้ข้อมูลป้อนกลับ และ 4) คุณภาพระบบการทดสอบแบบปรับเหมาะหลายขั้นตอนที่มีการให้ข้อมูลป้อนกลับ

ส่วนที่ 1 หลักการและแนวคิดในการพัฒนาระบบการทดสอบแบบปรับเหมาะหลายขั้นตอนที่มีการให้ข้อมูลป้อนกลับ ประกอบด้วยหลักการและแนวคิดของการทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์ (CAT) การทดสอบแบบปรับเหมาะโดยใช้คอมพิวเตอร์แบบหลายขั้นตอน (MST) การให้ข้อมูลป้อนกลับ และการทดสอบระดับชาติ (O-NET) วิชาภาษาอังกฤษสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6 ส่วนที่ 2 การพัฒนาระบบการทดสอบแบบปรับเหมาะหลายขั้นตอนที่มีการให้ข้อมูลป้อนกลับ ประกอบด้วยศึกษากระบวนการทำงาน (algorithm) ของการทดสอบแบบปรับเหมาะหลายขั้นตอน และรูปแบบการรายงานผลการให้ข้อมูลป้อนกลับ การออกแบบและพัฒนา ระบบการทดสอบ การตรวจสอบคุณภาพของระบบการทดสอบและปรับปรุงระบบ และการนำระบบไปทดลองใช้

ส่วนที่ 3 ระบบการทดสอบแบบปรับเหมาะหลายขั้นตอนที่มีการให้ข้อมูลป้อนกลับมี 5 องค์ประกอบ ได้แก่ 1) จุดเริ่มต้นการทดสอบ ประกอบด้วย ชุดข้อสอบที่มีความยากปานกลาง 2) การประมาณค่าความสามารถและการคัดเลือกข้อสอบ ประกอบด้วย การประมาณค่าความสามารถของผู้สอบที่แท้จริงตามทฤษฎีการตอบสนองข้อสอบ และการคัดเลือกข้อสอบที่มีค่าความยาก

สอดคล้องกับระดับความสามารถของผู้สอบที่ประมาณค่าได้ 3) เกณฑ์การยุติ ประกอบด้วย การกำหนดระดับความคลาดเคลื่อนที่สามารถยอมรับได้ และ 5) รายงานการให้ข้อมูลป้อนกลับ รายงานผลการทดสอบที่ผู้ทดสอบสามารถนำข้อมูลไปพัฒนาตนเอง และส่วนที่ 4 คุณภาพระบบการทดสอบแบบปรับเหมาะหลายขั้นตอนที่มีการให้ข้อมูลป้อนกลับ ใช้การประเมินด้วยการประเมินตามมาตรฐานการประเมินระบบ ได้แก่ การใช้ประโยชน์ ความเป็นไปได้ ความเหมาะสม และความถูกต้อง และการประเมินจากผู้ใช้งานด้วยการประเมินความพึงพอใจของผู้สอบต่อการใช้งานระบบ โดยมีรายละเอียดดังภาพที่ 5



ภาพที่ 5 กรอบแนวคิดในการวิจัย

บทที่ 3 วิธีดำเนินการวิจัย

การพัฒนาระบบการทดสอบแบบปรับเหมาะหลายขั้นตอนที่มีการให้ข้อมูลป้อนกลับในการทดสอบระดับชาติ (O-NET) วิชาภาษาอังกฤษสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6 ใช้ระเบียบวิธีวิจัยและพัฒนา (research and development) ในการศึกษาโดยมีขั้นตอนการวิจัยทั้งสิ้น 3 ระยะ ได้แก่ ระยะที่ 1 การพัฒนาระบบการทดสอบแบบปรับเหมาะหลายขั้นตอนที่มีการให้ข้อมูลป้อนกลับในการทดสอบระดับชาติ (O-NET) วิชาภาษาอังกฤษสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6 ระยะที่ 2 การทดลองใช้ระบบการทดสอบแบบปรับเหมาะหลายขั้นตอนที่มีการให้ข้อมูลป้อนกลับในการทดสอบระดับชาติ (O-NET) วิชาภาษาอังกฤษสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6 และระยะที่ 3 การประเมินผลระบบการทดสอบแบบปรับเหมาะหลายขั้นตอนที่มีการให้ข้อมูลป้อนกลับในการทดสอบระดับชาติ (O-NET) วิชาภาษาอังกฤษสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6 โดยมีรายละเอียดแต่ละระยะดังนี้

ระยะที่ 1 การพัฒนาระบบการทดสอบแบบปรับเหมาะหลายขั้นตอนที่มีการให้ข้อมูลป้อนกลับในการทดสอบระดับชาติ (O-NET) วิชาภาษาอังกฤษสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6

สำหรับการพัฒนาระบบการทดสอบแบบปรับเหมาะหลายขั้นตอนที่มีการให้ข้อมูลป้อนกลับในการทดสอบระดับชาติ (O-NET) วิชาภาษาอังกฤษสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6 แบ่งการดำเนินงานออกเป็น 3 ตอนได้แก่ 1) ศึกษาองค์ประกอบของการทดสอบแบบปรับเหมาะหลายขั้นตอนและรูปแบบการรายงานผลข้อมูลป้อนกลับ 2) ออกแบบและพัฒนาระบบการทดสอบและรูปแบบการรายงานผลข้อมูลป้อนกลับ และ 3) ตรวจสอบคุณภาพของระบบการทดสอบที่มีการสะท้อนข้อมูลป้อนกลับ ซึ่งมีรายละเอียดของแต่ละตอนมีดังนี้

ตอนที่ 1 ศึกษาองค์ประกอบของการทดสอบแบบปรับเหมาะหลายขั้นตอนและรูปแบบการรายงานผลข้อมูลป้อนกลับ

โดยทำการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องกับการทดสอบแบบปรับเหมาะหลายขั้นตอน หลังจากนั้นเลือกองค์ประกอบที่มีความเหมาะสมสำหรับนำมาใช้ในการพัฒนาระบบ และศึกษาองค์ประกอบแต่ละองค์ประกอบให้มีความชัดเจนมากยิ่งขึ้น และดำเนินการศึกษารูปแบบการรายงานผลข้อมูลป้อนกลับที่มีความเหมาะสมกับการทดสอบระดับชาติ (O-NET) วิชาภาษาอังกฤษสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6

ตอนที่ 2 ออกแบบและพัฒนาระบบการทดสอบและรูปแบบการรายงานผลข้อมูลป้อนกลับ

การออกแบบและพัฒนาระบบแบ่งออกเป็น 7 ส่วน ได้แก่ ระบบการลงทะเบียน ระบบสร้างชุดข้อสอบ ระบบการประมาณค่าความสามารถ ระบบการคัดเลือกข้อสอบ ระบบการทดสอบ ระบบรายงานผลการทดสอบ และระบบประเมินระบบการทดสอบ ซึ่งใช้ภาษา PHP ในการพัฒนาเพื่อให้สามารถใช้งานผ่านระบบออนไลน์ โดยมีรายละเอียดดังนี้

- ระบบการลงทะเบียน ตัวอย่างในการวิจัยทุกคนต้องดำเนินการลงทะเบียนก่อนเข้าสู่ระบบ เพื่อเป็นการยืนยันตัวบุคคล ระบบจะนำข้อมูลของผู้เข้ารับการทดสอบลงทะเบียนไว้มาใช้ในการออกรายงานผลการทดสอบ

- ระบบสร้างชุดข้อสอบ ในการทดสอบแบบปรับเหมาะหลายขั้นตอนต้องมีการจัดชุดข้อสอบไว้ก่อนล่วงหน้า ดังนั้นจึงต้องมีระบบที่ใช้สำหรับจัดชุดข้อสอบ ดังนี้ ขั้นที่ 1 ประกอบด้วยข้อสอบจำนวน 25 ข้อ ขั้นที่ 2 ประกอบด้วยข้อสอบจำนวน 10 ข้อ และ ขั้นที่ 3 ประกอบด้วยข้อสอบจำนวน 15 ข้อ รวมข้อสอบที่ใช้ในการทดสอบทั้งสิ้น 50 ข้อ ซึ่งครอบคลุมตามเนื้อหาที่กำหนด และมีค่าความยาก (b) 3 ระดับ ได้แก่ ระดับง่าย (Easy) คือ มีค่าความยาก น้อยกว่า -1 ระดับปานกลาง (Moderate) คือ มีค่าความยากอยู่ระหว่าง -1 ถึง 1 และ ระดับยาก คือ มีค่าความยากมากกว่า 1 โดยข้อสอบที่นำมาใช้ผ่านการวิเคราะห์ข้อสอบตามทฤษฎีการตอบสนองข้อสอบ (IRT) ด้วยโปรแกรม R จากชุดคำสั่งสำเร็จรูป (package) ltm แบบ 3 พารามิเตอร์ได้แก่ ค่าความยาก (b) อำนาจจำแนก (a) และความน่าจะเป็นในการเดา (c) ซึ่งมีเกณฑ์ในการคัดเลือกข้อสอบในการจัดทำคลัง คือต้องมีค่าความยากอยู่ในช่วง -4 ถึง 4 ค่าอำนาจจำแนกอยู่ระหว่าง 0 ถึง ∞ และ ความน่าจะเป็นในการเดาไม่เกิน 0.50 (Reckase, 2003; ศิริชัย กาญจนวาสี, 2555)

- ระบบการประมาณค่าความสามารถ สำหรับขั้นตอนการออกแบบระบบการประมาณค่า ใช้การประมาณค่าร่วมกัน 2 แบบได้แก่ วิธีการประมาณค่าด้วยความเป็นไปได้สูงสุด (maximum likelihood estimation: MLE) ใช้สำหรับกรณีมีผลการตอบข้อสอบถูกและผิดรวมกัน และวิธีการประมาณค่าด้วยค่าคาดหวังภายหลัง (expected a posterior: EAP) ใช้สำหรับกรณีมีผลการตอบข้อสอบถูกหรือผิดทั้งหมด

- ระบบการคัดเลือกข้อสอบ สำหรับการออกแบบระบบการเลือกข้อสอบสำหรับการวิจัยนี้ ใช้วิธีการเลือกแบบให้สารสนเทศสูงสุด (maximum information) ในการเลือกชุดข้อสอบสำหรับขั้นถัดไป และควบคุมเนื้อหาของแบบทดสอบ (content balance) ให้เป็นไปตามเนื้อหาที่กำหนด

- ระบบการทดสอบ สำหรับหน้าจอการทดสอบจะมีองค์ประกอบหลักที่แสดงได้แก่ ชุดข้อสอบที่ผู้สอบกำลังดำเนินการทดสอบ ข้อสอบและตัวเลือกสำหรับข้อสอบแต่ละข้อ และระยะเวลาสำหรับการทำแบบทดสอบ

- ระบบรายงานผลการทดสอบ สำหรับรายงานผลการทดสอบจะแบ่งออกเป็น รายงานคะแนน สังเกตได้ และระดับความสามารถผู้สอบ โดยรูปแบบรายงานจะประกอบด้วยข้อมูลส่วนตัวของ ตัวอย่างในการวิจัย ได้แก่ ชื่อ นามสกุล สายการเรียน ระดับชั้น โรงเรียน และจังหวัดที่โรงเรียนสังกัด

- ระบบประเมินระบบการทดสอบ สำหรับการประเมินระบบการทดสอบจากตัวอย่างในการวิจัย จะใช้แบบประเมินความพึงพอใจของผู้สอบต่อการใช้งานระบบ โดยมีองค์ประกอบหลัก 5 องค์ประกอบ ได้แก่ 1) ความเห็นต่อการใช้งานระบบในภาพรวม 2) หน้าจอการแสดงผล (screen) 3) คำศัพท์และสารสนเทศของระบบ (terminology and system information) 4) การเรียนรู้ (learning) และ 5) ความสามารถระบบ (system capabilities)

ตอนที่ 3 ตรวจสอบคุณภาพของระบบการทดสอบที่มีการสะท้อนข้อมูลป้อนกลับ

เมื่อได้ระบบการทดสอบต้นแบบแล้ว จะดำเนินการส่งให้ผู้เชี่ยวชาญด้านการทดสอบและผู้เชี่ยวชาญด้านเทคโนโลยีสารสนเทศจำนวน 5 ท่าน ตรวจสอบความเหมาะสมการทำงานของระบบ โดยใช้แบบประเมินระบบแบบอิงมาตรฐาน (standard evaluation) หลังจากนั้นดำเนินการปรับปรุงตามคำแนะนำของผู้เชี่ยวชาญ และนำระบบไปทดลองใช้กับตัวอย่างในการวิจัยต่อไป

ระยะที่ 2 การทดลองใช้ระบบการทดสอบแบบปรับเหมาะหลายขั้นตอนที่มีการให้ข้อมูลป้อนกลับ ในการทดสอบระดับชาติ (O-NET) วิชาภาษาอังกฤษสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6

ในการทดลองใช้ระบบการทดสอบแบบปรับเหมาะหลายขั้นตอนที่มีการให้ข้อมูลป้อนกลับในการทดสอบระดับชาติ (O-NET) วิชาภาษาอังกฤษสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6 มีรายละเอียดในการดำเนินการดังนี้

2.1 การเตรียมระบบ ขั้นตอนในการเตรียมนำระบบไปทดลองใช้ มีการดำเนินงานดังนี้ 1) ทำหนังสือขอความอนุเคราะห์ทดลองใช้ระบบการทดสอบที่พัฒนาขึ้น พร้อมแนบคู่มือการใช้งานระบบไปยังกลุ่มตัวอย่าง 2) ประสานงานเพื่อดำเนินการเก็บรวบรวมข้อมูลในการทดลองใช้ระบบการทดสอบ พร้อมนัดหมาย วัน เวลาและสถานที่สำหรับจัดการทดสอบ และ 3) ติดต่อครูผู้ประสานงานกับตัวอย่างวิจัยเพื่อเข้ารับการทดสอบ โดยจัดเตรียมเอกสารคู่มือและตรวจสอบความพร้อมของระบบ

2.2 ตัวอย่างวิจัยในการทดลองใช้ระบบ ทำการคัดเลือกกลุ่มตัวอย่างในการวิจัยเพื่อทดลองใช้ระบบการทดสอบแบบปรับเหมาะหลายขั้นตอนที่มีการให้ข้อมูลป้อนกลับในการทดสอบระดับชาติ (O-NET) วิชาภาษาอังกฤษสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6 โดยใช้วิธีการเลือกแบบเฉพาะเจาะจง (purposive sampling) จากนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6 ในโรงเรียนที่มีความพร้อมสำหรับการทดสอบด้วยระบบคอมพิวเตอร์แบบออนไลน์ ภายในเขตกรุงเทพมหานคร ได้แก่ โรงเรียนสายปัญญา ในพระบรมราชินูปถัมภ์ จำนวน 50 คน

2.3 สถานที่ในการทดสอบ สำหรับสถานที่ในการทดสอบใช้ห้องเรียนของโรงเรียนที่ตัวอย่างวิจัยสังกัดซึ่งมีเครื่องคอมพิวเตอร์ที่ติดตั้งระบบอินเทอร์เน็ต หรือสถานที่ซึ่งตัวอย่างวิจัยสะดวกในการ

เข้าทดสอบ เพื่อให้ตัวอย่างวิจัยเข้ารับการทดสอบพร้อมกัน ครั้งละไม่เกิน 50 คน โดยใช้คอมพิวเตอร์ 1 คน ต่อ 1 เครื่อง

2.4 การชี้แจงก่อนการทดสอบ ก่อนการทดสอบ 10 นาที จะมีการชี้แจงการใช้งานระบบให้กับ ตัวอย่างวิจัยก่อนการดำเนินการทดสอบจริง โดยในขั้นตอนการชี้แจงจะให้ตัวอย่างวิจัยทำการ ลงทะเบียนพร้อมกัน เมื่อตัวอย่างวิจัยลงทะเบียนครบทุกคนแล้วจึงเริ่มต้นทำการทดสอบ

2.5 การดำเนินการทดสอบ ในขั้นตอนการทดสอบตัวอย่างวิจัยจะทำข้อสอบพร้อมกัน โดย ตัวอย่างวิจัยแต่ละคนจะได้รับชุดข้อสอบที่แตกต่างกัน ขึ้นอยู่กับระดับความสามารถของตัวอย่างวิจัย แต่ละคน หลังจากทำการทดสอบเสร็จสิ้น ระบบจะมีการแจ้งเตือนให้ตัวอย่างวิจัยทราบว่าได้ทำการ ทดสอบเสร็จสิ้นแล้ว และจะแสดงหน้าจอให้ตัวอย่างวิจัยประเมินความพึงพอใจในการใช้งานระบบ หลังจากนั้นระบบจะแสดงผลการตอบข้อสอบพร้อมทั้งแสดงรายงานสะท้อนข้อมูลป้อนกลับ เมื่อ ตัวอย่างวิจัยทุกคนในห้องสอบทำการทดสอบเสร็จสิ้น จึงจะอนุญาตให้ออกจากห้องสอบได้

ระยะที่ 3 การประเมินผลระบบการทดสอบแบบปรับเหมาะหลายขั้นตอนที่มีการให้ข้อมูล ป้อนกลับในการทดสอบระดับชาติ (O-NET) วิชาภาษาอังกฤษสำหรับนักเรียนระดับชั้น มัธยมศึกษาปีที่ 6

การประเมินผลระบบการทดสอบแบ่งออกเป็น 3 ตอน ได้แก่ ตอนที่ 1 การตรวจสอบคุณภาพ แบบประเมินระบบ ตอนที่ 2 การประเมินระบบแบบอิงมาตรฐาน (evaluation standard) และตอนที่ 3 การประเมินความพึงพอใจของผู้สอบต่อการใช้งานระบบ โดยมีรายละเอียดดังนี้

ตอนที่ 1 การตรวจสอบคุณภาพแบบประเมินระบบ ผู้วิจัยได้ดำเนินการวิเคราะห์ด้วยค่าดัชนี ความสอดคล้องระหว่างข้อคำถามหรือรายการตรวจสอบกับวัตถุประสงค์ (IOC) เป็นการตรวจสอบ คุณภาพแบบประเมินโดยอาศัยดุลพินิจจากผู้เชี่ยวชาญด้านการทดสอบจำนวน 5 ท่าน ซึ่งผู้เชี่ยวชาญ ได้ตัดสินรายการหรือข้อคำถาม (item) ในรูปแบบตัวเลข โดย 1 = เหมาะสม 0 = ไม่แน่ใจ และ -1 = ไม่เหมาะสม จากนั้นนำค่าที่ได้มาคำนวณหาค่า IOC (item objective congruency index: IOC) โดยมีสูตรในการคำนวณ ดังนี้

$$IOC = \frac{\sum R}{N}$$

โดยที่ IOC คือ ความสอดคล้องระหว่างรายการตรวจสอบกับวัตถุประสงค์

$\sum R$ คือ ผลรวมคะแนนตัดสินของผู้เชี่ยวชาญ

N คือ จำนวนผู้เชี่ยวชาญ

สำหรับเกณฑ์ที่ใช้ตัดสินของรายการประเมินคือค่า IOC ต้องมีค่าอยู่ในช่วงระหว่าง 0.5 -1.00 สำหรับรายการประเมินที่มีค่าดัชนี IOC ต่ำกว่า 0.50 ผู้วิจัยได้ดำเนินการปรับปรุงแก้ไขรายการ

ประเมินดังกล่าวให้มีความเหมาะสมตามข้อเสนอแนะ ในกรณีที่ค่า IOC มีค่ามากกว่า 0.50 แต่มีข้อเสนอแนะจากผู้เชี่ยวชาญ ผู้วิจัยก็จะดำเนินปรับปรุงแบบประเมินตามข้อเสนอแนะเพื่อให้แบบประเมินมีความสมบูรณ์มากยิ่งขึ้น จากนั้นจึงนำแบบประเมินไปใช้ในการประเมินคุณภาพระบบต่อไป

ตอนที่ 2 การประเมินระบบแบบอิงมาตรฐาน (evaluation standard) ผู้วิจัยได้ทำการประเมินระบบโดยผู้เชี่ยวชาญก่อนนำระบบไปใช้งานด้วยแบบประเมินแบบอิงมาตรฐานตามแนวคิดของ Stufflebeam ประกอบด้วยเนื้อหาหลักในการประเมินทั้งสิ้น 4 หมวด ได้แก่ ความมีประโยชน์ (utility) ความเป็นไปได้ (feasibility) ความเหมาะสม (propriety) และความถูกต้อง (accuracy) โดยแบบประเมินจะนำไปใช้ประเมินความเหมาะสมของระบบก่อนนำระบบไปใช้จริง ซึ่งมีรายละเอียดดังนี้ (ศิริชัย กาญจนวาสี, 2554)

- ความมีประโยชน์ (utility) เป็นการประกันความเป็นประโยชน์ของระบบในการตอบสนองต่อความต้องการในการใช้สารสนเทศของผู้ที่เกี่ยวข้องให้ครอบคลุม ทันเวลา และมีผลต่อการนำไปใช้

- ความเป็นไปได้ (feasibility) เป็นการประกันความสอดคล้องของระบบกับสภาพความเป็นจริง เหมาะสมกับสถานการณ์ ปฏิบัติได้ ยอมรับได้ ประหยัดและคุ้มค่า

- ความเหมาะสม (propriety) เป็นการประกันว่าระบบมีความเหมาะสมซึ่งเป็นไปตามกฎระเบียบ จรรยาบรรณ คำนึงถึงสวัสดิภาพของผู้ที่เกี่ยวข้องและผู้ที่ได้รับผลกระทบ

- ความถูกต้อง (accuracy) เป็นการประกันว่าระบบมีการใช้เทคนิคที่เหมาะสม ให้สารสนเทศที่เพียงพอ

เครื่องมือในการประเมินเป็นรูปแบบมาตรประมาณค่า (rating scale) 5 ระดับ ได้แก่ 1 = น้อยที่สุด 2= น้อย 3= ปานกลาง 4 = มาก และ 5 = มากที่สุด โดยมีเกณฑ์ในการแปลผลดังนี้

คะแนนเฉลี่ย 1.00 – 1.49 หมายถึง มีประโยชน์/ เป็นไปได้/ เหมาะสม/ ถูกต้อง/น้อยที่สุด

คะแนนเฉลี่ย 1.50 – 2.49 หมายถึง มีประโยชน์/ เป็นไปได้/ เหมาะสม/ ถูกต้อง/น้อย

คะแนนเฉลี่ย 2.50 – 3.49 หมายถึง มีประโยชน์/ เป็นไปได้/ เหมาะสม/ ถูกต้อง/ปานกลาง

คะแนนเฉลี่ย 3.50 – 4.49 หมายถึง มีประโยชน์/ เป็นไปได้/ เหมาะสม/ ถูกต้อง/มาก

คะแนนเฉลี่ย 4.50 – 5.00 หมายถึง มีประโยชน์/ เป็นไปได้/ เหมาะสม/ ถูกต้อง/มากที่สุด

ตอนที่ 3 การประเมินความพึงพอใจของผู้สอบต่อการใช้งานระบบ ประกอบด้วยรายการประเมินจำนวน 5 รายการ ได้แก่ 1) ความเห็นต่อการใช้งานระบบในภาพรวม 2) หน้าจอการแสดงผล (screen) 3) คำศัพท์และสารสนเทศของระบบ (terminology and system information) 4) การเรียนรู้ (learning) และ 5) ความสามารถระบบ (system capabilities) เครื่องมือในการประเมินเป็นมาตรการวัดตามแนว Osgood 10 ระดับ เริ่มตั้งแต่ 0 ถึง 9

บทที่ 4

ผลการวิเคราะห์ข้อมูล

การวิจัยในครั้งนี้มีวัตถุประสงค์ 2 ประการ 1) เพื่อพัฒนาระบบการทดสอบแบบปรับเหมาะหลายขั้นตอนที่มีการให้ข้อมูลป้อนกลับในการทดสอบระดับชาติ (O-NET) วิชาภาษาอังกฤษสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6 และ 2) เพื่อประเมินคุณภาพระบบการทดสอบแบบปรับเหมาะหลายขั้นตอนที่มีการให้ข้อมูลป้อนกลับในการทดสอบระดับชาติ (O-NET) วิชาภาษาอังกฤษสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6 โดยการดำเนินการวิจัยแบ่งออกเป็น 3 ระยะ ได้แก่ ระยะที่ 1 การพัฒนาระบบการทดสอบแบบปรับเหมาะหลายขั้นตอนที่มีการให้ข้อมูลป้อนกลับในการทดสอบระดับชาติ (O-NET) วิชาภาษาอังกฤษสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6 ระยะที่ 2 การทดลองใช้ระบบการทดสอบแบบปรับเหมาะหลายขั้นตอนที่มีการให้ข้อมูลป้อนกลับในการทดสอบระดับชาติ (O-NET) วิชาภาษาอังกฤษสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6 และระยะที่ 3 การประเมินผลระบบการทดสอบแบบปรับเหมาะหลายขั้นตอนที่มีการให้ข้อมูลป้อนกลับในการทดสอบระดับชาติ (O-NET) วิชาภาษาอังกฤษสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6 ซึ่งรายละเอียดผลการวิเคราะห์ข้อมูลแต่ละตอนมีดังนี้

ระยะที่ 1 ผลการพัฒนาระบบการทดสอบแบบปรับเหมาะหลายขั้นตอนที่มีการให้ข้อมูลป้อนกลับในการทดสอบระดับชาติ (O-NET) วิชาภาษาอังกฤษสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6

สำหรับผลการศึกษาในระยะที่ 1 แบ่งเนื้อหาออกเป็น 4 ตอน ได้แก่ 1) ผลการศึกษารายละเอียดองค์ประกอบของการทดสอบแบบปรับเหมาะหลายขั้นตอนที่มีการให้ข้อมูลป้อนกลับ 2) ผลการพัฒนาระบบการทดสอบและรูปแบบการรายงานข้อมูลป้อนกลับ และ 3) ผลการวิเคราะห์ข้อสอบวิชาภาษาอังกฤษของการทดสอบระดับชาติ (O-NET) สำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6 และ 4) ผลการตรวจสอบคุณภาพระบบก่อนนำระบบไปใช้งานจริง ซึ่งมีรายละเอียดของแต่ละขั้นตอนดังนี้

