

ประสิทธิภาพของโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์
สำหรับมาตรวัดความสุขของคนไทย
The Efficiency of Computerized Adaptive Testing Program
for Thai Happiness Scale

สุชาดา สกลกิจรุ่งโรจน์¹
เสรี ชัดรัมย์²
สมพร สุทัศน์ย์³

บทคัดย่อ

โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย เป็นโปรแกรมที่พัฒนาขึ้นตามแนวคิดทฤษฎีการตอบสนองข้อสอบร่วมกับการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับวัดความอยู่ดีมีสุขเชิงอัตวิสัย การวิจัยนี้มีวัตถุประสงค์เพื่อตรวจสอบประสิทธิภาพของโปรแกรมในด้านความประหยัดและความรวดเร็วในการทดสอบ และศึกษาความสัมพันธ์ของผลการประมาณค่าความสุขที่ได้จากการทดสอบด้วยโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย กับการทดสอบด้วยข้อคำถามทั้งหมดในคลังข้อคำถาม เครื่องมือที่ใช้ในการวิจัย คือโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทยที่พัฒนาขึ้นในงานวิจัยก่อนหน้านี้ ในรูปแบบเว็บแอปพลิเคชัน ซึ่งสามารถเข้าใช้งานผ่านทางเว็บไซต์ <http://www.thscat.com/Test/> โดยสามารถเลือกทดสอบได้ 2 ลักษณะ ได้แก่ 1) การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ (THS-CAT) และ 2) การทดสอบด้วยข้อคำถามทั้งหมดในคลังข้อคำถาม (THS-ทุกข้อ) กลุ่มตัวอย่างที่ใช้ในการวิจัยมีจำนวนทั้งสิ้น 30 คน กลุ่มตัวอย่างจะได้รับการทดสอบทั้งสองรูปแบบ คือ การทดสอบด้วย THS-CAT และ THS-ทุกข้อ วิเคราะห์ข้อมูลด้วยสถิติบรรยาย และวิเคราะห์ความสัมพันธ์ของผลการประมาณค่าความสุขที่ได้จากการทดสอบทั้งสองรูปแบบ ด้วยการหาค่าสัมประสิทธิ์สหสัมพันธ์เพียร์สัน ผลการวิจัย ปรากฏว่า โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทยสามารถดำเนินการทดสอบโดยใช้จำนวนข้อคำถามและระยะเวลาในการทดสอบน้อยกว่าการทดสอบด้วยข้อคำถามทั้งหมด ในคลังข้อคำถาม อีกทั้งมีความแม่นยำในการประมาณค่าสูง ผลการประมาณค่าที่ได้จากการทดสอบทั้งสองรูปแบบมีความสัมพันธ์กันทางบวก

คำสำคัญ : โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์, มาตรวัดความสุขของคนไทย, ประสิทธิภาพ

¹ อาจารย์ ดร. วิทยาลัยวิทยาการวิจัยและวิทยาการปัญญา มหาวิทยาลัยบูรพา

² รองศาสตราจารย์ ดร. วิทยาลัยวิทยาการวิจัยและวิทยาการปัญญา มหาวิทยาลัยบูรพา

³ รองศาสตราจารย์ ดร. วิทยาลัยวิทยาการวิจัยและวิทยาการปัญญา มหาวิทยาลัยบูรพา

Abstract

The Computerized Adaptive Testing Program for Thai Happiness Scale was developed based on two concepts : Item Response Theory and Computerized Adaptive Testing. This program was designed to assess the individual level of subjective well-being while this study was aimed to investigate the efficiency of the Computerized Adaptive Testing Program for Thai Happiness Scale and study the relationship between the happiness theta derived from Computerized Adaptive Testing Program for Thai Happiness Scale (THS-CAT) and all items in bank (THS-All). The Computerized Adaptive Testing Program for Thai Happiness Scale was created in web application format which can access anywhere and anytime via <http://www.thscat.com/Test/>. The program include two options for testing: 1) THS-CAT and 2) THS-All. The sample consisted of 30 individuals who assigned for both testing options. The data analysis were conducted using descriptive statistic and correlation coefficient. The result revealed that THS-CAT used less time and items comparing with THS-All. Moreover, the happiness theta derived from both options were approximately similar.

Keywords : Computerized Adaptive Testing Program, Thai Happiness Scale, Efficiency

ความนำ

ความก้าวหน้าทางด้านเทคโนโลยีคอมพิวเตอร์และทฤษฎีการทดสอบ ก่อให้เกิดทางเลือกของการบริหารการทดสอบรูปแบบหนึ่ง เรียกว่า การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ (Computerized Adaptive Testing: CAT) ซึ่งเป็นการบริหารการทดสอบที่นำหลักการของทฤษฎีการตอบสนองข้อสอบ (Item Response Theory: IRT) มาเป็นแนวคิดในการวิเคราะห์ข้อคำถามและใช้ในการประมาณค่าความสามารถ (Ability) หรือคุณลักษณะแฝงของบุคคล (Latent Trait) เพื่อเพิ่มความแม่นยำในการทดสอบ (Precision) รวมทั้งช่วยลดระยะเวลาและจำนวนข้อสอบหรือข้อคำถามที่ใช้ในการทดสอบ (Gibbons et al., 2008)

การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์มีส่วนประกอบสำคัญ 5 ส่วน ได้แก่ 1) คลังข้อคำถาม ซึ่งเป็นแหล่งจัดเก็บข้อคำถามที่มีคุณภาพซึ่งผ่านการวิเคราะห์ด้วยทฤษฎีการตอบสนองข้อสอบ 2) การกำหนดจุดเริ่มต้นการทดสอบ ซึ่งเป็นการกำหนดวิธีการเลือกข้อคำถามข้อแรก 3) การกำหนดวิธีการประมาณค่าคุณลักษณะแฝงภายหลังการตอบข้อคำถามของผู้ทดสอบ 4) การกำหนดวิธีการคัดเลือกข้อคำถามในการทดสอบถัดไป และ 5) การกำหนดเกณฑ์ยุติการทดสอบ (De Ayala, 2009; Thompson & Weiss, 2011)

การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์เป็นที่รู้จักอย่างแพร่หลายในแวดวงการศึกษาซึ่งพัฒนาแบบทดสอบวัดความสามารถด้านต่างๆ แต่ในช่วง 2 ทศวรรษที่ผ่านมาได้มีการนำการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์มาใช้ในตัวแปรอื่นๆ มากยิ่งขึ้น ตัวอย่างเช่น การประเมินภาวะซึมเศร้า การประเมินความวิตกกังวล การประเมินความเครียด การประเมินปัญหาทางด้านสุขภาพ และการประเมินปัญหาพฤติกรรมทางสังคม การตรวจสอบประสิทธิภาพของมาตรวัดที่พัฒนาขึ้นในรูปแบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ทั้งในสถานการณ์จำลองและในสถานการณ์จริง ผลปรากฏว่า มาตรวัดรูปแบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สามารถช่วยลดจำนวนข้อคำถามและระยะเวลาที่ใช้ในการทดสอบได้อย่างมาก รวมทั้งมีความแม่นยำในการทดสอบสูง และได้รับการยอมรับจากผู้ทดสอบเป็นอย่างดี (Becker et al., 2008;

Kocalevent et al., 2009; Smits, Cuijpers, & Straten, 2011; Vogels, Jacobusse, & Reijneveld, 2011; Zheng, Chang, & Chang, 2013)

ในประเทศไทย การพัฒนาการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ส่วนใหญ่ยังอยู่ในขอบเขตของการวัดความสามารถ ความก้าวหน้าทางด้านเทคโนโลยีและทฤษฎีการทดสอบจึงถูกนำไปใช้ในบริบทของแวดวงการศึกษาเป็นหลัก ทั้งที่จริงแล้วความก้าวหน้าดังกล่าวสามารถนำมาประยุกต์กับการพัฒนามาตรวัดตัวแปรทางด้านจิตใจอื่นๆ ได้ ซึ่งตัวแปรทางด้านจิตใจที่กำลังได้รับความสนใจเป็นอย่างมากทั้งในและต่างประเทศ คือ ความสุข (Happiness) โดยมีผู้คิดค้นและนิยามความสุขอย่างหลากหลาย มาตราวัดที่ใช้ในการวัดความสุขล้วนสร้างขึ้นจากแนวคิดและทฤษฎีที่แตกต่างกัน ตัวอย่างของมาตราวัดความสุขทั้งในประเทศและต่างประเทศ ได้แก่ แบบวัดสุขภาวะองค์กรรวมแนวพุทธ (ชูฤทธิ์ เติ่งไตรสรณ์ พระครุชิต คุณวโร และ อรุณี วิชาพรทิพย์, 2551) แบบประเมินความสุขคนไทย (ประเวช ตันติพิวัฒนสกุล และ เอกอนงค์ สีตลาภินันท์, 2554) แบบสำรวจความสุขด้วยตนเอง : Happinometer (ศิรินันท์ กิตติสุขสถิต และคณะ, 2555) Satisfaction with Life Scale (Diener, Emmons, Larsen, & Griffin, 1985) The Positive and Negative Affect Schedule (Watson, Clark, & Tellegen, 1988) Quality of Life Inventory (Frisch, Cornell, Villanueva, & Retslaff, 1992) Psychological Well-Being Scales (Ryff & Singer, 1998) Subjective Happiness Scale (Lyubomirsky & Lepper, 1999) Flourishing Scale และ Scale of Positive and Negative Experience (Diener et al., 2009)

