

## การพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับการจัดสอบ O-NET

สุชาดา กรเพชรปาดิ<sup>1</sup> ปิยะทิพย์ ตินวร<sup>1\*</sup> โสฬส สุขานนท์สวัสดิ์<sup>2</sup>

<sup>1</sup> ศูนย์การวัดผลประยุกต์ วิทยาลัยวิทยาการวิจัยและวิทยาการปัญญา มหาวิทยาลัยบูรพา

<sup>2</sup> สำนักมาตรฐานและมั่นคงปลอดภัยสารสนเทศ การไฟฟ้านครหลวง

### บทคัดย่อ

การวิจัยนี้มีวัตถุประสงค์เพื่อจัดทำคลังข้อสอบ O-NET และพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการจัดสอบ O-NET ระดับชั้นประถมศึกษาปีที่ 6 ชั้นมัธยมศึกษาปีที่ 3 และชั้นมัธยมศึกษาปีที่ 6 ระดับชั้นละ 8 กลุ่มสาระการเรียนรู้ การจัดทำคลังข้อสอบใช้ฐานข้อมูล MySQL และคัดเลือกข้อสอบ O-NET ของสถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) ระหว่างปี พ.ศ. 2551-2553 ที่ผ่านการวิเคราะห์ข้อสอบตามทฤษฎีการตอบสนองข้อสอบ แบบ 3 พารามิเตอร์ ในระดับชั้นประถมศึกษาปีที่ 6 จำนวน 258 ข้อ ชั้นมัธยมศึกษาปีที่ 3 จำนวน 469 ข้อ และชั้นมัธยมศึกษาปีที่ 6 จำนวน 1,197 ข้อ โดยข้อสอบในคลังข้อสอบ O-NET อยู่ในระดับค่อนข้างยาก พัฒนาโปรแกรมในรูปแบบของ Web application โดยใช้ภาษา PHP พัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ตามแนวคิดของ Thompson and Weiss (2011) ผู้ใช้สามารถเข้าใช้โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับการจัดสอบ O-NET ใน [website://www.onetcat.net](http://www.onetcat.net)

การประเมินความเหมาะสมในการนำไปใช้ของโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ในเดือนมิถุนายน พ.ศ. 2557 โดยนักเรียนระดับชั้นประถมศึกษาปีที่ 6 จำนวน 224 คน ชั้นมัธยมศึกษาปีที่ 3 จำนวน 432 คน และชั้นมัธยมศึกษาปีที่ 6 จำนวน 435 คน ปรากฏว่า การใช้งานของโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับการจัดสอบ O-NET อยู่ในเกณฑ์ดี เป็นที่พึงพอใจของนักเรียน

**คำสำคัญ:** การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์, คลังข้อสอบ O-NET, ทฤษฎีการตอบสนองข้อสอบ

\*Corresponding author. E-mail: piyatip10@hotmail.com

## Development of a computerized adaptive testing program for O-NET

Suchada Kornpetpanee<sup>1</sup>, Piyathip Tinnaworn<sup>1\*</sup>, Soros Sukhanonsawat<sup>2</sup>

<sup>1</sup> *Centre for Applied Measurement  
College of Research Methodology and Cognitive Science  
Burapha University, Thailand*

<sup>2</sup> *Information Technology Standard and Security Office,  
Metropolitan Electricity Authority, Thailand*

### Abstract

The goal of this research was to develop a Computerized Adaptive Testing (CAT) system to deliver the Ordinary National Education Test (O-NET) examination items at three Grade levels: Six, Nine, and Twelve. Each Grade level was comprised of eight learning substance groups. The item bank was developed by calibrating O-NET items, used in years 2008-2010, with the 3-PL item response model, a process which resulted in 258 items for Grade Six, 469 for Grade Nine, and 1,197 for Grade Twelve. These results indicated that the items in the O-NET item bank were in a difficult level. The PHP-based CAT system was developed based on the framework published by Thompson and Weiss (2011), and was delivered over a ONETCAT.NET server.

The CAT system was evaluated in June, 2014, with samples that included 224 Grade Six students, 432 Grade Nine students, and 435 Grade Twelve students. The CAT system was found to perform well, and a feedback from students was overwhelmingly positive.

**Keywords:** Computerized adaptive testing, O-NET item bank, Item response theory

---

\*Corresponding author. E-mail: piyatip10@hotmail.com

## ความนำ

สถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) หรือ สทศ. เป็นหน่วยงานทดสอบระดับชาติ อยู่ในกำกับของกระทรวงศึกษาธิการ จัดตั้งขึ้นตามพระราชกฤษฎีกาจัดตั้งสถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) เมื่อวันที่ 3 กันยายน พ.ศ. 2548 โดยสถาบันมีวัตถุประสงค์เพื่อบริหารจัดการและดำเนินการเกี่ยวกับการศึกษา วิจัย พัฒนาและให้บริการทางการประเมินผลการศึกษาและทดสอบทางการศึกษาอย่างต่อเนื่อง รวมทั้งเป็นศูนย์กลางความร่วมมือด้านการทดสอบทางการศึกษาในระดับชาติและระดับนานาชาติ

สถาบันทดสอบทางการศึกษาแห่งชาติ ทำหน้าที่ในการทดสอบทางการศึกษาในระดับต่าง ๆ เช่น การจัดสอบ O-NET (Ordinary National Educational Testing: O-NET) ในระดับชั้นประถมศึกษาปีที่ 6 ระดับชั้นมัธยมศึกษาปีที่ 3 และระดับชั้นมัธยมศึกษาปีที่ 6 และการจัดการทดสอบอื่นๆ เช่น การจัดสอบ I-NET, N-NET, V-NET, U-NET, GAT/PAT และการจัดสอบเพื่อวัดระดับความรู้ของครู ในการดำเนินงานที่ผ่านมาส่วนใหญ่ใช้การทดสอบแบบดั้งเดิม คือ การทดสอบด้วยกระดาษ - ดินสอ (Paper – pencil testing) ซึ่งปรากฏว่ามีปัญหาหลายประการที่ควรได้รับการแก้ไข สถาบันฯ จึงได้ดำเนินการพัฒนาปรับปรุงรูปแบบการทดสอบจากการทดสอบแบบดั้งเดิมเป็นการทดสอบแบบออนไลน์ (Online testing) อย่างไรก็ตาม การทดสอบแบบออนไลน์ที่กำลังอยู่ในระหว่างการดำเนินการนี้ ยังคงใช้หลักการทดสอบเช่นเดียวกับการทดสอบแบบดั้งเดิม คือ ผู้เข้าสอบทุกคนทำแบบทดสอบชุดเดียวกันและทดสอบในเวลาเดียวกัน ทำให้ยังคงมีปัญหาที่เกี่ยวข้องหลายประการ เช่น ความยุ่งยากในการจัดสอบ ความคลาดเคลื่อนของคะแนนสอบ และปัญหาการทุจริตในการทดสอบ

วงการศึกษานำเทคโนโลยีทางคอมพิวเตอร์มาใช้ในการทดสอบ เพื่อช่วยเพิ่มประสิทธิภาพในการตรวจข้อสอบ การวิเคราะห์ข้อสอบ และการรายงานผลการทดสอบ รวมทั้งการนำมาใช้สร้างคลังข้อสอบ (Item Bank) เพื่อให้สามารถผลิตข้อสอบได้จำนวนมากเก็บไว้ในคลังข้อสอบ ในเวลาต่อมาได้นำคอมพิวเตอร์มาช่วยในการดำเนินการสอบ ใช้พิมพ์แบบทดสอบเก็บไว้ในหน่วยความจำและเขียนโปรแกรมให้สามารถแสดงข้อสอบบนจอภาพสามารถทำการทดสอบบนเครื่องคอมพิวเตอร์ (Computer Based Testing: CBT) แทนการทดสอบด้วยกระดาษ - คำตอบ พร้อมทั้งสามารถตรวจให้คะแนนและแปลผลการทดสอบได้อย่างอัตโนมัติ (Bunderson, Inouye, & Olsen, 1989)