ตอนที่ 1 ผลการศึกษารายละเอียดองค์ประกอบของการทดสอบแบบปรับเหมาะหลายขั้นตอนที่มีการให้ข้อมูลป้อนกลับ

การศึกษารายละเอียดองค์ประกอบของการทดสอบแบบปรับเหมาะหลายขั้นตอนที่มีการให้ข้อมูลป้อนกลับพบว่า องค์ประกอบของระบบประกอบด้วย 4 องค์ประกอบ ได้แก่ 1) จุดเริ่มต้นในการทดสอบ (initial stage) 2) การประมาณค่าความสามารถของผู้สอบ (ability estimation) 3) การคัดเลือกข้อสอบ (item selection) และ 4) จุดยุติการทดสอบ (stopping criteria) และรูปแบบรายงานผลที่สะท้อนข้อมูลป้อนกลับ ได้แก่ รายงานคะแนนสังเกตได้และรายงานระดับความสามารถของผู้สอบ โดยรายละเอียดของแต่ละขั้นตอนมีดังนี้

1) จุดเริ่มต้นการทดสอบ (initial stage) จากการสังเคราะห์งานวิจัยที่เกี่ยวข้องพบว่า จุดเริ่มต้นในการทำข้อสอบจากข้อสอบข้อเดียวจะทำให้การประมาณค่าความสามารถสูงหรือต่ำกว่าจริง (over and under estimate) ดังนั้น จึงได้มีการพัฒนาการทดสอบแบบปรับเหมาะหลายขั้นตอนปรับเหมาะหลายขั้นตอน (MST) ขึ้นเพื่อช่วยลดความผิดพลาดในการประมาณค่าที่สูงหรือต่ำกว่าจริง (Zheng & Chang, 2015; Zheng et al., 2012) สำหรับการวิจัยในครั้งนี้จะใช้การทดสอบจำนวน 3 ขั้น ในรูปแบบ 1-2-3 กล่าวคือ ข้อสอบในขั้นแรกประกอบด้วยข้อสอบระดับปานกลาง (Moderate: M) ข้อสอบในขั้นที่สองประกอบด้วยข้อสอบระดับยาก (Hard: H) และง่าย (Easy: E) และข้อสอบในขั้นที่สามประกอบด้วยข้อสอบระดับง่าย (E) ปานกลาง (M) และยาก (H) โดยข้อสอบในขั้นแรก มีจำนวน 25 ข้อ โดยผู้วิจัยเป็นผู้ดำเนินการจัดชุดข้อสอบไว้ จำนวน 5 ชุด ซึ่งชุดข้อสอบในชุดแรกจะถูกเรียกขึ้นมาใช้อย่างสุ่ม (random) จากทั้งหมด 5 ชุด โดยข้อสอบในชุดแรกเป็นข้อสอบที่มีค่าความยาก (b) อยู่ในระดับปานกลาง (ค่า b อยู่ระหว่าง -1 ถึง 1) ซึ่งเป็นไปตามแนวทางการทดสอบแบบปรับเหมาะหลายขั้นตอน เพื่อให้การประมาณค่าความสามารถมีความแม่นยำมากยิ่งขึ้น ซึ่งสอดคล้องกับแนวทางของ Zheng & Chang

2) การประมาณค่าความสามารถของผู้สอบ (ability estimation) สำหรับการประมาณค่าความสามารถของผู้สอบจากการศึกษาพบว่า การประมาณค่าด้วยความเป็นไปได้สูงสุด (maximum likelihood estimation: MLE) เป็นวิธีการประมาณค่าซึ่งมีความแม่นยำมากที่สุด แต่จะมีความคลาดเคลื่อนสูงเมื่อผู้สอบตอบข้อสอบถูกหรือผิดหมด ซึ่งวิธีที่ประมาณค่าได้แม่นยำสำหรับกรณีผู้สอบตอบข้อสอบถูกหรือผิดหมด คือวิธีการประมาณค่าด้วยค่าคาดหวังภายหลัง (expected a posterior: EAP) ดังนั้นจึงมีการรวมวิธีจะใช้การประมาณค่าทั้ง 2 วิธีเข้าด้วยกัน ได้แก่ วิธีการประมาณค่าด้วยความเป็นไปได้สูงสุด (maximum likelihood estimation: MLE) และวิธีการประมาณค่าด้วยค่าคาดหวังภายหลัง (expected a posterior: EAP) ซึ่งวิธีการประมาณค่าด้วยค่าคาดหวังภายหลังจะใช้สำหรับกรณีผู้สอบตอบข้อสอบถูกหรือผิดทุกข้อ หากผู้สอบตอบข้อสอบถูกและผิดผสมกันจะใช้วิธีการประมาณค่าด้วยวิธีการประมาณค่าด้วยความเป็นไปได้สูงสุด ดังนั้น ในการวิจัยนี้จึงใช้วิธีการประมาณค่าความสามารถทั้งสองวิธีร่วมกันเพื่อให้ค่าความสามารถที่ได้มีความถูกต้องและแม่นยำมากยิ่งขึ้น (Bock & Mislevy, 1982; Hambleton & Swaminathan, 1985; Zheng & Chang, 2015)

วิธีการประมาณค่าด้วยความเป็นไปได้สูงสุด (maximum likelihood estimation: MLE) จะใช้แนวทางของ Hambleton & Swaminathan (1985) โดยมีสมการที่ใช้ในการคำนวณดังนี้

$$P(U_{ij} = 1 | \theta_i, a_j, b_j, c_j) = c_j + (1 - c_j) \frac{\exp(a_j(\theta_i - b_j))}{1 + \exp(a_j(\theta_i - b_j))}$$

$$L_J(\theta) = \prod_{j=1}^J [P_j(\theta)]^{x_j} [1 - P_j(\theta)]^{1-x_j}$$

$$h_m = \frac{D[r_m - \sum P_i(\theta_m)]}{-D^2 \sum P_i(\theta_m) Q_i(\theta_m)}$$

$$\theta_{m+1} = \theta_m - h_m$$

โดยสมการที่ใช้ในการคำนวณจะไม่นำค่าปรับสเกล (D=1.7) มาใช้เนื่องจาก Camilli (1994) ได้ทำการศึกษาการใช้ค่าปรับสเกลพบว่าให้ผลการประมาณค่าความสามารถไม่แตกต่างจากการใช้ค่าปรับสเกล

วิธีการประมาณค่าความสามารถด้วยค่าคาดหวังภายหลัง (expected a posterior: EAP) ใช้แนวทางของ Bock and Mislevy (1982) โดยมีสมการที่ใช้ในการคำนวณดังนี้

$$\bar{\theta}_j = \frac{\sum_{k=1}^q X_k L_j(X_k) \cdot W(X_k)}{\sum_{k=1}^q L_j(X_k) \cdot W(X_k)}$$

3) การคัดเลือกข้อสอบ (item selection) สำหรับการคัดเลือกข้อสอบเป็นสิ่งที่สำคัญที่สุดในกระบวนการทดสอบแบบปรับเหมาะ ซึ่งในงานวิจัยนี้ใช้การเลือกข้อสอบตามระดับความสามารถที่สอดคล้องกับค่าความยาก นอกจากนี้ควรคำนึงถึงความครอบคลุมของเนื้อหาภายในข้อสอบและเงื่อนไขอื่นร่วมด้วย เช่น การควบคุมการเปิดเผยของข้อสอบไม่ให้มีจำนวนมาก (Cheng & Chang, 2009; Zheng & Chang, 2015)

4) จุดยุติการทดสอบ (stopping criteria) ในการวิจัยครั้งนี้กำหนดความยาวของแบบสอบแบบคงที่ (Fixed length) เนื่องจากเป็นการทดสอบแบบหลายขั้นตอนซึ่งต้องกำหนดชุดข้อสอบไว้ล่วงหน้า โดยใช้ข้อสอบทั้งสิ้น จำนวน 50 ข้อ ต่อผู้สอบ 1 คน เมื่อยุติการทดสอบระบบจะรายงานเป็นคะแนนร้อยละ และคะแนนมาตรฐาน พร้อมสรุปการผ่านเกณฑ์ ดังตารางที่ 3 และ 4

ตารางที่ 3 รูปแบบรายงานคะแนนจากการสอบ

เนื้อหา	จำนวนข้อสอบ	จำนวนข้อ ที่ตอบถูก	ร้อยละ
Part I: Language Use	25	25	100
- Dialog Completion	5	5	100
- Situational Dialog	5	5	100
- Sentence Completion	5	5	100
- Error Correction	10	10	100
Part II: Writing Ability	10	10	100
Part III: Reading Ability	15	15	100
- Vocabulary	5	5	100
- Reading Comprehension	10	10	100

ตารางที่ 4 รูปแบบรายงานระดับความสามารถของผู้สอบ

ค่า Theta (Θ)	ค่าคะแนนมาตรฐาน (Scale Score)	ผลการทดสอบ
1.5	65	ผ่านการทดสอบ

เกณฑ์การผ่านพิจารณาจากค่าคะแนนมาตรฐาน

** ค่าคะแนนมาตรฐานฐาน ≥ 50 หมายถึง ผ่านการทดสอบ

***ค่าคะแนนมาตรฐานฐาน ≤ 50 หมายถึง ไม่ผ่านการทดสอบ

ตอนที่ 2 ผลการพัฒนาระบบการทดสอบและรูปแบบการรายงานข้อมูลย้อนกลับ

หลังจากได้องค์ประกอบของระบบการทดสอบและรูปแบบการรายงานข้อมูลย้อนกลับเรียบร้อยแล้ว นำองค์ประกอบที่ได้มาพัฒนาระบบการทดสอบ โดยพัฒนาระบบการทดสอบแบ่งออกเป็น 6 ขั้นตอน ได้แก่ 1) การลงทะเบียนสำหรับตัวอย่างวิจัย 2) การเข้าสู่ระบบการทดสอบ 3) การดำเนินการทดสอบ 4) การยุติการทดสอบ 5) การประเมินคุณภาพระบบการทดสอบ และ 6) การรายงานผลการทดสอบ โดยแต่ละขั้นตอนมีรายละเอียดดังนี้โดยมีรายละเอียดของแต่ละขั้นตอนดังนี้

ขั้นที่ 1 การลงทะเบียนสำหรับตัวอย่างในการวิจัย

ตัวอย่างวิจัยทุกคนต้องดำเนินการลงทะเบียนก่อนเข้ารับการทดสอบ โดยเข้าเว็บไซต์ <https://www.nhabhat.com> หลังจากนั้นจะปรากฏข้อมูลดังภาพที่ 6 ทำการเลือกแถบเมนู “ลงทะเบียน” เพื่อลงทะเบียน โดยตัวอย่างวิจัยต้องกรอก อีเมล (e-mail) กำหนดรหัสผ่าน (password) และกรอกข้อมูลพื้นฐานของตนเองให้ครบถ้วน ดังภาพที่ 7 เมื่อกรอกข้อมูลเรียบร้อยแล้วให้กดปุ่ม “ลงทะเบียน” ระบบจะแสดงหน้าจอข้อมูลพื้นฐานของผู้ลงทะเบียน ซึ่งตัวอย่างวิจัยสามารถแก้ไขข้อมูลส่วนบุคคลของตนเองได้ โดยกดปุ่ม “EDIT” ที่มุมบนด้านขวามือ แต่หากต้องการเริ่มทดสอบ สามารถกดที่ปุ่ม “เข้าห้องสอบ” ทางด้านซ้ายมือของหน้าจอ รายละเอียดดังภาพที่ 8

ระบบการทดสอบแบบปรับเหมาะหลายขั้นตอนที่มีการให้ข้อมูลป้อนกลับ
ในการทดสอบระดับชาติ (O-NET)
วิชา ภาษาอังกฤษสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6

เข้าใช้งานระบบ
ลงทะเบียน

อีเมลสำหรับล็อกอิน

รหัสผ่าน

[Forgot Password?](#)

ภาพที่ 6 หน้าจอหลักกระบบการทดสอบ

ระบบการทดสอบแบบปรับเหมาะหลายขั้นตอนที่มีการให้ข้อมูลป้อนกลับ
ในการทดสอบระดับชาติ (O-NET)
วิชา ภาษาอังกฤษสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6

เข้าใช้งานระบบ
ลงทะเบียน

อีเมล : (ใช้สำหรับล็อกอินเข้าระบบและรีเซตรหัสผ่าน)

รหัสผ่านสำหรับเข้าระบบสอบ : ยืนยันอีกครั้ง :

ชื่อ :

นามสกุล :

อายุ : เพศ : สายการเรียน : ระดับชั้น :

จังหวัด : โรงเรียน :

ภาพที่ 7 หน้าจอการลงทะเบียน

Welcome : onet@onet.com EDIT

ข้อมูลพื้นฐาน

ชื่อ : ตัวอย่าง
นามสกุล : ทดสอบ
เพศ : ชาย

ข้อมูลการศึกษา

ชั้นเรียน : ม. 6
โรงเรียน : ตัวอย่าง
จังหวัด : กรุงเทพมหานคร

อายุ : 18

[เข้าสู่ห้องสอบ](#)

[เปลี่ยนรหัสผ่านเข้าสู่ระบบ](#)

ภาพที่ 8 การแสดงสถานะเมื่อกดปุ่มลงทะเบียน

ขั้นที่ 2 การเข้าสู่ระบบการทดสอบ

การเข้าสู่ระบบการทดสอบในหน้าแรก ตัวอย่างวิจัยดำเนินการกรอกอีเมล (e-mail) และรหัส (password) ที่ลงทะเบียนไว้ หลังจากนั้นกดปุ่ม “เข้าสู่ระบบ” เพื่อเข้าสู่ระบบ รายละเอียดแสดงดังภาพที่ 9

ระบบการทดสอบแบบปรับเหมาะหลายขั้นตอนที่มีการให้ข้อมูลป้อนกลับ
ในการทดสอบระดับชาติ (O-NET)
วิชา ภาษาอังกฤษสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6

[เข้าใช้งานระบบ](#) [ลงทะเบียน](#)

อีเมลสำหรับล็อกอิน
onet@onet.com

รหัสผ่าน
....

[เข้าสู่ระบบ](#)

[Forgot Password?](#)

ภาพที่ 9 การเข้าสู่ระบบการทดสอบ

เมื่อตัวอย่างวิจัยเข้าสู่ระบบการทดสอบเรียบร้อยแล้ว ระบบจะแสดงหน้าจอข้อมูลพื้นฐานของผู้ใช้ระบบ ซึ่งจะมีลักษณะเช่นเดียวกับที่แสดงในภาพที่ 8 ซึ่งตัวอย่างวิจัยสามารถแก้ไขข้อมูลโดยกดปุ่ม “EDIT” ที่มุมบนด้านขวามือ เมื่อแก้ไขเรียบร้อยแล้วให้กด “บันทึก” เพื่อบันทึกการแก้ไขข้อมูลดังแสดงในภาพที่ 10 หรือกดเลือก “เข้าสู่ห้องสอบ” เพื่อเข้าสู่ระบบการทดสอบ

ภาพที่ 10 การแก้ไขข้อมูลส่วนตัว

ขั้นที่ 3 การดำเนินการทดสอบ

เมื่อเข้าสู่ระบบเรียบร้อยแล้วหน้าจอจะแสดงคำชี้แจงรายละเอียดเกี่ยวกับการทดสอบ ซึ่งระบบการทดสอบนี้เป็นระบบที่จัดทำขึ้นเพื่อเตรียมความพร้อมในการทดสอบระดับชาติ (O-NET) วิชาภาษาอังกฤษสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6 โดยรายละเอียดหน้าจอคำชี้แจงแสดงดังภาพที่ 11 หลังจากอ่านคำชี้แจงเรียบร้อยแล้วให้กดเลือก “ยอมรับ-รับทราบขั้นตอน” และกด “เข้าสู่ห้องสอบ” เพื่อเริ่มการทดสอบ

ภาพที่ 11 หน้าจอคำชี้แจง

หลังจากกด “เข้าสู่ห้องสอบ” หน้าจอจะแสดงข้อสอบข้อแรก ดังภาพที่ 12 แถบข้อความด้านบนจะแสดงชื่อผู้เข้าสอบ และเวลาคงเหลือในการทำแบบทดสอบ

ผู้เข้าสอบ : ตัวอย่าง ทดสอบ ท่านเหลือเวลาทำข้อสอบอีก 01:59:40

ใช้เวลาในการทดสอบจำนวน 2 ชั่วโมง เลือกคำตอบที่ถูกต้องและคลิกปุ่ม [ไปขั้นตอนถัดไป >>](#) เมื่อครบเวลาในการทดสอบ ระบบจะหยุดการทดสอบทันที
ท่านไม่สามารถย้อนกลับไปยังคำตอบในขั้นก่อนหน้าได้

Part I: Language Use
Dialog Completion Directions: Read the dialogs and choose the expression that BEST completes each missing part.

Situation: Sue is moving into a new apartment. Now she is asking the clerk for the key to her room. Sue: Hello. My name is Susan Clayton. I'm moving in today. 1. _____ Clerk: Good afternoon, Ms. Clayton. I'll show you to your room. 2. _____ Sue: That's very kind of you. Clerk: 3. _____ Er.....Ms. Clayton, may I remind you not to hang clothes on the balcony? Sue: 4. _____ Clerk: I'm afraid not. I hope you will enjoy staying here. 5. _____ Sue: Thank you.

1

1. I do want to go to my room now.
2. Did you get the key for the room?
3. May I have the key to my room, please?
4. Can't you give me my room key now?
5. I should have my key now, you know.

2

1. Do you bring along your things?
2. Will you take your things to your room?
3. Don't forget to take all of your things.
4. Won't you take all your things with you?
5. Let me give you a hand with your things.

สถานการณ์ทำข้อสอบ ข้อที่ 1 ข้อที่ 2 ข้อที่ 3 ข้อที่ 4 ข้อที่ 5 [ไปขั้นตอนถัดไป >>](#)

ภาพที่ 12 หน้าจอการทดสอบ

ในการตอบคำถามแต่ละข้อนั้น ทำได้โดยการนำเมาส์ชี้ที่คำตอบที่ต้องการ ซึ่งหน้าจอจะแสดงแถบสีเขียวอ่อน และเมื่อกดเลือกคำตอบแล้ว แถบสีนี้จะเปลี่ยนเป็นสีเขียวเข้ม นอกจากนี้ หน้าจอด้านล่างยังแสดงแถบ “สถานการณ์ทำข้อสอบ” ให้ผู้สอบทราบว่าได้ทำข้อสอบในตอนนี้ครบถ้วนหรือยัง รายละเอียดดังภาพที่ 13

Part I: Language Use
Situational Dialogs Directions: Read each situation and choose the best alternative.

1 Situation: A foreign tourist wants to take a bus to Kanchanaburi, so he asks a man for directions to the bus terminal. The man says: _____

1. Let's take a taxi to Kanchanaburi.
2. There are many buses going to the bus terminal.
3. You need to go by train because it's faster.
4. It's on your left, just after you pass the next intersection.
5. Kanchanaburi's about 120 kilometers from Bangkok.

2 Situation: Dorothy is at the Tourist Information Center. She wants to get directions to the train station. She says: _____

1. How can I get to the train station?
2. Do you mind taking me to the train station?
3. Excuse me, have you been to the train station?
4. I'm a stranger around here. What can you do for me?
5. Can you show me the train station, please?

3 Situation: Linda is answering the phone. Mr. Hughes, Marketing Director of the ABC Company, would like to talk to Mr. Jenkins, Head of the Personnel Department. She says: _____

1. Mr. Hughes? Yes, Mr. Jenkins is expecting your call.
2. Oh, hi! The meeting's just started. Don't call again.

สถานการณ์ทำข้อสอบ ข้อที่ 1 ข้อที่ 2 ข้อที่ 3 ข้อที่ 4 ข้อที่ 5 [ไปขั้นตอนถัดไป >>](#)

ภาพที่ 13 หน้าจอการกดเลือกคำตอบ

เมื่อผู้สอบตอบคำถามครบทุกข้อของตอนนั้น ๆ แล้ว แถบสถานการณ์ทำข้อสอบจะเป็นสีเขียวทุกข้อ และปุ่ม “ไปขั้นตอนถัดไป” จะปรากฏเป็นสีฟ้าเข้ม และสามารถกดเพื่อไปทำการทดสอบในขั้นถัดไปได้ ดังภาพที่ 14

4 Situation: Mark has not had enough sleep. When he meets John, John notices it and says: ____

1. I can't believe it, Mark. You haven't changed at all. 2. Hi, Mark. Will you come to my birthday party next Sunday?

3. You look so tired today! Why don't you get some coffee? 4. What's up? You never answer my calls.

5. I heard that you went to Spain last year. Did you have a good time there?

5 Situation: Laura is unhappy with her roommate, Lucy. She says to Rose, " ____ "

1. Lucy's just moved in and she is very well-organized. I must annoy her. 2. You know what? Lucy bought me a box of chocolate for my birthday.

3. Do you know Lucy? She had a fight with her ex-boyfriend. **4. I hate to say this, but I can't stand Lucy any longer. I want to move out.**

5. Everyone in this apartment knows Lucy. She has a reputation as a good negotiator.

สถานะการทำข้อสอบ [ข้อที่ 1](#) [ข้อที่ 2](#) [ข้อที่ 3](#) [ข้อที่ 4](#) [ข้อที่ 5](#) [ไปที่ข้อต่อไป >>](#)


ภาพที่ 14 หน้าจอเมื่อตอบคำถามครบในขั้นนั้น ๆ

ขั้นที่ 4 การยุติการทดสอบ

ในการทดสอบตัวอย่างวิจัยต้องดำเนินการทดสอบจำนวน 3 ขั้น (stage) หรือเท่ากับ 50 ข้อ ประกอบด้วยขั้นที่ 1 จำนวน 25 ข้อ ขั้นที่ 2 จำนวน 10 ข้อ และขั้นที่ 3 จำนวน 15 ข้อ เมื่อตัวอย่างวิจัยตอบข้อสอบครบทุกข้อ ระบบจะทำการยุติการทดสอบเนื่องจากการทดสอบแบบความยาวคงที่ (fixed length) โดยแสดงหน้าจอดังภาพที่ 15 และจะแสดงหน้าการประเมินความพึงพอใจต่อการใช้งานระบบ โดยตัวอย่างวิจัยต้องทำการประเมินทุกคน หากไม่ประเมิน ระบบจะไม่แสดงรายงานผลและข้อมูลย้อนกลับให้ตัวอย่างวิจัยทราบ

ขอแสดงความยินดี

การทดสอบเสร็จสิ้น



ระบบกำลังจะนำท่านไปยังหน้าสำหรับตอบแบบประเมินความพึงพอใจ
หลังจากท่านทำแบบประเมินฯ ครบแล้ว ท่านจะได้รับผลการสอบในลำดับต่อไป

[เพื่อไปยังหน้าตอบแบบประเมินความพึงพอใจ](#)

ภาพที่ 15 หน้าจอเมื่อตอบคำถามครบทุกขั้น

ขั้นที่ 5 การประเมินคุณภาพระบบการทดสอบ

เมื่อตัวอย่างวิจัยทำข้อสอบเสร็จสิ้นครบจำนวน 50 ข้อตามเกณฑ์ที่กำหนดไว้ ระบบจะทำการแสดงหน้าการประเมินความพึงพอใจต่อการใช้งานระบบให้ตัวอย่างวิจัยประเมิน ตัวอย่างวิจัยต้องทำการประเมินทุกข้อ มิเช่นนั้นระบบจะไม่แสดงรายงานผลและข้อมูลย้อนกลับให้ตัวอย่างวิจัยทราบ หลังตัวอย่างวิจัยกดส่งผลการประเมินเรียบร้อยแล้ว ระบบจะแสดงผลรายงานการทดสอบและข้อมูลย้อนกลับให้ตัวอย่างวิจัยทราบ สำหรับหน้าจอกการประเมินความพึงพอใจต่อการใช้งานระบบแสดงดังภาพที่ 16 และเมื่อตัวอย่างวิจัยประเมินครบทุกข้อ จะปรากฏปุ่ม “ส่งแบบสอบถาม & ดูรายงาน” ให้สามารถกดเพื่อแสดงรายงานผลการทดสอบที่ได้ทำเสร็จสิ้นไปแล้วได้ ดังภาพที่ 17

ONET System : Survey												
แบบประเมินความพึงพอใจของผู้สอบต่อการใช้งานระบบ												
คำชี้แจง: โปรดคลิกให้คะแนนที่ ๑ ตามความเต็มใจของท่าน ดังนี้ ***หมายเหตุ: 10 ระดับ 0 1 2 3 4 5 6 7 8 9												
ความเห็นต่อการใช้งานระบบในภาพรวม												
1. ความเห็นต่อการใช้งานระบบในภาพรวม	แย่มาก	0	1	2	3	4	5	6	7	8	9	ดีเยี่ยม
	ไม่พอใจ	0	1	2	3	4	5	6	7	8	9	พอใจ
	น่าเบื่อ	0	1	2	3	4	5	6	7	8	9	น่าสนใจ
หน้าจอกการแสดงผล (screen)												
2. การแสดงผลของหน้าจอ	อ่านยาก	0	1	2	3	4	5	6	7	8	9	อ่านง่าย
3. การใช้งานของเครื่องมือ (tools) บนหน้าจอกการแสดงผล	ยาก	0	1	2	3	4	5	6	7	8	9	ง่าย
4. การเรียงลำดับเครื่องมือบนหน้าจอกการแสดงผล	สับสน	0	1	2	3	4	5	6	7	8	9	ชัดเจน
5. การเรียงลำดับเนื้อหาบนหน้าจอกการแสดงผล	สับสน	0	1	2	3	4	5	6	7	8	9	ชัดเจน
คำศัพท์และสารสนเทศระบบ (terminology and system information)												
6. การใช้คำศัพท์ในระบบ	เข้าใจยาก	0	1	2	3	4	5	6	7	8	9	เข้าใจง่าย
7. ศัพท์ในระบบมีความสัมพันธ์กับเครื่องมือ (tools) ที่แสดงในระบบ	น้อย	0	1	2	3	4	5	6	7	8	9	มาก

