


การประยุกต์โมเดลการตอบสนองข้อสอบในการพัฒนามาตรวัดความสุขของคนไทย:
การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์

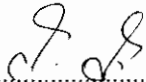
สุชาดา สกลกิจรุ่งโรจน์

ดุษฎีนิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปรัชญาดุษฎีบัณฑิต
สาขาวิชาการวัดและเทคโนโลยีทางวิทยาการปัญญา
วิทยาลัยวิทยาการวิจัยและวิทยาการปัญญา มหาวิทยาลัยบูรพา
เมษายน 2558
ลิขสิทธิ์เป็นของมหาวิทยาลัยบูรพา

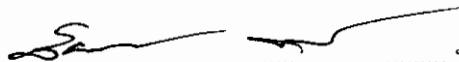
คณะกรรมการควบคุมดุขฎีนิพนธ์และคณะกรรมการสอบดุขฎีนิพนธ์ ได้พิจารณาดุขฎีนิพนธ์ของ สุชาดา สกลกิจรุ่งโรจน์ ฉบับนี้แล้ว เห็นสมควรรับเป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาดุขฎีบัณฑิต สาขาวิชาการวัดและเทคโนโลยีทางวิทยาการปัญญา ของมหาวิทยาลัยบูรพาได้

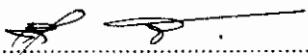
คณะกรรมการควบคุมดุขฎีนิพนธ์

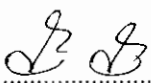

.....อาจารย์ที่ปรึกษาหลัก
(รองศาสตราจารย์ ดร.เสรี ชัดแฉ่ม)

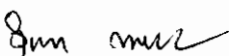

.....อาจารย์ที่ปรึกษาร่วม
(รองศาสตราจารย์ ดร. ม.ร.ว. สมพร สุทัศน์ีย์)

คณะกรรมการสอบดุขฎีนิพนธ์


.....ประธาน
(รองศาสตราจารย์ ดร.สัมพันธ์ พันธุ์พฤกษ์)


.....กรรมการ
(รองศาสตราจารย์ ดร.เสรี ชัดแฉ่ม)


.....กรรมการ
(รองศาสตราจารย์ ดร. ม.ร.ว. สมพร สุทัศน์ีย์)


.....กรรมการ
(ผู้ช่วยศาสตราจารย์ ดร.สุชาดา กรเพชรปानी)

วิทยาลัยวิทยาการวิจัยและวิทยาการปัญญาอนุมัติให้รับดุขฎีนิพนธ์ฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาดุขฎีบัณฑิต สาขาวิชาการวัดและเทคโนโลยีทางวิทยาการปัญญา ของมหาวิทยาลัยบูรพา


.....คณบดีวิทยาลัยวิทยาการวิจัย
และวิทยาการปัญญา

(ผู้ช่วยศาสตราจารย์ ดร.สุชาดา กรเพชรปानी)

วันที่ 16 เดือน สิงหาคม พ.ศ. 2558

ประกาศคุณูปการ

ดุชฎินิพนธ์ฉบับนี้สำเร็จลุล่วงได้ด้วยความรู้ความกรุณาจากรองศาสตราจารย์ ดร.เสรี ชัดเข้ม อาจารย์ที่ปรึกษาหลัก ซึ่งเฝ้าอบรมบ่มเพาะ ถ่ายทอดความรู้ ประสิทธิ์ประสาทวิชา รวมทั้งให้คำแนะนำเป็นอย่างดี ทั้งในส่วนของเนื้อหาดุชฎินิพนธ์และวิธีการปรับตัว ตลอดระยะเวลาที่ดำเนินการวิจัย ทำให้ผู้วิจัยสามารถฟันฝ่าอุปสรรคต่าง ๆ มาได้ด้วยดี นอกจากนี้ ผู้วิจัยขอกราบขอบพระคุณรองศาสตราจารย์ ดร. ม.ร.ว. สมพร สุทัศนีย์ สำหรับการให้คำแนะนำ ตรวจสอบข้อบกพร่องต่าง ๆ ด้วยความละเอียดถี่ถ้วน รวมถึงความเอาใจใส่ ช่วยเหลือในทุกขั้นตอนของการทำดุชฎินิพนธ์ ผู้วิจัยรู้สึกซาบซึ้งและระลึกถึงพระคุณอันหาที่เปรียบมิได้ในครั้งนี้เสมอ จึงขอกราบขอบพระคุณเป็นอย่างสูงไว้ ณ โอกาสนี้

ขอขอบพระคุณผู้ช่วยศาสตราจารย์ ดร.สุชาติดา กรเพชรปानी สำหรับการสนับสนุน และการให้แนวคิดในการทำงานอย่างเป็นระบบ ทำให้ผู้วิจัยมีความมุ่งมั่นในการทำงานจนประสบความสำเร็จ ขอขอบพระคุณ Professor Dr. Larry R. Nelson สำหรับคำปรึกษาและการให้ความช่วยเหลือ สนับสนุน และผลักดันให้ผู้วิจัยสามารถดึงศักยภาพของตนมาใช้ได้อย่างเต็มที่ อีกทั้งเมตตาอนุเคราะห์โปรแกรม LERTAP ให้ใช้สำหรับการวิเคราะห์ข้อมูลโดยไม่เสียค่าใช้จ่าย นอกจากนี้ ผู้วิจัยขอขอบพระคุณ ผู้ช่วยศาสตราจารย์ ดร.จตุภัทร เมฆพ่ายพ สำหรับการให้คำแนะนำช่วยเหลือทางด้านสถิติและด้านวิชาการอื่น ๆ ขอขอบพระคุณผู้เชี่ยวชาญและผู้ทรงคุณวุฒิทุกท่าน ที่ให้ความอนุเคราะห์ในการตรวจสอบ และให้คำแนะนำ ทั้งในส่วนของเครื่องมือที่ใช้ในการวิจัย และโปรแกรมที่พัฒนาขึ้น ทำให้มั่นใจในคุณภาพของผลงานวิจัยที่ได้มากยิ่งขึ้น นอกจากนี้ ขอขอบพระคุณบุคลากรทุกท่านและหน่วยงานที่เกี่ยวข้องทุกแห่งสำหรับการอำนวยความสะดวกตลอดการดำเนินการวิจัย และที่ขาดไม่ได้ คือ ผู้เข้าร่วมการวิจัยทุกท่านสำหรับการเสียสละเวลาอันมีค่าเพื่อการให้ข้อมูลสำคัญสำหรับการศึกษาในครั้งนี้