การทดสอบโดยใช้คอมพิวเตอร์ (Computer Based Testing: CBT) เป็นการนำคอมพิวเตอร์มาใช้ในการบริหารการทดสอบ แบ่งออกเป็น 4 ยุค คือ 1) การทดสอบด้วยคอมพิวเตอร์ (Computerized Testing: CT) เป็นยุคที่นำคอมพิวเตอร์มาใช้ในการทดสอบแทนการใช้กระดาษคำตอบ โดยผู้เข้าสอบทำแบบทดสอบเดียวกัน ใช้เวลาเท่ากันและสอบในสถานที่เดียวกัน 2) การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ (Computerized Adaptive Testing: CAT) เป็นยุคที่นำคอมพิวเตอร์มาใช้ในการทดสอบ โดยผู้เข้าสอบทำแบบทดสอบที่เหมาะสมกับระดับความสามารถของผู้เข้าสอบ ซึ่งแต่ละคนอาจจะทำข้อสอบไม่เหมือนกันและจำนวนข้อไม่เท่ากัน 3) การวัดต่อเนื่องกัน (Continuous Measurement: CM) เป็นยุคที่มีการนำการเรียนการสอนและการทดสอบมาใช้ในการระบบคอมพิวเตอร์เดียวกัน ซึ่งสลับกันไปมาตลอดบทเรียน กล่าวคือ เมื่อเรียนเรื่องใดจบก็ทำการทดสอบเรื่องนั้นทันที และ 4) การวัดอย่างเฉลียวฉลาด (Intelligent Measurement: IM) เป็นยุคที่นำคอมพิวเตอร์มาใช้ในการบริหารการทดสอบครบทุกขั้นตอน กล่าวคือ เป็นการนำคอมพิวเตอร์มาตรวจให้คะแนนแบบทดสอบประเภทเขียนตอบที่มีคำตอบที่แน่นอน และสามารถวิเคราะห์จุดแข็งและจุดอ่อน รวมทั้งให้ข้อเสนอแนะกับผู้เรียนหรือผู้เข้าสอบได้ (Bunderson, Inouye, & Olsen, 1989)

การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ (Computerized Adaptive Testing: CAT) เป็นการทดสอบที่จัดข้อสอบให้เหมาะสมกับความสามารถของผู้เข้าสอบ ซึ่งอยู่บนพื้นฐานของการตอบข้อสอบข้อแรกหรือข้อที่ผ่านมาของผู้เข้าสอบ กล่าวคือ เมื่อผู้เข้าสอบทำข้อสอบข้อเริ่มต้นหรือชุดเริ่มต้นเรียบร้อยแล้ว ก็จะนำผลการตอบข้อสอบมาวิเคราะห์หรือประเมินระดับความสามารถของผู้เข้าสอบ เพื่อที่จะคัดเลือกข้อสอบข้อถัดไปที่เหมาะสม โดยอาศัยทฤษฎีการตอบสนองข้อสอบ (Item Response Theory: IRT) เป็นพื้นฐาน และจะสิ้นสุดการทดสอบเมื่อผู้เข้าสอบทำการทดสอบตามเงื่อนไขหรือเกณฑ์ที่กำหนด สำหรับขั้นตอนของการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ประกอบด้วย 5 ขั้นตอน ดังนี้ ขั้นตอนที่ 1 การสร้างคลังข้อสอบ ขั้นตอนนี้เป็นการคัดเลือกข้อสอบที่ผ่านการวิเคราะห์โดยใช้ทฤษฎีการตอบสนองข้อสอบ แบบ 3 พารามิเตอร์ ประกอบด้วย ค่าอำนาจจำแนกของข้อสอบ ( $d$ ) ค่าความยากของข้อสอบ ( $b$ ) และค่าการเดาของข้อสอบ ( $c$ ) ขั้นตอนที่ 2 การคัดเลือกข้อสอบข้อเริ่มต้น ขั้นตอนนี้มักจะคัดเลือกข้อสอบที่มีค่าความยากของข้อสอบ ( $b$ ) อยู่ในระดับปานกลาง ขั้นตอนที่ 3 การคัดเลือกข้อสอบข้อถัดไป เป็นขั้นตอนที่ต้องคัดเลือกข้อสอบให้ใกล้เคียงกับความสามารถของผู้เข้าสอบในขณะนั้น โดยพิจารณาจากผลการตอบข้อสอบข้อก่อนหน้านี้ ขั้นตอนที่ 4 การประมาณค่าความสามารถของผู้เข้าสอบ เป็นขั้นตอนที่จะประมาณค่าความสามารถของผู้เข้าสอบ หลังจากที่ได้ตอบข้อสอบข้อนั้น ๆ แล้ว และขั้นตอนที่ 5 การยุติการทดสอบ ขั้นตอนนี้เป็นการสิ้นสุดการทดสอบ เมื่อผู้เข้าสอบทำข้อสอบได้ครบตรงตามเกณฑ์ที่กำหนด (Thompson, & Weiss, 2011)

สถาบันทดสอบทางการศึกษาแห่งชาติ ได้เล็งเห็นความสำคัญของการใช้เทคโนโลยีกับการวัดผลการศึกษา จึงต้องการพัฒนาระบบการจัดการการทดสอบต่างๆ ให้มีประสิทธิภาพยิ่งขึ้น จึงมอบหมายให้วิทยาลัยวิทยาการวิจัยและวิทยาการปัญญา มหาวิทยาลัยบูรพา ศึกษาการนำเทคโนโลยีทางด้านคอมพิวเตอร์มาประยุกต์กับการทดสอบที่เรียกว่า การทดสอบแบบปรับเหมาะตามระดับความสามารถของผู้เข้าสอบ (Adaptive testing) ด้วยคอมพิวเตอร์ ซึ่งเป็นการทดสอบที่ผู้เข้าสอบแต่ละคนใช้แบบทดสอบแตกต่างกัน โดยมีการคัดเลือกข้อสอบที่มีความเหมาะสมกับระดับความสามารถของผู้เข้าสอบ (Weiss, 1974) เพื่อประมาณค่าความสามารถของผู้เข้าสอบได้อย่างแม่นยำและมีความคลาดเคลื่อนต่ำ จึงยุติการทดสอบ ทำให้จำนวนข้อสอบที่ใช้ในการทดสอบน้อยลง ประหยัดเวลาในการทดสอบ นอกจากนั้นยังช่วยผ่อนคลายความเครียดของผู้เข้าสอบ เช่น ผู้เข้าสอบที่มีความสามารถต่ำ จะไม่รู้สึกเครียดกับการทำข้อสอบที่ยากเกินระดับความสามารถของตนเอง ส่วนผู้เข้าสอบที่มีความสามารถสูง ไม่ต้องเสียเวลาในการทำข้อสอบที่ง่ายเกินไป (Sakolkijrungrroj, Chadcham, & Sudhasani, 2015) ในโครงการวิจัยเรื่อง การพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับการจัดสอบ O-NET ในระดับชั้นประถมศึกษาปีที่ 6 ระดับชั้นมัธยมศึกษาปีที่ 3 และระดับชั้นมัธยมศึกษาปีที่ 6 เพื่อช่วยให้ผู้เข้าสอบมีความรู้ ความเข้าใจเกี่ยวกับวิธีการทดสอบแบบปรับเหมาะตามระดับความสามารถของผู้สอบด้วยคอมพิวเตอร์ และได้มีโอกาสฝึกฝนการทำข้อสอบเก่าก่อนที่สถาบันทดสอบทางการศึกษาแห่งชาติ จะได้พิจารณานำวิธีการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ มาใช้ในการจัดสอบต่อไป