ภาพที่ 16 หน้าจอกการประเมินความพึงพอใจต่อการใช้งานระบบ

9. ข้อความที่แสดงให้กับผู้สอบเพื่อป้อนข้อมูลบนหน้าจอกการแสดงผล	สับสน	0	1	2	3	4	5	6	7	8	9	ชัดเจน
10. การแจ้งเตือนสถานะระหว่างการใช้งาน	ไม่มีเลย	0	1	2	3	4	5	6	7	8	9	มี
11. ข้อความแจ้งเตือนความคิดพลาด	เข้าใจยาก	0	1	2	3	4	5	6	7	8	9	เข้าใจง่าย
การเรียนรู้ (learning)												
12. การเรียนรู้การใช้งานของระบบ	ยาก	0	1	2	3	4	5	6	7	8	9	ง่าย
13. ความสามารถในการจดจำข้อมูลและคำสั่งที่เคยป้อน	ทำไม่ได้	0	1	2	3	4	5	6	7	8	9	ทำได้
14. ผู้สอบสามารถใช้งานเครื่องมือต่าง ๆ ในระบบ	ไม่ต่อเนื่อง	0	1	2	3	4	5	6	7	8	9	ต่อเนื่อง
15. ข้อความแจ้งเตือนความคิดพลาดช่วยให้เข้าใจการทำงานของระบบ	ไม่ช่วย	0	1	2	3	4	5	6	7	8	9	ช่วย
16. เอกสารคู่มือการใช้งาน	สับสน	0	1	2	3	4	5	6	7	8	9	ชัดเจน
ความสามารถระบบ (system capabilities)												
17. ความเร็วในการตอบสนองต่อการใช้งานระบบ	ช้า	0	1	2	3	4	5	6	7	8	9	เร็ว
18. ความต่อเนื่องในการใช้งานระบบ	ติดขัด	0	1	2	3	4	5	6	7	8	9	ราบรื่น
19. แนวโน้มความต่อเนื่องในการทำงานของระบบ	ติดขัด	0	1	2	3	4	5	6	7	8	9	ราบรื่น
20. ข้อขัดข้องในการใช้งานระบบ	มี	0	1	2	3	4	5	6	7	8	9	ไม่มี

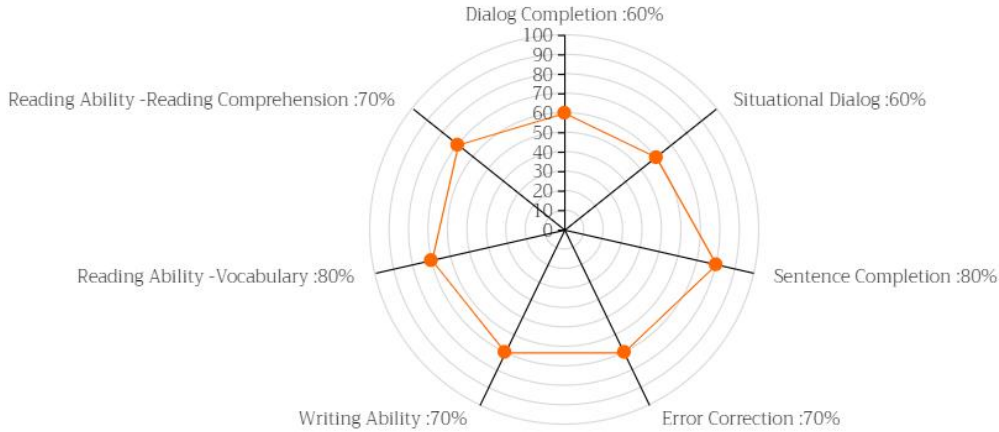
ส่งแบบสอบถาม & ดูรายงาน

ภาพที่ 17 หน้าจอกเมื่อประเมินความพึงพอใจต่อการใช้งานระบบครบทุกข้อ

ขั้นที่ 6 การรายงานผลการทดสอบ

สำหรับการออกผลรายงานผลการทดสอบ ตัวอย่างวิจัยต้องดำเนินการประเมินคุณภาพระบบ ด้วยแบบประเมินความพึงพอใจของผู้สอบต่อการใช้งานระบบให้เรียบร้อยก่อน เมื่อดำเนินการประเมินผลเรียบร้อยแล้ว ระบบจะรายงานผลการทดสอบให้ตัวอย่างวิจัยทราบดังภาพที่ 18





แผนภูมิแสดงผลการตอบแบบทดสอบ

ภาพที่ 18 รายงานผลการทดสอบและข้อมูลย้อนกลับ

ระยะที่ 2 ผลการทดลองใช้ระบบการทดสอบแบบปรับเหมาะหลายขั้นตอนที่มีการให้ข้อมูลย้อนกลับในการทดสอบระดับชาติ (O-NET) วิชาภาษาอังกฤษสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6

สำหรับระยะที่ 2 เป็นการนำระบบไปทดลองใช้กับตัวอย่างในการวิจัย ซึ่งแบ่งการนำเสนอออกเป็น 2 ตอน ได้แก่ 1) ข้อมูลทั่วไปของตัวอย่างในการวิจัย 2) ข้อมูลผลการตอบข้อสอบ

ตอนที่ 1 ข้อมูลทั่วไปของตัวอย่างในการวิจัย

หลังจากนำระบบที่พัฒนาขึ้นเรียบร้อยแล้วไปให้ผู้เชี่ยวชาญตรวจสอบคุณภาพ พบว่า ระบบที่พัฒนาขึ้นมีคุณภาพและสามารถนำไปทดลองใช้ได้ โดยตัวอย่างวิจัยที่ทดลองใช้ระบบการทดสอบเป็นนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6 ในเขตกรุงเทพมหานคร จำนวน 53 คน โดยเป็นเพศหญิงและเรียนสายวิทยาศาสตร์ทั้งหมด จำนวน 53 คน สำหรับอายุของตัวอย่างในการวิจัย พบว่า มีอายุ 17 ปี จำนวน 14 คน คิดเป็นร้อยละ 26.41 และอายุ 18 ปี จำนวน 39 คิดเป็นร้อยละ 73.53 โดยมีรายละเอียดข้อมูลอายุของตัวอย่างวิจัย ดังตารางที่ 5

ตารางที่ 5 ข้อมูลอายุของตัวอย่างในการวิจัย

ข้อมูล	จำนวน	ร้อยละ
อายุ		
- 17 ปี	14	26.41
- 18 ปี	39	73.59
รวม	53	100.00

ตอนที่ 2 ข้อมูลผลการตอบข้อสอบ

สำหรับข้อมูลผลการตอบข้อสอบของตัวอย่างวิจัย แบ่งการนำเสนอออกเป็น 3 ส่วน ได้แก่ 1) ผลการสอบจากระบบจำแนกตามอายุ 2) ร้อยละของคะแนนการทดสอบและการให้ข้อมูลป้อนกลับ และ 3) ค่าเฉลี่ยความสามารถและความคลาดเคลื่อนมาตรฐาน โดยมีรายละเอียดดังนี้

2.1 ผลการสอบจากระบบจำแนกตามอายุ

สำหรับผลการทดสอบจากระบบที่พัฒนาขึ้นจำแนกตามอายุ พบว่า กลุ่มอายุที่ผ่านการทดสอบมากที่สุด คือ กลุ่มอายุ 18 ปี ซึ่งมีจำนวนผ่านการทดสอบ 2 คน คิดเป็นร้อยละ 3.78 โดยรายละเอียดแสดงดังตารางที่ 6

ตารางที่ 6 ผลการทดสอบจากระบบจำแนกตามอายุ

รายการ	ผลการทดสอบจากระบบ (จำนวน(ร้อยละ))		รวม จำนวน (ร้อยละ)
	ผ่าน	ไม่ผ่าน	
อายุ	17 ปี	0 (0.00)	14 (26.41)
	18 ปี	2 (3.78)	37 (69.81)
รวม	2 (2.78)	51 (96.22)	53 (100.00)

2.2 คะแนนรวมและคะแนนการให้ข้อมูลป้อนกลับรายด้าน

สำหรับข้อมูลคะแนนรวมและคะแนนการให้ข้อมูลป้อนกลับรายด้าน พบว่า คะแนนรวมมีค่าเฉลี่ยเท่ากับ 10.58 ($SD=2.53$) สำหรับคะแนนรายด้านสำหรับการให้ข้อมูลย้อนกลับความสามารถในการใช้งานมีค่าเฉลี่ยเท่ากับ 5.20 ($SD=1.92$) ด้านความสามารถในการเขียนมีค่าเฉลี่ยเท่ากับ 2.18 ($SD=1.57$) และด้านความสามารถในการอ่านมีค่าเฉลี่ยเท่ากับ 3.18 ($SD=1.66$) โดยรายละเอียดแสดงดังตารางที่ 7

ตารางที่ 7 คะแนนรวมและคะแนนการให้ข้อมูลป้อนกลับรายด้าน

	คะแนนรวม (เต็ม 50)		คะแนนรายด้านสำหรับการให้ข้อมูลย้อนกลับ					
	M	SD	ความสามารถในการใช้งาน (เต็ม 25)		ความสามารถในการเขียน (เต็ม 10)		ความสามารถในการอ่าน (เต็ม 15)	
			M	SD	M	SD	M	SD
			การทดสอบระดับชาติ (O-NET) วิชาภาษาอังกฤษ สำหรับนักเรียนระดับชั้น มัธยมศึกษาปีที่ 6	10.58	2.53	5.20	1.92	2.18

2.3 ค่าเฉลี่ยความสามารถและความคลาดเคลื่อนมาตรฐาน

สำหรับค่าเฉลี่ยความสามารถและความคลาดเคลื่อนมาตรฐานที่ได้จากการนำระบบไปทดลองใช้ พบว่า มีค่าความสามารถ (θ) เฉลี่ย เท่ากับ -1.28 และค่าความคลาดเคลื่อนมาตรฐาน (standard error) เฉลี่ย เท่ากับ 0.88 รายละเอียดดังตารางที่ 8

ตารางที่ 8 ค่าเฉลี่ยความสามารถและความคลาดเคลื่อนมาตรฐาน

Theta (θ)	SE
-1.28	0.88

ระยะที่ 3 ผลการประเมินระบบการทดสอบแบบปรับเหมาะหลายขั้นตอนที่มีการให้ข้อมูลป้อนกลับในการทดสอบระดับชาติ (O-NET) วิชาภาษาอังกฤษสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6

สำหรับผลการประเมินระบบการทดสอบแบบปรับเหมาะหลายขั้นตอนที่มีการให้ข้อมูลป้อนกลับในการทดสอบระดับชาติ (O-NET) วิชาภาษาอังกฤษสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6 แบ่งการนำเสนอของเป็น 2 ตอนได้แก่ 1) การตรวจสอบคุณภาพแบบประเมินระบบ และ 2) ผลการประเมินระบบการทดสอบ

ตอนที่ 1 การตรวจสอบคุณภาพแบบประเมินระบบ

แบบประเมินระบบที่นำมาใช้ในการวิจัยครั้งนี้ประกอบด้วยแบบประเมินจำนวน 2 ฉบับ ได้แก่ 1) แบบประเมินระบบแบบอิงมาตรฐาน (standard evaluation) และ 2) แบบประเมินความพึงพอใจของผู้สอต่อการใช้งานระบบ ซึ่งในการตรวจสอบใช้ผู้เชี่ยวชาญด้านการวัดและประเมินผลจำนวน 5 ท่าน และดำเนินการวิเคราะห์แบบประเมินด้วยค่าดัชนีความสอดคล้อง (IOC) โดยค่า IOC ต้องมีค่าอยู่ในช่วงระหว่าง 0.5 -1.00 ขึ้นไปจึงถือว่าคำถามในแบบประเมินสามารถนำไปใช้งานได้ โดยผลการตรวจสอบคุณภาพแบบประเมินระบบแต่ละฉบับมีรายละเอียดดังนี้

1.1 แบบประเมินระบบแบบอิงมาตรฐาน (standard evaluation)

สำหรับแบบประเมินระบบแบบอิงมาตรฐานที่พัฒนาขึ้นมีจำนวนทั้งสิ้น 20 ข้อ ผลการวิเคราะห์ค่าดัชนีความสอดคล้อง (IOC) พบว่า คำถามประเมินที่มีค่าดัชนีความสอดคล้องมากกว่า 0.5 มีจำนวน 14 ข้อ ทั้งนี้ผู้เชี่ยวชาญได้แนะนำให้ดำเนินการปรับแก้แบบประเมิน มีรายละเอียดผลการวิเคราะห์ค่าดัชนีความสอดคล้อง ดังภาคผนวก ค และผู้วิจัยได้ดำเนินการปรับแก้แบบประเมินเป็นฉบับสมบูรณ์ก่อนนำไปใช้งานดังภาคผนวก ง

1.2 แบบประเมินความพึงพอใจแบบประเมินความพึงพอใจของผู้สอกับการใช้งานระบบ

สำหรับแบบประเมินความพึงพอใจของผู้สอกับการใช้งานระบบที่พัฒนาขึ้นมีจำนวน 22 ข้อ ผลการวิเคราะห์ค่าดัชนีความสอดคล้อง (IOC) พบว่า คำถามประเมินที่มีค่าดัชนีความสอดคล้องมากกว่า 0.5 มีจำนวน 22 ข้อ ทั้งนี้ผู้เชี่ยวชาญได้แนะนำให้ดำเนินการปรับแก้แบบประเมิน มีรายละเอียดผลการวิเคราะห์ค่าดัชนีความสอดคล้อง ดังภาคผนวก จ และผู้วิจัยได้ดำเนินการปรับแก้แบบประเมินเป็นฉบับสมบูรณ์ก่อนนำไปใช้งานดังภาคผนวก ฉ

ตอนที่ 2 ผลการประเมินระบบการทดสอบ

สำหรับผลการประเมินระบบการทดสอบแบบปรับเหมาะหลายขั้นตอนที่มีการให้ข้อมูลป้อนกลับในการทดสอบระดับชาติ (O-NET) วิชาภาษาอังกฤษสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6 ประกอบด้วยผลการประเมิน 2 ส่วน ได้แก่ 1) ผลการประเมินระบบก่อนนำระบบไปใช้จริงด้วยแบบประเมินระบบแบบอิงมาตรฐาน และ 2) ผลการประเมินความพึงพอใจของผู้สอกับการใช้งานระบบ สำหรับการตรวจสอบคุณภาพระบบก่อนนำระบบไปใช้จริงใช้การประเมินจากผู้เชี่ยวชาญ ซึ่งเป็นผู้เชี่ยวชาญที่มีความเชี่ยวชาญด้านเทคโนโลยี จำนวน 2 ท่านและผู้เชี่ยวชาญด้านการวัดและประเมินผลการศึกษา จำนวน 3 ท่าน รวมทั้งสิ้น จำนวน 5 ท่าน และการประเมินความพึงพอใจมาจากตัวอย่างวิจัยที่ทดลองใช้งานระบบ โดยรายละเอียดผลการประเมินมีดังนี้

1) ผลการประเมินระบบก่อนนำระบบไปใช้จริงด้วยแบบประเมินระบบแบบอิงมาตรฐาน

สำหรับการตรวจสอบคุณภาพระบบจากผู้เชี่ยวชาญก่อนนำระบบไปใช้จริงใช้แบบประเมินระบบแบบอิงมาตรฐาน (standard evaluation) โดยแบ่งประเด็นในการตรวจสอบออกเป็น 4 ด้าน

ได้แก่ ด้านความมีประโยชน์ (utility) ด้านความเป็นไปได้ในการนำไปใช้ (feasibility) ด้านความเหมาะสม (propriety) และด้านความถูกต้อง (accuracy) โดยการประเมินใช้ผู้เชี่ยวชาญทางด้านการวัดและประเมินผลจำนวน 3 ท่าน และผู้เชี่ยวชาญทางด้านเทคโนโลยีสารสนเทศจำนวน 2 ท่าน

ผลการประเมิน พบว่า ผู้เชี่ยวชาญมีความเห็นว่าระบบมีความเหมาะสมทั้ง 4 ด้าน โดยอันดับหนึ่งได้แก่ ด้านความเป็นไปได้ในการนำไปใช้ (feasibility) ($M=4.96$, $SD=0.15$) รองลงมาคือ ด้านความถูกต้อง (accuracy) ($M=4.88$, $SD=0.18$) ด้านความมีประโยชน์ (utility) ($M=4.84$, $SD=0.22$) และด้านความเหมาะสม (propriety) ($M=4.83$, $SD=0.14$) นอกจากนี้ ผู้เชี่ยวชาญได้ให้ข้อคิดเห็นว่าระบบมีความชัดเจนเข้าใจง่ายและเป็นการพัฒนาระบบที่คำนึงถึงมาตรฐาน สามารถเข้าถึงผู้ใช้งาน รวมถึงให้ข้อมูลย้อนกลับเพื่อการพัฒนา ก่อให้เกิดประโยชน์ต่อส่วนรวม สำหรับส่วนที่ควรปรับปรุงเพิ่มเติม คือ ควรปรับรายละเอียดการลงทะเบียนเข้าใช้งานของผู้สอบ โดยให้ผู้สอบกรอกข้อมูลส่วนตัวก่อนเข้าสู่ระบบเพื่อให้เบ็ดเสร็จในจุดเดียว นอกจากนี้ ควรมีระบบการเรียกคืนรหัสผ่านอีเมลกรณีผู้ใช้งานลืมรหัสผ่าน (forget password) และควรให้ผู้ใช้งานสามารถอัปเดตรูปภาพตนเองในระบบได้ สำหรับในอนาคตอาจเชื่อมโยงระบบเข้ากับสื่อโซเชียลมีเดียต่าง ๆ เช่น เฟซบุ๊ก ไลน์ เพื่อให้ผู้ใช้งานเข้าถึงระบบการทดสอบได้ง่ายขึ้นและอาจออกแบบระบบการทดสอบให้รองรับอุปกรณ์การใช้งานอื่นนอกจากเครื่องคอมพิวเตอร์ เช่น แท็บเล็ต สมาร์ทโฟน โดยรายละเอียดผลการประเมินแสดงดังตารางที่ 9

ตารางที่ 9 ผลการประเมินระบบด้วยแบบประเมินระบบแบบอิงมาตรฐาน

ประเด็นการประเมิน	<i>M</i>	<i>SD</i>	ความหมาย
1. ด้านความมีประโยชน์ (utility)	4.84	0.22	มากที่สุด
1.1 ระบบการทดสอบสามารถตอบสนองความต้องการในการเตรียมความพร้อมเพื่อการทดสอบระดับชาติของผู้สอบ	5.00	0.00	มากที่สุด
1.2 ระบบการทดสอบสามารถนำไปใช้เตรียมความพร้อมให้กับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6 เพื่อทดสอบระดับชาติ (O-NET) วิชาภาษาอังกฤษ	5.00	0.00	มากที่สุด
1.3 ระบบการทดสอบมีรายละเอียดเนื้อหาเหมือนกับการทดสอบระดับชาติ	4.60	0.44	มากที่สุด
1.4 รายงานผลการทดสอบสามารถแสดงระดับความสามารถของผู้สอบได้	5.00	0.00	มากที่สุด
1.5 รายงานข้อมูลย้อนกลับสามารถ	4.60	0.44	มากที่สุด

ประเด็นการประเมิน	<i>M</i>	<i>SD</i>	ความหมาย
นำไปใช้ประโยชน์ได้อย่างเป็นรูปธรรม			
2. ด้านความเป็นไปได้ในการนำไปใช้ (feasibility)	4.96	0.15	มากที่สุด
2.1 ระบบการทดสอบสามารถนำไปใช้ งานในการเตรียมความพร้อมเพื่อการทดสอบ ระดับชาติได้จริง	5.00	0.00	มากที่สุด
2.2 ระบบการทดสอบสามารถตอบสนอง การใช้งานของผู้สอบได้อย่างรวดเร็ว	5.00	0.00	มากที่สุด
2.3 รายงานผลการทดสอบสามารถแสดง ผลได้ทันทีหลังจบการทดสอบ	5.00	0.00	มากที่สุด
2.4 รายงานข้อมูลป้อนกลับมีสารสนเทศ ที่เหมาะสมสำหรับการพัฒนาผู้สอบ	4.80	0.36	มากที่สุด
2.5 การพัฒนาระบบการทดสอบมีความ คุ้มค่า	5.00	0.00	มากที่สุด
3. ด้านความเหมาะสม (propriety)	4.83	0.14	มากที่สุด
3.1 ระบบการทดสอบมีคำชี้แจงก่อนการ สอบชัดเจน	4.80	0.36	มากที่สุด
3.2 ขอบเขตเนื้อหาในการทดสอบมีความ เหมาะสม	4.80	0.36	มากที่สุด
3.3 ระยะเวลาในการทดสอบมีความ เหมาะสม	4.80	0.36	มากที่สุด
3.4 รายงานผลการทดสอบมีสารสนเทศ ครบถ้วน	5.00	0.00	มากที่สุด
3.5 รายงานผลข้อมูลป้อนกลับมี รายละเอียดที่จำเป็นเพื่อการพัฒนา	4.80	0.36	มากที่สุด
3.6 ระบบการทดสอบมีความสอดคล้อง กับแนวทางการเพิ่มระดับความสามารถด้าน ภาษาอังกฤษของผู้เรียน	4.80	0.36	มากที่สุด
4. ด้านความถูกต้อง (accuracy)	4.88	0.18	มากที่สุด
4.1 การอธิบายวัตถุประสงค์ของการ ทดสอบถูกต้องชัดเจน	5.00	0.00	มากที่สุด

ประเด็นการประเมิน	<i>M</i>	<i>SD</i>	ความหมาย
4.2 การอธิบายขั้นตอนในการทดสอบ ถูกต้องชัดเจน	4.80	0.36	มากที่สุด
4.3 กระบวนการคัดเลือกข้อสอบมีความ สอดคล้องกับเนื้อหาในการทดสอบระดับชาติ	4.80	0.36	มากที่สุด
4.4 รายงานผลการทดสอบมีความ ถูกต้อง	4.80	0.36	มากที่สุด
4.5 รายงานผลข้อมูลป้อนกลับสอดคล้อง กับผลการทดสอบ	5.00	0.00	มากที่สุด

2) ผลการประเมินความพึงพอใจของผู้สอบกับการใช้งานระบบ

สำหรับการตรวจสอบคุณภาพระบบด้วยแบบประเมินความพึงพอใจของผู้สอบกับการใช้งานระบบ มีประเด็นในการสอบถามทั้งสิ้น 5 ประเด็น ได้แก่ 1) ความเห็นต่อการใช้งานระบบในภาพรวม 2) หน้าจอการแสดงผล (screen) 3) คำศัพท์และสารสนเทศของระบบ (terminology and system information) 4) การเรียนรู้ (learning) และ 5) ความสามารถระบบ (system capabilities) ผลการประเมิน พบว่า ความเห็นต่อการใช้งานระบบในภาพรวม ค่อนข้างดี ($M=5.88$, $SD=1.97$) ด้านหน้าจอการแสดงผล (screen) ประเด็นการใช้งานของเครื่องมือ (tools) ต่าง ๆ บนหน้าจอมีค่าสูงสุด หมายความว่า การใช้งานค่อนข้างง่าย ($M=6.10$, $SD=2.03$) ด้านคำศัพท์และสารสนเทศของระบบ ประเด็นข้อความบนหน้าจอที่แสดงแก่ผู้ใช้งานเพื่อป้อนข้อมูลมีค่าสูงสุด หมายความว่า ระบบข้อความบนหน้าจอมีความชัดเจน ($M=6.17$, $SD=1.67$) ด้านการเรียนรู้ประเด็นผู้ใช้งานสามารถจำชื่อและการใช้งานคำสั่งมีค่าสูงสุด หมายความว่า ผู้ใช้งานสามารถจำชื่อและการใช้งานคำสั่งได้อย่างง่าย ($M=6.43$, $SD=2.13$) และด้านความสามารถของระบบประเด็นความเที่ยง (reliability) ของระบบมีค่าสูงสุด หมายความว่า การทำงานของระบบมีความคงเส้นคงวา ($M=6.33$, $SD=2.20$) โดยรายละเอียดแสดงดังตารางที่ 10

ตารางที่ 10 ผลการประเมินความพึงพอใจของผู้สอบกับการใช้งานระบบ

ข้อ	รายการประเมิน	มาตรวัด 10 ระดับ (0 – 9)		<i>M</i>	<i>SD</i>
1	ภาพรวมปฏิสัมพันธ์ของระบบ	ไม่ดี	ดีเยี่ยม	5.88	1.97
หน้าจอการแสดงผล (screen)					
2	รูปแบบของข้อความและ ส่วนประกอบต่าง ๆ บนหน้าจอ	ยากต่อการ อ่าน	ง่ายต่อการ อ่าน	5.93	1.98
3	การใช้งานของเครื่องมือ (tools) ต่าง ๆ บนหน้าจอ	ยาก	ง่าย	6.10	2.03
4	การจัดการสารสนเทศบน	สับสน	ชัดเจน	5.52	2.42

ข้อ	รายการประเมิน	มาตรวัด 10 ระดับ (0 – 9)		M	SD
	หน้าจอ				
5	ลำดับการแสดงผลและส่วนประกอบต่าง ๆ บนหน้าจอ	สับสน	ชัดเจน	6.10	2.42
คำศัพท์และสารสนเทศของระบบ (terminology and system information)					
6	การใช้คำศัพท์เฉพาะในส่วนต่าง ๆ ของระบบ	ไม่ถูกต้อง	ถูกต้อง	5.83	2.41
7	คำศัพท์ทางคอมพิวเตอร์มีความสัมพันธ์กับเครื่องมือที่ใช้	ไม่สัมพันธ์	สัมพันธ์	5.69	2.53
8	ตำแหน่งของข้อความบนหน้าจอ	ไม่เหมาะสม	เหมาะสม	6.14	1.86
9	ข้อความบนหน้าจอที่แสดงแก่ผู้ใช้งานเพื่อป้องกันข้อมูล	สับสน	ชัดเจน	6.17	1.67
10	การแจ้งเตือนสถานะว่ากำลังดำเนินการอะไรอยู่	ไม่มี	มี	5.79	2.30
11	การมีข้อความแจ้งเตือนความผิดพลาด ช่วยให้ใช้งานระบบได้ง่ายขึ้น	ไม่ใช่	ใช่	5.83	2.29
การเรียนรู้ (learning)					
12	ผู้ใช้งานสามารถเรียนรู้การใช้งานและการทำงานของระบบ	ยาก	ง่าย	6.14	2.19
13	การแสดงองค์ประกอบใหม่ (new features) ด้วยการแสดงตัวอย่างและข้อผิดพลาด ทำให้ผู้ใช้งานเข้าใจการทำงานของระบบได้ดีขึ้น	ไม่ใช่	ใช่	5.81	2.22
14	ผู้ใช้งานสามารถจำชื่อและการทำงานของคำสั่ง	ยาก	ง่าย	6.43	2.13
15	เครื่องมือสามารถทำงานได้อย่างต่อเนื่อง	ไม่ใช่	ใช่	6.40	1.84
16	ข้อความช่วยเหลือบนหน้าจอทำให้ผู้ใช้งานเข้าใจการทำงานของระบบได้ดีขึ้น	ไม่ใช่	ใช่	6.14	1.92
17	เอกสารอ้างอิงเพิ่มเติม	สับสน	ชัดเจน	5.98	1.86
ความสามารถของระบบ (system capabilities)					
18	ความเร็วระบบ	ช้า	เร็ว	5.98	2.14