ขอกราบขอบพระคุณคุณแม่ พี่สาว และครอบครัวสำหรับการสนับสนุนทุนการศึกษา รวมถึงการให้กำลังใจและความเข้าใจอย่างสม่ำเสมอ ทำให้ผู้วิจัยสามารถทำสิ่งที่ฝันสำเร็จโดยสมบูรณ์ นอกจากนี้ ขอขอบคุณ คุณปิยะพงษ์ ต้นศิริ ที่เสียสละเวลา เป็นที่ปรึกษา และช่วยเหลือด้านโปรแกรมคอมพิวเตอร์ซึ่งเป็นหัวใจสำคัญของงานวิจัยครั้งนี้ ขอขอบคุณมิตรภาพของเพื่อนทุกกลุ่ม สำหรับการสนับสนุนและกระตุ้นให้ผู้วิจัยมีแรงกาย แรงใจมากเพียงพอที่จะดำเนินการวิจัยจนผ่านพ้นไปได้ด้วยดี ผู้วิจัยขอขอบคุณในมิตรภาพที่มอบให้เสมอมา

ขอขอบคุณการสนับสนุนทุนสำหรับโปรแกรมวิเคราะห์ข้อมูลทางการวัดและการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์จาก Applied Psychological Measurement Inc.

คุณค่าและประโยชน์ของดุชฎินิพนธ์ฉบับนี้ ผู้วิจัยขอมอบแด่บุพการี บุรพาจารย์ และผู้มีพระคุณทุกท่านทั้งในอดีตและปัจจุบัน ที่ทำให้ข้าพเจ้าเป็นผู้มีการศึกษา และประสบความสำเร็จมาจนตราบเท่าทุกวันนี้

สุชาติดา สกลกิจรุ่งโรจน์

54810038: สาขาวิชา: การวัดและเทคโนโลยีทางวิทยาการปัญญา

ปร.ด. (การวัดและเทคโนโลยีทางวิทยาการปัญญา)

คำสำคัญ: มาตรฐานวัดความสุขของคนไทย/ การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์/
โมเดลการตอบสนองข้อสอบ

สุชาดา สกลกิจรุ่งโรจน์: การประยุกต์โมเดลการตอบสนองข้อสอบในการพัฒนามาตรวัด
ความสุขของคนไทย: การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ (APPLYING ITEM RESPONSE
MODEL FOR DEVELOPING THE THAI HAPPINESS SCALE: COMPUTERIZED ADAPTIVE
TESTING) คณะกรรมการควบคุมคุชณินิพนธ์: เสรี ชัดเข้ม, ค.ด., ม.ร.ว. สมพร สุทัศน์ีย์, ค.ด.
278 หน้า. ปี พ.ศ. 2558.

การวิจัยนี้มีวัตถุประสงค์เพื่อสร้างคลังข้อคำถามที่มีคุณภาพ สำหรับวัดความสุขของคนไทย
พัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับมาตรฐานวัดความสุขของคนไทย
และศึกษาความสัมพันธ์ของผลการประมาณค่าความสุขที่ได้จากการทดสอบแบบปรับเหมาะด้วย
คอมพิวเตอร์กับการทดสอบด้วยข้อคำถามทั้งหมดในคลัง การดำเนินการวิจัยแบ่งเป็น 3 ขั้นตอน
ได้แก่ 1) สร้างคลังข้อคำถามสำหรับมาตรฐานวัดความสุขของคนไทย โดยตรวจสอบความตรงเชิงเนื้อหา
วิเคราะห์คุณภาพรายข้อตามทฤษฎีการตอบสนองข้อสอบ และวิเคราะห์ความตรงเชิงโครงสร้างของ
ข้อคำถามในคลังข้อคำถาม โดยศึกษาในกลุ่มตัวอย่างที่เป็นคนไทย อายุ 18 ปีขึ้นไป จำนวนทั้งสิ้น
934 คน 2) พัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับมาตรฐานวัดความสุข
ของคนไทย ในรูปแบบเว็บแอปพลิเคชัน และ 3) ศึกษาความสัมพันธ์ของผลการประมาณค่าความสุข
ที่ได้จากโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรฐานวัดความสุขของคนไทย
กับการทดสอบด้วยข้อคำถามทั้งหมดในคลัง โดยการคำนวณค่าสัมประสิทธิ์สหสัมพันธ์เพียร์สัน

ผลการวิจัย ปรากฏว่า คลังข้อคำถามสำหรับมาตรฐานวัดความสุขของคนไทยมีข้อคำถาม
จำนวนทั้งสิ้น 61 ข้อ (มาตรฐานวัดย่อยความพึงพอใจในชีวิต 13 ข้อ อารมณ์ความรู้สึกเชิงบวก 22 ข้อ
และอารมณ์ความรู้สึกเชิงลบ 26 ข้อ) ข้อคำถามทุกข้อผ่านเกณฑ์การตรวจสอบความตรงเชิงเนื้อหา
และผ่านการวิเคราะห์คุณภาพรายข้อตามทฤษฎีการตอบสนองข้อสอบ การตรวจสอบความตรง
เชิงโครงสร้าง ปรากฏว่า ข้อคำถามวัดองค์ประกอบตามทฤษฎีความอยู่ดีมีสุขเชิงอัตวิสัย โปรแกรม
การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับมาตรฐานวัดความสุขของคนไทยช่วยลดระยะเวลา
และจำนวนข้อคำถามที่ใช้ในการทดสอบได้ ค่าความสุขที่ประมาณค่าได้จากโปรแกรมการทดสอบ
แบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรฐานวัดความสุขของคนไทย กับการทดสอบด้วยข้อคำถาม
ทั้งหมดในคลังมีความสัมพันธ์กันทางบวก มีค่าสัมประสิทธิ์สหสัมพันธ์ เท่ากับ 0.85 ที่ระดับนัยสำคัญ
ทางสถิติ .01 ซึ่งให้เห็นว่า มาตรฐานวัดความสุขของคนไทย รูปแบบการทดสอบแบบปรับเหมาะด้วย
คอมพิวเตอร์เป็นอีกทางเลือกหนึ่งในการทดสอบได้

54810038: MAJOR: MEASUREMENT AND TECHNOLOGY IN COGNITIVE SCIENCE
Ph.D. (MEASUREMENT AND TECHNOLOGY IN COGNITIVE SCIENCE)

KEYWORDS: THAI HAPPINESS SCALE/ COMPUTERIZED ADAPTIVE TESTING/
ITEM RESPONSE MODEL

SUCHADA SAKOLKIJRUNGROJ: APPLYING ITEM RESPONSE MODEL FOR
DEVELOPING THE THAI HAPPINESS SCALE: COMPUTERIZED ADAPTIVE TESTING.
ADVISORY COMMITTEE: SEREE CHADCHAM, Ph.D., M.R. SOMPORN SUDHASANI, Ph.D.
278 P. 2015.