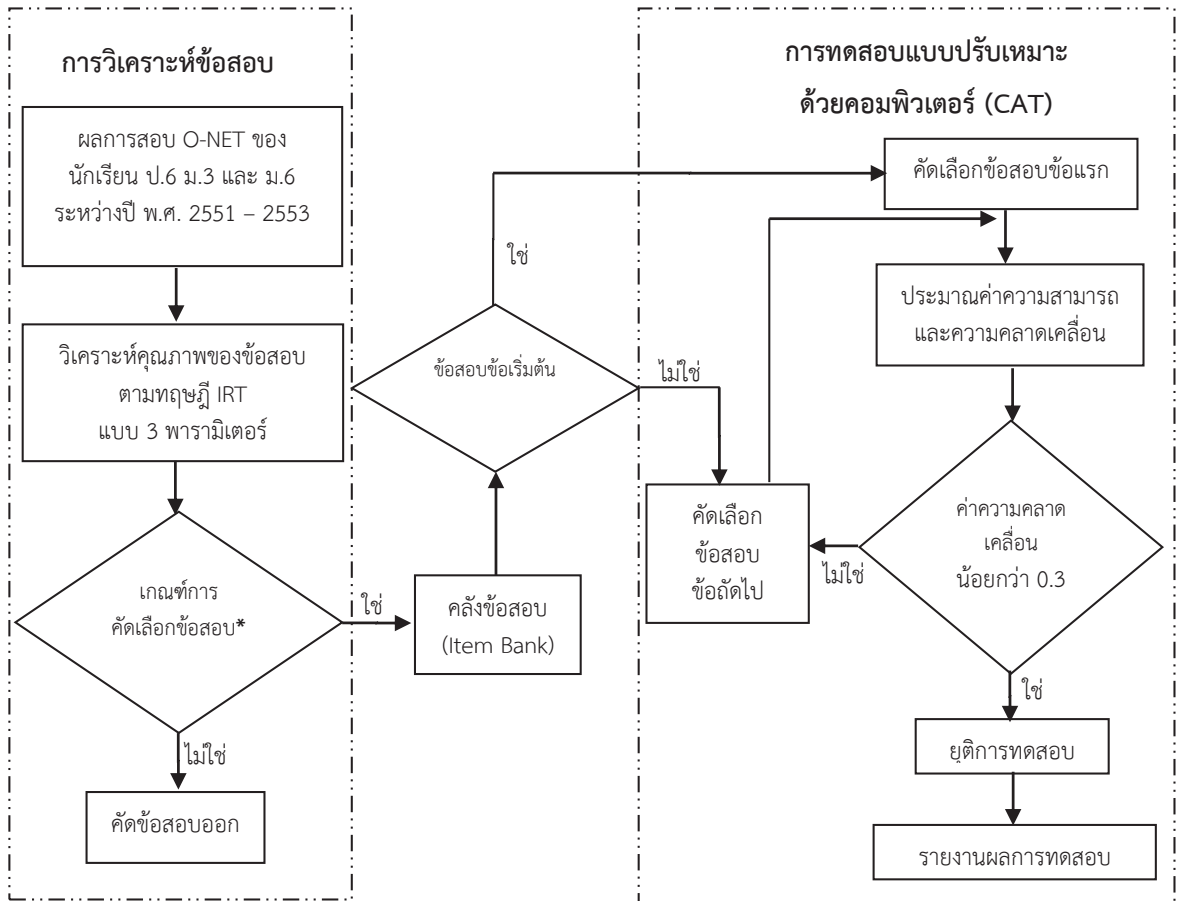
## วัตถุประสงค์ของการวิจัย

1. เพื่อจัดทำคลังข้อสอบ O-NET ระดับชั้นประถมศึกษาปีที่ 6 ระดับชั้นมัธยมศึกษาปีที่ 3 และระดับชั้นมัธยมศึกษาปีที่ 6 ระดับชั้นละ 8 กลุ่มสาระการเรียนรู้ ที่ใช้สำหรับการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์
2. เพื่อพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับการจัดสอบ O-NET ระดับชั้นประถมศึกษาปีที่ 6 ระดับชั้นมัธยมศึกษาปีที่ 3 และระดับชั้นมัธยมศึกษาปีที่ 6 ระดับชั้นละ 8 กลุ่มสาระการเรียนรู้

## กรอบแนวคิดการวิจัย

การพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับการจัดสอบ O-NET ระดับชั้นประถมศึกษาปีที่ 6 ระดับชั้นมัธยมศึกษาปีที่ 3 และระดับชั้นมัธยมศึกษาปีที่ 6 ระดับชั้นละ 8 กลุ่มสาระการเรียนรู้ ข้อสอบที่นำมาใช้ในงานวิจัยระดับชั้นละ 8 กลุ่มสาระการเรียนรู้ ได้แก่ 1) วิชาภาษาไทย 2) วิชาคณิตศาสตร์ 3) วิทยาศาสตร์ 4) วิชาสังคมศึกษา ศาสนาและวัฒนธรรม 5) วิชาภาษาอังกฤษ 6) วิชาสุขศึกษาและพลศึกษา 7) วิชาศิลปะ และ 8) วิชาการงานพื้นฐานอาชีพและเทคโนโลยี ซึ่งผู้วิจัยได้นำมาวิเคราะห์หาคุณภาพของข้อสอบโดยใช้ทฤษฎีการตอบสนองข้อสอบ (IRT) แบบ 3 พารามิเตอร์ ประกอบด้วย ค่าความยากของข้อสอบ ( $b$ ) ค่าอำนาจจำแนกของข้อสอบ ( $a$ ) และค่าการเดาของข้อสอบ ( $c$ )

ผู้วิจัยได้พัฒนาโปรแกรมการทดสอบตามแนวคิดของ Thompson and Weiss (2011) ซึ่งได้กำหนดขั้นตอนของการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ (CAT) ดังภาพที่ 1



หมายเหตุ: หมายถึง ค่าอำนาจจำแนกของข้อสอบ (a) มีค่าตั้งแต่ 0.50 ถึง 2.50  
 ค่าความยากของข้อสอบ (b) มีค่าตั้งแต่ -2.50 ถึง 2.50  
 ค่าการเดาของข้อสอบ (c) มีค่าไม่เกิน 0.30

ภาพที่ 1 กรอบแนวคิดการพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับการจัดสอบ O-NET

### วิธีดำเนินการวิจัย

การจัดทำคลังข้อสอบ O-NET และพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับการจัดสอบ O-NET ระดับชั้นประถมศึกษาปีที่ 6 ระดับชั้นมัธยมศึกษาปีที่ 3 และระดับชั้นมัธยมศึกษาปีที่ 6 ระดับชั้นละ 8 กลุ่มสาระการเรียนรู้ จำนวน 24 กลุ่มสาระการเรียนรู้ ระหว่างปี พ.ศ. 2551 - 2553 มีขั้นตอนการดำเนินการ ดังนี้

## ตอนที่ 1 การจัดทำคลังข้อสอบ O-NET

1. การรวบรวมข้อสอบ O-NET ผู้วิจัยรวบรวมข้อสอบ O-NET ระหว่างปี พ.ศ. 2551-2553 ระดับชั้นประถมศึกษาปีที่ 6 ระดับชั้นมัธยมศึกษาปีที่ 3 และระดับชั้นมัธยมศึกษาปีที่ 6 ระดับชั้นละ 8 กลุ่มสาระการเรียนรู้ ประกอบด้วย 1) ภาษาไทย 2) คณิตศาสตร์ 3) วิทยาศาสตร์ 4) ภาษาต่างประเทศ 5) ศิลปะ 6) สุขศึกษาและพลศึกษา 7) สังคมศึกษา ศาสนาและวัฒนธรรม และ 8) การงานอาชีพและเทคโนโลยี ที่ได้จากสถาบันทดสอบทางการศึกษาแห่งชาติ

2. การวิเคราะห์คุณภาพของข้อสอบโดยใช้ทฤษฎีการตอบสนองข้อสอบ มีขั้นตอน ดังนี้

2.1 ตรวจสอบความสมบูรณ์ของข้อสอบ O-NET ทั้งข้อคำถาม ตัวเลือก และเฉลย ถ้าข้อสอบข้อใดไม่มีความสมบูรณ์ จะคัดข้อสอบข้อนั้นออก

2.2 จัดเตรียมข้อมูลจากผลการตอบข้อสอบ พร้อมกำหนดเงื่อนไขของการวิเคราะห์คุณภาพของข้อสอบตามหลักการของทฤษฎีการตอบสนองข้อสอบ แบบ 3 พารามิเตอร์

2.3 วิเคราะห์คุณภาพของข้อสอบตามหลักการของทฤษฎีการตอบสนองข้อสอบ โดยใช้โปรแกรมสำเร็จรูป Xcalibre version 4.1.7