ข้อ	รายการประเมิน	มาตรวัด 10 ระดับ (0 – 9)		<i>M</i>	<i>SD</i>
19	ความเที่ยง (reliability) ของระบบ	ไม่เที่ยง	เที่ยง	6.33	2.20
20	แนวโน้มการทำงานของระบบ	ติดขัด	ราบรื่น	6.21	2.33
21	การแก้ไขข้อผิดพลาด	ยาก	ง่าย	6.07	2.35
22	การพิจารณาความต้องการของผู้ใช้งานจากพฤติกรรมการใช้งานระบบ	ไม่มี	มี	5.21	2.65

บทที่ 5

สรุปผลการวิจัย อภิปรายผล และข้อเสนอแนะ

การวิจัยเรื่องการพัฒนากระบวนการทดสอบแบบปรับเหมาะหลายขั้นตอนที่มีการให้ข้อมูลป้อนกลับในการทดสอบระดับชาติ (O-NET) วิชาภาษาอังกฤษสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6 มีวัตถุประสงค์ คือ 1) เพื่อพัฒนากระบวนการทดสอบแบบปรับเหมาะหลายขั้นตอนที่มีการให้ข้อมูลป้อนกลับในการทดสอบระดับชาติ (O-NET) วิชาภาษาอังกฤษสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6 และ 2) เพื่อประเมินคุณภาพระบบการทดสอบแบบปรับเหมาะหลายขั้นตอนที่มีการให้ข้อมูลป้อนกลับในการทดสอบระดับชาติ (O-NET) วิชาภาษาอังกฤษสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6

ในการดำเนินการวิจัยแบ่งการดำเนินงานออกเป็น 3 ระยะ ได้แก่ ระยะที่ 1 การพัฒนากระบวนการทดสอบแบบปรับเหมาะหลายขั้นตอนที่มีการให้ข้อมูลป้อนกลับในการทดสอบระดับชาติ (O-NET) วิชาภาษาอังกฤษสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6 ระยะที่ 2 การทดลองใช้ระบบการทดสอบแบบปรับเหมาะหลายขั้นตอนที่มีการให้ข้อมูลป้อนกลับในการทดสอบระดับชาติ (O-NET) วิชาภาษาอังกฤษสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6 และระยะที่ 3 การประเมินผลระบบการทดสอบแบบปรับเหมาะหลายขั้นตอนที่มีการให้ข้อมูลป้อนกลับในการทดสอบระดับชาติ (O-NET) วิชาภาษาอังกฤษสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6 สำหรับระยะที่ 1 แบ่งการดำเนินงานออกเป็น 3 ตอนได้แก่ 1) ศึกษาองค์ประกอบของการทดสอบแบบปรับเหมาะหลายขั้นตอนและรูปแบบการรายงานผลข้อมูลป้อนกลับ 2) ออกแบบและพัฒนากระบวนการทดสอบและรูปแบบการรายงานผลข้อมูลป้อนกลับ และ 3) ตรวจสอบคุณภาพของระบบการทดสอบที่มีการสะท้อนข้อมูลป้อนกลับ

สำหรับการวิเคราะห์ข้อสอบเพื่อจัดทำคลังข้อสอบ ดำเนินการนำผลการตอบข้อสอบระดับชาติ (O-NET) วิชาภาษาอังกฤษสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6 ระหว่างปี 2558-2561 มาทำการวิเคราะห์ข้อสอบตามทฤษฎีการตอบสนองข้อสอบ (IRT) แบบ 3 พารามิเตอร์ ได้แก่ ค่าความยาก (b) อำนาจจำแนก (a) และการเดา (c) ด้วยโปรแกรม R ซึ่งการคัดเลือกข้อสอบเพื่อจัดทำคลังมีเกณฑ์ คือ ค่าความยากอยู่ในช่วง -4 ถึง 4 ค่าอำนาจจำแนกอยู่ระหว่าง 0 ถึง ∞ และความน่าจะเป็นในการเดาไม่เกิน 0.50 หากข้อสอบมีค่าพารามิเตอร์ของข้อสอบไม่เป็นไปตามเกณฑ์จะตัดข้อสอบข้อดังกล่าวทิ้งโดยไม่นำมาจัดทำคลังข้อสอบ สำหรับการวิเคราะห์ข้อมูลเพื่อตรวจสอบคุณภาพของระบบก่อนนำระบบไปใช้งาน จะดำเนินการวิเคราะห์ด้วยแบบประเมินระบบแบบอิงมาตรฐาน (standard evaluation) โดยในการตรวจสอบคุณภาพระบบก่อนการนำไปทดลองใช้อาศัยดุลพินิจจากผู้เชี่ยวชาญด้านการทดสอบและผู้เชี่ยวชาญด้านเทคโนโลยีสารสนเทศ จำนวน 5 ท่าน

และนำไปทดลองใช้เบื้องต้นกับกลุ่มตัวอย่างจำนวน 10 คน หลังจากทำการตรวจสอบคุณภาพระบบเบื้องต้นเรียบร้อยแล้วจึงนำระบบไปทดลองใช้ในระยะเวลาที่ 2 ต่อไป

ระยะที่ 2 เป็นการนำไประบบไปทดลองใช้กับตัวอย่างในการวิจัย จำนวน 53 คน ซึ่งเป็นนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6 ในเขตกรุงเทพมหานครโดยเป็นเพศหญิงและเรียนสายวิทยาศาสตร์ทั้งหมด และในระยะเวลาที่ 3 เป็นการประเมินระบบด้วยการประเมินความพึงพอใจของผู้สอบกับการใช้งานระบบ

สรุปผลการวิจัย

สรุปผลการวิจัยมีดังนี้

1. ผลการพัฒนาระบบการทดสอบแบบปรับเหมาะหลายขั้นตอนที่มีการให้ข้อมูลป้อนกลับในการทดสอบระดับชาติ (O-NET) วิชาภาษาอังกฤษสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6 ประกอบด้วยกระบวนการในการศึกษา 4 ส่วน 1) ผลการศึกษาองค์ประกอบของการทดสอบแบบปรับเหมาะหลายขั้นตอนที่มีการให้ข้อมูลป้อนกลับ 2) ผลการพัฒนาระบบการทดสอบและรูปแบบการรายงานข้อมูลป้อนกลับ และ 3) ผลการวิเคราะห์ข้อสอบวิชาภาษาอังกฤษของการทดสอบระดับชาติ (O-NET) สำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6 และ 4) ผลการตรวจสอบคุณภาพระบบก่อนนำระบบไปใช้งานจริง

ผลการศึกษาองค์ประกอบของการทดสอบแบบปรับเหมาะหลายขั้นตอนที่มีการให้ข้อมูลป้อนกลับ พบว่า องค์ประกอบของระบบประกอบด้วย 4 องค์ประกอบ ได้แก่ 1) จุดเริ่มต้นในการทดสอบ (initial stage) 2) การประมาณค่าความสามารถของผู้สอบ (ability estimation) 3) การคัดเลือกข้อสอบ (item selection) และ 4) การยุติการทดสอบ (stopping criteria) และรูปแบบรายงานผลที่สะท้อนข้อมูลป้อนกลับ ได้แก่ รายงานคะแนนสังเกตได้และรายงานระดับความสามารถของผู้สอบ

ผลการพัฒนาระบบการทดสอบแบ่งออกเป็น 6 ขั้นตอน ได้แก่ 1) การลงทะเบียนสำหรับตัวอย่างวิจัย 2) การเข้าสู่ระบบการทดสอบ 3) การดำเนินการทดสอบ 4) การยุติการทดสอบ 5) การประเมินคุณภาพระบบการทดสอบ และ 6) การรายงานผลการทดสอบ สำหรับผลการวิเคราะห์ข้อสอบวิชาภาษาอังกฤษของการทดสอบระดับชาติ (O-NET) สำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6 พบว่าข้อสอบ จำนวน 320 ข้อ ซึ่งใช้ในการทดสอบระหว่างปี 2558-2561 เมื่อนำมาวิเคราะห์คุณภาพข้อสอบตามเกณฑ์ คือ ค่าความยาก (b) อยู่ในช่วง -4 ถึง 4 ค่าอำนาจจำแนก (a) อยู่ในช่วง 0 ถึง α และความน่าจะเป็นในการเดา (c) ไม่เกิน 0.50 ตามทฤษฎีการตอบสนองข้อสอบ (item response theory: IRT) พบว่า มีข้อสอบที่ผ่านเกณฑ์ จำนวน 179 ข้อ

ผลการตรวจสอบคุณภาพระบบก่อนนำระบบไปใช้งานจริงโดยผู้เชี่ยวชาญด้วยแบบประเมินระบบแบบอิงมาตรฐาน (standard evaluation) ซึ่งแบ่งประเด็นในการตรวจสอบออกเป็น 4 ด้าน

ได้แก่ ด้านความมีประโยชน์ (utility) ด้านความเป็นไปได้ในการนำไปใช้ (feasibility) ด้านความเหมาะสม (propriety) และด้านความถูกต้อง (accuracy) ผลการประเมิน พบว่า ระบบมีความเหมาะสมทั้ง 4 ด้าน โดยอันดับหนึ่งได้แก่ ด้านความเป็นไปได้ในการนำไปใช้ (feasibility) ($M=4.96$, $SD=0.15$) รองลงมาคือ ด้านความถูกต้อง (accuracy) ($M=4.88$, $SD=0.18$) ด้านความมีประโยชน์ (utility) ($M=4.84$, $SD=0.22$) และด้านความเหมาะสม (propriety) ($M=4.83$, $SD=0.14$) นอกจากนี้ ผู้เชี่ยวชาญได้ให้ข้อคิดเห็นว่า ระบบมีความชัดเจนเข้าใจง่ายและเป็นการพัฒนาระบบที่คำนึงถึงมาตรฐาน สามารถเข้าถึงผู้ใช้งาน รวมถึงให้ข้อมูลย้อนกลับเพื่อการพัฒนา ก่อให้เกิดประโยชน์ต่อส่วนรวม สำหรับส่วนที่ควรปรับปรุงเพิ่มเติม คือ ควรปรับรายละเอียดการลงทะเบียนเข้าใช้งานของผู้สอบ โดยให้ผู้สอบกรอกข้อมูลส่วนตัวก่อนเข้าสู่ระบบเพื่อให้เกิดเสร็จในจุดเดียว นอกจากนี้ ควรมีระบบการเรียกคืนรหัสผ่านอีเมลกรณีผู้ใช้งานลืมรหัสผ่าน (forget password) และควรให้ผู้ใช้งานสามารถอัปเดตรูปภาพตนเองในระบบได้ สำหรับในอนาคตอาจเชื่อมโยงระบบเข้ากับสื่อโซเชียลมีเดียต่าง ๆ เช่น เฟซบุ๊ก ไลน์ เพื่อให้ผู้ใช้งานเข้าถึงระบบการทดสอบได้ง่ายขึ้นและอาจออกแบบระบบการทดสอบให้รองรับอุปกรณ์การใช้งานอื่นนอกจากเครื่องคอมพิวเตอร์ เช่น แท็บเล็ต สมาร์ทโฟน

2. ผลวิเคราะห์การทดลองใช้ระบบการทดสอบแบบปรับเหมาะหลายขั้นตอนที่มีการให้ข้อมูลย้อนกลับในการทดสอบระดับชาติ (O-NET) วิชาภาษาอังกฤษสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6 ดำเนินการเก็บรวบรวมข้อมูลจากตัวอย่างวิจัย จำนวน 53 คน เป็นนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6 ในเขตกรุงเทพมหานคร จำนวน 53 คน โดยเป็นเพศหญิง และเรียนสายวิทยาศาสตร์ทั้งหมด สำหรับอายุของตัวอย่างในการวิจัย พบว่า มีอายุ 17 ปี จำนวน 14 คน คิดเป็นร้อยละ 26.41 และอายุ 18 ปี จำนวน 39 คิดเป็นร้อยละ 73.53

สำหรับข้อมูลที่ได้จากการนำระบบไปทดลองใช้ พบว่า กลุ่มอายุที่ผ่านการทดสอบมากที่สุด คือ กลุ่มอายุ 18 ปี ซึ่งมีจำนวนผ่านการทดสอบ 2 คน คิดเป็นร้อยละ 3.78 ในส่วนของข้อมูลคะแนนรวมและคะแนนการให้ข้อมูลย้อนกลับรายด้าน พบว่า คะแนนรวมมีค่าเฉลี่ยเท่ากับ 10.58 ($SD=2.53$) สำหรับคะแนนรายด้านสำหรับการให้ข้อมูลย้อนกลับ ความสามารถในการใช้งานมีค่าเฉลี่ยเท่ากับ 5.20 ($SD=1.92$) ด้านความสามารถในการเขียนมีค่าเฉลี่ยเท่ากับ 2.18 ($SD=1.57$) และด้านความสามารถในการอ่านมีค่าเฉลี่ยเท่ากับ 3.18 ($SD=1.66$) และเมื่อพิจารณาค่าเฉลี่ยความสามารถและความคลาดเคลื่อนมาตรฐานที่ได้จากการนำระบบไปทดลองใช้ พบว่า มีค่าความสามารถ (θ) เฉลี่ย เท่ากับ -1.28 และค่าความคลาดเคลื่อนมาตรฐาน (standard error) เฉลี่ย เท่ากับ 0.88

3. ผลการประเมินระบบการทดสอบแบบปรับเหมาะหลายขั้นตอนที่มีการให้ข้อมูลย้อนกลับในการทดสอบระดับชาติ (O-NET) วิชาภาษาอังกฤษสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6 แบ่งเป็น 2 ส่วน ได้แก่ 1) การตรวจสอบคุณภาพแบบประเมินระบบ และ 2) ผลการประเมินระบบการทดสอบ ผลการตรวจสอบคุณภาพแบบประเมินระบบสำหรับแบบประเมินระบบแบบอิงมาตรฐาน

(standard evaluation) ที่พัฒนาขึ้นจำนวน 20 ข้อ พบว่า คำถามประเมินที่มีค่าดัชนีความสอดคล้องมากกว่า 0.5 มีจำนวน 14 ข้อ และผู้เชี่ยวชาญได้แนะนำให้ดำเนินการปรับแก้รายการประเมินบางส่วนซึ่งผู้วิจัยได้ดำเนินการปรับแก้ก่อนนำไปใช้งานเรียบร้อยแล้ว สำหรับแบบประเมินความพึงพอใจของผู้สอบกับการใช้งานระบบที่พัฒนาขึ้นมีจำนวน 22 ข้อ ผลการวิเคราะห์ค่าดัชนีความสอดคล้อง (IOC) พบว่า คำถามประเมินที่มีค่าดัชนีความสอดคล้องมากกว่า 0.5 มีจำนวน 22 ข้อ ทั้งนี้ผู้เชี่ยวชาญได้แนะนำให้ดำเนินการปรับแก้แบบประเมินบางส่วน และผู้วิจัยได้ดำเนินการปรับแก้ก่อนนำไปใช้เรียบร้อยแล้ว

ผลการประเมินระบบก่อนนำระบบไปใช้จริงด้วยแบบประเมินระบบแบบอิงมาตรฐาน พบว่าผู้เชี่ยวชาญมีความเห็นว่าระบบมีความเหมาะสมทั้ง 4 ด้าน โดยอันดับหนึ่งได้แก่ ด้านความเป็นไปได้ในการนำไปใช้ (feasibility) ($M=4.96$, $SD=0.15$) รองลงมาคือ ด้านความถูกต้อง (accuracy) ($M=4.88$, $SD=0.18$) ด้านความมีประโยชน์ (utility) ($M=4.84$, $SD=0.22$) และด้านความเหมาะสม (propriety) ($M=4.83$, $SD=0.14$) นอกจากนี้ ผู้เชี่ยวชาญได้ให้ข้อคิดเห็นว่า ระบบมีความชัดเจน เข้าใจง่ายและเป็นการพัฒนาระบบที่คำนึงถึงมาตรฐาน สามารถเข้าถึงผู้ใช้งาน รวมถึงให้ข้อมูลย้อนกลับเพื่อการพัฒนา ก่อให้เกิดประโยชน์ต่อส่วนรวม สำหรับส่วนที่ควรปรับปรุงเพิ่มเติม คือ ควรปรับรายละเอียดการลงทะเบียนเข้าใช้งานของผู้สอบ โดยให้ผู้สอบกรอกข้อมูลส่วนตัวก่อนเข้าสู่ระบบ เพื่อให้เปิดเสร็จในจุดเดียว นอกจากนี้ ควรมีระบบการเรียกคืนรหัสผ่านอีเมลกรณีผู้ใช้งานลืมรหัสผ่าน (forget password) และควรให้ผู้ใช้งานสามารถอัปเดตรูปภาพตนเองในระบบได้ สำหรับในอนาคตอาจเชื่อมโยงระบบเข้ากับสื่อโซเชียลมีเดียต่าง ๆ เช่น เฟซบุ๊ก ไลน์ เพื่อให้ผู้ใช้งานเข้าถึงระบบการทดสอบได้ง่ายขึ้นและอาจออกแบบระบบการทดสอบให้รองรับอุปกรณ์การใช้งานอื่นนอกจากเครื่องคอมพิวเตอร์ เช่น แท็บเล็ต สมาร์ทโฟน

ผลการประเมินความพึงพอใจของผู้สอบกับการใช้งานระบบ พบว่า ความเห็นต่อการใช้งานระบบในภาพรวม ค่อนข้างดี ($M=5.88$, $SD=1.97$) ด้านหน้าจอการแสดงผล (screen) ประเด็นการใช้งานของเครื่องมือ (tools) ต่าง ๆ บนหน้าจอมีค่าสูงสุด หมายความว่า การใช้งานค่อนข้างง่าย ($M=6.10$, $SD=2.03$) ด้านคำศัพท์และสารสนเทศของระบบประเด็นข้อความบนหน้าจอที่แสดงแก่ผู้ใช้งานเพื่อป้อนข้อมูลมีค่าสูงสุด หมายความว่าระบบข้อความบนหน้าจอมีความชัดเจน ($M=6.17$, $SD=1.67$) ด้านการเรียนรู้ประเด็นผู้ใช้งานสามารถจำชื่อและการใช้งานคำสั่งมีค่าสูงสุด หมายความว่าผู้ใช้งานสามารถจำชื่อและการใช้งานคำสั่งได้อย่างง่าย ($M=6.43$, $SD=2.13$) และด้านความสามารถของระบบประเด็นความเที่ยง (reliability) ของระบบมีค่าสูงสุด หมายความว่าการทำงานของระบบมีความคงเส้นคงวา ($M=6.33$, $SD=2.20$)

อภิปรายผลการวิจัย

จากผลการวิจัยสามารถนำมาอภิปรายได้ 4 ประเด็น ได้แก่ 1) การพัฒนาระบบการทดสอบแบบปรับเหมาะหลายขั้นตอน 2) การทดลองใช้ระบบการทดสอบแบบปรับเหมาะหลายขั้นตอน 3) รายงานการสะท้อนข้อมูลป้อนกลับ และ 4) การประเมินระบบการทดสอบแบบปรับเหมาะหลายขั้นตอน

1. การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ เป็นการทดสอบที่ผู้สอบทำแบบสอบต่างฉบับตามระดับความสามารถแต่ยังคงให้ความแม่นยำจากผลการสอบ โดยมีหลักการคือผู้สอบจะเริ่มทำข้อสอบจากข้อที่มีระดับความยากปานกลาง หากผู้สอบตอบข้อสอบผิดข้อสอบ ข้อถัดไปจะง่ายลงในทางตรงข้าม หากผู้สอบตอบข้อสอบถูกข้อสอบข้อถัดไปจะยากขึ้น ผู้สอบจะทำการสอบไปเรื่อย ๆ จนกระทั่งระดับความสามารถของผู้สอบคงที่หรืออยู่ในเกณฑ์ความคลาดเคลื่อนที่ยอมรับได้ การทดสอบจะยุติลง (Cisar et al., 2010; ศิริชัย กาญจนวาสี, 2555) ต่อมาพบว่า การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ (computerize adaptive testing: CAT) มีการประมาณค่าความสามารถที่สูงและต่ำกว่าจริง เนื่องประมาณค่าข้อสอบเริ่มต้นจากข้อสอบเพียงข้อเดียว จึงได้มีการพัฒนาการทดสอบแบบปรับเหมาะหลายขั้นตอน (multistage adaptive testing: MST) ขึ้น โดยผู้ที่นำแนวคิดดังกล่าวมาใช้ คือ Lord และ Novick ซึ่งพบว่าสามารถประมาณค่าความสามารถได้แม่นยำกว่าการทดสอบแบบปรับเหมาะแบบดั้งเดิม (Zenisky, Hambleton, & Luecht, 2010) จากผลการวิจัยพบว่า การทดสอบแบบปรับเหมาะหลายขั้นตอนที่มีการให้ข้อมูลป้อนกลับในการทดสอบระดับชาติ (O-NET) วิชาภาษาอังกฤษสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6 มีองค์ประกอบที่สำคัญ ได้แก่ 1) จุดเริ่มต้นในการทดสอบ (initial stage) 2) การประมาณค่าความสามารถของผู้สอบ (ability estimation) 3) การคัดเลือกข้อสอบ (item selection) และ 4) การยุติการทดสอบ (stopping criteria) ซึ่งสอดคล้องกับ Chaimongkol, Pasiphol, and Kanjanawasee (2016)

2. สำหรับการนำระบบการทดสอบแบบปรับเหมาะหลายขั้นตอนที่มีการให้ข้อมูลป้อนกลับในการทดสอบระดับชาติ (O-NET) วิชาภาษาอังกฤษสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6 ไปทดลองใช้กับตัวอย่างวิจัย จำนวน 53 คน พบว่า มีผู้สอบผ่านเพียง 2 คน อาจเนื่องมาจากตัวอย่างวิจัยมิได้เตรียมความพร้อมก่อนการสอบ เนื่องจากผลการสอบไม่มีผลได้ผลเสีย (high stake) กับตัวอย่างวิจัย ทำให้ตัวอย่างวิจัยอาจไม่ได้ใช้ความพยายามเต็มที่ในการทำข้อสอบ ซึ่งสอดคล้องกับแนวคิดของ Amrein and Beliner (2003) และ Mesler (2008) ว่าข้อสอบที่มีผลได้ผลเสียกับผู้สอบ จะทำให้ผู้สอบมีความพยายามในการทำข้อสอบมากกว่าปกติ

3. การให้ข้อมูลป้อนกลับเป็นองค์ประกอบที่สำคัญในการบวนการจัดการเรียนการสอนและเป็นส่วนหนึ่งในการประเมินความก้าวหน้าระหว่างเรียน (formative assessment) เพื่อระบุให้ผู้เรียนทราบจุดเด่นและข้อจำกัด และนำข้อจำกัดดังกล่าวไปพัฒนาตนเองจนนำไปสู่การบรรลุผลสัมฤทธิ์ตามศักยภาพของผู้เรียน โดยข้อมูลป้อนกลับแบบทั่วไป (feedback) เป็นการให้ข้อความที่

สามารถอธิบายผลงานของนักเรียนว่าถูกหรือผิดและดีหรือไม่ดีอย่างไร โดยอิงจากเกณฑ์ในการประเมิน (โชติกา ภาชีผล ประกอบ กรณีกิจ และพิทักษ์ โสถถยาคม, 2558) และยังพบว่า การนำการสะท้อนข้อมูลป้อนกลับมาใช้ระหว่างครูฝึกสอนกับครูพี่เลี้ยงช่วยเพิ่มประสิทธิภาพการพัฒนาทักษะการให้เหตุผลทางคลินิกของผู้เรียนได้ (Damar, 2013; Wojcikowski & Brownie, 2013)

ในการวิจัยนี้ใช้การให้ข้อมูลป้อนกลับแบบทั่วไป (feedback) โดยรูปแบบรายงานจะประกอบด้วยข้อมูลส่วนบุคคลของตัวอย่างวิจัย ได้แก่ ชื่อ นามสกุล สายการเรียน ระดับชั้น โรงเรียน และจังหวัด และข้อมูลในส่วนของการทดสอบ จะแสดงจำนวนข้อสอบทั้งหมดที่ตัวอย่างวิจัยทำการสอบ พร้อมทั้งแสดงคะแนนผลการตอบข้อสอบทั้งหมด หลังจากนั้นปรับข้อมูลให้เป็นร้อยละเพื่อนำไปใช้แสดงผลการผ่านการทดสอบรายด้าน โดยหากมีคะแนนมากกว่าร้อยละ 50 ถือว่าผ่านการทดสอบรายด้าน สำหรับการผ่านการทดสอบจะไม่ใช้การพิจารณาจากคะแนนดิบ (raw score) แต่จะพิจารณาจากคะแนนในมาตร (scale score) ซึ่งมีการปรับเทียบให้เป็นคะแนนมาตรฐานที่ (standardized t-score) โดยอิงจากค่าความสามารถผู้สอบ (Θ) หากค่าคะแนนมาตรมากกว่าร้อยละ 50 จึงจะถือว่าผ่านการทดสอบ