The goals of this research were to construct a good quality item bank for a Thai Happiness Scale, to develop a Computerized Adaptive Testing Program for the Thai Happiness Scale (THS-CAT), and to study the relationship between happiness scores derived from THS-CAT and the test comprising the whole item bank. The research process comprised-3 phases: 1) the happiness item bank was constructed and all items were analyzed using the Item Response Theory (IRT). The sample comprised 934 Thai aged 18 and above, 2) the development of the THS-CAT in web application format, and 3) the study of the relationship between happiness scores derived from the THS-CAT and the test that used the whole item bank using Pearson correlation coefficient.

The results reveal that the final item bank includes 61 items: 13 items measuring life satisfaction, 22 items about positive affect, and 26 items measuring negative affect. All items showed good quality. The item bank for the Thai Happiness Scale consisted of 3 domains as suggested by subjective well-being theory. The THS-CAT consumed less time and items for testing compared to the test using the whole item bank. The correlation coefficients between happiness scores derived from the THS-CAT and the test using the whole item bank equaled 0.85 ($p < .01$) showing a positive relationship. Hence, the THS-CAT is a good alternative way of testing that can assess happiness effectively.

สารบัญ

	หน้า
บทคัดย่อภาษาไทย.....	ง
บทคัดย่อภาษาอังกฤษ.....	จ
สารบัญ.....	ฉ
สารบัญตาราง.....	ช
สารบัญภาพ.....	ฎ
บทที่	
1 บทนำ.....	1
ความเป็นมาและความสำคัญของปัญหา.....	1
วัตถุประสงค์ของการวิจัย.....	7
สมมติฐานการวิจัย.....	7
กรอบแนวคิดในการวิจัย.....	7
ประโยชน์ที่ได้รับจากการวิจัย.....	8
ขอบเขตของการวิจัย.....	9
นิยามศัพท์เฉพาะ.....	9
2 เอกสารและงานวิจัยที่เกี่ยวข้อง.....	13
ตอนที่ 1 ทฤษฎีเกี่ยวกับความสุขและงานวิจัยที่เกี่ยวข้อง.....	13
ตอนที่ 2 ทฤษฎีการตอบสนองข้อสอบและงานวิจัยที่เกี่ยวข้อง.....	23
ตอนที่ 3 การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์และงานวิจัยที่เกี่ยวข้อง.....	54
3 วิธีดำเนินการวิจัย.....	76
ขั้นตอนที่ 1 การจัดทำคลังข้อคำถามสำหรับมาตรวัดความสุขของคนไทย รูปแบบ การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์.....	77
ขั้นตอนที่ 2 การพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย.....	88
ขั้นตอนที่ 3 การศึกษาความสัมพันธ์ของผลการประมาณค่าความสุขที่ได้จาก การทดสอบด้วยโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับ มาตรวัดความสุขของคนไทย กักับการทดสอบด้วยข้อคำถามทั้งหมดในคลังข้อคำถาม...	107
4 ผลการวิจัย.....	109
ตอนที่ 1 ผลการสร้างคลังข้อคำถามสำหรับมาตรวัดความสุขของคนไทย รูปแบบ การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์.....	110
ตอนที่ 2 ผลการพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย.....	129

สารบัญ (ต่อ)

บทที่	หน้า
ตอนที่ 3 ผลการศึกษาความสัมพันธ์ของผลการประมาณค่าความสุขที่ได้จากการทดสอบด้วยโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย กับ การทดสอบด้วยข้อคำถามทั้งหมดในคลังข้อคำถาม...	144
5 สรุปผล และอภิปรายผล.....	147
สรุปผลการวิจัย.....	147
อภิปรายผล.....	149
ข้อเสนอแนะในการนำผลการวิจัยไปใช้.....	156
ข้อเสนอแนะสำหรับการวิจัยต่อไป.....	156
บรรณานุกรม.....	158
ภาคผนวก.....	164
ภาคผนวก ก รายชื่อผู้เชี่ยวชาญ.....	165
ภาคผนวก ข การพัฒนามาตรวัดความสุขของคนไทย รูปแบบกระดาษ-ดินสอ.....	167
ภาคผนวก ค ผลการสร้างคลังข้อคำถาม สำหรับมาตรวัดความสุขของคนไทย.....	213
ภาคผนวก ง คู่มือการใช้โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับมาตรวัดความสุขของคนไทย.....	243
ภาคผนวก จ แบบประเมินความเหมาะสมของโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย.....	269
ภาคผนวก ฉ ผลการประเมินความเหมาะสมของโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย โดยผู้เชี่ยวชาญเป็นรายข้อ.....	274
ภาคผนวก ช ผลการประเมินความเหมาะสมของโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย โดยผู้ใช้งานเป็นรายข้อ.....	276
ประวัติย่อของผู้วิจัย.....	278

สารบัญตาราง

ตารางที่	หน้า
2-1 การเทียบเคียงรายละเอียดเฉพาะของ Two-Parameter Logistic Model (2PL) กับ Graded Response Model (GRM) และ Generalized Partial Credit Model (G-PCM).....	45
2-2 จำนวนข้อคำถามในคลังข้อคำถามของมาตรวัดต่าง ๆ ที่ใช้กับโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์.....	59
2-3 วิธีการประมาณค่าคุณลักษณะแฝงของมาตรวัดต่าง ๆ ที่ใช้กับโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์.....	65
2-4 เกณฑ์การพิจารณาการสิ้นสุดการทดสอบของมาตรวัดต่าง ๆ ที่ใช้กับโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์.....	66
3-1 ความหมายและสัญลักษณ์ที่ใช้ในการเขียนแผนภาพกระแสข้อมูล.....	89
3-2 เอนทิตี แอตทริบิวต์ และคีย์หลักของโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์.....	98
4-1 จำนวนข้อคำถามที่สร้างขึ้นตามทฤษฎีความอยู่ดีมีสุขเชิงอัตวิสัย.....	110
4-2 ผลการคัดเลือกข้อคำถามจากการประเมินความตรงเชิงเนื้อหาโดยผู้เชี่ยวชาญ.....	111
4-3 ผลการคัดเลือกข้อคำถามจากการทดลองใช้.....	112
4-4 ลักษณะของกลุ่มตัวอย่าง.....	112
4-5 ค่า Item-Total Correlations ของข้อคำถาม ในแต่ละมาตรวัดย่อย และมาตรวัดความสุขโดยรวม.....	114
4-6 ผลการคัดเลือกข้อคำถาม จากค่า Item-Total Correlations.....	114
4-7 ดัชนีความสอดคล้องของข้อมูลเชิงประจักษ์กับโมเดลองค์ประกอบเชิงยืนยัน ในการตรวจสอบความเป็นเอกมิตี.....	115
4-8 ค่าสหสัมพันธ์พิเศษเหลือจากการวิเคราะห์องค์ประกอบเชิงยืนยันมาตรวัดความสุขของคนไทย.....	116
4-9 ผลการคัดเลือกข้อคำถาม ตามความสอดคล้องของข้อมูลเชิงประจักษ์ที่ได้จากการตอบข้อคำถามกับ Grade Response Model.....	116
4-10 ผลการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองของมาตรวัดความสุข ข้อคำถาม 61 ข้อ (ขั้นที่ 1).....	118
4-11 ผลการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองของมาตรวัดความสุข ข้อคำถาม 61 ข้อ (ขั้นที่ 2).....	120
4-12 ข้อคำถาม ค่าพารามิเตอร์ความชัน และค่าพารามิเตอร์เทรซโฮลด์ของข้อคำถามในคลังข้อคำถามสุดท้าย.....	121
4-13 เกณฑ์การแปลความหมายคะแนนที่ ได้จากโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์.....	135