3. การคัดเลือกข้อสอบ คัดเลือกข้อสอบจากผลการวิเคราะห์คุณภาพของข้อสอบ โดยคัดเลือกข้อสอบที่ตรงตามเงื่อนไขที่กำหนดไว้ ดังนี้

เกณฑ์การคัดเลือกข้อสอบ (Urry, 1977)

ก. ค่าอำนาจจำแนกของข้อสอบ ( $a$ ) มีค่าตั้งแต่ 0.50 ถึง 2.50

ข. ค่าความยากของข้อสอบ ( $b$ ) มีค่าตั้งแต่ -2.50 ถึง 2.50

ค. ค่าการเดาของข้อสอบ ( $c$ ) มีค่าไม่เกิน 0.30

4. การจัดทำคลังข้อสอบ O-NET ได้นำข้อคำถาม ตัวเลือก เฉลย ค่าความยากของข้อสอบ ( $b$ ) ค่าอำนาจจำแนกของข้อสอบ ( $a$ ) และค่าการเดาของข้อสอบ ( $c$ ) บันทึกลงในคลังข้อสอบที่พัฒนาขึ้นโดยใช้ฐานข้อมูล MySQL ซึ่งเป็นฐานข้อมูลที่ไม่เสียค่าใช้จ่าย

## ตอนที่ 2 การพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับการจัดสอบ O-NET

1. การวิเคราะห์และออกแบบโปรแกรมการทดสอบ มีขั้นตอนการดำเนินการ ดังนี้

1.1 วิเคราะห์แผนผังบริบท (Context diagram) เป็นการวิเคราะห์เพื่อหา Source destination ที่เกี่ยวข้องกับโปรแกรมการทดสอบ

1.2 จัดทำผังการไหลของข้อมูล (Data Flow Diagram: DFD) เป็นการวิเคราะห์ให้เห็นภาพรวมของโปรแกรมทั้งข้อมูลและการทำงาน โดยจัดทำในทุกขั้นตอนของกระบวนการของโปรแกรมการทดสอบ ซึ่งทำให้ง่ายต่อการศึกษาของผู้ที่สนใจ

2. การออกแบบฐานข้อมูลโปรแกรมการทดสอบ มี 10 แฟ้มข้อมูล คือ 1) แฟ้มข้อมูลข้อสอบ 2) แฟ้มข้อมูลตัวเลือก 3) แฟ้มข้อมูลระดับชั้นการศึกษา 4) แฟ้มข้อมูลกลุ่มสาระการเรียนรู้ 5) แฟ้มข้อมูลจังหวัด 6) แฟ้มข้อมูลผู้เข้าสอบ 7) แฟ้มข้อมูลวันที่สอบ 8) แฟ้มข้อมูลเวลาที่ใช้สอบ 9) แฟ้มข้อมูลความสามารถของผู้เข้าสอบ และ 10) แฟ้มข้อมูลค่าสถิติ

3. การพัฒนาโปรแกรมการทดสอบ ใช้ภาษา SQL เป็นคำสั่งเชื่อมโยงข้อมูลในฐานข้อมูล พัฒนาโปรแกรมคอมพิวเตอร์ตามหลักการของวงจรการพัฒนาซอฟต์แวร์ (Software Development Life Circle: SDLC)

(Elliott, 2004) และขั้นตอนการดำเนินการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ของ Thompson and Weiss (2011) ที่ใช้กับทฤษฎีการตอบสนองข้อสอบ แบบ 3 พารามิเตอร์

4. ทดสอบและแก้ไขข้อผิดพลาดของโปรแกรมการทดสอบ เป็นขั้นตอนการตรวจสอบหาจุดบกพร่องเบื้องต้นของโปรแกรม เพื่อนำไปปรับปรุงแก้ไขต่อไป โดยนำไปทดสอบกับอาจารย์และนิสิต มหาวิทยาลัยบูรพา จังหวัดชลบุรี

5. จัดทำคู่มือการใช้โปรแกรมการทดสอบ เป็นการจัดทำคู่มือการใช้โปรแกรมเพื่อเป็นแนวทางการนำไปใช้จริง

6. การประเมินความเหมาะสมของโปรแกรมการทดสอบ เป็นการนำโปรแกรมและคู่มือการใช้งานไปให้ผู้เชี่ยวชาญประเมินความเหมาะสมของโปรแกรม โดยผู้เชี่ยวชาญที่มีความรู้ทางด้านการวัดผลการศึกษาและด้านการพัฒนาโปรแกรมคอมพิวเตอร์ มีวุฒิการศึกษาไม่ต่ำกว่าระดับปริญญาโท และมีประสบการณ์ไม่น้อยกว่า 10 ปี

### ตอนที่ 3 การประเมินความเหมาะสมในการนำไปใช้ของโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์

การประเมินความเหมาะสมในการนำไปใช้เป็นการประเมินผลความพึงพอใจของนักเรียนที่ทดลองใช้โปรแกรมการทดสอบในด้านความสะดวกในการนำไปใช้และด้านลักษณะทั่วไปของโปรแกรม

#### ประชากรและกลุ่มตัวอย่าง

ประชากรที่ใช้ในการวิจัยเป็นนักเรียนที่ศึกษาในภาคเรียนที่ 1 ปีการศึกษา 2557 จำนวน 6,281 คน แบ่งเป็น 3 ระดับชั้น ดังนี้ นักเรียนระดับชั้นประถมศึกษาปีที่ 6 จำนวน 485 คน นักเรียนระดับชั้นมัธยมศึกษาปีที่ 3 จำนวน 3,648 คน และนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6 จำนวน 2,148 คน

กลุ่มตัวอย่างได้มาโดยวิธีการสุ่มอย่างง่าย จำนวน 1,091 คน แบ่งเป็น 3 ระดับชั้น ดังนี้

1) นักเรียนระดับชั้นประถมศึกษาปีที่ 6 จำนวน 3 โรงเรียน ประกอบด้วย 1) โรงเรียนวัดปทุมวนาราม ในพระราชูปถัมภ์สมเด็จพระเทพรัตนราชสุดาฯ สยามบรมราชกุมารี จำนวน 74 คน 2) โรงเรียนราชวินิต จำนวน 100 คน และ 3) โรงเรียนบ้านอ้อไพล จำนวน 50 คน จำนวนทั้งสิ้น 224 คน

2) นักเรียนระดับชั้นมัธยมศึกษาปีที่ 3 จำนวน 6 โรงเรียน ประกอบด้วย 1) โรงเรียนวัดปทุมวนารามในพระราชูปถัมภ์สมเด็จพระเทพรัตนราชสุดาฯ สยามบรมราชกุมารี จำนวน 65 คน 2) โรงเรียนสวนกุหลาบวิทยาลัย จำนวน 68 คน 3) โรงเรียนทวีธาภิเศก จำนวน 71 คน 4) โรงเรียนศึกษานารี จำนวน 100 คน 5) โรงเรียนมัธยมวัดดุสิตาราม จำนวน 78 คน และ 6) โรงเรียนมทิสราธิบดี จำนวน 50 คน จำนวนทั้งสิ้น 432 คน

3) นักเรียนระดับชั้นมัธยมศึกษาปีที่ 6 จำนวน 5 โรงเรียน ประกอบด้วย 1) โรงเรียนสวนกุหลาบวิทยาลัย จำนวน 96 คน 2) โรงเรียนทวีธาภิเศก จำนวน 86 คน 3) โรงเรียนศึกษานารี จำนวน 103 คน 4) โรงเรียนมัธยมวัดดุสิตาราม จำนวน 100 คน และ 5) โรงเรียนมทิสราธิบดี จำนวน 50 คน จำนวนทั้งสิ้น 435 คน

#### การเก็บรวบรวมข้อมูล

1. จัดเตรียมเอกสารชี้แจงการใช้งานโปรแกรมการทดสอบ ที่พัฒนาขึ้น พร้อมแบบประเมินความพึงพอใจในการใช้งานของโปรแกรมที่ผ่านการพิจารณาจากผู้เชี่ยวชาญแล้ว ซึ่งมีลักษณะเป็นมาตราประมาณค่า (Rating scale) 5 ระดับ