จากข้อมูลข้างต้นเห็นได้ว่าการสะท้อนข้อมูลป้อนกลับเป็นสิ่งจำเป็นต่อการพัฒนาผู้เรียน หากผู้เรียนไม่ได้รับข้อมูลย้อนกลับที่รวดเร็วและถูกต้อง อาจทำให้การพัฒนาผู้เรียนเป็นอย่างล่าช้า โดยการให้ข้อมูลป้อนกลับควรมีความสะดวกกับผู้รับข้อมูล ซึ่งในปัจจุบันมีการนำระบบออนไลน์มาใช้ในการให้ข้อมูลป้อนกลับเพิ่มมากขึ้น ดังนั้น การพัฒนาระบบการทดสอบที่มีการสะท้อนข้อมูลป้อนกลับในครั้งนี้จึงมีความสอดคล้องและเป็นไปตามแนวโน้มการเปลี่ยนแปลงของโลก เพื่อให้ผู้รับข้อมูลป้อนกลับสามารถนำข้อมูลที่ได้ไปใช้ในการพัฒนาตนเองได้ทันที

4. ประเด็นสำคัญในการพัฒนาระบบ คือ การตรวจสอบคุณภาพ ดังนั้น จึงต้องพิจารณากระบวนการประเมินให้มีความเหมาะสมกับสิ่งที่มุ่งประเมิน โดยการวิจัยครั้งนี้ได้นำวิธีการประเมินจำนวน 2 วิธี มาใช้ในการตรวจสอบคุณภาพระบบ โดยวิธีแรกใช้การประเมินคุณภาพระบบก่อนนำระบบไปใช้งานจริง ซึ่งใช้แบบประเมินระบบแบบอิงมาตรฐาน (standard evaluation) ตามแนวคิดของ Stufflebeam โดยครอบคลุม 4 ประเด็น ได้แก่ ด้านความมีประโยชน์ (utility) ด้านความเป็นไปได้ในการนำไปใช้ (feasibility) ด้านความเหมาะสม (propriety) และด้านความถูกต้อง (accuracy) (ศิริชัย กาญจนวาสิ, 2554) ซึ่งจะช่วยประเมินเบื้องต้นได้ว่าระบบที่พัฒนาขึ้นสามารถนำไปใช้งานได้จริงหรือไม่

หลังจากนำระบบไปทดลองใช้เรียบร้อยแล้ว การประเมินคุณภาพวิธีที่ 2 คือ การประเมินความพึงพอใจของผู้สอบต่อการใช้งานระบบซึ่งพัฒนามาจากแบบประเมินความพึงพอใจผู้ใช้งานของอินเตอร์เฟซระหว่างผู้ใช้งานกับคอมพิวเตอร์ (user satisfaction of the human-computer interface) โดยเป้าหมายหลักของการประเมิน คือ ต้องการทราบว่าระบบทดสอบที่พัฒนาขึ้นสามารถตอบสนองการใช้งานของตัวอย่างวิจัยได้มากน้อยเพียงใด เนื่องจากผลการประเมินที่ได้จะ

นำไปใช้เป็นแนวทางในการพัฒนาระบบที่ตอบสนองการใช้งานของผู้สอบในอนาคตต่อไป (Chin et al., 1988; Navas et al., 2007) จากข้อมูลข้างต้นจึงเชื่อมั่นได้ว่าระบบการทดสอบแบบปรับเหมาะหลายขั้นตอนที่มีการให้ข้อมูลป้อนกลับในการทดสอบระดับชาติ (O-NET) วิชาภาษาอังกฤษสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6 มีคุณภาพ ได้มาตรฐาน พร้อมทั้งสามารถนำไปใช้งานได้จริง

ข้อเสนอแนะในการนำผลวิจัยไปใช้

1. ผู้เรียนที่มีความสนใจเตรียมความพร้อมในการทดสอบระดับชาติ (O-NET) วิชาภาษาอังกฤษสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6 เพื่อประเมินความสามารถของตนเองในวิชาภาษาอังกฤษว่าอยู่ในระดับใดและควรพัฒนาความสามารถด้านใดเพิ่มเติม โดยสามารถทราบผลข้อมูลป้อนกลับได้จากระบบทันที
2. สถาบันทดสอบการศึกษาแห่งชาติ (องค์การมหาชน) ซึ่งเป็นองค์กรด้านการทดสอบระดับชาติสามารถนำไปใช้เผยแพร่ให้ผู้เรียนเตรียมความพร้อมในการทดสอบระดับชาติ (O-NET) วิชาภาษาอังกฤษสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6 เพื่อยกระดับความสามารถของผู้เรียนในวิชาภาษาอังกฤษของประเทศให้เป็นไปตามมาตรฐาน

ข้อเสนอแนะในการวิจัยครั้งต่อไป

1. การวิจัยครั้งนี้มีข้อสอบที่ผ่านเกณฑ์เพียง 179 ข้อ ซึ่งข้อสอบส่วนใหญ่เป็นข้อสอบชุด (testlets) จึงทำให้โอกาสที่ผู้สอบจะพบข้อสอบซ้ำมีอัตราสูง ด้วยเหตุนี้ งานวิจัยในอนาคตจึงควรเพิ่มจำนวนข้อสอบเพื่อบรรจุลงในคลังข้อสอบ เพื่อให้เนื้อหาในการทดสอบแบบปรับเหมาะมีระดับความยากที่หลากหลายและครอบคลุมเนื้อหาการทดสอบระดับชาติวิชาภาษาอังกฤษระดับชั้นมัธยมศึกษาปีที่ 6 มากยิ่งขึ้น
2. การวิจัยครั้งนี้ยังไม่ได้ดำเนินการปรับเทียบคะแนนสอบของแบบสอบแต่ละฉบับเนื่องจากข้อจำกัดของจำนวนข้อสอบที่ผ่านเกณฑ์ ดังนั้น งานวิจัยครั้งต่อไปควรมีการปรับเทียบคะแนนของแบบสอบแต่ละปีเพื่อให้คลังข้อสอบของระบบการทดสอบแบบปรับเหมาะมีความเท่าเทียมกันมากยิ่งขึ้น
3. การวิจัยครั้งนี้มีการจัดชุดข้อสอบ (testlets) ซึ่งอาจมีสัดส่วนที่แตกต่างกันมาก เนื่องมาจากข้อจำกัดของข้อสอบที่ผ่านเกณฑ์ ดังนั้น ในการวิจัยครั้งต่อไปควรมีการจัดชุดข้อสอบที่มีสัดส่วนของชุดข้อสอบที่ใกล้เคียงกับข้อสอบระดับชาติวิชาภาษาอังกฤษระดับชั้นมัธยมศึกษาปีที่ 6 ให้มากยิ่งขึ้น
4. ปัจจุบันการวินิจฉัยผู้เรียน (diagnosis) เป็นประเด็นที่เข้ามามีบทบาทสำคัญในการพัฒนาผู้เรียน ดังนั้น ในการวิจัยครั้งต่อไปควรพัฒนาระบบการทดสอบแบบปรับเหมาะเพื่อวินิจฉัยผู้เรียนเพื่อให้ผู้เรียนได้รับข้อสอบสำหรับการวินิจฉัยที่ตรงกับความสามารถกับผู้เรียน และหลังจากการทดสอบเสร็จสิ้น สามารถให้ข้อมูลป้อนกลับให้กับผู้เรียนเพื่อนำไปใช้ในการพัฒนาตนเองได้ทันที

บรรณานุกรม

- โชติกา ภาชีผล, ประกอบ กรณীগิจ, & พิทักษ์ โสตถยาคม. (2558). การพัฒนารูปแบบแฟ้มสะสมงานอิเล็กทรอนิกส์ที่สะท้อนข้อมูลย้อนกลับในการประเมินพัฒนาการเรียนรู้ของนักเรียนในสังกัดสำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน. *วารสารวิธีวิทยาการวิจัย*, 28(1), 1-25.
- ทัศน์ศิริรินทร์ สว่างบุญ, ศิริเดช สุชีวะ, ศิริชัย กาญจนวาสี, & Muraki, E. (2555). ประสิทธิภาพการวัดของการทดสอบแบบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์. *วารสารวิธีวิทยาการวิจัย*, 25(1), 27-45.
- ศิริชัย กาญจนวาสี. (2538). *การทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบ (Adaptive Testing)*. กรุงเทพมหานคร: ศูนย์ตำราและเอกสารทางวิชาการคณะครุศาสตร์.
- ศิริชัย กาญจนวาสี. (2554). *ทฤษฎีการประเมิน (พิมพ์ครั้งที่ 7)*. กรุงเทพมหานคร: สำนักพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.
- ศิริชัย กาญจนวาสี. (2555). *ทฤษฎีการทดสอบแนวใหม่ (Modern Test Theory) (พิมพ์ครั้งที่ 4)*. กรุงเทพมหานคร: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.
- สถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน). (2562). *ค่าสถิติพื้นฐานผลการทดสอบ O-NET ชั้นมัธยมศึกษาปีที่ 6 ปีการศึกษา 2557 – 2560*. สืบค้นเมื่อวันที่ 13 มิถุนายน 2562, จากเว็บไซต์ <https://www.niets.or.th/uploads/editor/files/Download/ช่วงคะแนน%20ม.6.pdf>
- Andjelic, S., & Cekerevac, Z. (2014). CAT model with personalized algorithm for evaluation of estimated student knowledge. *Educational Information Technology*, 2014(19), 173–191.
- Baranczyk, M. C., & Best, C. A. (2019). How does instructor feedback help and hurt students' academic performance? *Scholarship of Teaching and Learning in Psychology*, 1-14. doi: <http://dx.doi.org/10.1037/stl0000169>
- Batool, T., & Akhter, M. (2019). Feedback effect on secondary school students achievement in mathematics. *Global Social Sciences Review*, 4(3), 123-134. doi: 10.31703/gssr.2019(IV-III).12
- Bock, R. D., & Mislevy, R. J. (1982). Adaptive EAP Estimaion of Ability in a Microcomputer Environment. *Applied Psychological Measurement*, 6(4), 431-444.
- Brossman, B. G., & Guille, R. A. (2014). A Comparison of Multi-Stage and Linear Test Designs for Medium-Size Licensure and Certification Examinations. *Journal of Computerized Adaptive Testing*, 2(2).

- Camilli, G. (1994). Origin of the Scaling Constant $d = 1.7$ in Item Response Theory. *Journal of Educational and Behavioral Statistics*, 19(3), 293-295.
- Cheng, Y., & Chang, H.-H. (2009). The maximum priority index method for severely constrained item selection in computerized adaptive testing *British Journal of Mathematical and Statistical Psychology*, 2009(62), 369-383.
- 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.
- Čisar, S. M., Radosav, D., Markoski, B., Pinter, R., & Čisar, P. (2010). Computer Adaptive Testing of Student Knowledge. *Acta Polytechnica Hungarica*, 7(4), 139-152.
- Dallas, A. (2014). *The Effects of Routing and Scoring Within a Computer Adaptive Multi-Stage Framework*. (Doctor of Philosophy), The University of North Carolina.
- Damar, E. A. (2013). ELT Teacher Trainees' Reflective Feedback to Their Cooperating Teachers. *Journal of Educational and Social Research*, 3(7), 235-242.
- Embretson, S. E., & Reise, S. P. (2000). *Item Response Theory for Psychologists*. New Jersey: Lawrence Erlbaum Associates.
- Fithriani, R. (2019). ZPD and the benefits of written feedback in L2 writing: Focusing on students' perceptions *Reading Matrix: An International Online Journal*, 19(1).
- Goldsmith, M. (2003). Leadership Development : Try Feedforward Instead of Feedback. *Journal of Excellence*, 8, 15-19.
- Hambleton, R. K., & Swaminathan, H. (1985). *Item Response Theory: Principles and Applications*. MA, U.S.A.: Kluwer Academic Publishers.
- Hattie, J., & Timperley, H. (2007). The Power of Feedback. *Review of Educational Research*, 77(1), 81-112.
- Hatziapostolou, T., & Paraskakis, I. (2010). Enhancing the impact of formative feedback on student learning through an online feedback system. *Electronic Journal of e-Learning*, 8(2), 111-122.
- Iron, A. (2008). *Enhancing Learning through Formative Assessment and Feedback*. NY: Routledge.

- Jahin, J. H. (2012). The effect of peer reviewing on writing apprehension and essay writing ability of prospective EFL teachers. *Australian Journal of Teacher Education*, 37(11), 60-84. doi: <http://dx.doi.org/10.14221/ajte.2012v37n11.3>
- Kamberi, L. (2013). The significance of teacher feedback in EFL writing for tertiary level foreign language learners. *Procedia - Social and Behavioral Sciences*, 70, 1686-1690. doi: 10.1016/j.sbspro.2013.01.241
- 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.
- Linden, W. J. v. d. (2000). Optimal assembly of tests with item sets. *Applied Psychological Measurement*, 24, 225-240.
- Ludvigsen, K., Krumsvik, R., & Furnes, B. (2015). Creating formative feedback spaces in large lectures. *Computers & Education*, 88(2015).
- Martin, A. J., & Lazendic, G. (2018). Computer-Adaptive Testing: Implications for Students' Achievement, Motivation, Engagement, and Subjective Test Experience. *Journal of Educational Psychology*, 110(1), 27-45. doi: <http://dx.doi.org/10.1037/edu0000205>
- Mead, A. D. (2006). An Introduction to Multistage Testing. *APPLIED MEASUREMENT IN EDUCATION*, 19(3), 185-187. doi: 10.1207/s15324818ame1903_1
- Navas, H., Osornio, A. L., Baum, A., Gomez, A., Luna, D., & Quiros, F. G. B. d. (2007). *Creation and Evaluation of a Terminology Server for the Interactive Coding of Discharge Summaries*. Paper presented at the MEDINFO 2007, Brisbane, Australia.
- 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
- Sari, H. I., Sari, H. Y., & Huggins-Manley, A. C. (2016). Computer Adaptive Multistage Testing: Practical Issues, Challenges and Principles. *Journal of Measurement and Evaluation in Education and Psychology*, 7(2), 388-406. doi: 10.21031/epod.280183

- Wang, K. (2017). *A FAIR COMPARISON OF THE PERFORMANCE OF COMPUTERIZED ADAPTIVE TESTING AND MULTISTAGE ADAPTIVE TESTING*. (Doctor of Philosophy), Michigan State University.
- Wojcikowski, K., & Brownie, S. (2013). Generic reflective feedback: An effective approach to developing clinical reasoning skills. *Journal of Computer Assisted Learning, 29*, 371–382.
- Yamamoto, K., Khorramdel, L., & Shin, H. J. (2018). Introducing multistage adaptive testing into international large-scale assessments designs using the example of PIAAC. *Psychological Test and Assessment Modeling, 60*(3), 347-368.
- 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 Research Report Series: ACT*.
- Zheng, Y., Nozawa, Y., Zhu, R., & Gao, X. (2014). Automated Top-Down Heuristic Assembly of a Classification Multistage Test.
https://www.academia.edu/10825094/Automated_top-down_heuristic_assembly_of_a_classification_multistage_test

ภาคผนวก

ภาคผนวก ก

รายนามผู้ทรงคุณวุฒิในการตรวจสอบคุณภาพแบบประเมินระบบ

1. รองศาสตราจารย์ ดร.โชติกา ภาชีผล
ค.ด. (การวัดและประเมินผลการศึกษา) จุฬาลงกรณ์มหาวิทยาลัย
อาจารย์ประจำสาขาวิชาการวัดและประเมินผลการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
2. รองศาสตราจารย์ ดร.ณัฐภรณ์ หลาวทอง
ค.ด. (การวัดและประเมินผลการศึกษา) จุฬาลงกรณ์มหาวิทยาลัย
อาจารย์ประจำสาขาวิชาการวัดและประเมินผลการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
สาขาวิชาการวัดและประเมินผลการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
3. รองศาสตราจารย์ ดร.กมลวรรณ ตังธนากานนท์
ค.ด. (การวัดและประเมินผลการศึกษา) จุฬาลงกรณ์มหาวิทยาลัย
อาจารย์ประจำสาขาวิชาการวัดและประเมินผลการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
สาขาวิชาการวัดและประเมินผลการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
4. อาจารย์ ดร.วรัญญา รุมนแสง
ค.ด. (การวัดและประเมินผลการศึกษา) จุฬาลงกรณ์มหาวิทยาลัย
ผู้จัดการบริษัท อบรมและทดสอบ พีทีเอส จำกัด
5. อาจารย์ ดร.สุรเดช อนันตสวัสดิ์
ค.ด. (การวัดและประเมินผลการศึกษา) จุฬาลงกรณ์มหาวิทยาลัย
ครู อันดับ คศ. 2 โรงเรียนดลิ่งชั้นวิทยา

ภาคผนวก ข

รายนามผู้ทรงคุณวุฒิในการประเมินคุณภาพระบบ

1. รองศาสตราจารย์ ดร.สังวรณัฏ ังคกระโทก
Ph.D. (Measurement and Quantitative Methods) Michigan State University
ผู้อำนวยการสถาบันวิจัย มหาวิทยาลัยสุโขทัยธรรมาธิราช
2. รองศาสตราจารย์ ดร.นลินี ฒ นคร
ค.ด. (การวัดและประเมินผลการศึกษา) จุฬาลงกรณ์มหาวิทยาลัย
อาจารย์ประจำสาขาวิชาศึกษาศาสตร์ มหาวิทยาลัยสุโขทัยธรรมาธิราช
3. ผู้ช่วยศาสตราจารย์ ดร.ณัฐวี ุตตฤกษ์
Ph.D. (Information Technology) Edith Cowan University
อาจารย์ประจำคณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
4. อาจารย์ ดร.ชุติวัดมน์ สุวัตถิพงษ์
ค.ด. (เทคโนโลยีและสื่อสารการศึกษา) จุฬาลงกรณ์มหาวิทยาลัย
อาจารย์ประจำศูนย์วิชาการเทคโนโลยีและสื่อสารการศึกษา สำนักเทคโนโลยีการศึกษา
มหาวิทยาลัยสุโขทัยธรรมาธิราช
5. อาจารย์ ดร.อนุสรณ์ เกิดศรี
ค.ด. (การวัดและประเมินผลการศึกษา) จุฬาลงกรณ์มหาวิทยาลัย
อาจารย์ประจำศูนย์วิชาการประเมินผล สำนักทะเบียนและวัดผล มหาวิทยาลัยสุโขทัยธรรมาธิราช

ภาคผนวก ค

ผลการพิจารณาความสอดคล้องระหว่างรายการประเมินกับวัตถุประสงค์/เนื้อหาที่ต้องการ
ประเมินของแบบประเมินระบบแบบอิงมาตรฐาน

ข้อ	รายการประเมิน	จำนวนผลการตัดสิน			IOC	ผลการพิจารณา	ข้อเสนอแนะ
		สอดคล้อง	ไม่แน่ใจ	ไม่สอดคล้อง			
ด้านความมีประโยชน์ (utility) หมายถึง ประโยชน์ของระบบในการตอบสนองต่อความต้องการใช้สารสนเทศของผู้ที่เกี่ยวข้องได้ครอบคลุมและสอดคล้องกับวัตถุประสงค์ในการทดสอบ							
1	ระบบการทดสอบสามารถตอบสนองความต้องการของผู้สอบ	5	-	-	1.0	สอดคล้อง	หากระบุได้ว่าสนองความต้องการของผู้สอบในด้านใดจะชัดเจนมากขึ้น
2	ระบบการทดสอบสามารถนำไปใช้เตรียมความพร้อมให้กับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6 เพื่อทดสอบระดับชาติ (O-NET) วิชาภาษาอังกฤษ	5	-	-	1.0	สอดคล้อง	
3	ระบบการทดสอบมีรายละเอียดเนื้อหาเหมือนกับการทดสอบระดับชาติ	3	2	-	0.6	สอดคล้อง	- น่าจะอยู่ในด้านความถูกต้อง (accuracy) - ไม่มีในนิยาม
4	รายงานผลการทดสอบสามารถแสดงระดับความสามารถของผู้สอบได้	5	-	-	1.0	สอดคล้อง	
5	รายงานข้อมูลป้อนกลับสามารถนำไปใช้ประโยชน์ได้อย่างเป็นรูปธรรม	5	-	-	1.0	สอดคล้อง	
ด้านความเป็นไปได้ในการนำไปใช้ (feasibility) หมายถึง ความสอดคล้องของระบบกับสภาพความเป็นจริง มีความเหมาะสมกับสถานการณ์ สามารถนำไปปฏิบัติได้จริง และมีความคุ้มค่า							
6	ระบบการทดสอบสามารถนำไปใช้งานได้จริง	5	-	-	1.0	สอดคล้อง	- ควรปรับเป็น "ใช้งานในการ..." ในการทำอะไรควรระบุให้ชัดเจน - ข้อ 6 คล้ายกับข้อ 7
7	ระบบการทดสอบสามารถตอบสนองการใช้งานของผู้สอบได้	4	1	-	0.8	สอดคล้อง	- ใกล้เคียงกับด้านความมีประโยชน์ ข้อ 1 หากหมายความว่า quick responsive ของระบบ อาจปรับเป็น "ระบบการทดสอบสามารถตอบสนองการใช้งานของผู้สอบได้อย่างรวดเร็ว"
8	รายงานผลการทดสอบ	5	-	-	1.0	สอดคล้อง	

ข้อ	รายการประเมิน	จำนวนผลการตัดสิน			IOC	ผลการพิจารณา	ข้อเสนอแนะ
		สอดคล้อง	ไม่แน่ใจ	ไม่สอดคล้อง			
	สามารถแสดงผลได้ทันทีหลังจบการทดสอบ						
9	รายงานข้อมูลป้อนกลับสามารถแสดงผลเพื่อการพัฒนาผู้สอบ	4	1	-	0.8	สอดคล้อง	- ควรปรับเป็น “รายงานข้อมูลป้อนกลับมีสารสนเทศที่เหมาะสมสำหรับการพัฒนาผู้สอบ” - ไม่มีในนิยาม
10	การพัฒนากระบวนการทดสอบมีความคุ้มค่า	4	1	-	0.8	สอดคล้อง	- ไม่น่าเกี่ยวข้องความเป็นไปได้ในการนำไปใช้ควรปรับเป็นผู้ใช้งานระบบการทดสอบสามารถเข้าใช้งานระบบได้อย่างสะดวก และทั่วถึง
ด้านความเหมาะสม (propriety) หมายถึง การดำเนินงานที่เป็นไปตามขั้นตอนการทำงานของระบบ ความเหมาะสมของระยะเวลาสารสนเทศที่ได้จากรายงานผลการทดสอบ และรายงานผลข้อมูลป้อนกลับ							
11	ระบบการทดสอบมีค่าใช้จ่ายก่อนการสอบชัดเจน	5	-	-	1.0	สอดคล้อง	
12	ขอบเขตเนื้อหาในการทดสอบมีความเหมาะสม	5	-	-	1.0	สอดคล้อง	
13	ระยะเวลาในการทดสอบมีความเหมาะสม	5	-	-	1.0	สอดคล้อง	
14	รายงานผลการทดสอบมีสารสนเทศครบถ้วน	5	-	-	1.0	สอดคล้อง	
15	รายงานผลข้อมูลป้อนกลับมีรายละเอียดที่จำเป็นเพื่อการพัฒนา	4	1	-	0.8	สอดคล้อง	- ใกล้เคียงกับด้านความเป็นไปได้ในการนำไปใช้ ข้อ 9
16	ระบบการทดสอบมีความสอดคล้องกับแนวทางการเพิ่มระดับความสามารถด้านภาษาอังกฤษของผู้เรียน	5	-	-	1.0	สอดคล้อง	
ด้านความถูกต้อง (accuracy) หมายถึง ความถูกต้องของขั้นตอนการทดสอบ กระบวนการคัดเลือกข้อสอบ การประมาณค่าความสามารถของผู้สอบ การรายงานผลการทดสอบ และการรายงานผลข้อมูลป้อนกลับ							
17	มีการอธิบายวัตถุประสงค์ของการทดสอบไว้ครบถ้วน	4	1	-	0.8	สอดคล้อง	- ครบถ้วนแต่อาจไม่ถูกต้อง - ควรตัดคำว่า “มี” ออก - ควรปรับคำว่า “ไว้ครบถ้วน” เป็น “ได้อย่างชัดเจน” - ควรสอบถามในเชิงความสอดคล้องของคำอธิบายกับการทำงานของระบบ ควร

ข้อ	รายการประเมิน	จำนวนผลการตัดสิน			IOC	ผลการพิจารณา	ข้อเสนอแนะ
		สอดคล้อง	ไม่แน่ใจ	ไม่สอดคล้อง			
							ปรับเป็น "มีการอธิบายวัตถุประสงค์ของการทดสอบไว้ครบถ้วน และสอดคล้องกับการทำงานของระบบ"
18	มีการอธิบายขั้นตอนในการทดสอบไว้ครบถ้วน	4	1	-	0.8	สอดคล้อง	<ul style="list-style-type: none"> - ครบถ้วนแต่อาจไม่ถูกต้อง - ควรปรับคำว่า "ไว้ครบถ้วน" เป็น "ได้อย่างชัดเจน" - ควรสอบถามในเชิงความสอดคล้องของคำอธิบายกับการทำงานของระบบ อาจปรับเป็น "มีการอธิบายขั้นตอนในการทดสอบไว้ครบถ้วน และสอดคล้องกับการทำงานของระบบ"
19	กระบวนการคัดเลือกข้อสอบมีความสอดคล้องกับเนื้อหาในการทดสอบระดับชาติ	5	-	-	1.0	สอดคล้อง	<ul style="list-style-type: none"> - หากถามกระบวนการคัดเลือก ผู้ประเมินอาจตอบได้ยาก แต่หากถามว่า "ข้อสอบที่ถูกคัดเลือกไว้ใช้ในระบบสอดคล้องกับเนื้อหาในการทดสอบระดับชาติ" อาจไปเข้าซ้อนกับด้านความมีประโยชน์ ข้อ 3 จึงไม่แน่ใจว่าต้องการสอบถามในประเด็นเดียวกันหรือไม่
20	รายงานผลการทดสอบมีความถูกต้อง	5	-	-	1.0	สอดคล้อง	
21	รายงานผลข้อมูลป้อนกลับสอดคล้องกับผลการทดสอบ	5	-	-	1.0	สอดคล้อง	

ภาคผนวก ง

แบบประเมินระบบแบบอิงมาตรฐาน (standard evaluation)