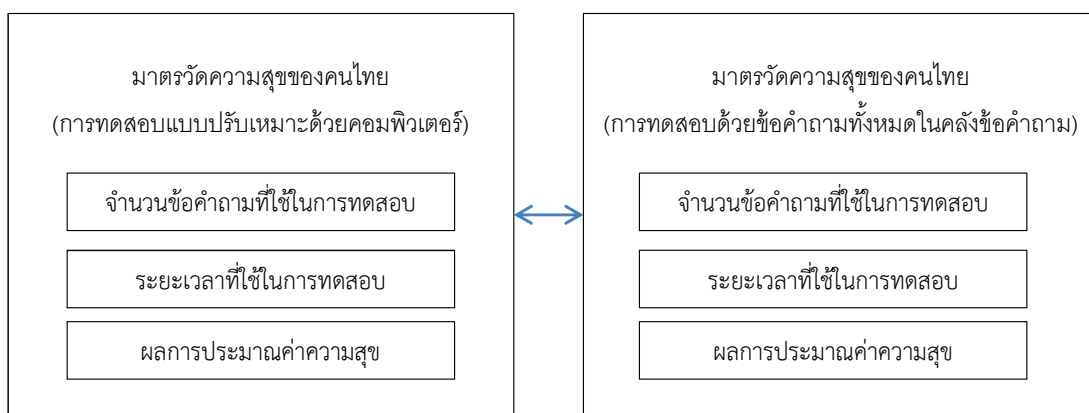
มาตราวัดความสุขส่วนใหญ่พัฒนาขึ้นตามแนวคิดทฤษฎีการทดสอบแบบดั้งเดิม โดยมีรูปแบบการบริหารการทดสอบในรูปแบบกระดาษ-ดินสอหรือมาตราวัดออนไลน์ การประมาณค่าคุณลักษณะแฝงของบุคคลยังจำกัดอยู่เพียงการใช้ผลคะแนนรวมที่ได้จากข้อคำถามทุกข้อ ดังนั้น ผู้รับการทดสอบไม่ว่ามีคุณลักษณะแฝงอยู่ในระดับใดจึงจำเป็นต้องทดสอบด้วยข้อคำถามทั้งหมดในมาตราวัดนั้นอย่างหลีกเลี่ยงไม่ได้ ทั้งที่ข้อคำถามบางข้ออาจไม่ได้ให้สารสนเทศทางการทดสอบที่ดีสำหรับผู้รับการทดสอบคนนั้นก็ตาม ปัจจุบันได้มีโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตราวัดความสุขของคนไทย ซึ่งมุ่งวัดความสุขตามแนวคิดทฤษฎีความอยู่ดีมีสุขเชิงอัตวิสัย (Pavot & Diener, 2013) ผู้วิจัยได้ตรวจสอบคุณภาพของข้อคำถามตามแนวคิดทฤษฎีการตอบสนองข้อสอบเรียบร้อยแล้ว (สุชาติ สกลกิจรุ่งโรจน์, เสรี ชัดเข้ม และม.ร.ว.สมพร สุทัศนีย์, ม.ป.ป. ; Sakolkijrungrroj, Chadcham, & Sudhasani, 2015) อย่างไรก็ตาม โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตราวัดความสุขของคนไทยดังกล่าวยังมิได้ตรวจสอบประสิทธิภาพและผลการประมาณค่าที่ได้จากโปรแกรมการทดสอบ ผู้วิจัยจึงสนใจตรวจสอบประสิทธิภาพของโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตราวัดความสุขของคนไทยในด้านความประหยัดและความรวดเร็วในการทดสอบ รวมทั้งศึกษาความสัมพันธ์ของผลการประมาณค่าที่ได้จากโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตราวัดความสุขของคนไทย กับการทดสอบด้วยข้อคำถามทั้งหมดในคลังข้อคำถาม

วัตถุประสงค์ของการวิจัย

1. เพื่อตรวจสอบประสิทธิภาพของโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตราวัดความสุขของคนไทยในด้านความประหยัดและความรวดเร็วในการทดสอบ
2. เพื่อศึกษาความสัมพันธ์ของผลการประมาณค่าความสุขที่ได้จากการทดสอบด้วยโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตราวัดความสุขของคนไทย กับการทดสอบด้วยข้อคำถามทั้งหมดในคลังข้อคำถาม

กรอบแนวคิดการวิจัย

การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์เป็นกระบวนการทดสอบที่ใช้ผลการตอบข้อคำถามก่อนหน้าเป็นข้อมูลในการเลือกข้อคำถามข้อถัดไปให้เหมาะสมกับระดับคุณลักษณะแฝงของผู้รับการทดสอบ โดยการทดสอบจะใช้ชุดของข้อคำถามเพียงบางส่วนจากคลังข้อคำถามทั้งหมด แต่ละบุคคลอาจได้ชุดของข้อคำถามที่แตกต่างกัน แต่ข้อคำถามที่ถูกเลือกนั้นจะเป็นข้อคำถามที่เหมาะสมกับระดับคุณลักษณะแฝงของผู้รับการทดสอบคนนั้นๆ ทำให้การดำเนินการทดสอบใช้จำนวนข้อคำถามน้อยกว่า และใช้เวลาในการทดสอบลดลง การวิจัยครั้งนี้ต้องการตรวจสอบประสิทธิภาพของโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย (สุชาติ สกลกิจรุ่งโรจน์, เสรี ชัดเข้ม และ ม.ร.ว. สมพร สุทัศน์, ม.ป.ป.; Sakolkijrunroj, Chadcham, & Sudhasani, 2015) ในด้านความประหยัดและความรวดเร็วในการทดสอบโดยเปรียบเทียบระยะเวลาและข้อคำถามที่ใช้ในการทดสอบด้วยโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์กับการทดสอบด้วยข้อคำถามทั้งหมดในคลังข้อคำถาม นอกจากนี้ ปกติแล้วการประมาณค่าความสุขจากการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์เป็นการประมาณค่าความสุขที่ใช้ผลการตอบข้อคำถามเพียงจำนวนหนึ่งจากข้อคำถามทั้งหมดเป็นข้อมูลในการประมาณค่า ข้อคำถามที่ใช้เป็นข้อคำถามที่ผ่านกระบวนการเลือกและตรวจสอบด้วยกลไกของโปรแกรมแล้วว่ามีความเหมาะสมกับผู้รับการทดสอบ การดำเนินการทดสอบจะดำเนินอย่างต่อเนื่องจนกระทั่งถึงเกณฑ์ยุติการทดสอบ ซึ่งทั่วไปโปรแกรมจะยุติการทดสอบด้วยเกณฑ์ความคลาดเคลื่อนมาตรฐานของการประมาณค่าที่ยอมรับได้ก่อนที่ข้อคำถามจะหมดคลัง อย่างไรก็ตาม บางแนวคิดเชื่อว่าหากข้อคำถามที่ใช้ในการทดสอบมีจำนวนมากขึ้นอาจทำให้ได้ข้อมูลการทดสอบที่ละเอียดและแม่นยำเพิ่มขึ้น การศึกษาความสัมพันธ์ของผลการประมาณค่า ความสุขที่ได้จากการทดสอบ ทั้งสองรูปแบบนี้จึงเป็นการตรวจสอบว่าการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ซึ่งใช้ข้อคำถามในการประมาณค่าความสุขน้อยกว่าการทดสอบด้วยข้อคำถามทั้งหมดในคลังข้อคำถามสามารถประมาณค่าความสุขได้ใกล้เคียงกับการทดสอบด้วยข้อคำถามทั้งหมดในคลังข้อคำถามหรือไม่ ด้วยการศึกษาผลการประมาณค่าที่ได้จากการทดสอบทั้งสองรูปแบบ จากแนวคิดดังกล่าวสามารถสรุปเป็นกรอบแนวคิดการวิจัยดังภาพที่ 1