สารบัญตาราง (ต่อ)

ตารางที่	หน้า
4-14 ผลการประเมินความเหมาะสมของโปรแกรมการทดสอบแบบปรับเหมาะด้วย คอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย โดยผู้เชี่ยวชาญ.....	142
4-15 ผลการประเมินความเหมาะสมของโปรแกรมการทดสอบแบบปรับเหมาะด้วย คอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย โดยผู้ใช้งาน.....	143
4-16 ระยะเวลาที่ใช้ในการทดสอบ (วินาที) จำนวนข้อคำถามที่ใช้ในการทดสอบ (ข้อ) ค่าความสุขที่ประมาณค่าได้ และคะแนนที่ได้จากการทดสอบแบบปรับเหมาะด้วย คอมพิวเตอร์ กับการทดสอบด้วยข้อคำถามทั้งหมดในคลังข้อคำถามของกลุ่มตัวอย่าง (n=30).....	144
4-17 ค่าสัมประสิทธิ์สหสัมพันธ์เพียร์สันระหว่างผลการประมาณค่าความสุข (Theta) ที่ได้ จากการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ กับการทดสอบด้วยข้อคำถาม ทั้งหมดในคลังข้อคำถาม.....	146
ข-1 การเปรียบเทียบคะแนนสเตโนกับตำแหน่งเปอร์เซนไทล์.....	169
ข-2 ผลการวิเคราะห์ค่าสถิติพื้นฐานของมาตรวัดความสุขของคนไทย (61 ข้อ).....	169
ข-3 ความเที่ยงของมาตรวัดย่อยและมาตรวัดความสุขโดยรวม.....	170
ข-4 เกณฑ์สำหรับเปรียบเทียบคะแนนดิบกับตำแหน่งเปอร์เซนไทล์ คะแนนซี คะแนนที และคะแนนสเตโนของมาตรวัดย่อยความพึงพอใจในชีวิต (n=934).....	171
ข-5 เกณฑ์การเทียบช่วงคะแนนดิบ กับระดับความพึงพอใจในชีวิต.....	172
ข-6 เกณฑ์สำหรับเปรียบเทียบคะแนนดิบกับตำแหน่งเปอร์เซนไทล์ คะแนนซี คะแนนที และคะแนนสเตโนของมาตรวัดย่อยอารมณ์ความรู้สึกเชิงบวก (n=934).....	173
ข-7 เกณฑ์การเทียบช่วงคะแนนดิบ กับระดับอารมณ์ความรู้สึกเชิงบวก.....	175
ข-8 เกณฑ์สำหรับเปรียบเทียบคะแนนดิบกับตำแหน่งเปอร์เซนไทล์ คะแนนซี คะแนนที และคะแนนสเตโนของมาตรวัดย่อยอารมณ์ความรู้สึกเชิงลบ (n=934).....	176
ข-9 เกณฑ์การเทียบช่วงคะแนนดิบ กับระดับอารมณ์ความรู้สึกเชิงลบ.....	179
ข-10 เกณฑ์สำหรับเปรียบเทียบคะแนนดิบกับตำแหน่งเปอร์เซนไทล์ คะแนนซี คะแนนที และคะแนนสเตโนของมาตรวัดความสุขโดยรวม (n=934).....	178
ข-11 เกณฑ์การเทียบช่วงคะแนนดิบ กับระดับความสุขโดยรวม.....	184
ค-1 ค่าเฉลี่ยความสอดคล้องของข้อคำถามกับนิยามเชิงปฏิบัติการ.....	214
ค-2 ค่า Item-Total Correlations รายข้อ จากการทดลองใช้ (30 คน).....	221
ค-3 ค่า Item-Total Correlations รายข้อ และความถี่ของการเลือกรายการคำตอบใน แต่ละข้อคำถามจากกลุ่มตัวอย่าง (934 คน).....	228
ค-4 ค่าพารามิเตอร์ของข้อคำถามและค่าความสอดคล้องของข้อมูลเชิงประจักษ์ที่ได้จาก การตอบข้อคำถามกับ Graded Response Model.....	235

สารบัญตาราง (ต่อ)

ตารางที่	หน้า
ฉ-1 ผลการประเมินความเหมาะสมของโปรแกรมการทดสอบแบบปรับเหมาะด้วย คอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย โดยผู้เชี่ยวชาญ เป็นรายข้อ.....	275
ช-1 ผลการประเมินความเหมาะสมของโปรแกรมการทดสอบแบบปรับเหมาะด้วย คอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย โดยใช้งาน เป็นรายข้อ.....	277