คำชี้แจง: โปรดเติมเครื่องหมาย ✓ ในช่องตามระดับความคิดเห็นของท่านดังนี้

- 1 หมายถึง น้อยที่สุด
- 2 หมายถึง น้อย
- 3 หมายถึง ปานกลาง
- 4 หมายถึง มาก
- 5 หมายถึง มากที่สุด

ข้อ	รายการประเมิน	ระดับความคิดเห็น					ข้อเสนอแนะ
		1	2	3	4	5	
ด้านความมีประโยชน์ (utility) หมายถึง ประโยชน์ของระบบในการตอบสนองต่อความต้องการใช้สารสนเทศในการเตรียมความพร้อมเพื่อการทดสอบระดับชาติของผู้สอบได้ครอบคลุมและสอดคล้องกับวัตถุประสงค์ในการทดสอบ รวมถึงรายละเอียดเนื้อหาเหมือนกับการทดสอบระดับชาติ							
1	ระบบการทดสอบสามารถตอบสนองความต้องการในการเตรียมความพร้อมเพื่อการทดสอบระดับชาติของผู้สอบ						
2	ระบบการทดสอบสามารถนำไปใช้เตรียมความพร้อมให้กับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6 เพื่อทดสอบระดับชาติ (O-NET) วิชาภาษาอังกฤษ						
3	ระบบการทดสอบมีรายละเอียดเนื้อหาเหมือนกับการทดสอบระดับชาติ						
4	รายงานผลการทดสอบสามารถแสดงระดับความสามารถของผู้สอบได้						
5	รายงานข้อมูลป้อนกลับสามารถนำไปใช้ประโยชน์ได้อย่างเป็นรูปธรรม						
ด้านความเป็นไปได้ในการนำไปใช้ (feasibility) หมายถึง ความสอดคล้องของระบบกับสภาพความเป็นจริง มีความเหมาะสมกับสถานการณ์ สามารถนำไปปฏิบัติได้จริง และมีความคุ้มค่า							
6	ระบบการทดสอบสามารถนำไปใช้งานในการเตรียมความพร้อมเพื่อการทดสอบระดับชาติได้จริง						
7	ระบบการทดสอบสามารถตอบสนองการใช้งานของผู้สอบได้อย่างรวดเร็ว						
8	รายงานผลการทดสอบสามารถแสดงผลได้ทันทีหลังจบ						

ข้อ	รายการประเมิน	ระดับความคิดเห็น					ข้อเสนอแนะ
		1	2	3	4	5	
	การทดสอบ						
9	รายงานข้อมูลป้อนกลับมีสารสนเทศที่เหมาะสมสำหรับการพัฒนาผู้สอบ						
10	การพัฒนาระบบการทดสอบมีความคุ้มค่า						
ด้านความเหมาะสม (propriety) หมายถึง การดำเนินงานที่เป็นไปตามขั้นตอนการทำงานของระบบความเหมาะสมของระยะเวลา สารสนเทศที่ได้จากรายงานผลการทดสอบ และรายงานผลข้อมูลป้อนกลับ							
11	ระบบการทดสอบมีค่าใช้จ่ายก่อนการสอบชัดเจน						
12	ขอบเขตเนื้อหาในการทดสอบมีความเหมาะสม						
13	ระยะเวลาในการทดสอบมีความเหมาะสม						
14	รายงานผลการทดสอบมีสารสนเทศครบถ้วน						
15	รายงานผลข้อมูลป้อนกลับมีรายละเอียดที่จำเป็นเพื่อการพัฒนา						
16	ระบบการทดสอบมีความสอดคล้องกับแนวทางการเพิ่มระดับความสามารถด้านภาษาอังกฤษของผู้เรียน						
ด้านความถูกต้อง (accuracy) หมายถึง ความถูกต้องของขั้นตอนการทดสอบ กระบวนการคัดเลือกข้อสอบ การประมาณค่าความสามารถของผู้สอบ การรายงานผลการทดสอบ และการรายงานผลข้อมูลป้อนกลับ							
17	การอธิบายวัตถุประสงค์ของการทดสอบถูกต้องชัดเจน						
18	การอธิบายขั้นตอนในการทดสอบถูกต้องชัดเจน						
19	กระบวนการคัดเลือกข้อสอบมีความสอดคล้องกับเนื้อหาในการทดสอบระดับชาติ						
20	รายงานผลการทดสอบมีความถูกต้อง						
21	รายงานผลข้อมูลป้อนกลับสอดคล้องกับผลการทดสอบ						

ภาคผนวก จ

ผลการพิจารณาความสอดคล้องระหว่างรายการประเมินกับวัตถุประสงค์/เนื้อหาที่ต้องการ
ประเมินของ “แบบประเมินความพึงพอใจของผู้สอบกับการใช้งานระบบ”

ข้อ	รายการประเมิน	มาตรใน การวัด (0-9)	จำนวนผลการตัดสิน			IOC	ผลการ พิจารณา	ข้อเสนอแนะ
			สอดคล้อง	ไม่แน่ใจ	ไม่สอดคล้อง			
1.1	การใช้งานในภาพรวม ต่อระบบ	ไม่ดี -ดีเยี่ยม	4	1	-	0.8	สอดคล้อง	- ควรปรับรายการ ประเมินเป็น “ความเห็น ต่อการใช้งานระบบใน ภาพรวม” - มาตร “ไม่ดี” ควรปรับ เป็น “แย่มาก”
1.2	การใช้งานในภาพรวม ต่อระบบ	ผิดหวัง -พอใจ	5	-	-	1.0	สอดคล้อง	- มาตร “ผิดหวัง” ควร เปลี่ยนเป็น “ไม่น่าพึง พอใจ”
1.3	การใช้งานในภาพรวม ต่อระบบ	ธรรมดา- น่าสนใจ	4	1	-	0.8	สอดคล้อง	- มาตร “ธรรมดา” ควร ปรับเป็น “ไม่น่าสนใจ” หรือ “น่าเบื่อ”
การแสดงผล (screen)								
2	คุณลักษณะของการ แสดงผล	ยากต่อ การอ่าน -ง่ายต่อ การอ่าน	5	-	-	1.0	สอดคล้อง	-ควรปรับเป็น “หน้าจอ การแสดงผล” เพื่อให้ เข้าใจง่ายขึ้น - ควรปรับรายการ ประเมินเป็น “การ แสดงผลของ...”
3	เครื่องมือ (tools) บน หน้าการแสดงผล สามารถใช้งานง่าย	บางส่วน- ทั้งหมด	4	1	-	0.8	สอดคล้อง	-มาตร “บางส่วน” ควร ปรับเป็น “ไม่เลย” หรือ ปรับมาตรเป็น “ง่าย- ยาก” - ควรปรับรายการ ประเมินเป็น “การใช้ งานของเครื่องมือ (tools) บนหน้าการ แสดงผล”
4	การเรียงลำดับ เครื่องมือบนหน้าการ แสดงผล	สับสน -ชัดเจน	5	-	-	1.0	สอดคล้อง	
5	การเรียงลำดับเนื้อหา บนหน้าการแสดงผล	สับสน -ชัดเจน	5	-	-	1.0	สอดคล้อง	

ข้อ	รายการประเมิน	มาตรฐานในการวัด (0-9)	จำนวนผลการตัดสิน			IOC	ผลการพิจารณา	ข้อเสนอแนะ
			สอดคล้อง	ไม่แน่ใจ	ไม่สอดคล้อง			
คำศัพท์และสารสนเทศของระบบ (terminology and system information)								
6	การใช้คำศัพท์ในระบบ	ไม่สอดคล้อง - สอดคล้อง	4	1	-	0.8	สอดคล้อง	-ควรปรับมาตรฐานเป็น “เข้าใจยาก – เข้าใจง่าย” หรือ “ชัดเจน-สอดคล้อง”
7	คำศัพท์ในระบบมีความสัมพันธ์กับเครื่องมือที่ใช้	ไม่สัมพันธ์-สัมพันธ์	4	1	-	0.8	สอดคล้อง	-ควรปรับ “เครื่องมือที่ใช้” เป็น “เครื่องมือ (tools) ที่แสดงในระบบ” -ควรปรับมาตรฐานเป็น “น้อย-มาก”
8	ตำแหน่งของข้อความบนหน้าการแสดงผล	ไม่สอดคล้อง - สอดคล้อง	4	1	-	0.8	สอดคล้อง	-ควรปรับมาตรฐานเป็น “ไม่เหมาะสม-เหมาะสม” -ควรปรับมาตรฐาน “ไม่สอดคล้อง” เป็น “ชัดเจน”
9	ข้อความที่แสดงให้กับผู้สอบเพื่อป้อนข้อมูลบนหน้าการแสดงผล	สับสน -ชัดเจน	5	-	-	1.0	สอดคล้อง	
10	การแจ้งเตือนสถานะว่ากำลังอยู่ในกระบวนการใด	ไม่มี-มี สม่ำเสมอ	4	-	1	0.6	สอดคล้อง	-ควรปรับมาตรฐานเป็น “ไม่มีเลย” -ควรปรับเป็นประโยคบอกเล่า
11	ข้อความแจ้งเตือนความผิดพลาด	ไม่ช่วยเหลือ - ช่วยเหลือ	4	1	-	0.8	สอดคล้อง	- ไม่ค่อยเกี่ยวกับด้านคำศัพท์และสารสนเทศของระบบ - ควรปรับมาตรฐานเป็น “เข้าใจยาก – เข้าใจง่าย” -ควรปรับมาตรฐาน “ไม่ช่วยเหลือ” เป็น “ชัดเจน”
การเรียนรู้ (learning)								
12	การเรียนรู้การใช้งานของระบบ	ยาก-ง่าย	5	-	-	1.0	สอดคล้อง	
13	สามารถจดจำข้อมูลและคำสั่งที่เคยป้อน	ทำไม่ได้-ทำได้	4	1	-	0.8	สอดคล้อง	-ควรปรับเป็น “จดจำข้อมูลผู้สอบ” เพื่อความชัดเจน - ควรปรับเป็น “ความสามารถในการ

ข้อ	รายการประเมิน	มาตรในการวัด (0-9)	จำนวนผลการตัดสิน			IOC	ผลการพิจารณา	ข้อเสนอแนะ
			สอดคล้อง	ไม่แน่ใจ	ไม่สอดคล้อง			
								จดจำข้อมูลและคำสั่งที่เคยป้อน”
14	การทำงานของเครื่องมือบนระบบสามารถใช้งานได้ อย่างต่อเนื่อง	ไม่ต่อเนื่อง-ต่อเนื่อง	3	2	-	0.6	สอดคล้อง	- ควรอยู่ด้านความสามารถระบบและควรปรับเป็น “ผู้สอบสามารถใช้งานเครื่องมือต่าง ๆ ในระบบได้อย่างต่อเนื่อง” - ควรตัดคำว่า “สามารถใช้งานได้ต่อเนื่อง”
15	การปรากฏของข้อความช่วยเหลือบนหน้าการแสดงผล	ไม่ช่วยเหลือ - ช่วยเหลือ	4	1	-	0.8	สอดคล้อง	- ข้างกับด้านคำศัพท์และสารสนเทศของระบบ ข้อ 11 ควรปรับเป็น “ข้อความแจ้งเตือนความผิดพลาด ช่วยให้เข้าใจการทำงานของระบบ”
16	เอกสารคู่มือการใช้งาน	สับสน-ชัดเจน	5	-	-	1.0	สอดคล้อง	
ความสามารถระบบ (system capabilities)								
17	ความเร็วในการใช้งานระบบ	ช้า-เร็ว	5	-	-	1.0	สอดคล้อง	- ควรปรับเป็นการตอบสนองของระบบ
18	ความราบรื่นในการใช้งานระบบ	ติดขัด-ราบรื่น	4	1	-	0.8	สอดคล้อง	-คำว่า “ราบรื่น” ไม่ควรนำมาใช้เป็นมาตรฐานเนื่องจากใช้ในข้อรายการ
19	แนวโน้มการทำงานของระบบ	ติดขัด-ราบรื่น	4	1	-	0.8	สอดคล้อง	- ต้องการทราบประเด็นในควรระบุให้ชัดเจน - ข้อ 18 และ 19 ค่อนข้างคล้ายกัน
20	การแก้ไขข้อผิดพลาดจากการใช้งาน	ยาก-ง่าย	5	-	-	1.0	สอดคล้อง	- ควรเพิ่มข้อความอีก 1 ข้อ คือ “ระบบมีข้อขัดข้องขณะใช้งาน” โดยใช้มาตร “มีบางจุด-ไม่มี”

ภาคผนวก ฉ

แบบประเมินความพึงพอใจของผู้สอบต่อการใช้งานระบบ

คำชี้แจง: โปรดวงกลม บนตัวเลขตามระดับความคิดเห็นของท่านดังนี้

ข้อ	รายการประเมิน	สเกลในการวัด		
			0 1 2 3 4 5 6 7 8 9	
1	ความเห็นต่อการใช้งานระบบในภาพรวม	แย่มาก		ดีเยี่ยม
		ไม่พอใจ		พอใจ
		น่าเบื่อ		น่าสนใจ
หน้าจอการแสดงผล (screen)				
2	การแสดงผลของหน้าจอ	ยากต่อ การอ่าน	0 1 2 3 4 5 6 7 8 9	ง่ายต่อ การอ่าน
3	การใช้งานของเครื่องมือ (tools) บนหน้า การแสดงผล	ยาก	0 1 2 3 4 5 6 7 8 9	ง่าย
4	การเรียงลำดับเครื่องมือบนหน้าการ แสดงผล	สับสน	0 1 2 3 4 5 6 7 8 9	ชัดเจน
5	การเรียงลำดับเนื้อหาบนหน้าการ แสดงผล	สับสน	0 1 2 3 4 5 6 7 8 9	ชัดเจน
คำศัพท์และสารสนเทศของระบบ (terminology and system information)				
6	การใช้คำศัพท์ในระบบ	เข้าใจ ยาก	0 1 2 3 4 5 6 7 8 9	เข้าใจ ง่าย
7	คำศัพท์ในระบบมีความสัมพันธ์กับ เครื่องมือ (tools) ที่แสดงในระบบ	น้อย	0 1 2 3 4 5 6 7 8 9	มาก
8	ตำแหน่งของข้อความบนหน้าการ แสดงผล	ไม่ เหมาะสม	0 1 2 3 4 5 6 7 8 9	เหมาะสม
9	ข้อความที่แสดงให้กับผู้สอบเพื่อป้อน ข้อมูลบนหน้าการแสดงผล	สับสน	0 1 2 3 4 5 6 7 8 9	ชัดเจน
10	การแจ้งเตือนสถานะระหว่างการใช้งาน	ไม่มีเลย	0 1 2 3 4 5 6 7 8 9	มี
11	ข้อความแจ้งเตือนความผิดพลาด	เข้าใจ	0 1 2 3 4 5 6 7 8 9	เข้าใจ

ข้อ	รายการประเมิน	สเกลในการวัด		
		ยาก	9	ง่าย
การเรียนรู้ (learning)				
12	การเรียนรู้การใช้งานของระบบ	ยาก	0 1 2 3 4 5 6 7 8 9	ง่าย
13	ความสามารถในการจดจำข้อมูลและคำสั่งที่เคยป้อน	ทำไม่ได้	0 1 2 3 4 5 6 7 8 9	ทำได้
14	ผู้สอบสามารถใช้งานเครื่องมือต่าง ๆ ในระบบ	ไม่ต่อเนื่อง	0 1 2 3 4 5 6 7 8 9	ต่อเนื่อง
15	ข้อความแจ้งเตือนความผิดพลาดช่วยให้เข้าใจการทำงานของระบบ	ไม่ช่วย	0 1 2 3 4 5 6 7 8 9	ช่วย
16	เอกสารคู่มือการใช้งาน	สับสน	0 1 2 3 4 5 6 7 8 9	ชัดเจน
ความสามารถระบบ (system capabilities)				
17	ความเร็วในการตอบสนองต่อการใช้งานระบบ	ช้า	0 1 2 3 4 5 6 7 8 9	เร็ว
18	ความต่อเนื่องในการใช้งานระบบ	ติดขัด	0 1 2 3 4 5 6 7 8 9	ราบรื่น
19	แนวโน้มความต่อเนื่องในการทำงานของระบบ	ติดขัด	0 1 2 3 4 5 6 7 8 9	ราบรื่น
20	ข้อขัดข้องในการใช้งานระบบ	มี	0 1 2 3 4 5 6 7 8 9	ไม่มี

ภาคผนวก ข
ค่าพารามิเตอร์ของข้อสอบ

ข้อ	เนื้อหา	a	b	c	หมายเหตุ
ปี 2558					
1	Dialog Completion	1.6004	0.8500	0.1820	ผ่านเกณฑ์
2	Dialog Completion	1.9197	1.2189	0.2020	ผ่านเกณฑ์
3	Dialog Completion	1.0394	0.8034	0.1212	ผ่านเกณฑ์
4	Dialog Completion	1.0303	0.5736	0.0926	ผ่านเกณฑ์
5	Dialog Completion	1.2806	2.0028	0.2014	ผ่านเกณฑ์
6	Dialog Completion	1.4392	0.6958	0.1199	ผ่านเกณฑ์
7	Dialog Completion	1.9097	1.3684	0.2236	ผ่านเกณฑ์
8	Dialog Completion	1.1947	0.7768	0.1144	ผ่านเกณฑ์
9	Dialog Completion	0.9675	0.3859	0.0000	ผ่านเกณฑ์
10	Dialog Completion	1.3937	1.9851	0.1975	ผ่านเกณฑ์
11	Situational Dialog	0.8241	0.5617	0.0008	ผ่านเกณฑ์
12	Situational Dialog	1.7523	1.1660	0.2158	ผ่านเกณฑ์
13	Situational Dialog	0.9216	1.1787	0.1416	ผ่านเกณฑ์
14	Situational Dialog	1.4891	1.4757	0.1916	ผ่านเกณฑ์
15	Situational Dialog	2.0769	1.5727	0.2087	ผ่านเกณฑ์
16	Sentence Completion	1.4498	2.1890	0.1440	ผ่านเกณฑ์
17	Sentence Completion	1.3256	2.5455	0.1973	ผ่านเกณฑ์
18	Sentence Completion	2.4703	2.7152	0.1101	ผ่านเกณฑ์
19	Sentence Completion	2.0185	2.1018	0.1489	ผ่านเกณฑ์
20	Sentence Completion	1.2280	2.0935	0.1379	ผ่านเกณฑ์
21	Sentence Completion	1.0525	1.7344	0.0543	ผ่านเกณฑ์
22	Sentence Completion	0.8925	3.6873	0.1845	ผ่านเกณฑ์
23	Sentence Completion	2.2940	2.1033	0.1711	ผ่านเกณฑ์

ข้อ	เนื้อหา	a	b	c	หมายเหตุ
24	Sentence Completion	1.8731	2.7324	0.1775	ผ่านเกณฑ์
25	Sentence Completion	2.6556	2.9591	0.2034	ผ่านเกณฑ์
26	Sentence Completion	1.7095	1.8740	0.1540	ผ่านเกณฑ์
27	Sentence Completion	2.0303	2.5904	0.1716	ผ่านเกณฑ์
28	Sentence Completion	1.5874	2.3376	0.1755	ผ่านเกณฑ์
29	Sentence Completion	1.4813	2.3449	0.1667	ผ่านเกณฑ์
30	Sentence Completion	1.4355	2.0953	0.2059	ผ่านเกณฑ์
31	Writing Ability	1.4127	2.6520	0.1827	ชุดเดียวกับ ข้อไม่ผ่านเกณฑ์
32	Writing Ability	0.0931	25.5857	0.2541	ไม่ผ่านเกณฑ์
33	Writing Ability	-2.9540	-2.7440	0.1472	ไม่ผ่านเกณฑ์
34	Writing Ability	2.7529	2.5688	0.1885	ชุดเดียวกับ ข้อไม่ผ่านเกณฑ์
35	Writing Ability	3.1365	2.2843	0.2177	ชุดเดียวกับ ข้อไม่ผ่านเกณฑ์
36	Writing Ability	2.8666	2.0060	0.2154	ชุดเดียวกับ ข้อไม่ผ่านเกณฑ์
37	Writing Ability	3.4920	2.3048	0.1942	ชุดเดียวกับ ข้อไม่ผ่านเกณฑ์
38	Writing Ability	2.0278	1.7716	0.1895	ชุดเดียวกับ ข้อไม่ผ่านเกณฑ์
39	Writing Ability	2.4213	2.8136	0.2267	ชุดเดียวกับ ข้อไม่ผ่านเกณฑ์
40	Writing Ability	2.2218	2.0319	0.1666	ชุดเดียวกับ ข้อไม่ผ่านเกณฑ์
41	Reading Ability -Vocabulary	1.6566	1.8785	0.1197	ชุดเดียวกับ ข้อไม่ผ่านเกณฑ์
42	Reading Ability -Vocabulary	1.4139	1.3739	0.1756	ชุดเดียวกับ ข้อไม่ผ่านเกณฑ์

ข้อ	เนื้อหา	a	b	c	หมายเหตุ
43	Reading Ability -Vocabulary	0.4280	2.7409	0.0025	ชุดเดียวกับข้อไม่ผ่านเกณฑ์
44	Reading Ability -Vocabulary	0.8755	2.9972	0.1551	ชุดเดียวกับข้อไม่ผ่านเกณฑ์
45	Reading Ability -Vocabulary	2.0260	1.7255	0.2845	ชุดเดียวกับข้อไม่ผ่านเกณฑ์
46	Reading Ability -Vocabulary	-2.6570	-3.2921	0.1723	ไม่ผ่านเกณฑ์
47	Reading Ability -Vocabulary	4.5310	2.7434	0.2205	ชุดเดียวกับข้อไม่ผ่านเกณฑ์
48	Reading Ability -Vocabulary	2.6959	1.5052	0.2391	ชุดเดียวกับข้อไม่ผ่านเกณฑ์
49	Reading Ability -Vocabulary	2.0865	1.0914	0.1510	ชุดเดียวกับข้อไม่ผ่านเกณฑ์
50	Reading Ability -Vocabulary	3.3199	2.4001	0.1616	ชุดเดียวกับข้อไม่ผ่านเกณฑ์
51	Reading Ability -Reading Comprehension	1.3033	0.1722	0.0816	ผ่านเกณฑ์
52	Reading Ability -Reading Comprehension	0.3226	2.2333	0.0180	ผ่านเกณฑ์
53	Reading Ability -Reading Comprehension	0.8267	0.1254	0.0001	ผ่านเกณฑ์
54	Reading Ability -Reading Comprehension	1.1004	1.9154	0.1682	ผ่านเกณฑ์
55	Reading Ability -Reading Comprehension	0.4685	0.9920	0.0000	ผ่านเกณฑ์
56	Reading Ability -Reading Comprehension	3.1489	2.1789	0.1363	ข้อสอบเกินชุด (5ข้อ)
57	Reading Ability -Reading Comprehension	3.1201	2.2907	0.1967	ข้อสอบเกินชุด (5ข้อ)
58	Reading Ability -Reading Comprehension	2.1075	2.0472	0.1560	ข้อสอบเกินชุด (5ข้อ)

ข้อ	เนื้อหา	a	b	c	หมายเหตุ
59	Reading Ability -Reading Comprehension	3.1262	1.6886	0.2470	ข้อสอบเกินชุด (5ข้อ)
60	Reading Ability -Reading Comprehension	2.3522	2.6982	0.1853	ข้อสอบเกินชุด (5ข้อ)
61	Reading Ability -Reading Comprehension	2.6094	1.9304	0.1987	ข้อสอบเกินชุด (5ข้อ)
62	Reading Ability -Reading Comprehension	1.5749	1.4765	0.1951	ข้อสอบเกินชุด (5ข้อ)
63	Reading Ability -Reading Comprehension	1.8115	1.9648	0.1654	ข้อสอบเกินชุด (5ข้อ)
64	Reading Ability -Reading Comprehension	2.2187	1.4917	0.1331	ข้อสอบเกินชุด (5ข้อ)
65	Reading Ability -Reading Comprehension	2.3942	1.6771	0.1590	ข้อสอบเกินชุด (5ข้อ)
66	Reading Ability -Reading Comprehension	2.0655	1.2683	0.1780	ข้อสอบเกินชุด (5ข้อ)
67	Reading Ability -Reading Comprehension	2.6930	2.0603	0.1198	ข้อสอบเกินชุด (5ข้อ)
68	Reading Ability -Reading Comprehension	2.5347	2.6656	0.1684	ข้อสอบเกินชุด (5ข้อ)
69	Reading Ability -Reading Comprehension	1.9107	0.9203	0.1917	ข้อสอบเกินชุด (5ข้อ)
70	Reading Ability -Reading Comprehension	1.0568	2.0578	0.1486	ข้อสอบเกินชุด (5ข้อ)
71	Reading Ability -Reading Comprehension	3.2273	1.3345	0.1966	ข้อสอบไม่ครบชุด (5ข้อ)
72	Reading Ability -Reading Comprehension	1.2064	2.8244	0.1606	ข้อสอบไม่ครบชุด (5ข้อ)
73	Reading Ability -Reading Comprehension	1.7671	1.4667	0.1722	ข้อสอบไม่ครบชุด (5ข้อ)
74	Reading Ability -Reading	2.6736	2.0818	0.1345	ข้อสอบเกินชุด (5ข้อ)