2. เก็บข้อมูลจากนักเรียนระดับชั้นประถมศึกษาปีที่ 6 ระดับชั้นมัธยมศึกษาปีที่ 3 และระดับชั้นมัธยมศึกษาปีที่ 6 ภาคเรียนที่ 1 ปีการศึกษา 2557



3. อธิบายหลักการของการทดสอบด้วยโปรแกรมการทดสอบกับนักเรียนกลุ่มตัวอย่างอย่างละเอียด พร้อมกับแจกวิธีการดำเนินการทดสอบให้กับนักเรียนกลุ่มตัวอย่าง

4. ดำเนินการทดสอบด้วยโปรแกรมการทดสอบกับนักเรียนทั้ง 3 กลุ่มข้างต้น ซึ่งจัดการทดสอบผ่านทาง Website (www.onetcat.net)

หลังจากเสร็จสิ้นการทดสอบตามวันที่กำหนด ได้นำผลการทดสอบของนักเรียนที่ทดลองใช้โปรแกรม มาประเมินความเหมาะสมในการใช้โปรแกรม โดยวิเคราะห์ความพึงพอใจของนักเรียนที่ทดลองใช้โปรแกรมในด้านความสะดวกในนำไปใช้และด้านลักษณะทั่วไปของโปรแกรม

#### การวิเคราะห์ข้อมูล

การวิเคราะห์คะแนนความพึงพอใจของนักเรียนที่ทดลองใช้โปรแกรมการทดสอบ ได้แก่ ค่าเฉลี่ย และ ส่วนเบี่ยงเบนมาตรฐาน

### ผลการวิจัย

การจัดทำคลังข้อสอบ O-NET ในการพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ และการประเมินความเหมาะสมในการใช้โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ มีดังนี้

#### ตอนที่ 1 ผลการจัดทำคลังข้อสอบ O-NET

ผลการวิเคราะห์คุณภาพข้อสอบ O-NET สำหรับคลังข้อสอบ แสดงดังตารางที่ 1 - 3

ตารางที่ 1 ผลการวิเคราะห์คุณภาพข้อสอบ O-NET ระดับชั้นประถมศึกษาปีที่ 6 ตั้งแต่ปี พ.ศ. 2551-2553

ลำดับ	กลุ่มสาระการเรียนรู้	จำนวน ข้อสอบ (ผ่านเกณฑ์)	ค่าความ ยากเฉลี่ย	ค่าอำนาจ จำแนกเฉลี่ย	ค่าการ เดาเฉลี่ย	ความหมาย
1	ภาษาไทย	62	1.0253	1.0235	0.2048	ยาก
2	คณิตศาสตร์	51	0.5932	1.2292	0.2047	ค่อนข้างยาก
3	วิทยาศาสตร์	74	0.7127	0.7622	0.2067	ค่อนข้างยาก
4	ภาษาต่างประเทศ	19	1.6347	1.1798	0.2215	ยาก
5	ศิลปะ	5	0.7384	1.3386	0.2478	ค่อนข้างยาก
6	สุขศึกษา และพลศึกษา	4	0.9653	1.2618	0.1895	ค่อนข้างยาก
7	สังคมศึกษา ศาสนาและวัฒนธรรม	39	0.5913	0.7472	0.2006	ค่อนข้างยาก
8	การงานอาชีพ และเทคโนโลยี	4	0.3278	1.5150	0.1775	ปานกลาง
	รวม	258	0.7998	0.9764	0.2061	ค่อนข้างยาก

จากตารางที่ 1 แสดงผลการวิเคราะห์คุณภาพข้อสอบ O-NET ระดับชั้นประถมศึกษาปีที่ 6 มีข้อสอบผ่านเกณฑ์ที่กำหนด จำนวน 258 ข้อ มีค่าความยากของข้อสอบเฉลี่ยเท่ากับ 0.7998 ค่าอำนาจจำแนกของข้อสอบเฉลี่ยเท่ากับ 0.9764 และค่าการเดาของข้อสอบเฉลี่ยเท่ากับ 0.2061 แสดงให้เห็นว่า ข้อสอบ O-NET ระดับชั้น

ประถมศึกษาปีที่ 6 มีค่าความยากของข้อสอบเฉลี่ยค่อนข้างยาก (Muninsakorn, Tinnaworn, & Sukhanonsawat, 2015)

ตารางที่ 2 ผลการวิเคราะห์คุณภาพข้อสอบ O-NET ระดับชั้นมัธยมศึกษาปีที่ 3 ตั้งแต่ปี พ.ศ. 2551-2553

ลำดับ	กลุ่มสาระการเรียนรู้	จำนวน ข้อสอบ (ผ่านเกณฑ์)	ค่าความ ยากเฉลี่ย	ค่าอำนาจ จำแนกเฉลี่ย	ค่าการเดา เฉลี่ย	ความหมาย
1	ภาษาไทย	95	0.9149	0.9583	0.2078	ค่อนข้างยาก
2	คณิตศาสตร์	61	1.3047	1.8749	0.2068	ยาก
3	วิทยาศาสตร์	78	1.0990	1.4075	0.2189	ยาก
4	ภาษาต่างประเทศ	58	1.1436	1.9238	0.2089	ยาก
5	ศิลปะ	25	1.9253	1.4490	0.0959	ยาก
6	สุขศึกษา และพลศึกษา	41	0.2307	0.9308	0.2110	ปานกลาง
7	สังคมศึกษา ศาสนาและวัฒนธรรม	85	0.7742	0.9145	0.2193	ค่อนข้างยาก
8	การงานอาชีพ และเทคโนโลยี	26	0.9588	0.8464	0.1974	ค่อนข้างยาก
	รวม	469	0.9958	1.2826	0.1974	ค่อนข้างยาก

จากตารางที่ 2 แสดงผลการวิเคราะห์คุณภาพข้อสอบ O-NET ระดับชั้นมัธยมศึกษาปีที่ 3 มีข้อสอบผ่านเกณฑ์ที่กำหนด จำนวน 469 ข้อ มีค่าความยากของข้อสอบเฉลี่ยเท่ากับ 0.9958 ค่าอำนาจจำแนกของข้อสอบเฉลี่ยเท่ากับ 1.2826 และค่าการเดาของข้อสอบเฉลี่ยเท่ากับ 0.1974 แสดงให้เห็นว่า ข้อสอบ O-NET ระดับชั้นมัธยมศึกษาปีที่ 3 มีค่าความยากของข้อสอบเฉลี่ยค่อนข้างยาก (จารุจิตร สิทธิประยูร, ปิยะทิพย์ ดินวร และโสฬส สุขานนท์สวัสดิ์, 2559)

**ตารางที่ 3 ผลการวิเคราะห์คุณภาพข้อสอบ O-NET ระดับชั้นมัธยมศึกษาปีที่ 6 ตั้งแต่ปี พ.ศ. 2551-2553**

ลำดับ	กลุ่มสาระการเรียนรู้	จำนวน ข้อสอบ (ผ่านเกณฑ์)	ค่าความ ยากเฉลี่ย	ค่าอำนาจ จำแนกเฉลี่ย	ค่าการ เดาเฉลี่ย	ความหมาย
1	ภาษาไทย	265	0.5161	1.1383	0.2051	ค่อนข้างยาก
2	คณิตศาสตร์	92	1.0478	2.2979	0.2155	ยาก
3	วิทยาศาสตร์	218	1.1369	1.7023	0.2055	ยาก
4	ภาษาต่างประเทศ	166	1.0967	2.0389	0.2213	ยาก
5	ศิลปะ	79	0.7377	0.7875	0.1968	ค่อนข้างยาก
6	สุขศึกษา และพลศึกษา	109	0.2136	0.7953	0.1897	ปานกลาง
7	สังคมศึกษา ศาสนาและวัฒนธรรม	181	1.1569	1.0732	0.1955	ยาก
8	การงานอาชีพ และเทคโนโลยี	87	0.8990	0.8422	0.1727	ค่อนข้างยาก
รวม		1,197	0.8624	1.3693	0.2024	ค่อนข้างยาก