ภาพที่ 1 กรอบแนวคิดการวิจัย

สมมติฐานการวิจัย

สมมติฐานการวิจัยจากวัตถุประสงค์ของการวิจัยข้อที่ 2 มีดังนี้

ผลการประมาณค่าความสุขที่ได้จากโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัด ความสุขของคนไทย กับผลการประมาณค่าความสุขที่ได้จากการทดสอบด้วยข้อคำถามทั้งหมดในคลังข้อคำถามมี ความสัมพันธ์กันทางบวก

วิธีดำเนินการวิจัย

กลุ่มตัวอย่าง

กลุ่มตัวอย่างที่ใช้ในการวิจัย เป็นประชาชนชาวไทย อายุตั้งแต่ 18 ปีขึ้นไป ใช้ภาษาไทยในการติดต่อสื่อสาร สามารถใช้คอมพิวเตอร์เบื้องต้นได้ กลุ่มตัวอย่างได้มาด้วยวิธีการเลือกตัวอย่างแบบตามสะดวก จำนวน 30 คน

เครื่องมือที่ใช้ในการวิจัย

ประกอบด้วยโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย (สุชาดา สกolkijrungrroj, เสรี ชัดเข้ม และม.ร.ว.สมพร สุทัศนีย์, ม.ป.ป.; Sakolkijrungrroj, Chadcham, & Sudhasani, 2015) พัฒนาขึ้นในรูปแบบเว็บแอปพลิเคชัน สามารถเข้าใช้งานผ่านทางเว็บไซต์ <http://www.thscat.com/Test/> โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย มุ่งวัดความสุขตามทฤษฎี ความอยู่ดีมีสุขเชิงอัตวิสัย ซึ่งสามารถเลือกทดสอบได้ 2 ลักษณะ ได้แก่

- 1) การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ (THS-CAT)
- 2) การทดสอบด้วยข้อคำถามทั้งหมดในคลังข้อคำถาม (THS-ทุกข้อ)

เมื่อเริ่มการทดสอบโปรแกรมจะแสดงข้อคำถามทีละข้อ ผู้ทดสอบต้องตอบคำถามข้อปัจจุบันก่อนจึงจะสามารถตอบคำถามข้อถัดไปได้ และไม่สามารถย้อนกลับมาเปลี่ยนแปลงคำตอบในคำถามข้อที่ผ่านไปแล้วได้ การทดสอบจะเริ่มต้นด้วยข้อคำถามที่ให้ค่าสารสนเทศสูงที่สุด (Maximum Information) ณ ตำแหน่งค่าคุณลักษณะแฝงเริ่มต้นเท่ากับ 0.0 เมื่อผู้ทดสอบตอบข้อคำถามแล้ว โปรแกรมจะประมาณค่าคุณลักษณะแฝงของผู้ทดสอบอีกครั้ง โดยใช้วิธีการ Expected a Posteriori (EAP) จากนั้นจึงคัดเลือกข้อคำถามข้อถัดไปซึ่งมีค่าสารสนเทศสูงที่สุด ณ ตำแหน่งค่าคุณลักษณะแฝงใหม่ที่ประมาณค่าได้ การทดสอบจะดำเนินการไปเรื่อยๆ จนกระทั่งถึงเกณฑ์การยุติการทดสอบ คือ เมื่อความคลาดเคลื่อนมาตรฐานของการประมาณค่าคุณลักษณะแฝงมีค่าต่ำกว่า 0.30 หรือเมื่อข้อคำถามในคลังข้อคำถามหมด แล้วจึงรายงานผลการทดสอบ

การเก็บรวบรวมข้อมูล

ดำเนินการเก็บรวบรวมข้อมูลด้วยตัวเองตามสถานที่ต่างๆ เพื่อให้ได้กลุ่มตัวอย่างที่มีความหลากหลายทางด้านเพศ และอายุ การเก็บข้อมูลใช้โน้ตบุ๊กส่วนตัวของผู้วิจัยและเชื่อมต่ออินเทอร์เน็ตผ่านเครือข่ายโทรศัพท์มือถือที่ความเร็ว 3G-4G ทำให้สามารถเก็บข้อมูลได้ทุกที่ทุกเวลา การเก็บข้อมูลดำเนินการแบบรายบุคคล แต่ละบุคคลจะได้รับการทดสอบทั้งสองรูปแบบ โดยกลุ่มตัวอย่าง 15 คนแรกจะได้รับการทดสอบด้วย THS-CAT ก่อนแล้วตามด้วย THS-ทุกข้อ ขณะที่กลุ่มตัวอย่าง 15 คนหลังจะได้รับการทดสอบด้วย THS-ทุกข้อ ก่อนแล้วตามด้วย THS-CAT ทั้งนี้เพื่อจัดการปัญหาเรื่องลำดับของการทดสอบ (Order Effect) ที่อาจส่งผลกระทบต่อคะแนนการทดสอบได้