สารบัญญภาพ

ภาพที่	หน้า
1-1 กรอบแนวคิดการประยุกต์โมเดลการตอบสนองข้อสอบในการพัฒนาโปรแกรม การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย.....	8
2-1 โมเดลการทดสอบแบบดั้งเดิม.....	25
2-2 โมเดลการตอบสนองข้อสอบ	25
2-3 โค้งลักษณะข้อสอบของโมเดลโลจิสแบบหนึ่งพารามิเตอร์หรือราส์ชโมเดล.....	30
2-4 โค้งลักษณะข้อสอบของโมเดลโลจิสแบบสองพารามิเตอร์.....	32
2-5 โค้งลักษณะข้อสอบของโมเดลโลจิสแบบสามพารามิเตอร์.....	34
2-6 โค้งลักษณะปฏิบัติการของข้อคำถามที่มีตัวเลือกรายการคำตอบ 5 รายการ.....	41
2-7 โค้งการเลือกรายการคำตอบของข้อคำถามที่มีตัวเลือกรายการคำตอบ 5 รายการ.....	41
2-8 ตัวอย่างสารสนเทศของข้อคำถาม 3 ข้อ จากมาตรวัดภาวะซึมเศร้า.....	50
2-9 ฟังก์ชันสารสนเทศของแบบสอบและความคลาดเคลื่อนมาตรฐานในการประมาณค่า.....	52
2-10 กระบวนการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์.....	56
2-11 ขั้นตอนในการพัฒนามาตรวัดตัวแปรทางจิตวิทยาในรูปแบบการทดสอบแบบ ปรับเหมาะด้วยคอมพิวเตอร์.....	72
3-1 สรุปขั้นตอนหลักของการดำเนินการวิจัย.....	77
3-2 ขั้นตอนการสร้างข้อคำถามสำหรับวัดความสุขของคนไทย.....	80
3-3 ขั้นตอนการตรวจสอบคุณภาพรายข้อของข้อคำถาม.....	81
3-4 ขั้นตอนการตรวจสอบความตรงเชิงเนื้อหา.....	82
3-5 ขั้นตอนการทดลองใช้.....	83
3-6 ขั้นตอนการเก็บรวบรวมข้อมูล.....	85
3-7 ขั้นตอนการวิเคราะห์ข้อมูล.....	87
3-8 วงจรการพัฒนาาระบบ (System Development Life Cycle: SDLC).....	88
3-9 แผนภาพบริบทของโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับ มาตรวัดความสุขของคนไทย.....	90
3-10 แผนภาพกระแสข้อมูล ระดับที่ 0 ของโปรแกรมการทดสอบแบบปรับเหมาะด้วย คอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย.....	91
3-11 แผนภาพกระแสข้อมูล ระดับที่ 1 ของกระบวนการจัดการคลังข้อคำถาม.....	93
3-12 แผนภาพกระแสข้อมูล ระดับที่ 1 ของกระบวนการจัดการการทดสอบแบบปรับเหมาะ ด้วยคอมพิวเตอร์.....	95
3-13 แผนภาพ E-R (Entity Relationship Diagram) ของโปรแกรมการทดสอบแบบ ปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย.....	97
3-14 โครงสร้างหน้าจอรายงานผลการทดสอบสำหรับผู้ทดสอบ.....	100
3-15 โครงสร้างหน้าจอรายงานผลการทดสอบสำหรับผู้ดูแลระบบ หน้าจอที่ 1	100

สารบัญภาพ (ต่อ)

ภาพที่	หน้า
3-16	101
3-17	101
3-18	102
3-19	103
3-20	103
3-21	108
4-1	117
4-2	124
4-3	125
4-4	126
4-5	126
4-6	127
4-7	128
4-8	130
4-9	130
4-10	132
4-11	133
4-12	134
4-13	136
4-14	136
4-15	137
4-16	138

สารบัญภาพ (ต่อ)

ภาพที่	หน้า
4-17 หน้าจอการจัดการคลังข้อความ.....	138
4-18 หน้าจอการจัดการมาตรวัดย่อย.....	139
4-19 หน้าจอการจัดการข้อความ.....	140
4-20 หน้าจอการจัดการประวัติการทดสอบ.....	141

บทที่ 1

บทนำ

ความเป็นมาและความสำคัญของปัญหา

Seligman (1998) อดีตนักจิตวิทยาอเมริกัน (American Psychological Association) ได้กล่าวถึงวงการจิตวิทยาในช่วงก่อนการเกิดสงครามโลกครั้งที่ 2 ว่า วงการจิตวิทยามีพันธกิจที่สำคัญ 3 ประการ ได้แก่ 1) การบำบัดรักษาภาวะเจ็บป่วยทางด้านจิตใจ (Curing Mental Illness) 2) การเติมเต็มชีวิตความเป็นอยู่ของมนุษย์ทุกคนให้มีความสมบูรณ์มากยิ่งขึ้น (Making the Lives of all People More Fulfilling) และ 3) การค้นหาและส่งเสริมทักษะความสามารถของบุคคลที่มีความสามารถพิเศษ (Identifying and Nurturing High Talent) แต่หลังจากเกิดสงครามโลกครั้งที่ 2 ความสนใจหลักของวงการจิตวิทยาได้มุ่งไปที่ การบำบัดรักษาเป็นส่วนใหญ่ เนื่องจากผู้คนได้รับผลกระทบทางด้านจิตใจจากสงคราม ทำให้การศึกษาวิจัยที่เกิดขึ้นในช่วงเวลาดังกล่าวมุ่งเน้นที่ การทำความเข้าใจความผิดปกติของจิตใจ และค้นหาแนวทางการช่วยเหลือที่เหมาะสม จนกระทั่งพันธกิจที่สำคัญอีกสองประการได้ถูกหลงลืมไป ทำให้สังคมมีความเข้าใจคลาดเคลื่อนว่า จิตวิทยาเป็นศาสตร์ที่ศึกษาเกี่ยวกับความเจ็บป่วยและความบกพร่องของมนุษย์ ทั้งที่ในความเป็นจริงแล้ว จิตวิทยาเป็นศาสตร์ที่ศึกษาการเสริมสร้างสิ่งที่ดีในตัวตนของมนุษย์ รวมถึงส่งเสริมศักยภาพของมนุษย์ด้วย ดังนั้น การส่งเสริมป้องกันทางด้านสุขภาพจิตจึงมีความสำคัญไม่น้อยกว่าการบำบัดรักษา จากทัศนะดังกล่าวก่อให้เกิดการเปลี่ยนแปลงในวงการจิตวิทยาอย่างกว้างขวาง การศึกษาวิจัยในระยะต่อมาจึงให้ความสำคัญกับจิตวิทยาเชิงบวก (Positive Psychology) มากยิ่งขึ้น จิตวิทยาเชิงบวกเป็นศาสตร์ที่ศึกษาเกี่ยวกับความเข้มแข็งและจิตวิญญาณของมนุษย์ ซึ่งมุ่งศึกษาและทำความเข้าใจบุคคลปกติ เพื่อค้นหาปัจจัยที่ทำให้บุคคลเหล่านั้นมีคุณภาพชีวิตที่ดีและเหมาะสม รวมถึงค้นหาปัจจัยที่ส่งผลให้บุคคลสามารถพัฒนาคุณภาพชีวิตของตนให้ดียิ่งขึ้น จิตวิทยาเชิงบวกต้องการศึกษาศักยภาพของมนุษย์ในการปรับตัว การสร้างทักษะการเรียนรู้สำหรับก้าวข้ามความยากลำบากของชีวิต และการมุ่งไปสู่เป้าหมายที่ตนเองตั้งไว้ได้ ดังนั้น จิตวิทยาเชิงบวกจึงเป็นมุมมองที่เปิดกว้างและให้ความสำคัญต่อศักยภาพ แรงจูงใจ และความสามารถของมนุษย์ ประเด็นที่เกี่ยวข้องกับจิตวิทยาเชิงบวกมีหลากหลาย เช่น ความเสียสละ (Altruism) ความเห็นอกเห็นใจ (Empathy) ความคิดสร้างสรรค์ (Creativity) การให้อภัย (Forgiveness) (Compton & Hoffman, 2013, pp. 1-2) อีกหนึ่งประเด็นที่เกี่ยวข้องกับจิตวิทยาเชิงบวกซึ่งได้รับความสนใจจากนักวิจัยและนักจิตวิทยา คือ ความสุข (Happiness)