จากตารางที่ 3 แสดงผลการวิเคราะห์คุณภาพข้อสอบ O-NET ระดับชั้นมัธยมศึกษาปีที่ 6 มีข้อสอบผ่านเกณฑ์ที่กำหนด จำนวน 1,197 ข้อ มีค่าความยากของข้อสอบเฉลี่ยเท่ากับ 0.8624 ค่าอำนาจจำแนกของข้อสอบเฉลี่ยเท่ากับ 1.3693 และค่าการเดาของข้อสอบเฉลี่ยเท่ากับ 0.2024 แสดงให้เห็นว่า ข้อสอบ O-NET ระดับชั้นมัธยมศึกษาปีที่ 3 มีค่าความยากของข้อสอบเฉลี่ยค่อนข้างยาก (Pluemjai, Tinnaworn, & Sukhanonsawat, 2015)

**ตอนที่ 2 ผลการพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์**

การพัฒนาโปรแกรมการทดสอบในรูปแบบ Web application (<http://www.onetcat.net>) ซึ่งเป็นโปรแกรมทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน 8 กลุ่มสาระการเรียนรู้ ได้แก่ 1) ภาษาไทย 2) คณิตศาสตร์ 3) วิทยาศาสตร์ 4) ภาษาต่างประเทศ 5) ศิลปะ 6) สุขศึกษาและพลศึกษา 7) สังคมศึกษา ศาสนาและวัฒนธรรม และ 8) การงานอาชีพ และเทคโนโลยี

ผลการพัฒนาโปรแกรมการทดสอบ มีดังนี้

1. รูปแบบโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์
  - 1.1 หน้าจอแรกของโปรแกรมการทดสอบ แสดงดังภาพที่ 2

## การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการทดสอบทางการศึกษาระดับชาตินิยม Computerized Adaptive Testing (CAT) for Ordinary National Education Test (O-NET)

หน้าแรก	การลงทะเบียนสอบ	การทดสอบ O-NET	ผลการทดสอบ	สำหรับเจ้าหน้าที่	คู่มือการใช้งาน
---------	-----------------	----------------	------------	-------------------	-----------------

### การทดสอบ O-NET



การทดสอบความรู้ในกลุ่มสาระการเรียนรู้ 8 กลุ่มสาระการเรียนรู้ ได้แก่ 1) ภาษาไทย 2) คณิตศาสตร์ 3) วิทยาศาสตร์ 4) ภาษาต่างประเทศ 5) ศิลปะ 6) สุขศึกษาและพลศึกษา 7) สังคมศึกษา ศาสนาและวัฒนธรรม และ 8) การงานอาชีพและเทคโนโลยี

### การลงทะเบียนสอบ



การลงทะเบียนเพื่อทดสอบความรู้ในกลุ่มสาระการเรียนรู้ 8 กลุ่มสาระการเรียนรู้ ได้แก่ 1) ภาษาไทย 2) คณิตศาสตร์ 3) วิทยาศาสตร์ 4) ภาษาอังกฤษ 5) ศิลปะ 6) สุขศึกษาและพลศึกษา 7) สังคมศึกษา ศาสนาและวัฒนธรรม และ 8) การงานอาชีพและเทคโนโลยี

### ผลการทดสอบ

ผลการทดสอบความรู้ในกลุ่มสาระการเรียนรู้ 8 กลุ่มสาระการเรียนรู้ ได้แก่ 1) ภาษาไทย 2) คณิตศาสตร์ 3) วิทยาศาสตร์ 4) ภาษาต่างประเทศ 5) ศิลปะ 6) สุขศึกษาและพลศึกษา 7) สังคมศึกษา ศาสนาและวัฒนธรรม และ 8) การงานอาชีพและเทคโนโลยี



### คู่มือการทดสอบ

คู่มือที่ไม่ทดสอบความรู้ในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการทดสอบทางการศึกษาระดับชาตินิยมพื้นฐาน Computerized Adaptive Testing (CAT) for Ordinary National Education Test (O-NET)



## ภาพที่ 2 หน้าจอแรกของโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับการจัดสอบ O-NET

จากภาพที่ 2 โปรแกรมการทดสอบ ประกอบด้วย เมนูหลัก และส่วนรายละเอียดของโปรแกรมการทดสอบ สำหรับส่วนของเมนูหลัก ประกอบด้วย 6 เมนู ดังนี้

- 1) หน้าแรก เป็นเมนูที่แสดงหน้าแรกของโปรแกรมการทดสอบ
- 2) การลงทะเบียนสอบ เป็นเมนูสำหรับผู้ที่จะเข้าสอบ O-NET ใน 8 กลุ่มสาระการเรียนรู้
- 3) การทดสอบ O-NET เป็นเมนูสำหรับการทดสอบ O-NET 8 กลุ่มสาระการเรียนรู้
- 4) ผลการทดสอบ เป็นเมนูที่แสดงส่วนตรวจสอบผลการทดสอบ O-NET 8 กลุ่มสาระการเรียนรู้
- 5) สำหรับเจ้าหน้าที่ เป็นเมนูส่วนบริหารการทดสอบ
- 6) ติดต่อเรา เป็นเมนูที่แสดงข้อมูล แผนที่ ที่อยู่ของสำนักทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) และข้อมูลของผู้วิจัย

1.2 รายละเอียดของโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์นั้น แบ่งออกเป็น 4 ส่วน คือ 1) การลงทะเบียนสอบ 2) การทดสอบ O-NET 3) รายงานผลการทดสอบ และ 4) การบริหารการทดสอบ มีรายละเอียด ดังภาพที่ 3



จากตารางที่ 4 แสดงผลการประเมินความพึงพอใจในภาพรวมของนักเรียนที่ทดลองใช้โปรแกรมด้านความสะดวกในการนำไปใช้ แสดงให้เห็นว่า นักเรียนที่ทดลองใช้โปรแกรมการทดสอบ มีความพึงพอใจด้านความสะดวกในการนำไปใช้ ในระดับดี (Mean = 3.84)

**ตารางที่ 5** ผลการประเมินความพึงพอใจในภาพรวมของนักเรียนที่ทดลองใช้โปรแกรม ด้านลักษณะทั่วไปของโปรแกรม

ลำดับที่	รายการ	Mean	SD	ความหมาย
1	การออกแบบโปรแกรมมีความน่าสนใจ	3.81	1.08	ดี
2	การจัดรูปแบบหน้าจอต่อการใช้งาน	3.89	1.05	ดี
3	มีความเร็วในการแสดงผลข้อมูล	3.73	1.10	ดี
4	การเรียกใช้งานโปรแกรมง่าย	3.86	1.04	ดี
5	โปรแกรมมีระบบป้องกันการดำเนินงานผิดพลาดของผู้ใช้ในทุกระดับ	3.66	1.14	ดี
6	สีสันทของโปรแกรมดึงดูดความสนใจ	3.46	1.21	ปานกลาง
	สรุปผลด้านลักษณะทั่วไปของโปรแกรม	3.74	1.11	ดี

หมายเหตุ: Mean คือ ค่าเฉลี่ย, SD คือ ส่วนเบี่ยงเบนมาตรฐาน

จากตารางที่ 5 แสดงผลการประเมินความพึงพอใจในภาพรวมของนักเรียนที่ทดลองใช้โปรแกรมด้านลักษณะทั่วไปของโปรแกรม แสดงให้เห็นว่า นักเรียนที่ทดลองใช้โปรแกรมการทดสอบ มีความพึงพอใจด้านลักษณะทั่วไปของโปรแกรม ในระดับดี (Mean = 3.74)