การวิเคราะห์ข้อมูล

1) วิเคราะห์ข้อมูลการทดสอบทั่วไป ด้วยสถิติบรรยาย (Descriptive Statistic) ได้แก่ ความถี่ ร้อยละ ค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐาน

2) วิเคราะห์ความสัมพันธ์ของผลการประมาณค่าความสุขที่ได้จากการทดสอบทั้งสองรูปแบบ ด้วยการหาค่าสัมประสิทธิ์สหสัมพันธ์เพียร์สัน (Pearson Product-Moment Correlation Coefficient) โดยผลการประมาณค่าความสุขที่นำมาใช้ในการศึกษา เป็นค่าเฉลี่ยคุณลักษณะแฝงที่ประมาณค่าได้จากมาตรวัดย่อยทั้งสามมาตรวัดย่อย ในรูปแบบของค่า Theta

การวิเคราะห์ข้อมูลทั้งสองส่วนดำเนินการด้วยโปรแกรมสำเร็จรูป SPSS

ผลการวิจัย

กลุ่มตัวอย่างสำหรับการตรวจสอบประสิทธิภาพของโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย และศึกษาความสัมพันธ์ระหว่างผลการประมาณค่าความสุขที่ได้จากโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย กับการทดสอบด้วยข้อคำถามทั้งหมดในคลังข้อคำถาม มีจำนวน 30 คน ปรากฏว่าส่วนใหญ่เป็นเพศหญิง (ร้อยละ 83.30) และมีอายุเฉลี่ย 30.83 ปี กลุ่มตัวอย่างแต่ละคนได้รับการทดสอบทั้งสองรูปแบบ ซึ่งปรากฏผลการทดสอบดังตารางที่ 1

ตารางที่ 1 ระยะเวลาที่ใช้ในการทดสอบ (วินาที) จำนวนข้อคำถามที่ใช้ในการทดสอบ (ข้อ) ค่าความสุขที่ประมาณค่าได้ และคะแนนที่ได้จากการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ กับการทดสอบด้วยข้อคำถามทั้งหมดในคลังข้อคำถามของกลุ่มตัวอย่าง (n=30)

คนที่	ระยะเวลาที่ใช้ในการทดสอบ(วินาที)		จำนวนข้อคำถามที่ใช้ในการทดสอบ(ข้อ)		ค่าความสุขที่ประมาณค่าได้	
	CAT	ทุกข้อ	CAT	ทุกข้อ	CAT	ทุกข้อ
1	79	460	18	61	1.21	1.24
2	78	365	15	61	1.32	1.44
3	123	254	16	61	1.71	1.12
4	86	172	16	61	1.89	2.42
5	107	233	19	61	1.24	1.38
6	109	112	15	61	1.31	1.23
7	57	194	15	61	1.84	1.73
8	67	162	16	61	1.16	1.05
9	60	239	15	61	1.58	1.48
10	122	281	18	61	0.73	1.20
11	65	198	17	61	1.36	1.58
12	77	239	18	61	1.90	2.32
13	57	211	16	61	1.22	1.12
14	128	341	22	61	0.85	1.21
15	57	202	17	61	0.79	1.21
16	85	337	20	61	0.97	1.62
17	72	238	18	61	1.44	1.64
18	90	261	19	61	1.55	1.69
19	55	217	15	61	1.62	1.67

ตารางที่ 1 (ต่อ)

คนที่	ระยะเวลาที่ใช้ในการทดสอบ(วินาที)		จำนวนข้อคำถามที่ใช้ในการทดสอบ(ข้อ)		ค่าความสุขที่ประมาณค่าได้	
	CAT	ทุกข้อ	CAT	ทุกข้อ	CAT	ทุกข้อ
20	58	167	18	61	1.49	1.74
21	109	289	17	61	-0.18	-0.14
22	131	278	17	61	0.87	0.69
23	74	232	16	61	1.63	1.35
24	55	321	17	61	0.89	0.97
25	121	474	15	61	0.78	0.75
26	143	357	18	61	1.50	1.99
27	68	425	16	61	1.22	1.39
28	89	236	15	61	0.37	0.51
29	67	223	16	61	1.78	1.62
30	114	181	20	61	1.29	1.55
Min	55.00	112.00	15.00	61.00	-0.18	-0.14
Max	143.00	474.00	22.00	61.00	1.90	2.42
M	86.77	263.30	17.00	61.00	1.24	1.36
SD	27.00	88.23	1.78	0.00	0.47	0.51

จากตารางที่ 1 ปรากฏว่าการทดสอบด้วยข้อคำถามทั้งหมดในคลังใช้ระยะเวลาในการทดสอบเฉลี่ย 263.30 วินาที ขณะที่การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ใช้ระยะเวลาในการทดสอบเฉลี่ย 86.77 วินาที แสดงให้เห็นว่าโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทยใช้ระยะเวลาในการทดสอบ ลดลงจากการทดสอบด้วยข้อคำถามทั้งหมดในคลัง คิดเป็นร้อยละ 67.05