ข้อ	เนื้อหา	a	b	c	หมายเหตุ
	Comprehension				
75	Reading Ability -Reading Comprehension	3.6418	2.7844	0.2503	ข้อสอบเกินชุด (5ข้อ)
76	Reading Ability -Reading Comprehension	1.2780	1.7472	0.1521	ข้อสอบเกินชุด (5ข้อ)
77	Reading Ability -Reading Comprehension	2.2365	2.1428	0.2427	ข้อสอบเกินชุด (5ข้อ)
78	Reading Ability -Reading Comprehension	2.1151	2.2368	0.2276	ข้อสอบเกินชุด (5ข้อ)
79	Reading Ability -Reading Comprehension	1.5252	2.5952	0.1956	ข้อสอบเกินชุด (5ข้อ)
80	Reading Ability -Reading Comprehension	2.4119	1.9978	0.1365	ข้อสอบเกินชุด (5ข้อ)
ปี 2559					
1	Dialog Completion	1.5740	0.8240	0.2541	ผ่านเกณฑ์
2	Dialog Completion	3.2291	1.1011	0.1633	ผ่านเกณฑ์
3	Dialog Completion	2.9056	1.6217	0.1602	ผ่านเกณฑ์
4	Dialog Completion	2.6925	1.8819	0.1221	ผ่านเกณฑ์
5	Dialog Completion	2.3396	0.7515	0.2047	ผ่านเกณฑ์
6	Dialog Completion	1.7101	0.8383	0.2502	ผ่านเกณฑ์
7	Dialog Completion	3.1332	1.0824	0.1572	ผ่านเกณฑ์
8	Dialog Completion	2.8003	1.6489	0.1537	ผ่านเกณฑ์
9	Dialog Completion	2.3316	1.9078	0.1238	ผ่านเกณฑ์
10	Dialog Completion	2.3635	0.7304	0.1966	ผ่านเกณฑ์
11	Situational Dialog	1.9875	1.0170	0.1688	ผ่านเกณฑ์
12	Situational Dialog	1.7775	1.7553	0.1846	ผ่านเกณฑ์
13	Situational Dialog	2.3640	1.0666	0.1528	ผ่านเกณฑ์
14	Situational Dialog	1.1866	1.5826	0.1681	ผ่านเกณฑ์

ข้อ	เนื้อหา	a	b	c	หมายเหตุ
15	Situational Dialog	2.9063	1.3068	0.1666	ผ่านเกณฑ์
16	Sentence Completion	1.7875	2.1510	0.1812	ผ่านเกณฑ์
17	Sentence Completion	3.4154	2.7374	0.1971	ผ่านเกณฑ์
18	Sentence Completion	2.7386	2.4343	0.2028	ผ่านเกณฑ์
19	Sentence Completion	3.8784	2.8974	0.2086	ผ่านเกณฑ์
20	Sentence Completion	1.9603	2.7253	0.2060	ผ่านเกณฑ์
21	Sentence Completion	2.8746	2.5988	0.1826	ผ่านเกณฑ์
22	Sentence Completion	3.0595	2.7662	0.2144	ผ่านเกณฑ์
23	Sentence Completion	2.2078	1.6002	0.1912	ผ่านเกณฑ์
24	Sentence Completion	1.3122	2.2951	0.1701	ผ่านเกณฑ์
25	Sentence Completion	1.0533	2.2345	0.1705	ผ่านเกณฑ์
26	Error Correction	-0.0091	- 105.7836	0.0678	ไม่ผ่านเกณฑ์
27	Error Correction	1.1339	1.9616	0.2795	ชุดเดียวกับ ข้อไม่ผ่านเกณฑ์
28	Error Correction	0.8337	1.9956	0.2025	ชุดเดียวกับ ข้อไม่ผ่านเกณฑ์
29	Error Correction	2.7039	2.2734	0.1729	ชุดเดียวกับ ข้อไม่ผ่านเกณฑ์
30	Error Correction	2.2860	2.1373	0.1668	ชุดเดียวกับ ข้อไม่ผ่านเกณฑ์
31	Error Correction	4.0298	2.6211	0.3035	ข้อสอบไม่ครบชุด (10 ข้อ)
32	Error Correction	0.9716	1.7445	0.2292	ข้อสอบไม่ครบชุด (10 ข้อ)
33	Error Correction	0.7612	1.9969	0.1847	ข้อสอบไม่ครบชุด (10 ข้อ)
34	Error Correction	2.8218	2.2704	0.1755	ข้อสอบไม่ครบชุด (10 ข้อ)

ข้อ	เนื้อหา	a	b	c	หมายเหตุ
35	Error Correction	2.4875	2.1071	0.1687	ข้อสอบไม่ครบชุด (10 ข้อ)
36	Writing Ability	3.0015	2.1825	0.1404	ผ่านเกณฑ์
37	Writing Ability	2.2122	2.0790	0.1771	ผ่านเกณฑ์
38	Writing Ability	0.6381	3.1036	0.1816	ผ่านเกณฑ์
39	Writing Ability	3.2579	1.6841	0.1161	ผ่านเกณฑ์
40	Writing Ability	2.3547	1.7942	0.1700	ผ่านเกณฑ์
41	Writing Ability	2.0375	1.9305	0.1490	ผ่านเกณฑ์
42	Writing Ability	3.0884	1.6455	0.1594	ผ่านเกณฑ์
43	Writing Ability	2.0875	1.8464	0.1612	ผ่านเกณฑ์
44	Writing Ability	2.7751	1.9101	0.1770	ผ่านเกณฑ์
45	Writing Ability	1.9703	1.6145	0.2187	ผ่านเกณฑ์
46	Reading Ability -Vocabulary	1.4054	1.6397	0.1643	ผ่านเกณฑ์
47	Reading Ability -Vocabulary	3.3739	1.9596	0.1986	ผ่านเกณฑ์
48	Reading Ability -Vocabulary	3.8920	2.6448	0.1327	ผ่านเกณฑ์
49	Reading Ability -Vocabulary	2.2516	2.6537	0.2097	ผ่านเกณฑ์
50	Reading Ability -Vocabulary	3.2194	1.6766	0.1947	ผ่านเกณฑ์
51	Reading Ability -Reading Comprehension	1.4718	0.6304	0.1487	ผ่านเกณฑ์
52	Reading Ability -Reading Comprehension	0.7464	0.8009	0.0000	ผ่านเกณฑ์
53	Reading Ability -Reading Comprehension	1.4633	0.2436	0.0268	ผ่านเกณฑ์
54	Reading Ability -Reading Comprehension	1.2913	-0.0232	0.0112	ผ่านเกณฑ์
55	Reading Ability -Reading Comprehension	1.0399	1.2959	0.1208	ผ่านเกณฑ์
56	Reading Ability -Reading Comprehension	1.3020	2.3808	0.1452	ข้อสอบเกินชุด (5ข้อ)

ข้อ	เนื้อหา	a	b	c	หมายเหตุ
57	Reading Ability -Reading Comprehension	1.4595	3.2533	0.1735	ข้อสอบเกินชุด (5ข้อ)
58	Reading Ability -Reading Comprehension	2.0453	1.7655	0.1875	ข้อสอบเกินชุด (5ข้อ)
59	Reading Ability -Reading Comprehension	2.8155	1.8465	0.2948	ข้อสอบเกินชุด (5ข้อ)
60	Reading Ability -Reading Comprehension	1.8221	2.1663	0.1622	ข้อสอบเกินชุด (5ข้อ)
61	Reading Ability -Reading Comprehension	1.1707	2.4450	0.1376	ข้อสอบเกินชุด (5ข้อ)
62	Reading Ability -Reading Comprehension	2.0364	1.6317	0.1766	ข้อสอบเกินชุด (5ข้อ)
63	Reading Ability -Reading Comprehension	2.1486	1.9002	0.2127	ข้อสอบไม่ครบชุด (5ข้อ)
64	Reading Ability -Reading Comprehension	1.6463	1.8633	0.1283	ข้อสอบไม่ครบชุด (5ข้อ)
65	Reading Ability -Reading Comprehension	2.8457	1.5858	0.1786	ข้อสอบไม่ครบชุด (5ข้อ)
66	Reading Ability -Reading Comprehension	1.8476	1.6794	0.2268	ข้อสอบไม่ครบชุด (5ข้อ)
67	Reading Ability -Reading Comprehension	1.8895	2.4481	0.2226	ข้อสอบเกินชุด (5ข้อ)
68	Reading Ability -Reading Comprehension	1.8385	2.0927	0.2073	ข้อสอบเกินชุด (5ข้อ)
69	Reading Ability -Reading Comprehension	2.6563	2.0512	0.2144	ข้อสอบเกินชุด (5ข้อ)
70	Reading Ability -Reading Comprehension	3.5811	2.1332	0.2150	ข้อสอบเกินชุด (5ข้อ)
71	Reading Ability -Reading Comprehension	1.8424	2.7233	0.1900	ข้อสอบเกินชุด (5ข้อ)
72	Reading Ability -Reading	2.5046	2.3228	0.2494	ข้อสอบเกินชุด (5ข้อ)

ข้อ	เนื้อหา	a	b	c	หมายเหตุ
	Comprehension				
73	Reading Ability -Reading Comprehension	2.2735	2.8967	0.2245	ข้อสอบเกินชุด (5ข้อ)
74	Reading Ability -Reading Comprehension	1.4392	2.5969	0.2106	ผ่านเกณฑ์
75	Reading Ability -Reading Comprehension	1.7977	2.0823	0.2022	ผ่านเกณฑ์
76	Reading Ability -Reading Comprehension	2.7670	2.0120	0.2130	ผ่านเกณฑ์
77	Reading Ability -Reading Comprehension	3.5568	2.1064	0.2209	ผ่านเกณฑ์
78	Reading Ability -Reading Comprehension	1.8946	2.6787	0.1852	ผ่านเกณฑ์
79	Reading Ability -Reading Comprehension	2.1008	2.4212	0.2419	ผ่านเกณฑ์
80	Reading Ability -Reading Comprehension	2.1406	2.9524	0.2232	ผ่านเกณฑ์
ปี 2560					
1	Dialog Completion	2.0388	0.5680	0.1068	ผ่านเกณฑ์
2	Dialog Completion	2.5558	1.8519	0.2636	ผ่านเกณฑ์
3	Dialog Completion	2.0289	0.3419	0.1702	ผ่านเกณฑ์
4	Dialog Completion	2.2534	0.5137	0.0900	ผ่านเกณฑ์
5	Dialog Completion	1.4516	-0.0351	0.0000	ผ่านเกณฑ์
6	Dialog Completion	2.6807	0.5991	0.1272	ผ่านเกณฑ์
7	Dialog Completion	1.4436	1.5881	0.1889	ผ่านเกณฑ์
8	Dialog Completion	2.1353	0.5694	0.2555	ผ่านเกณฑ์
9	Dialog Completion	2.6184	0.5655	0.1228	ผ่านเกณฑ์
10	Dialog Completion	1.7630	0.3973	0.1574	ผ่านเกณฑ์
11	Situational Dialog	1.9557	0.6566	0.1245	ผ่านเกณฑ์

ข้อ	เนื้อหา	a	b	c	หมายเหตุ
12	Situational Dialog	2.1074	0.9050	0.0966	ผ่านเกณฑ์
13	Situational Dialog	1.2971	1.1566	0.1321	ผ่านเกณฑ์
14	Situational Dialog	3.0160	1.0366	0.1549	ผ่านเกณฑ์
15	Situational Dialog	1.2379	1.3657	0.0536	ผ่านเกณฑ์
16	Sentence Completion	0.5784	1.9481	0.0203	ผ่านเกณฑ์
17	Sentence Completion	1.5078	2.3934	0.1693	ผ่านเกณฑ์
18	Sentence Completion	0.4939	3.0428	0.0329	ผ่านเกณฑ์
19	Sentence Completion	2.2037	2.6917	0.1906	ผ่านเกณฑ์
20	Sentence Completion	3.7097	2.7164	0.1644	ผ่านเกณฑ์
21	Sentence Completion	1.4600	2.3766	0.1877	ผ่านเกณฑ์
22	Sentence Completion	2.5217	2.5772	0.1312	ผ่านเกณฑ์
23	Sentence Completion	1.5226	2.5386	0.1615	ผ่านเกณฑ์
24	Sentence Completion	1.6710	3.3522	0.1924	ผ่านเกณฑ์
25	Sentence Completion	1.7345	2.3506	0.1490	ผ่านเกณฑ์
26	Error Correction	2.0435	1.6149	0.1057	ผ่านเกณฑ์
27	Error Correction	1.0117	2.0972	0.1613	ผ่านเกณฑ์
28	Error Correction	1.8713	2.2115	0.3340	ผ่านเกณฑ์
29	Error Correction	1.4000	1.3150	0.1567	ผ่านเกณฑ์
30	Error Correction	1.7112	1.4347	0.1454	ผ่านเกณฑ์
31	Error Correction	1.9498	2.1047	0.1672	ผ่านเกณฑ์
32	Error Correction	2.6441	2.8988	0.1272	ผ่านเกณฑ์
33	Error Correction	2.1987	2.1560	0.2658	ผ่านเกณฑ์
34	Error Correction	1.2693	1.3585	0.1987	ผ่านเกณฑ์
35	Error Correction	2.9522	1.9761	0.1516	ผ่านเกณฑ์
36	Writing Ability	3.7459	2.4048	0.2371	ชุดเดียวกับ ข้อไม่ผ่านเกณฑ์
37	Writing Ability	4.6412	2.0511	0.1204	ชุดเดียวกับ

ข้อ	เนื้อหา	a	b	c	หมายเหตุ
					ข้อไม่ผ่านเกณฑ์
38	Writing Ability	0.7754	4.4733	0.1294	ไม่ผ่านเกณฑ์
39	Writing Ability	1.7936	1.7155	0.1876	ชุดเดียวกับ ข้อไม่ผ่านเกณฑ์
40	Writing Ability	1.7383	1.9161	0.1875	ชุดเดียวกับ ข้อไม่ผ่านเกณฑ์
41	Writing Ability	2.8747	2.1146	0.2665	ชุดเดียวกับ ข้อไม่ผ่านเกณฑ์
42	Writing Ability	2.2536	1.2639	0.1700	ชุดเดียวกับ ข้อไม่ผ่านเกณฑ์
43	Writing Ability	1.9382	1.7446	0.2051	ชุดเดียวกับ ข้อไม่ผ่านเกณฑ์
44	Writing Ability	3.3026	2.1132	0.1859	ชุดเดียวกับ ข้อไม่ผ่านเกณฑ์
45	Writing Ability	4.1590	1.6965	0.1520	ชุดเดียวกับ ข้อไม่ผ่านเกณฑ์
46	Reading Ability -Vocabulary	1.0592	1.1584	0.0088	ผ่านเกณฑ์
47	Reading Ability -Vocabulary	3.7754	1.8307	0.2340	ผ่านเกณฑ์
48	Reading Ability -Vocabulary	2.2931	2.1929	0.1591	ผ่านเกณฑ์
49	Reading Ability -Vocabulary	2.4389	1.5399	0.1573	ผ่านเกณฑ์
50	Reading Ability -Vocabulary	2.6945	2.1077	0.3294	ผ่านเกณฑ์
51	Reading Ability -Reading Comprehension	1.2922	1.3363	0.1369	ผ่านเกณฑ์
52	Reading Ability -Reading Comprehension	1.4938	1.0961	0.1860	ผ่านเกณฑ์
53	Reading Ability -Reading Comprehension	2.3416	1.0240	0.1679	ผ่านเกณฑ์
54	Reading Ability -Reading Comprehension	2.0943	1.8016	0.1838	ผ่านเกณฑ์
55	Reading Ability -Reading	1.8364	0.7998	0.1463	ผ่านเกณฑ์

ข้อ	เนื้อหา	a	b	c	หมายเหตุ
	Comprehension				
56	Reading Ability -Reading Comprehension	2.1264	1.9294	0.1965	ข้อสอบเกินชุด (5ข้อ)
57	Reading Ability -Reading Comprehension	1.8225	1.9234	0.2292	ข้อสอบเกินชุด (5ข้อ)
58	Reading Ability -Reading Comprehension	1.9384	1.4359	0.1722	ข้อสอบเกินชุด (5ข้อ)
59	Reading Ability -Reading Comprehension	3.4327	1.7458	0.2291	ข้อสอบเกินชุด (5ข้อ)
60	Reading Ability -Reading Comprehension	3.3904	1.8963	0.2416	ข้อสอบเกินชุด (5ข้อ)
61	Reading Ability -Reading Comprehension	2.6902	1.6833	0.1175	ข้อสอบเกินชุด (5ข้อ)
62	Reading Ability -Reading Comprehension	2.5645	1.9821	0.1761	ข้อสอบเกินชุด (5ข้อ)
63	Reading Ability -Reading Comprehension	3.0882	2.1437	0.1844	ข้อสอบเกินชุด (5ข้อ)
64	Reading Ability -Reading Comprehension	2.8766	2.8974	0.2293	ข้อสอบเกินชุด (5ข้อ)
65	Reading Ability -Reading Comprehension	1.6821	2.3523	0.2056	ข้อสอบเกินชุด (5ข้อ)
66	Reading Ability -Reading Comprehension	1.5837	1.9153	0.1371	ข้อสอบเกินชุด (5ข้อ)
67	Reading Ability -Reading Comprehension	1.8247	2.0667	0.2234	ข้อสอบเกินชุด (5ข้อ)
68	Reading Ability -Reading Comprehension	3.0058	2.3408	0.2551	ข้อสอบเกินชุด (5ข้อ)
69	Reading Ability -Reading Comprehension	2.7275	1.2706	0.2076	ข้อสอบเกินชุด (5ข้อ)
70	Reading Ability -Reading Comprehension	2.5385	1.9826	0.2130	ข้อสอบเกินชุด (5ข้อ)

ข้อ	เนื้อหา	a	b	c	หมายเหตุ
71	Reading Ability -Reading Comprehension	1.6434	1.8810	0.1546	ข้อสอบเกินชุด (5ข้อ)
72	Reading Ability -Reading Comprehension	1.9510	1.9289	0.1394	ข้อสอบเกินชุด (5ข้อ)
73	Reading Ability -Reading Comprehension	3.2474	1.8735	0.2015	ข้อสอบเกินชุด (5ข้อ)
74	Reading Ability -Reading Comprehension	3.4600	2.1532	0.2363	ข้อสอบเกินชุด (5ข้อ)
75	Reading Ability -Reading Comprehension	2.6800	2.5311	0.2352	ข้อสอบเกินชุด (5ข้อ)
76	Reading Ability -Reading Comprehension	3.9147	2.4276	0.1440	ผ่านเกณฑ์
77	Reading Ability -Reading Comprehension	2.1206	1.9135	0.1877	ผ่านเกณฑ์
78	Reading Ability -Reading Comprehension	1.6556	2.2832	0.1866	ผ่านเกณฑ์
79	Reading Ability -Reading Comprehension	2.0484	1.6804	0.2263	ผ่านเกณฑ์
80	Reading Ability -Reading Comprehension	1.9546	1.9565	0.1684	ผ่านเกณฑ์

ปี 2561

1	Dialog Completion	3.1120	0.5260	0.2580	ผ่านเกณฑ์
2	Dialog Completion	3.2140	0.8970	0.1610	ผ่านเกณฑ์
3	Dialog Completion	1.9670	0.6930	0.1490	ผ่านเกณฑ์
4	Dialog Completion	3.1210	0.8030	0.2080	ผ่านเกณฑ์
5	Dialog Completion	2.1310	0.2460	0.2980	ผ่านเกณฑ์
6	Dialog Completion	2.8930	0.5490	0.2560	ผ่านเกณฑ์
7	Dialog Completion	3.5480	0.8050	0.1640	ผ่านเกณฑ์
8	Dialog Completion	2.2550	0.6430	0.1590	ผ่านเกณฑ์

ข้อ	เนื้อหา	a	b	c	หมายเหตุ
9	Dialog Completion	3.1870	0.8300	0.2090	ผ่านเกณฑ์
10	Dialog Completion	2.4060	0.2930	0.3270	ผ่านเกณฑ์
11	Situational Dialog	2.1230	1.2970	0.2080	ผ่านเกณฑ์
12	Situational Dialog	2.4100	0.8710	0.1930	ผ่านเกณฑ์
13	Situational Dialog	1.6340	0.8220	0.1490	ผ่านเกณฑ์
14	Situational Dialog	2.4850	0.7710	0.2030	ผ่านเกณฑ์
15	Situational Dialog	2.8190	0.8860	0.1650	ผ่านเกณฑ์
16	Sentence Completion	0.7250	2.7790	0.1430	ผ่านเกณฑ์
17	Sentence Completion	1.0510	2.1070	0.2060	ผ่านเกณฑ์
18	Sentence Completion	2.5280	2.5580	0.1870	ผ่านเกณฑ์
19	Sentence Completion	2.0420	2.7010	0.2150	ผ่านเกณฑ์
20	Sentence Completion	0.4870	3.4060	0.0710	ผ่านเกณฑ์
21	Sentence Completion	4.1050	2.2030	0.1670	ผ่านเกณฑ์
22	Sentence Completion	1.2020	2.8650	0.2090	ผ่านเกณฑ์
23	Sentence Completion	0.2350	4.5710	0.0400	ไม่ผ่านเกณฑ์
24	Sentence Completion	1.2250	1.7250	0.1040	ผ่านเกณฑ์
25	Sentence Completion	2.2020	2.3360	0.1990	ผ่านเกณฑ์
26	Error Correction	1.1870	1.4950	0.2230	ชุดเดียวกับ ข้อไม่ผ่านเกณฑ์
27	Error Correction	-67.9440	-1.5120	0.0900	ไม่ผ่านเกณฑ์
28	Error Correction	0.3850	0.5590	0.0050	ชุดเดียวกับ ข้อไม่ผ่านเกณฑ์
29	Error Correction	3.0540	1.5400	0.2170	ชุดเดียวกับ ข้อไม่ผ่านเกณฑ์
30	Error Correction	2.5040	2.1460	0.1170	ชุดเดียวกับ ข้อไม่ผ่านเกณฑ์
31	Error Correction	3.1740	1.8840	0.2080	ชุดเดียวกับ ข้อไม่ผ่านเกณฑ์

ข้อ	เนื้อหา	a	b	c	หมายเหตุ
32	Error Correction	3.7000	2.2520	0.1000	ชุดเดียวกับข้อไม่ผ่านเกณฑ์
33	Error Correction	4.4460	2.2120	0.2300	ชุดเดียวกับข้อไม่ผ่านเกณฑ์
34	Error Correction	5.2420	1.1290	0.2060	ชุดเดียวกับข้อไม่ผ่านเกณฑ์
35	Error Correction	2.8890	2.4850	0.2490	ชุดเดียวกับข้อไม่ผ่านเกณฑ์
36	Writing Ability	1.2210	0.9400	0.1870	ผ่านเกณฑ์
37	Writing Ability	2.2410	1.7560	0.1570	ผ่านเกณฑ์
38	Writing Ability	3.1200	1.6850	0.1700	ผ่านเกณฑ์
39	Writing Ability	1.8760	2.0810	0.2320	ผ่านเกณฑ์
40	Writing Ability	3.2250	1.5190	0.2560	ผ่านเกณฑ์
41	Writing Ability	2.6390	1.8020	0.0990	ผ่านเกณฑ์
42	Writing Ability	2.3220	1.8810	0.1630	ผ่านเกณฑ์
43	Writing Ability	2.4460	1.5020	0.2260	ผ่านเกณฑ์
44	Writing Ability	4.0400	2.6510	0.1260	ผ่านเกณฑ์
45	Writing Ability	3.0330	1.7830	0.1820	ผ่านเกณฑ์
46	Reading Ability -Vocabulary	2.0690	1.8500	0.1390	ผ่านเกณฑ์
47	Reading Ability -Vocabulary	2.5860	2.5780	0.1420	ผ่านเกณฑ์
48	Reading Ability -Vocabulary	1.5080	1.4360	0.1870	ผ่านเกณฑ์
49	Reading Ability -Vocabulary	3.6660	1.9510	0.1610	ผ่านเกณฑ์
50	Reading Ability -Vocabulary	3.5160	2.4510	0.1150	ผ่านเกณฑ์
51	Reading Ability -Reading Comprehension	1.8350	-0.1240	0.1160	ผ่านเกณฑ์
52	Reading Ability -Reading Comprehension	2.0700	-0.5080	0.0000	ผ่านเกณฑ์
53	Reading Ability -Reading Comprehension	2.2010	-0.5690	0.0000	ผ่านเกณฑ์

ข้อ	เนื้อหา	a	b	c	หมายเหตุ
54	Reading Ability -Reading Comprehension	2.1250	-0.3180	0.0000	ผ่านเกณฑ์
55	Reading Ability -Reading Comprehension	3.1690	-0.3530	0.1280	ผ่านเกณฑ์
56	Reading Ability -Reading Comprehension	1.6890	1.7460	0.1360	ผ่านเกณฑ์
57	Reading Ability -Reading Comprehension	1.5630	1.3690	0.1500	ผ่านเกณฑ์
58	Reading Ability -Reading Comprehension	2.5700	1.6950	0.1500	ผ่านเกณฑ์
59	Reading Ability -Reading Comprehension	2.2330	1.5360	0.1780	ผ่านเกณฑ์
60	Reading Ability -Reading Comprehension	2.0760	1.7680	0.2120	ผ่านเกณฑ์
61	Reading Ability -Reading Comprehension	2.3000	1.9650	0.2340	ข้อสอบเกินชุด (5ข้อ)
62	Reading Ability -Reading Comprehension	1.9880	1.6480	0.1980	ข้อสอบเกินชุด (5ข้อ)
63	Reading Ability -Reading Comprehension	1.1760	1.8660	0.1960	ข้อสอบเกินชุด (5ข้อ)
64	Reading Ability -Reading Comprehension	3.0230	1.6660	0.1950	ข้อสอบเกินชุด (5ข้อ)
65	Reading Ability -Reading Comprehension	3.4020	2.3990	0.1810	ข้อสอบเกินชุด (5ข้อ)
66	Reading Ability -Reading Comprehension	1.7620	1.5330	0.1960	ข้อสอบเกินชุด (5ข้อ)
67	Reading Ability -Reading Comprehension	3.5190	2.8410	0.1680	ข้อสอบเกินชุด (5ข้อ)
68	Reading Ability -Reading Comprehension	0.8550	2.7290	0.1470	ข้อสอบเกินชุด (5ข้อ)
69	Reading Ability -Reading	1.9180	2.1240	0.1740	ข้อสอบเกินชุด (5ข้อ)