### อภิปรายผลการวิจัย

การจัดทำคลังข้อสอบ O-NET เป็นขั้นตอนหนึ่งที่มีความสำคัญ เนื่องจากคลังข้อสอบจะส่งผลต่อการจัดการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ดังนั้น ควรจัดทำคลังข้อสอบให้มีคุณภาพ ซึ่งข้อสอบที่นำมาบรรจุในคลังข้อสอบต้องผ่านการวิเคราะห์คุณภาพของข้อสอบตามหลักเกณฑ์และเงื่อนไขของทฤษฎีการตอบสนองข้อสอบ (Khomyoo, Chadcham, & Chinnasarn, 2013) ตามหลักการของทฤษฎีการตอบสนองข้อสอบนั้น มีวิธีการวิเคราะห์คุณภาพของข้อสอบหลายวิธี เช่น การวิเคราะห์คุณภาพของข้อสอบ ด้วยโมเดลการตอบสนองข้อสอบ แบบ 1 พารามิเตอร์ แบบ 2 พารามิเตอร์ และแบบ 3 พารามิเตอร์ สำหรับข้อสอบที่นำมาใช้กับคลังข้อสอบในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ควรมีความละเอียดมากที่สุด (Mejang, Wongdee, & Konta, 2010) ดังนั้น ผู้วิจัยจึงใช้วิธีการวิเคราะห์คุณภาพของข้อสอบ โดยใช้โมเดลของทฤษฎีการตอบสนองข้อสอบแบบ 3 พารามิเตอร์ เนื่องจากโมเดลของทฤษฎีการตอบสนองข้อสอบ แบบ 3 พารามิเตอร์ สามารถวิเคราะห์คุณภาพของข้อสอบได้ทั้งค่าอำนาจจำแนกของข้อสอบ (a) ค่าความยากของข้อสอบ (b) และค่าการเดาของข้อสอบ (c)

คลังข้อสอบที่นำมาใช้กับโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ต้องมีขนาดใหญ่พอสมควร โดยปกติต้องมีข้อสอบในคลังข้อสอบอย่างน้อย 100-200 ข้อ โดยข้อสอบต้องครอบคลุมคุณลักษณะ

และเนื้อหาที่ต้องการวัดความสามารถของผู้เข้าสอบ (Weiss, 1988) และข้อสอบควรมีค่าอำนาจจำแนกของข้อสอบ (a) ค่าความยากของข้อสอบ (b) และค่าการเดาของข้อสอบ (c) ที่ครอบคลุมช่วงระดับความสามารถที่เหมาะสมด้วย (Urry, 1977)

การพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ในครั้งนี้ เป็นการพัฒนาโปรแกรมในรูปแบบของ Web Application โดยใช้ภาษา PHP ในการพัฒนาโปรแกรม ใช้ MySQL เป็นระบบจัดการฐานข้อมูล และใช้ภาษา SQL เป็นคำสั่ง ในการเชื่อมโยงข้อมูลในฐานข้อมูล (Sukhanonsawat, Chadcham, & Chinnasarn, 2013) ซึ่งผู้วิจัยพัฒนาโปรแกรมคอมพิวเตอร์ตามหลักการของวงจรการพัฒนาซอฟต์แวร์ (Software Development Life Cycle: SDLC) (Elliott, 2004) โดยมีขั้นตอนการพัฒนาโปรแกรม 6 ขั้นตอน คือ 1) วิเคราะห์และออกแบบโปรแกรม เป็นการวิเคราะห์รูปแบบ วัตถุประสงค์ของโปรแกรม รวมทั้งวิเคราะห์แผนผังบริบท (Context diagram) เพื่อหา Source destination และการจัดทำผังการไหลของข้อมูล (Data Flow Diagram: DFD) เพื่อแสดงให้เห็นภาพรวมของโปรแกรมทั้งข้อมูลและการทำงานของโปรแกรม การทดสอบ 2) ออกแบบฐานข้อมูล เป็นการออกแบบลักษณะของข้อมูลที่ต้องการจัดเก็บลงในโปรแกรมที่พัฒนาขึ้น 3) พัฒนาโปรแกรม เป็นการนำรูปแบบโครงสร้างของหน้าจอต่าง ๆ ที่ออกแบบไว้ และข้อสอบที่วิเคราะห์ไว้แล้ว มาพัฒนาโปรแกรมตามวัตถุประสงค์ 4) ทดสอบและแก้ไขโปรแกรม เป็นการทดลองใช้โปรแกรมที่พัฒนาขึ้น เพื่อค้นหาและแก้ไขข้อผิดพลาดที่เกิดขึ้นจากการพัฒนา 5) จัดทำคู่มือการใช้โปรแกรม เป็นการจัดทำเอกสารการใช้งานของโปรแกรมที่พัฒนาขึ้น และ 6) ประเมินความเหมาะสมกับการนำไปใช้งาน เป็นการสำรวจความพึงพอใจของผู้ใช้โปรแกรม

การประเมินความเหมาะสมในการใช้งานของโปรแกรม จากการทดลองใช้โปรแกรมการทดสอบในการวิจัยนี้มี 3 ขั้นตอน คือ 1) ผู้วิจัยตรวจสอบโปรแกรมด้วยตนเอง เพื่อหาจุดบกพร่องเบื้องต้น 2) การตรวจสอบความเหมาะสมของโปรแกรมจากผู้เชี่ยวชาญ 4 คน โดยผู้เชี่ยวชาญมีความรู้ทางด้านการศึกษาหรือการพัฒนาโปรแกรมคอมพิวเตอร์ วุฒิกศศึกษาไม่ต่ำกว่าปริญญาโท และมีประสบการณ์ในด้านนี้ไม่น้อยกว่า 10 ปี และ 3) การตรวจสอบโปรแกรม โดยผู้ทดลองใช้โปรแกรม เพื่อหาจุดบกพร่องในการใช้โปรแกรม ผู้ทดลองใช้ที่ประเมินความเหมาะสมในการใช้โปรแกรมครั้งนี้ เป็นนักเรียนระดับชั้นประถมศึกษาปีที่ 6 ระดับชั้นมัธยมศึกษาปีที่ 3 และระดับชั้นมัธยมศึกษาปีที่ 6 ซึ่งผู้วิจัยได้แก้ไข ตามข้อเสนอแนะจากผู้เชี่ยวชาญและผู้ทดลองใช้แล้ว

โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่พัฒนาขึ้น เป็นโปรแกรมในรูปแบบ Web application (www.onetcat.net) ซึ่งผู้เข้าสอบสามารถเข้ามาทดสอบที่ใด เวลาใดก็ได้ โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับการจัดสอบ O-NET ระดับชั้นประถมศึกษาปีที่ 6 ชั้นมัธยมศึกษาปีที่ 3 และชั้นมัธยมศึกษาปีที่ 6 ระดับชั้นละ 8 กลุ่มสาระการเรียนรู้ ได้แก่ 1) วิชาภาษาไทย 2) วิชาคณิตศาสตร์ 3) วิชาวิทยาศาสตร์ 4) วิชาภาษาต่างประเทศ 5) วิชาศิลปะ 6) วิชาสุขศึกษาและพลศึกษา 7) วิชาสังคมศึกษา ศาสนา และวัฒนธรรม และ 8) วิชาการงานอาชีพและเทคโนโลยี โดยแบ่งหน้าจอโปรแกรม ออกเป็น 5 ส่วน คือ 1) การลงทะเบียนสอบ เป็นส่วนที่ผู้เข้าสอบลงทะเบียนการทดสอบก่อนเข้าสอบ 2) การจัดการทดสอบ โดยใช้ทฤษฎีการตอบสนองข้อสอบในรูปแบบของโมเดลแยกทางแบบแปรผัน (Variable branching model) ซึ่งเป็นส่วนที่ให้ผู้เข้าสอบเข้าทดสอบระดับความรู้ขั้นพื้นฐานของตนเอง 3) การรายงานผลการทดสอบ เป็นส่วนที่แสดงผลการทดสอบให้ ผู้เข้าสอบทราบ 4) การบริหารการทดสอบ ซึ่งเป็นส่วนที่ให้ครู อาจารย์ หรือผู้ที่เกี่ยวข้องใน



การจัดการข้อสอบใส่ข้อสอบที่มีคุณภาพตามเกณฑ์ที่กำหนดลงในโปรแกรม เพื่อใช้เป็นคลังข้อสอบที่จะนำมาใช้ในการทดสอบต่อไป และ 5) คู่มือการใช้งาน เป็นส่วนที่บอกการใช้งานของโปรแกรมการทดสอบ