การทดสอบด้วยข้อคำถามทั้งหมดในคลัง ใช้ข้อคำถามจำนวนทั้งสิ้น 61 ข้อ ขณะที่การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ใช้จำนวนข้อคำถามเฉลี่ย 17 ข้อ ส่งผลให้โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทยใช้จำนวนข้อคำถามในการทดสอบลดลงจากการทดสอบด้วยข้อคำถามทั้งหมดในคลัง คิดเป็นร้อยละ 72.13

ค่าความสุขเฉลี่ยที่ประมาณค่าได้จากการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ (ค่า Theta) มีค่าเท่ากับ 1.24 ขณะที่ค่าความสุขเฉลี่ยที่ประมาณค่าได้จากการทดสอบด้วยข้อคำถามทั้งหมดในคลัง มีค่าเท่ากับ 1.36 ซึ่งมีค่าแตกต่างกันเพียงเล็กน้อย

ผลการวิเคราะห์ค่าสัมประสิทธิ์สหสัมพันธ์เพียร์สันระหว่างผลการประมาณค่าความสุข (Theta) ที่ได้จากการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ กับการทดสอบด้วยข้อคำถามทั้งหมดในคลังข้อคำถาม ปรากฏว่า ผลการประมาณค่าความสุข (Theta) ที่ได้จากการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ กับการทดสอบด้วยข้อคำถามทั้งหมดในคลังข้อคำถามมีความสัมพันธ์กันทางบวก โดยมีค่าสัมประสิทธิ์สหสัมพันธ์เพียร์สันเท่ากับ 0.85 ที่ระดับนัยสำคัญทางสถิติ .01 ซึ่งเป็นไปตามสมมติฐานการวิจัยที่ตั้งไว้

อภิปรายผลการวิจัย

การทดสอบด้วยโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย ช่วยลดจำนวนข้อคำถามที่ใช้ในการทดสอบคิดเป็นร้อยละ 72.13 และช่วยลดระยะเวลาในการทดสอบลง คิดเป็นร้อยละ 67.05 เมื่อเทียบกับการทดสอบด้วยข้อคำถามทั้งหมดในคลัง ผลดังกล่าวแสดงให้เห็นถึงประสิทธิภาพของโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทยได้เป็นอย่างดี ซึ่งโดยทั่วไปแล้วการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์มีความต้องการใช้จำนวนข้อคำถามและระยะเวลาในการทดสอบน้อยกว่าการทดสอบแบบดั้งเดิม เนื่องจากการดำเนินการทดสอบนั้นเลือกใช้เพียงข้อคำถามที่เหมาะสมกับระดับคุณลักษณะแฝงของผู้รับการทดสอบเท่านั้น ผลการศึกษาที่ได้ในครั้งนี้ สอดคล้องกับผลการวิจัยของ Rose, Bjorner, Becker, Fries, and Ware (2006) ซึ่งพัฒนาแบบประเมินทางด้านสุขภาพที่มีข้อคำถามรวมทั้งสิ้น 136 ข้อ หลังจากประเมินประสิทธิภาพของการทดสอบ โดยวิธีการจำลองสถานการณ์การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ปรากฏว่า การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ใช้จำนวนข้อคำถามเฉลี่ย 10 ข้อ ซึ่งมีจำนวนน้อยลงกว่าเดิมมาก เช่นเดียวกับ Walter et al. (2007) ซึ่งพัฒนามาตรวัดความวิตกกังวลในรูปแบบการทดสอบปรับเหมาะด้วยคอมพิวเตอร์ โดยคลังข้อคำถามมีจำนวนทั้งสิ้น 50 ข้อ เมื่อจำลองสถานการณ์ (Simulation) การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์กับกลุ่มตัวอย่าง พบว่า การวัดความวิตกกังวลโดยการกำหนดให้มีความเคลื่อนไหวมาตรฐานของการวัดน้อยกว่า 0.32 นั้นใช้ข้อคำถามในการทดสอบเฉลี่ยเพียง 6-8 ข้อ ขณะที่ Gibbons et al. (2008) ได้ประยุกต์ทฤษฎีการตอบสนองข้อสอบ กับการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับสร้างมาตรวัด Mood and Anxiety Spectrum Scales (MASS) ในรูปแบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ โดยตรวจสอบประสิทธิภาพของโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ใน 2 ลักษณะ ได้แก่ 1) ศึกษาด้วยวิธีการจำลองสถานการณ์แบบโพสท์ฮอก (Post Hoc Simulation) ซึ่งเป็นการจำลองสถานการณ์จากข้อมูลจริง และ 2) การศึกษาในสถานการณ์จริง ผลการศึกษาปรากฏว่า การทดสอบรูปแบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ด้วยวิธีการจำลองสถานการณ์ ใช้ข้อคำถามในการทดสอบเฉลี่ย 24 ข้อ และการนำไปใช้ในสถานการณ์จริง ใช้ข้อคำถามเฉลี่ย 30 ข้อ ซึ่งข้อคำถามที่ใช้ในการทดสอบมีจำนวนลดลงร้อยละ 95 เมื่อเทียบกับการทดสอบรูปแบบเดิม นอกจากนี้ การศึกษาของ Kocalevent et al. (2009) ซึ่งตรวจสอบประสิทธิภาพทางการทดสอบของมาตรวัดความเครียด รูปแบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ทั้งในรูปแบบการจำลองสถานการณ์ และการนำไปใช้จริง ปรากฏว่าการตรวจสอบประสิทธิภาพการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ด้วยการจำลองสถานการณ์การทดสอบ ในการวัดปฏิกิริยาตอบสนองต่อความเครียดใช้ข้อคำถามในการทดสอบประมาณ 7 ข้อ ขณะที่การวัดการเผชิญกับความเครียดใช้ข้อคำถามในการทดสอบประมาณ 12 ข้อ ส่วนการตรวจสอบประสิทธิภาพการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ด้วยการนำไปใช้จริง ปรากฏว่า ในการวัดปฏิกิริยาตอบสนองต่อความเครียดใช้ข้อคำถามในการทดสอบประมาณ 6 ข้อ และการวัดการเผชิญกับความเครียดใช้ข้อคำถามในการทดสอบประมาณ 10 ข้อ (คลังข้อคำถามสำหรับวัดการเผชิญกับความเครียดมีจำนวนทั้งสิ้น 38 ข้อ และคลังข้อคำถามสำหรับวัดปฏิกิริยาตอบสนองต่อความเครียดมีจำนวนข้อคำถาม 31 ข้อ)