ข้อ	เนื้อหา	a	b	c	หมายเหตุ
	Comprehension				
70	Reading Ability -Reading Comprehension	2.4050	1.3060	0.1710	ข้อสอบเกินชุด (5ข้อ)
71	Reading Ability -Reading Comprehension	1.1310	2.3820	0.1750	ข้อสอบเกินชุด (5ข้อ)
72	Reading Ability -Reading Comprehension	2.4320	2.3500	0.1600	ข้อสอบเกินชุด (5ข้อ)
73	Reading Ability -Reading Comprehension	1.7280	2.0890	0.1970	ข้อสอบเกินชุด (5ข้อ)
74	Reading Ability -Reading Comprehension	2.7000	1.3260	0.2000	ข้อสอบเกินชุด (5ข้อ)
75	Reading Ability -Reading Comprehension	2.7510	1.5650	0.2100	ข้อสอบเกินชุด (5ข้อ)
76	Reading Ability -Reading Comprehension	1.8570	2.2390	0.2050	ข้อสอบเกินชุด (5ข้อ)
77	Reading Ability -Reading Comprehension	0.8090	2.3300	0.1640	ข้อสอบเกินชุด (5ข้อ)
78	Reading Ability -Reading Comprehension	1.8900	2.6200	0.2050	ข้อสอบเกินชุด (5ข้อ)
79	Reading Ability -Reading Comprehension	3.3960	2.0630	0.1710	ข้อสอบเกินชุด (5ข้อ)
80	Reading Ability -Reading Comprehension	1.7720	1.8440	0.1580	ข้อสอบเกินชุด (5ข้อ)

ภาคผนวก ข
คู่มือการใช้งาน
ระบบการทดสอบแบบปรับเหมาะหลายชั้นตอนที่มีการให้ข้อมูลป้อนกลับ
ในการทดสอบระดับชาติ (O-NET) วิชาภาษาอังกฤษสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6

การเข้าใช้งานระบบทดสอบแบบปรับเหมาะหลายชั้นตอนที่มีการให้ข้อมูลป้อนกลับ ในการทดสอบระดับชาติ (O-NET) วิชาภาษาอังกฤษสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6 มีขั้นตอนดังต่อไปนี้

ขั้นที่ 1 การลงทะเบียนสำหรับตัวอย่างในการวิจัย

ตัวอย่างวิจัยทุกคนต้องดำเนินการลงทะเบียนก่อนเข้ารับการทดสอบ โดยเข้าเว็บไซต์ <https://www.nhabhat.com> หลังจากนั้นจะปรากฏข้อมูลดังภาพที่ 1 ทำการเลือกแถบเมนู “ลงทะเบียน” เพื่อลงทะเบียน โดยตัวอย่างวิจัยต้องกรอก อีเมล (e-mail) กำหนดรหัสผ่าน (password) และกรอกข้อมูลพื้นฐานของตนเองให้ครบถ้วน ดังภาพที่ 2 เมื่อกรอกข้อมูลเรียบร้อยแล้ว ให้กดปุ่ม “ลงทะเบียน” ระบบจะแสดงหน้าจอข้อมูลพื้นฐานของผู้ลงทะเบียน ซึ่งตัวอย่างวิจัยสามารถแก้ไขข้อมูลส่วนบุคคลของตนเองได้ โดยกดปุ่ม “EDIT” ที่มุมบนด้านขวามือ แต่หากต้องการเริ่มทดสอบ สามารถกดที่ปุ่ม “เข้าห้องสอบ” ทางด้านซ้ายมือของหน้าจอ รายละเอียดดังภาพที่ 3

ภาพที่ 1 หน้าจอหลักระบบการทดสอบ

ระบบการทดสอบแบบปรับเหมาะหลายขั้นตอนที่มีการให้ข้อมูลป้อนกลับ
ในการทดสอบระดับชาติ (O-NET)
วิชา ภาษาอังกฤษสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6

เข้าใช้งานระบบ ลงทะเบียน

อีเมล : (ใช้สำหรับล็อกอินเข้าระบบและรีเซตรหัสผ่าน)
onet@onet.com

รหัสผ่านสำหรับเข้าระบบสอบ : ยืนยันอีกครั้ง :
**** ****

ชื่อ :
ตัวอย่าง

นามสกุล :
ทดสอบ


อายุ : เพศ : สายการเรียน : ระดับชั้น :
18 ชาย วิทยาศาสตร์ ม. 6

จังหวัด : โรงเรียน :
กรุงเทพมหานคร ตัวอย่าง

ลงทะเบียน

ภาพที่ 2 หน้าจอการลงทะเบียน

Welcome : onet@onet.com EDIT

 ยินดีต้อนรับ
ตัวอย่าง ทดสอบ
ลงทะเบียนเมื่อ 2020-03-17 09:47:11

[เข้าห้องสอบ](#)

[เปลี่ยนรหัสผ่านเข้าระบบ](#)

ข้อมูลพื้นฐาน

ชื่อ : ตัวอย่าง
นามสกุล : ทดสอบ
เพศ : ชาย อายุ : 18

ข้อมูลการศึกษา

ชั้นเรียน : ม. 6 สายการเรียน : วิทยาศาสตร์
โรงเรียน : ตัวอย่าง
จังหวัด : กรุงเทพมหานคร

ภาพที่ 3 การแสดงสถานะเมื่อกดปุ่มลงทะเบียน

ขั้นที่ 2 การเข้าสู่ระบบการทดสอบ

การเข้าสู่ระบบการทดสอบในหน้าแรก ตัวอย่างวิจัยดำเนินการกรอกอีเมล (e-mail) และรหัส (password) ที่ลงทะเบียนไว้ หลังจากนั้นกดปุ่ม “เข้าระบบ” เพื่อเข้าสู่ระบบ รายละเอียดแสดงดังภาพที่ 4

ภาพที่ 4 การเข้าสู่ระบบการทดสอบ

เมื่อตัวอย่างวิจัยเข้าสู่ระบบการทดสอบเรียบร้อยแล้ว ระบบจะแสดงหน้าจอข้อมูลพื้นฐานของผู้เข้าระบบ ซึ่งจะมีลักษณะเช่นเดียวกับที่แสดงในภาพที่ 8 ซึ่งตัวอย่างวิจัยสามารถแก้ไขข้อมูลโดยกดปุ่ม “EDIT” ที่มุมบนด้านขวามือ เมื่อแก้ไขเรียบร้อยแล้วให้กด “บันทึก” เพื่อบันทึกการแก้ไขข้อมูลดังแสดงในภาพที่ 5 หรือกดเลือก “เข้าห้องสอบ” เพื่อเข้าสู่ระบบการทดสอบ

ภาพที่ 5 การแก้ไขข้อมูลส่วนตัว

ขั้นที่ 3 การดำเนินการทดสอบ

เมื่อเข้าสู่ระบบเรียบร้อยแล้วหน้าจอจะแสดงคำชี้แจงรายละเอียดเกี่ยวกับการทดสอบ ซึ่งระบบการทดสอบนี้เป็นระบบที่จัดทำขึ้นเพื่อเตรียมความพร้อมในการทดสอบระดับชาติ (O-NET) วิชาภาษาอังกฤษสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6 โดยรายละเอียดหน้าจอคำชี้แจงแสดงดังภาพที่ 6 หลังจากอ่านคำชี้แจงเรียบร้อยแล้วให้กดเลือก “ยอมรับ-รับทราบขั้นตอน” และกด “เข้าสู่ห้องสอบ” เพื่อเริ่มการทดสอบ

**ระบบการทดสอบแบบปรับเหมาะหลายขั้นตอนที่มีการให้ข้อมูลป้อนกลับ
ในการทดสอบระดับชาติ (O-NET)
วิชา ภาษาอังกฤษสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6**

คำชี้แจง :
ระบบการทดสอบแบบปรับเหมาะหลายขั้นตอนที่มีการให้ข้อมูลป้อนกลับ
ในการทดสอบระดับชาติ (O-NET)
วิชา ภาษาอังกฤษสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6

ขั้นตอนการทำแบบทดสอบ:

1. เวลาในการทำแบบทดสอบจะเริ่มขึ้นเมื่อกำหนดปุ่ม **เข้าสู่ห้องสอบ**
2. ผู้สอบมีเวลาในการทำข้อสอบทั้งสิ้น 120 นาที และระบบจะยุติการทดสอบอัตโนมัติเมื่อหมดเวลาสอบ
3. ข้อสอบจะแสดงเป็นขั้น (stage) ตามระดับความสามารถ ซึ่งมีทั้งสิ้น 3 ขั้น โดยผู้สอบสามารถกลับไปแก้ไขคำตอบภายในขั้นเดียวกันได้ **หากผู้สอบกดปุ่ม **ไม่ทันต่อไป** จะไม่สามารถกลับมาแก้ไขข้อสอบในขั้นก่อนหน้าได้**
4. ผู้สอบแต่ละคนจะได้รับข้อสอบไม่เหมือนกัน ขึ้นอยู่กับระดับความสามารถของผู้สอบ
5. การทดสอบจะยุติลงเมื่อระดับความสามารถของผู้สอบคงที่ หลังจากนั้นผู้สอบต้องดำเนินการ ตอบแบบประเมินความพึงพอใจ ระบบจึงจะแสดงรายงานผลข้อมูลป้อนกลับให้ผู้สอบทราบ

✓ ยอมรับ - รับทราบในขั้นตอน **เข้าสู่ห้องสอบ**

ภาพที่ 6 หน้าจอคำชี้แจง

หลังจากกด “เข้าสู่ห้องสอบ” หน้าจอจะแสดงข้อสอบข้อแรก ดังภาพที่ 7 แถบข้อความด้านบนจะแสดงชื่อผู้เข้าสอบ และเวลาคงเหลือในการทำแบบทดสอบ

ผู้เข้าสอบ : ตัวอย่าง ทดสอบ ท่านเหลือเวลาทำข้อสอบอีก 01:59:40

ใช้เวลาในการทดสอบจำนวน 2 ชั่วโมง เลือกคำตอบที่ถูกต้องและคลิกปุ่ม [ไปขั้นตอนถัดไป >>](#) เมื่อครบเวลาในการทดสอบ ระบบจะหยุดการทดสอบทันที
ท่านไม่สามารถย้อนกลับไปยังคำตอบในขั้นตอนหน้าได้

Part I: Language Use
Dialog Completion Directions: Read the dialogs and choose the expression that BEST completes each missing part.

Situation: Sue is moving into a new apartment. Now she is asking the clerk for the key to her room. Sue: Hello. My name is Susan Clayton. I'm moving in today. 1. _____ Clerk: Good afternoon, Ms. Clayton. I'll show you to your room. 2. _____ Sue: That's very kind of you. Clerk: 3. _____ Er.....Ms. Clayton, may I remind you not to hang clothes on the balcony? Sue: 4. _____ Clerk: I'm afraid not. I hope you will enjoy staying here. 5. _____ Sue: Thank you.

1

1. I do want to go to my room now.
2. Did you get the key for the room?
3. May I have the key to my room, please?
4. Can't you give me my room key now?
5. I should have my key now, you know.

2

1. Do you bring along your things?
2. Will you take your things to your room?
3. Don't forget to take all of your things.
4. Won't you take all your things with you?
5. Let me give you a hand with your things.

สถานการณ์ทำข้อสอบ ข้อที่ 1 ข้อที่ 2 ข้อที่ 3 ข้อที่ 4 ข้อที่ 5 [ไปขั้นตอนถัดไป >>](#)

ภาพที่ 7 หน้าจอการทดสอบ

ในการตอบคำถามแต่ละข้อนั้น ทำได้โดยการนำเมาส์ชี้ที่คำตอบที่ต้องการ ซึ่งหน้าจอจะแสดงแถบสีเขียวอ่อน และเมื่อกดเลือกคำตอบแล้ว แถบสีนี้จะเปลี่ยนเป็นสีเขียวเข้ม นอกจากนี้ หน้าจอด้านล่างยังแสดงแถบ “สถานการณ์ทำข้อสอบ” ให้ผู้สอบทราบว่าได้ทำข้อสอบในตอนนี้ครบถ้วนหรือยัง รายละเอียดดังภาพที่ 8

Part I: Language Use
Situational Dialogs Directions: Read each situation and choose the best alternative.

1 Situation: A foreign tourist wants to take a bus to Kanchanaburi, so he asks a man for directions to the bus terminal. The man says: _____

1. Let's take a taxi to Kanchanaburi.
2. There are many buses going to the bus terminal.
3. You need to go by train because it's faster.
4. **It's on your left, just after you pass the next intersection.**
5. Kanchanaburi's about 120 kilometers from Bangkok.

2 Situation: Dorothy is at the Tourist Information Center. She wants to get directions to the train station. She says: _____

1. **How can I get to the train station?**
2. Do you mind taking me to the train station?
3. Excuse me, have you been to the train station?
4. I'm a stranger around here. What can you do for me?
5. Can you show me the train station, please?

3 Situation: Linda is answering the phone. Mr. Hughes, Marketing Director of the ABC Company, would like to talk to Mr. Jenkins, Head of the Personnel Department. She says: _____

1. Mr. Hughes? Yes, Mr. Jenkins is expecting your call.
2. Oh, hi! The meeting's just started. Don't call again.

สถานการณ์ทำข้อสอบ ข้อที่ 1 ข้อที่ 2 ข้อที่ 3 ข้อที่ 4 ข้อที่ 5 [ไปขั้นตอนถัดไป >>](#)

ภาพที่ 8 หน้าจอการกดเลือกคำตอบ

เมื่อผู้สอบตอบคำถามครบทุกข้อของตอนนั้น ๆ แล้ว แถบสถานการณ์ทำข้อสอบจะเป็นสีเขียวทุกข้อ และปุ่ม “ไปขั้นตอนถัดไป” จะปรากฏเป็นสีฟ้าเข้ม และสามารถกดเพื่อไปทำการทดสอบในขั้นถัดไปได้ ดังภาพที่ 9

4 Situation: Mark has not had enough sleep. When he meets John, John notices it and says: ____

1. I can't believe it, Mark. You haven't changed at all. 2. Hi, Mark. Will you come to my birthday party next Sunday?
3. You look so tired today! Why don't you get some coffee? 4. What's up? You never answer my calls.
 5. I heard that you went to Spain last year. Did you have a good time there?

5 Situation: Laura is unhappy with her roommate, Lucy. She says to Rose, " ____"

1. Lucy's just moved in and she is very well-organized. I must annoy her. 2. You know what? Lucy bought me a box of chocolate for my birthday.
 3. Do you know Lucy? She had a fight with her ex-boyfriend. **4. I hate to say this, but I can't stand Lucy any longer. I want to move out.**
 5. Everyone in this apartment knows Lucy. She has a reputation as a good negotiator.

สถานะการทำข้อสอบ [ข้อที่ 1](#) [ข้อที่ 2](#) [ข้อที่ 3](#) [ข้อที่ 4](#) [ข้อที่ 5](#) [ไปที่ข้อต่อไป >>](#)

ภาพที่ 9 หน้าจอเมื่อตอบคำถามครบในขั้นนั้น ๆ

ขั้นที่ 4 การยุติการทดสอบ

ในการทดสอบตัวอย่างวิจัยต้องดำเนินการทดสอบจำนวน 3 ขั้น (stage) หรือเท่ากับ 50 ข้อ ประกอบด้วยขั้นที่ 1 จำนวน 25 ข้อ ขั้นที่ 2 จำนวน 10 ข้อ และขั้นที่ 3 จำนวน 15 ข้อ เมื่อตัวอย่างวิจัยตอบข้อสอบครบทุกข้อ ระบบจะทำการยุติการทดสอบเนื่องจากเป็นการทดสอบแบบความยาวคงที่ (fixed length) โดยแสดงหน้าจอดังภาพที่ 10 และจะแสดงหน้าการประเมินความพึงพอใจต่อการใช้งานระบบ โดยตัวอย่างวิจัยต้องทำการประเมินทุกคน หากไม่ประเมิน ระบบจะไม่แสดงรายงานผลและข้อมูลย้อนกลับให้ตัวอย่างวิจัยทราบ

ขอแสดงความยินดี

การทดสอบเสร็จสิ้น



ระบบกำลังจะนำท่านไปยังหน้าสำหรับตอบแบบประเมินความพึงพอใจ
หลังจากท่านทำแบบประเมินฯ ครบแล้ว ท่านจะได้รับผลการสอบในลำดับต่อไป

[เพื่อไปยังหน้าตอบแบบประเมินความพึงพอใจ](#)

ภาพที่ 10 หน้าจอเมื่อตอบคำถามครบทุกขั้น

ขั้นที่ 5 การประเมินคุณภาพระบบการทดสอบ

เมื่อตัวอย่างวิจัยทำข้อสอบเสร็จสิ้นครบจำนวน 50 ข้อตามเกณฑ์ที่กำหนดไว้ ระบบจะทำการแสดงหน้าการประเมินความพึงพอใจต่อการใช้งานระบบให้ตัวอย่างวิจัยประเมิน ตัวอย่างวิจัยต้องทำการประเมินทุกข้อ มิเช่นนั้นระบบจะไม่แสดงรายงานผลและข้อมูลย้อนกลับให้ตัวอย่างวิจัยทราบ หลังตัวอย่างวิจัยกดส่งผลการประเมินเรียบร้อยแล้ว ระบบจะแสดงผลรายงานการทดสอบและข้อมูลย้อนกลับให้ตัวอย่างวิจัยทราบ สำหรับหน้าจอกการประเมินความพึงพอใจต่อการใช้งานระบบแสดงดังภาพที่ 11 และเมื่อตัวอย่างวิจัยประเมินครบทุกข้อ จะปรากฏปุ่ม “ส่งแบบสอบถาม & ดูรายงาน” ให้สามารถกดเพื่อแสดงรายงานผลการทดสอบที่ได้ทำเสร็จสิ้นไปแล้วได้ ดังภาพที่ 12

ONET System : Survey												
แบบประเมินความพึงพอใจของผู้สอบต่อการใช้งานระบบ												
คำชี้แจง: โปรดคลิกให้คะแนนที่ ๑ ตามความเห็นของท่าน ดังนี้ ***มาตรฐานในการวัด: 10 ระดับ 0 1 2 3 4 5 6 7 8 9												
ความเห็นต่อการใช้งานระบบในภาพรวม												
1. ความเห็นต่อการใช้งานระบบในภาพรวม	แย่มาก	0	1	2	3	4	5	6	7	8	9	ดีเยี่ยม
	ไม่พอใจ	0	1	2	3	4	5	6	7	8	9	พอใจ
	น่าเบื่อ	0	1	2	3	4	5	6	7	8	9	น่าสนใจ
หน้าจอกการแสดงผล (screen)												
2. การแสดงผลของหน้าจอ	อ่านยาก	0	1	2	3	4	5	6	7	8	9	อ่านง่าย
3. การใช้งานของเครื่องมือ (tools) บนหน้าจอกการแสดงผล	ยาก	0	1	2	3	4	5	6	7	8	9	ง่าย
4. การเรียงลำดับเครื่องมือบนหน้าจอกการแสดงผล	สับสน	0	1	2	3	4	5	6	7	8	9	ชัดเจน
5. การเรียงลำดับเนื้อหาของหน้าจอกการแสดงผล	สับสน	0	1	2	3	4	5	6	7	8	9	ชัดเจน
คำศัพท์และสารสนเทศระบบ (terminology and system information)												
6. การใช้คำศัพท์ในระบบ	เข้าใจยาก	0	1	2	3	4	5	6	7	8	9	เข้าใจง่าย
7. คำศัพท์ในระบบมีความสัมพันธ์กับเครื่องมือ (tools) ที่แสดงในระบบ	น้อย	0	1	2	3	4	5	6	7	8	9	มาก

ภาพที่ 11 หน้าจอกการประเมินความพึงพอใจต่อการใช้งานระบบ

9. ข้อความที่แสดงให้กับผู้สอบเพื่อป้อนข้อมูลบนหน้าจอกการแสดงผล	สับสน	0	1	2	3	4	5	6	7	8	9	ชัดเจน
10. การแจ้งเตือนสถานะระหว่างการใช้งาน	ไม่มีเลย	0	1	2	3	4	5	6	7	8	9	มี
11. ข้อความแจ้งเตือนความคิดพลาด	เข้าใจยาก	0	1	2	3	4	5	6	7	8	9	เข้าใจง่าย
การเรียนรู้ (learning)												
12. การเรียนรู้การใช้งานของระบบ	ยาก	0	1	2	3	4	5	6	7	8	9	ง่าย
13. ความสามารถในการจดจำข้อมูลและคำสั่งที่เคยมีขึ้น	ทำไม่ได้	0	1	2	3	4	5	6	7	8	9	ทำได้
14. ผู้สอบสามารถใช้งานเครื่องมือต่าง ๆ ในระบบ	ไม่ค่อย	0	1	2	3	4	5	6	7	8	9	ดี
15. ข้อความแจ้งเตือนความคิดพลาดช่วยให้เข้าใจการทำงานของระบบ	ไม่ช่วย	0	1	2	3	4	5	6	7	8	9	ช่วย
16. เอกสารคู่มือการใช้งาน	สับสน	0	1	2	3	4	5	6	7	8	9	ชัดเจน
ความสามารถระบบ (system capabilities)												
17. ความเร็วในการตอบสนองต่อการใช้งานระบบ	ช้า	0	1	2	3	4	5	6	7	8	9	เร็ว
18. ความต่อเนื่องในการใช้งานระบบ	ติดขัด	0	1	2	3	4	5	6	7	8	9	ราบรื่น
19. แนวโน้มความต่อเนื่องในการทำงานของระบบ	ติดขัด	0	1	2	3	4	5	6	7	8	9	ราบรื่น
20. ข้อขัดข้องในการใช้งานระบบ	มี	0	1	2	3	4	5	6	7	8	9	ไม่มี

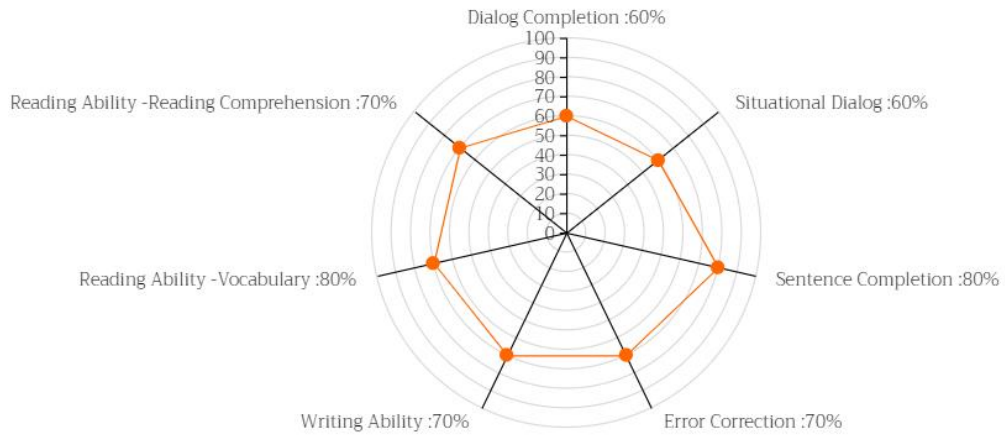
ส่งแบบสอบถาม & ดูรายงาน

ภาพที่ 12 หน้าจอกเมื่อประเมินความพึงพอใจต่อการใช้งานระบบครบทุกข้อ

ขั้นที่ 6 การรายงานผลการทดสอบ

สำหรับการออกผลรายงานผลการทดสอบ ตัวอย่างวิจัยต้องดำเนินการประเมินคุณภาพระบบ ด้วยแบบประเมินความพึงพอใจของผู้สอบต่อการใช้งานระบบให้เรียบร้อยก่อน เมื่อดำเนินการประเมินผลเรียบร้อยแล้ว ระบบจะรายงานผลการทดสอบให้ตัวอย่างวิจัยทราบดังภาพที่ 13





แผนภูมิแสดงผลการตอบแบบทดสอบ

ภาพที่ 13 รายงานผลการทดสอบและข้อมูลย้อนกลับ

ประวัติผู้วิจัย

อาจารย์ ดร.ณภัทร ชัยมงคล

การศึกษา:

ค.ต. (การวัดและประเมินผลการศึกษา)	จุฬาลงกรณ์มหาวิทยาลัย
วท.ม. (ระบบสารสนเทศเพื่อการจัดการ)	ม.พระจอมเกล้าพระนครเหนือ
บธ.บ. (การจัดการทั่วไป)	ม.สุโขทัยธรรมาธิราช
วท.บ. (สถิติประยุกต์)	ส.พระจอมเกล้าพระนครเหนือ

การทำงาน:

- อาจารย์ประจำภาควิชาวิจัยและจิตวิทยาการศึกษา คณะครุศาสตร์
จุฬาลงกรณ์มหาวิทยาลัย
- กรรมการและเลขานุการ คณะกรรมการบริหารหลักสูตรครุศาสตรมหาบัณฑิต
และดุขภูมิบัณฑิต สาขาวิชาการศึกษาวัดและประเมินผลการศึกษา คณะครุศาสตร์
จุฬาลงกรณ์มหาวิทยาลัย
- กรรมการและนายทะเบียน สมาคมวิจัยสังคมศาสตร์แห่งประเทศไทย
- บรรณาธิการวารสารการวิจัยสังคมศาสตร์

ผลงานทางวิชาการบางส่วน:

- ณภัทร ชัยมงคล และศิริชัย กาญจนวาสี. (2563). การประยุกต์ใช้โปรแกรมอาร์สำหรับการทดสอบ
ทางการศึกษา. *วารสารวิธีวิทยาการวิจัย*, 33(1). 43-67.
- ณภัทร ชัยมงคล. (2562). วิธีการทางสถิติสำหรับตรวจจับการทุจริตประเภทลอกคำตอบ
ในการทดสอบ. *วารสารการวัดผลการศึกษา*, 36(100). 16-28.
- สังวรรณ รัตตะโทก กาญจนนา วันสุนทร นลินี ณ นคร ศศิธร ชูตินันทกุล ปรรธนา พลอภิชาติ
อนุสรณ์ เกิดศรี ปิยะนาฏ สิทธิฤทธิ ณภัทร ชัยมงคล และนพรัตน์ ไบยา. คุณภาพข้อสอบ
ระดับชาติ (โอเน็ต) และแนวทางการพัฒนาคุณภาพการทดสอบ. (2561).
วารสารสุโขทัยธรรมาธิราช, 31(2). 110-123.
- โชติกา ภาษีผล, ศิริเดช สุชีวะ, ณัฐธรรณ หลาวทอง, กมลวรรณ ตั้งชนกานนท์ และ
ณภัทร ชัยมงคล. (2560). การวิจัยและพัฒนาฐานข้อมูลเครื่องมือวัดการเรียนรู้และ
คุณลักษณะอันพึงประสงค์ของผู้เรียน. *วารสารวิจัย มสค สาขามนุษยศาสตร์และสังคมศาสตร์*,
13(2). 207-224.