สมัชชาใหญ่แห่งสหประชาชาติ (United Nations General Assembly) ได้ประชุมร่วมกันในวันที่ 12 กรกฎาคม พ.ศ. 2555 และกำหนดให้วันที่ 20 มีนาคมของทุกปีเป็นวันความสุขสากล (International Day of Happiness) ด้วยเล็งเห็นว่า ความสุขเป็นเป้าหมายสำคัญในชีวิตของมนุษย์ จึงได้เชิญชวนประเทศสมาชิกให้ตระหนักถึงความสำคัญดังกล่าว ด้วยการจัดกิจกรรมต่าง ๆ เพื่อส่งเสริมให้บุคคลในสังคมมีความสุขมากยิ่งขึ้น อีกทั้งเพื่อสำรวจและเฝ้าระวังสถานการณ์ความสุข

ในประเทศของตน ประเทศไทยได้ตอบสนองต่อแนวคิดดังกล่าวด้วยการจัดทำโครงการรายงานสถานการณ์สุขภาพจิตคนไทย โดยความร่วมมือของ 4 หน่วยงาน ได้แก่ 1) กรมสุขภาพจิต 2) สำนักงานสถิติแห่งชาติ 3) สถาบันประชากรและสังคม มหาวิทยาลัยมหิดล และ 4) สำนักงานกองทุนสนับสนุนการสร้างเสริมสุขภาพ โครงการดังกล่าวมีวัตถุประสงค์สำคัญ คือ การพัฒนาข้อมูลความสุขคนไทยในระดับจังหวัด เพื่อเป็นกลไกในการขับเคลื่อนการสร้างความสุขทั้งในระดับประเทศและระดับจังหวัดให้มีพลังมากยิ่งขึ้น (ประเวช ตันติพิวัฒนสกุล และเอกอนงค์ สีสลาภินันท์, 2554, หน้า 13-17)

แม้เป็นที่ทราบกันดีว่าความสุขเป็นที่ปรารถนาของมนุษย์ทุกคน แต่การศึกษาวิจัยที่เกี่ยวข้องกับความสุขและอารมณ์เชิงบวกเพิ่งเริ่มต้นขึ้น เมื่อประมาณปี ค.ศ. 1990 ซึ่งนับได้ว่างานวิจัยทางด้านนี้เกิดขึ้นช้ามาก หากเทียบกับงานวิจัยที่เกี่ยวข้องกับอารมณ์เชิงลบ เช่น ความกลัว ความซึมเศร้า ความวิตกกังวล ความโกรธ สาเหตุสำคัญที่ทำให้งานวิจัยด้านความสุขเริ่มต้นช้า อาจเนื่องมาจากเหตุผลหลายประการ ประการแรกอาจเกี่ยวข้องกับมุมมองของนักวิจัยและนักจิตวิทยาที่เล็งเห็นว่า การช่วยเหลือผู้คนให้พ้นจากความรู้สึกเป็นทุกข์และความผิดปกติทางด้านจิตใจ เป็นภาระหน้าที่ที่สำคัญและเร่งด่วนกว่าการค้นหาแนวทางที่ทำให้คนมีความสุขเพิ่มมากขึ้น ประการที่สอง คือ ความสุขมีลักษณะเป็นนามธรรม ส่งผลให้การนิยามความหมายและการวัดเป็นเรื่องยาก ทำให้การศึกษาด้านนี้เกิดขึ้นน้อยกว่าการศึกษาอารมณ์เชิงลบที่สามารถวัดและให้ค่าจำกัดความได้ง่ายกว่า (Shiota & Kalat, 2012, pp. 230-231) อย่างไรก็ตาม ปัจจุบันแนวโน้มการศึกษาด้านความสุขมีจำนวนมากขึ้น โดยมีความพยายามศึกษาและอธิบายความหมายของความสุขในมุมมองที่แตกต่างกัน ซึ่งสามารถจัดได้เป็นกลุ่มทฤษฎีสองกลุ่ม ได้แก่ ทฤษฎีความสุขแบบเอกมิติ (Unidimensional Theories) และทฤษฎีความสุขแบบพหุมิติ (Multidimensional Theories)

ทฤษฎีความสุขแบบเอกมิติ ประกอบด้วยความสุขในสามมุมมองหลัก ได้แก่ 1) มุมมองความสุขแบบเฮโดนิคส์ (Hedonic Perspectives) 2) มุมมองความสุขแบบยูไดโมนิกส์ (Eudaimonic Perspectives) และ 3) มุมมองความสุขด้านความผูกพัน (Engagement Perspectives) ทฤษฎีความสุขแบบเฮโดนิคส์ให้ความสำคัญกับความรู้สึกพึงพอใจ และองค์ประกอบพื้นฐานที่ก่อให้เกิดชีวิตที่ดี (Compton & Hoffman, 2013, p. 10) โดยระบุว่าความสุข หรือ ความอยู่ดีมีสุขเชิงอัตวิสัย (Subjective Well-Being) เป็นความสุขที่ประกอบด้วยสามองค์ประกอบหลัก ได้แก่ ความพึงพอใจในชีวิต (Life Satisfaction) อารมณ์ความรู้สึกเชิงบวก (Positive Affects) และอารมณ์ความรู้สึกเชิงลบ (Negative Affects) โดยผู้ที่มีความสุข คือ ผู้ที่รู้สึกพึงพอใจในชีวิต มีอารมณ์เชิงบวกเกิดขึ้นบ่อยครั้ง ในขณะที่มีอารมณ์เชิงลบเกิดขึ้นน้อยครั้ง หรือแทบไม่มีเลย (Pavot & Diener, 2013) ทฤษฎีความสุขในกลุ่มที่สองเป็นมุมมองความสุขแบบยูไดโมนิกส์ ซึ่งให้ความสำคัญกับการเข้าถึงศักยภาพสูงสุดในตนเอง ความเจริญเติบโตทางจิตวิญญาณ และการตระหนักรู้ถึงตัวตนที่แท้จริงของตัวเอง ความพยายามศึกษาความสุขด้านนี้เป็นไปอย่างกว้างขวาง และมีการให้คำนิยามที่แตกต่างกันไป เช่น ชาวกรีกโบราณอธิบายความสุขในแง่ภูมิปัญญา (Wisdom) นักอัตถิภาวนิยม (Existentialists) ศึกษาความสุขที่จริงจังผ่านการค้นหาความหมายและความสุขของตนท่ามกลางการใช้ชีวิต นักจิตวิทยาแนวมนุษยนิยม (Humanistic Psychologists) อธิบายความสุขผ่าน