ผลการศึกษาความสัมพันธ์ของผลการประมาณค่าความสุขที่ได้จากโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย กับการทดสอบโดยใช้ข้อคำถามทั้งหมดในคลังข้อคำถาม โดยการหาค่าสัมประสิทธิ์สหสัมพันธ์เพียร์สัน ปรากฏว่า มีค่าสัมประสิทธิ์สหสัมพันธ์เพียร์สันเท่ากับ 0.85 ที่ระดับนัยสำคัญทางสถิติ .01 แสดงให้เห็นว่า ผลการประมาณค่าที่ได้จากการทดสอบทั้งสองรูปแบบมีความสัมพันธ์กันทางบวก โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุข

ของคนไทย สามารถใช้เป็นอีกทางเลือกหนึ่งของการทดสอบได้ อย่างไรก็ตาม โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทยสามารถดำเนินการทดสอบโดยใช้จำนวนข้อคำถามและระยะเวลาในการทดสอบน้อยกว่าการทดสอบด้วยข้อคำถามทั้งหมดในคลังข้อคำถาม อีกทั้งมีความแม่นยำในการประมาณค่าสูง เนื่องจากกำหนดค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าให้มีค่าน้อยกว่า 0.30 เป็นเกณฑ์ในการยุติการทดสอบ ดังนั้น การประมาณค่าคุณลักษณะแฝงของผู้ทดสอบทุกคน จึงมีความแม่นยำมากกว่า 0.90 โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย จึงเป็นโปรแกรมการทดสอบที่มีประสิทธิภาพ สามารถลดทรัพยากรในการทดสอบ (ข้อคำถาม) ประหยัดเวลา และประมาณค่าความสุขได้สอดคล้องกับการประมาณค่าความสุขโดยใช้ข้อคำถามทั้งหมดในคลังข้อคำถาม จึงสามารถใช้เป็นอีกทางเลือกหนึ่งของการวัดความสุขที่มีประสิทธิภาพ ไม่ต่างจากมาตรวัดอื่นๆ ที่พัฒนาขึ้นในรูปแบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ เช่น มาตรวัดความซึมเศร้า (Fliege et al., 2005) มาตรวัดความวิตกกังวล (Becker et al., 2008) มาตรวัดความเครียด (Kocalevent et al., 2009)

ข้อเสนอแนะ

โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทยสามารถใช้ประโยชน์ได้ ดังนี้

1. โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทยพัฒนาขึ้นในรูปแบบเว็บแอปพลิเคชัน ผู้สนใจสามารถเชื่อมต่อสัญญาณอินเทอร์เน็ตและเข้าใช้งานได้ที่ <http://www.thscat.com/Test/> เป็นประโยชน์สำหรับบุคคลทั่วไปในการเฝ้าระวังและสำรวจความสุขของตนเอง
2. นักวิจัยที่สนใจศึกษาวิจัยเกี่ยวกับความอยู่ดีมีสุขเชิงอัตวิสัย สามารถใช้โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทยเป็นเครื่องมือในการศึกษาวิจัยได้
3. บุคคลหรือองค์กรสามารถนำผลการทดสอบที่ได้จากมาตรวัดความสุขของคนไทย รูปแบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ มาเป็นแนวทางในการวางแผนการส่งเสริมให้บุคคลมีความสุขเพิ่มขึ้น เนื่องจากผลการทดสอบที่ได้ แสดงข้อมูลทั้งในส่วนองมาตรวัดย่อยและมาตรวัดความสุขโดยรวม ดังนั้นจึงสามารถพิจารณารายละเอียดที่ได้จากผลการทดสอบของแต่ละมาตรวัด ซึ่งช่วยให้เห็นภาพชัดเจนยิ่งขึ้นว่าบุคคลมีจุดเด่นหรือจุดที่ควรส่งเสริมในด้านใด

เอกสารอ้างอิง

ชูฤทธิ์ เต็งไตรสรณ์ พระครุฑิต คุณวโร และอรุณี วชิราพรทิพย์. “การพัฒนามาตรวัดสุขภาวะองค์กรรวมแนวพุทธ,” วารสารวิชาการสาธารณสุข. 17(6), 1650-1660, 2551.

ประเวช ตันติพิวัฒนสกุล และเอกอนงค์ สิตลาภินันท์. **คู่มือสร้างสุขระดับจังหวัด.** (พิมพ์ครั้งที่ 2).

กรุงเทพฯ : โรงพิมพ์ชุมนุมสหกรณ์การเกษตรแห่งประเทศไทย จำกัด, 2554.