จากผลการประเมินความเหมาะสมในการใช้งานของโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ปรากฏว่า การใช้งานอยู่ในเกณฑ์ดี ทั้งในด้านความสะดวกในการนำไปใช้ และด้านลักษณะทั่วไปของโปรแกรม เป็นที่พึงพอใจของผู้ใช้งาน ดังนั้น โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับการจัดสอบ O-NET สามารถนำไปใช้งานจริงได้

## ข้อเสนอแนะ

### ข้อเสนอแนะสำหรับการนำผลการวิจัยไปใช้

1. ข้อสอบในคลังข้อสอบของระดับชั้นประถมศึกษาปีที่ 6 มีจำนวนข้อสอบน้อยเกินไป ควรมีการบรรจุข้อสอบเพิ่มมากขึ้น เพื่อให้มีจำนวนข้อสอบเหมาะสมสำหรับการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์

2. ข้อสอบที่ผ่านเกณฑ์การวิเคราะห์คุณภาพเพื่อจัดเก็บไว้ในคลังข้อสอบในครั้งนี้ มีจำนวนค่อนข้างน้อย จึงควรมีการวิเคราะห์ข้อสอบ O-NET ทุกรายวิชาและทุกระดับชั้นเก็บไว้ในคลังข้อสอบให้มีจำนวนมาก ๆ เนื่องจากคลังข้อสอบที่จะนำมาใช้กับโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์จะต้องมีขนาดใหญ่คือมีข้อสอบจำนวนมาก โดยข้อสอบต้องครอบคลุมคุณลักษณะและเนื้อหาที่ต้องการวัดความสามารถของผู้เรียน (Weiss, 1988) คลังข้อสอบที่มีความเหมาะสมนั้น จะต้องมีความยาวข้อสอบเท่ากับ  $2^n$  ข้อ โดยที่ n คือจำนวนข้อสอบที่ต้องการใช้ในการทดสอบ เช่น ถ้าต้องการใช้ข้อสอบประมาณ 10 ข้อ ข้อสอบในคลังข้อสอบควรมีจำนวน  $2^{10}$  ข้อ หรือประมาณ 1,024 ข้อ

3. โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่พัฒนาขึ้นนี้เป็นโปรแกรมในรูปแบบ Web application ([www.onetcat.net](http://www.onetcat.net)) ซึ่งผู้สนใจสามารถเข้ามาทดสอบได้สะดวกทุกที่ ที่สามารถเชื่อมต่ออินเทอร์เน็ตได้และทุกเวลา อันจะเป็นประโยชน์กับนักเรียนในการทำควมคุ้นเคยกับข้อสอบ และเพื่อเตรียมตัวสำหรับการสอบ O-NET ได้ทุกระดับชั้นและทุกวิชา

4. การนำโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับการจัดสอบ O-NET ไปใช้ควรนำไปทดลองใช้ในบริบทเล็ก ๆ ก่อน เพื่อศึกษาปัญหาและความเหมาะสมของการนำไปใช้ในบริบทที่ศึกษา ซึ่งอาจมีปัญหาดังกล่าว เช่น ประสิทธิภาพของการเข้าถึงอินเทอร์เน็ตที่แตกต่างกันในแต่ละพื้นที่ ก่อนที่จะนำไปขยายผลต่อไป

### ข้อเสนอแนะสำหรับการวิจัยต่อไป

1. การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ควรมีการจัดสอบกับข้อสอบประเภทอื่น ๆ ด้วย เช่น การทดสอบทางการศึกษาระดับชาติด้านการศึกษานอกโรงเรียน (Non-Formal National Educational Testing: N-NET) การทดสอบทางการศึกษาระดับชาติด้านอาชีวศึกษา (Vocational National Educational Testing: V-NET) การทดสอบทางการศึกษาระดับชาติด้านพระพุทธศาสนา (Buddhism National Educational Testing: B-NET)

2. ข้อสอบ O-NET ที่นำมาใช้ในการวิจัยนี้ เป็นข้อสอบแบบเลือกตอบ ชนิดไม่เกิน 5 ตัวเลือก ซึ่งมึลักษณะการตรวจให้คะแนนแบบ 2 ค่า (ตอบถูกได้คะแนน 1 คะแนน และตอบผิดได้คะแนน 0 คะแนน) เท่านั้น ในอนาคตควรศึกษาวิจัยกับข้อสอบที่มีลักษณะอื่นๆ ด้วย เช่น ข้อสอบแบบเขียนตอบ



3. การพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับการจัดสอบ O-NET ครั้งนี้ ยังไม่มีการตรวจสอบอัตราส่วนการใช้ข้อสอบซ้ำ ในการทดสอบแต่ละครั้ง แต่ในการพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการเรียนการสอน ควรมีการทำวิจัยเพื่อตรวจสอบอัตราการใช้ข้อสอบซ้ำ ในการทดสอบแต่ละครั้งด้วย

4. ควรวิจัยเกี่ยวกับวิธีการคัดเลือกข้อสอบเริ่มต้น การคัดเลือกข้อสอบข้อต่อไป และเกณฑ์การหยุดที่เหมาะสมในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับการสอบ O-NET

### กิตติกรรมประกาศ

งานวิจัยนี้ได้รับการสนับสนุนทุนวิจัยจากสถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) ปีงบประมาณ 2557

### เอกสารอ้างอิง

- จารุจิตร สิริธิปุ ปิยะทิพย์ ดินวร และโสฬส สุขานนท์สวัสดิ์. (2559) การพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการจัดสอบ O-NET ระดับชั้นมัธยมศึกษาปีที่ 3. *วารสารการวัดผล การศึกษา มหาวิทยาลัยมหาสารคาม*, 22(1), (in press)
- Bunderson, V. C., Inouye, D. K., & Olsen, J. B. (1989). *The four generations of computerized educational measurement*. In R. L. Linn (Ed.), *Educational Measurement* (3<sup>rd</sup> ed., pp. 367-407). New York: Macmillan.
- Elliott, G. (2004). *Global Business Information Technology: An Integrated Systems Approach*. Ontario: Pearson Addison Wesley.
- Khumyoo, T., Chadcham, S., & Chinnasarn, K. (2013). Diagnostic of English reading skill by applying the attributes hierarchy model and computerized adaptive testing. *Research Methodology & Cognitive Science*, 10(2). 55-70. (in Thai)
- Mejang, S., Wongdee, P., & Konta, U. (2010). A comparison of results from equating O-NET scores with common items and with a three-parameter IRT model. *Research Methodology & Cognitive Science*, 7(2). 81-92. (in Thai)
- Muninsakorn, Y., Tinnaworn, P., & Sukhanonsawat, S. (2015). Development of the computerized adaptive testing program for O-NET at the Grade 6 level. In *Burapha University International Conference 2015: Moving forward to a Prosperous and Sustainable Community, July 10-12, 2015, Bangsaen Heritage Hotel, Chon Buri*. (pp. 290-298) Thailand: Burapha University.
- Pluemjai, N., Tinnaworn, P., & Sukhanonsawat, S. (2015). Development of the computerized adaptive testing program for O-NET at the Grade Twelve level. *Research Methodology & Cognitive Science*, 13(2), 109-125. (in Thai)

- Sakolkijrunroj, S., Chadcham, S., & Sudhasani, S. (2015). The development of computerized adaptive testing program for Thai happiness scale. *Research Methodology & Cognitive Science*, 13(1). 1-17. (in Thai)
- Sukhanonsawat, S., Chadcham, S., & Chinnasarn, K. (2013). Development of the next item selection procedure using decision theory in computerized adaptive testing. *Research Methodology & Cognitive Science*, 10(2). 71-85. (in Thai)
- Thompson, N. A., & Weiss, D. J. (2011). A framework for the development of computerized adaptive tests. *Practical Assessment, Research & Evaluation*, 16(1), 1-9.
- Urry, V. W. (1977). Tailored testing: A successful application of latent trait theory. *Journal of Educational Measurement*, 14, 181 – 196.
- Weiss, D. J. (1974). *Strategies of adaptive ability measurement*. Research Report. Minneapolis: University of Minnesota, Department of Psychology, Psychometric Methods Program.
- Weiss, R. S. (1988). Loss and Recovery. *Journal of Social Issues*, 44(3), 37-52.