การเข้าถึงความเป็นคนโดยสมบูรณ์ (Self-Actualization) และนักจิตวิทยาเชิงบวก (Positive Psychologists) ศึกษาความสุขในแง่ของความสำเร็จรุ่งเรือง (Flourishing) แม้ทฤษฎีส่วนใหญ่จะมีความเกี่ยวเนื่องกันและทับซ้อนกันในบางประเด็น แต่ทฤษฎีต่าง ๆ ก็ล้วนมีเอกลักษณ์ด้านแนวคิดที่แตกต่างกันไป ทฤษฎีความสุขในมุมมองที่สาม คือ มุมมองความสุขด้านความผูกพัน ซึ่งให้ความสำคัญกับอารมณ์ความรู้สึกน้อยกว่าทั้งสองทฤษฎีที่ผ่านมา แต่เน้นเรื่องความเกี่ยวข้องสัมพันธ์ของบุคคลที่มีต่อกิจกรรมที่กระทำ โดยมีหลักการว่า ความสุข คือ กระบวนการของการซึมซับและเรียนรู้สิ่งต่าง ๆ ผ่านสิ่งที่ทำ แต่ละบุคคลอาจมีกิจกรรมที่เกี่ยวข้องแตกต่างกัน ความสำคัญจึงไม่ได้อยู่ที่กิจกรรมที่ทำ แต่อยู่ที่กระบวนการเรียนรู้ที่ก่อให้เกิดการเติมเต็มตนเองที่แตกต่างกัน (Compton & Hoffman, 2013, pp. 42-44)

ทฤษฎีความสุขแบบพหุมิติ เป็นการขยายการอธิบายความสุขในมุมมองที่กว้างขึ้น เพื่อให้เกิดความเข้าใจความสุขอย่างสมบูรณ์ ทฤษฎีที่สำคัญ ได้แก่ ทฤษฎีอัตลิติต (Self-Determination Theory) อธิบายว่า มนุษย์มีความต้องการพื้นฐานที่สำคัญสามประการ ได้แก่ ต้องการเป็นตัวของตัวเอง (Autonomy) ต้องการมั่นใจว่าตนมีความสามารถ (Competency) และต้องการสร้างความสัมพันธ์กับผู้อื่น (Relatedness) เมื่อความต้องการทั้งสามประการนี้ได้รับการตอบสนองอย่างเหมาะสม มนุษย์จะมีแรงจูงใจ และมีความรู้สึกเป็นสุข (Deci & Ryan, 2008) ทฤษฎีความสุขที่แท้จริง (Authentic Happiness) อธิบายความสุขในสามองค์ประกอบหลัก ได้แก่ การมีชีวิตที่พึงพอใจ (Pleasant Life) การมีชีวิตที่ดี (Good Life) และการมีชีวิตที่มีความหมาย (Meaningful Life) (Seligman, 2002 cited in Compton & Hoffman, 2013, p. 45) ทฤษฎีความอยู่ดีมีสุข (Well-Being Theory) ประกอบด้วยองค์ประกอบสำคัญ 5 ประการ ได้แก่ อารมณ์เชิงบวก (Positive Emotion) ความผูกพัน (Engagement) สัมพันธภาพ (Relationships) ความหมาย (Meaning) และความสำเร็จ (Accomplishment) โดยใช้ตัวย่อแทนองค์ประกอบทั้งหมดว่า PERMA (Seligman, 2011 cited in Compton & Hoffman, 2013, pp. 45-46) ทฤษฎีความอยู่ดีมีสุขเชิงจิตวิทยา (Psychological Well-Being) อธิบายว่า ความสุขประกอบด้วย การยอมรับตนเอง (Self-Acceptance) การเติบโตส่วนบุคคล (Personal Growth) การมีจุดมุ่งหมายในชีวิต (Purpose in Life) การมีสัมพันธภาพที่ดีกับผู้อื่น (Positive Relations with Others) ความรอบรู้ด้านสิ่งแวดล้อม (Environmental Mastery) และความเป็นตัวของตัวเอง (Autonomy) (Ryff, 1989)

ปัจจุบันการศึกษาวิจัยที่เกี่ยวข้องกับความสุขมีจำนวนมาก แตกต่างกันไปตามบริบทที่ผู้วิจัยสนใจศึกษา แต่ไม่ว่างานวิจัยลักษณะใดต่างก็ต้องการเครื่องมือสำหรับวัดระดับความสุข ในต่างประเทศมีการพัฒนามาตรวัดความสุขและตัวแปรที่เกี่ยวข้องกับความสุขจำนวนมาก ได้แก่ 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) State-Trait-Cheerfulness Inventory (Ruch, Kohler, & Van Thriel, 1996) Psychological Well-Being Scales (Ryff & Singer, 1998) Subjective Happiness Scale (Lyubomirsky & Lepper, 1999) Gratitude Questionnaire – 6 (McCullough, Emmons, & Tsang, 2002) Inspiration Scale (Thrash & Elliot, 2003) Flourishing Scale และ Scale of Positive and Negative Experience (Diener et al., 2009)