ศิรินันท์ กิตติสุขสถิต, กาญจนา ตั้งชลทิพย์, สุภรต์ จรัสสิทธิ์, เฉลิมพล สายประเสริฐ, พอตตา บุญยธีระณะ และวรรณภา อารีย์. **คู่มือการวัดความสุขด้วยตนเอง HAPPINOMETER.** (พิมพ์ครั้งที่ 1). นครปฐม : สถาบันวิจัยประชากรและสังคม มหาวิทยาลัยมหิดล, 2555.

สุชาติ สกลกิจรุ่งโรจน์, เสรี ชัดแฉ่ม และม.ร.ว.สมพร สุทัศน์ีย์. “การพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับมาตรวัดความสุขของคนไทย,” **วิทยาการวิจัยและวิทยาการปัญญา.** (อยู่ระหว่างการตีพิมพ์).

- Becker, J., Fliege, H. Kocalevent, R. D., Bjorner, J. B, Rose, M., Walter, O. B., & Klapp, B. F. "Functioning and Validity of a Computerized Adaptive Test to Measure Anxiety (A-Cat)," **Depression and Anxiety**. 25, 182-194., 2008.
- De Ayala, R. J. **The Theory and Practice of Item Response Theory**. United States of America : The Guildford Press, 2009.
- Diener, E., Emmons, R. A., Larsen, R.J., & Griffin, S. "Satisfaction with Life Scale (SWLS)," **Journal of Personality Assessment**. 49, 71-75., 1985.
- Diener, E., Wirtz, D., Tov, W., Kim-Prieto, C., Choi, D., Oishi, S., & Biswas-Diener, R. "New measures of well-being: Flourishing and positive and negative feelings," **Social Indicators Research**. 39, 247-266, 2009.
- Fliege, H., Becker, J., Walter, O. B., Bjorner, J. B., Klapp, B. F., & Rose, M. "Development of a computer-adaptive test for depression (D-CAT)," **Quality of Life Research**. 14, 2277-2291, 2005.
- Frisch, M. B., Cornell, J., Villanueva, M., & Retzlaff, P. J. "Clinical validation of the Quality of Life Inventory: A measure of life satisfaction for use in treatment planning and outcome assessment," **Psychological Assessment : A Journal of Consulting and Clinical Psychology**. 4, 92-101., 1992.
- Gibbons, R. D., Weiss, D.J., Kupfer, D. J, Frank, E., Fagiolini, A. Grochocinski, V. J., Bhaumik, D. K., Stover, A., Bock, R. D., & Immekus, J. C. "Using Computerized Adaptive Testing to Reduce the Burden of Mental Health Assessment," **Psychiatric Services**. 59(4), 361-368, 2008.
- Kocalevent, R. D., Rose, M., Becker, J., Walter, O. B., Fliege, H., Bjorner, J. B., Kleiberg, D., & Klapp, B. F. (2009). "An evaluation of patient-reported outcomes found computerized adaptive testing was efficient in assessing stress perception," **Journal of Clinical Epidemiology**. 62, 278-287, 2009.
- Lyubomirsky, S., & Lepper, H. S. "A measure of subjective happiness: Preliminary reliability and construct validation," **Social Indicators Research**. 46, 137-155., 1999.
- Rose, M., Bjorner, J. B., Becker, J., Fries, J. F., & Ware, J. E. "Evaluation of a preliminary physical function item bank supported the expected advantages of the Patient-Reported Outcomes Measurement Information System (PROMIS)," **Journal of Clinical Epidemiology**. 61, 17-33., 2006.
- Ryff, C. D., & Singer, B. "The contours of positive human health," **Psychological Inquiry**. 9, 1-28, 1998.
- Smits, N., Cuijpers, P., & Straten, A. V. "Applying computerized adaptive testing to the CES-D scale: A simulation study," **Psychiatry Research**. 188, 147-155, 2011.

- Sakolkijrungrroj, S., Chadcham, S., & Sudhasani, S. (2015). The item bank construction for Thai happiness scale (computerized adaptive testing version). In Burapha University International Conference 2015 Committee (Eds.), *Burapha University International Conference 2015: Moving Forward to a Prosperous and Sustainable Community, July 10-12, 2015 Bangsaen Heritage Hotel Chon Buri* (pp.33-45). Thailand : Burapha University. Retrieved from <http://www.buuconference.buu.ac.th/buureg/getfile.php?file=file/BUU2015/proceedings.pdf>
- Thompson, N. A., & Weiss, D. J. (2011). A Framework for the Development of Computerized Adaptive Tests. **Practical Assessment, Research & Evaluation**. 16(1). Available online: <http://pareonline.net/getvn.asp?v=16&n=1>.
- Vogels, A. G., Jacobusse, G. W., & Reijneveld, S. A. “An accurate and efficient identification of children with psychosocial problems by means of computerized adaptive testing,” **BMC Medical Research Methodology**. 11, 1-9, 2011.
- Walter, O. B., Becker, J., Bjorner, J. B., Fliege, H., Klapp, B. F., & Rose M. “Development and evaluation of a computer adaptive test for anxiety (Anxiety-CAT),” **Quality of Life Research**. 16, 143-155, 2007.
- Watson, D., Clark, L. A., & Tellegen, A. “Development and validation of brief measures of positive and negative affect: The PANAS Scales,” **Journal of Personality and Social Psychology**. 47, 1063–1070, 1988.
- Zheng, Y., Chang, C. H., & Chang, H. H. “Content-balancing strategy in bifactor computerized adaptive patient-reported outcome measurement,” **Quality of life research**. 22(3), 491-499, 2013.