ในประเทศไทย ศูนย์สร้างเสริมสุขภาพองค์กร ภายใต้สำนักองค์กรสุขภาพ สำนักงานกองทุนสนับสนุนการสร้างเสริมสุขภาพ (สสส.) ได้ร่วมมือกับ สถาบันวิจัยประชากรและสังคม มหาวิทยาลัยมหิดล พัฒนา “แบบสำรวจความสุขด้วยตนเอง: Happinometer” เพื่อใช้วัดคุณภาพชีวิตและความสุขของคนทำงานในองค์กรต่าง ๆ โดยมองว่า ความสุขเกิดขึ้นจากองค์ประกอบ 9 ส่วน ได้แก่ ร่างกายดี ผ่อนคลายดี น้ำใจดี จิตวิญญาณดี ครอบครัวดี สังคมดี ใฝ่รู้ดี สุขภาพเงินดี และการทำงานดี แบบสำรวจนี้แบ่งเป็น 2 ชุด คือ ฉบับเต็มมีข้อคำถามจำนวน 56 ข้อ และฉบับสั้นมีข้อคำถามจำนวน 20 ข้อ (ศิริพันธ์ กิตติสุขสถิต และคณะ, 2555, หน้า 68-87) ในขณะที่ กรมสุขภาพจิตได้พัฒนาแบบประเมินสุขภาพจิตคนไทยสองฉบับ ฉบับเต็มมีข้อคำถามจำนวน 55 ข้อ และฉบับสั้นมีข้อคำถามจำนวน 15 ข้อ โดยเรียกแบบประเมินฉบับสั้นนี้ว่า แบบประเมินความสุขคนไทย โดยมีหลักคิดที่ว่าในบริบทของสังคมไทย สุขภาพจิตและความสุขเป็นเรื่องเดียวกัน ความสุขตามแนวคิดของกรมสุขภาพจิตประกอบด้วยองค์ประกอบด้านความพึงพอใจ ด้านสมรรถภาพจิตใจดี ด้านคุณภาพจิตใจดี และด้านปัจจัยสนับสนุนดี (ประเวช ตันติพิวัฒนสกุล และเอกอนงค์ สีตลาภินันท์, 2554, หน้า 44-48) นอกจากนี้ ชูฤทธิ์ เต็งไตรสรณ์ พระครุชิต คุณวโร และอรุณี วชิราพรทิพย์ (2551) ได้พัฒนาแบบวัดสุขภาพองค์กรรวมแนวพุทธ โดยมีวัตถุประสงค์เพื่อพัฒนาเครื่องมือวัดสุขภาพองค์กรรวมแนวพุทธสำหรับคนไทย จากแนวคิดภาวนา 4 ตามกรอบแนวคิดของพระพรหมคุณาภรณ์ (ป. อ. ปยุตโต) ซึ่งสำรวจภาวะที่สุขสมบูรณ์ทั้งทางร่างกาย จิตใจ สังคม และปัญญา จากตัวอย่างดังกล่าวจะเห็นได้ว่า ลักษณะการสร้างมาตรวัดในประเทศไทยบางมาตรวัดมีการสร้างเป็นสองฉบับ คือ ฉบับเต็มและฉบับสั้น เพื่อให้ผู้ใช้สามารถเลือกใช้ได้ตามความต้องการ มาตรวัดฉบับเต็มส่วนใหญ่มีเนื้อหาครอบคลุมสิ่งที่ต้องการวัดและสามารถแปลความหมายผลการวัดได้มากกว่าฉบับสั้น แต่ด้วยข้อจำกัดด้านเวลา ในบางบริบทจึงจำเป็นต้องใช้มาตรวัดฉบับสั้นที่สามารถวัดตัวแปรหลักที่ต้องการวัดได้สะดวกยิ่งขึ้น แต่มีข้อจำกัดด้านความครอบคลุมเนื้อหาและการแปลความหมายของผลการวัดตามมา (ศิริพันธ์ กิตติสุขสถิต และคณะ, 2555, หน้า 47)

อย่างไรก็ดี ด้วยความก้าวหน้าของเทคโนโลยีคอมพิวเตอร์ และทฤษฎีการตอบสนองข้อสอบ (Item Response Theory) จึงก่อให้เกิดทางเลือกของการวัดอีกรูปแบบหนึ่ง คือ การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์บนพื้นฐานทฤษฎีการตอบสนองข้อสอบ (IRT-based Computerized Adaptive Testing) การทดสอบรูปแบบนี้สามารถคัดเลือกชุดของข้อคำถามให้มีความเหมาะสมกับแต่ละบุคคล โดยใช้ระยะเวลาในการทดสอบลดลง แต่มีความแม่นยำ (Precision) ในการทดสอบตามที่กำหนด (Emberson & Reise, 2000, pp. 316-318)

ทฤษฎีการตอบสนองข้อสอบ (Item Response Theory: IRT) เป็นทฤษฎีการทดสอบแนวใหม่ซึ่งได้รับความนิยมและมีการนำไปใช้เพื่อพัฒนาคุณภาพของมาตรวัดอย่างแพร่หลาย ทฤษฎีนี้มุ่งศึกษาคุณลักษณะต่าง ๆ ที่ซ่อนอยู่ภายในตัวบุคคล ซึ่งเรียกว่า คุณลักษณะแฝง (Traits หรือ Latent traits) ด้วยโมเดลคณิตศาสตร์ที่แสดงความสัมพันธ์ระหว่างความน่าจะเป็น (Probability) ในการตอบข้อสอบถูก กับคุณลักษณะแฝงที่มีอิทธิพลอยู่เบื้องหลังการตอบข้อคำถามนั้น (Hambleton, Swaminathan, & Rogers, 1991, p. 7; Crocker & Algina, 2008, p. 339) ทฤษฎีการตอบสนองข้อสอบมีจุดเด่น ที่ได้เปรียบทฤษฎีการทดสอบแบบดั้งเดิม (Classical Test Theory) หลายประการ ได้แก่ 1) ข้อมูลที่ได้จากการตอบข้อคำถามแต่ละข้อมีความหมายและสามารถบ่งบอกถึงระดับ

คุณลักษณะแฝงของบุคคลได้โดยตรง ขณะที่ทฤษฎีการทดสอบแบบดั้งเดิมนั้น ต้องอาศัยผลรวมคะแนนของข้อคำถามทั้งหมด จึงจะสามารถอธิบายความหมายของผลคะแนนและเชื่อมโยงกับคุณลักษณะแฝงได้ คะแนนที่ได้จากข้อคำถามใดข้อคำถามหนึ่งไม่สามารถอธิบายถึงคุณลักษณะแฝงของบุคคลได้โดยตรง 2) ค่าพารามิเตอร์ของข้อคำถามไม่เปลี่ยนแปลงตามกลุ่มผู้ทดสอบที่เปลี่ยนไป 3) การประมาณค่าคุณลักษณะแฝงของบุคคลไม่เปลี่ยนแปลงตามมาตรวัดที่ใช้ และ 4) การตรวจสอบความเที่ยงในการทดสอบไม่จำกัดอยู่เพียงการสร้างแบบทดสอบคู่ขนาน แต่ความเที่ยงตามแนวคิดของทฤษฎีการตอบสนองข้อสอบนั้นอธิบายได้ด้วยฟังก์ชันสารสนเทศของมาตรวัด (Test Information Function) (Hambleton, Swaminathan & Rogers, 1991, pp. 2-5; Thomas, 2011)

การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ (Computerized Adaptive Testing: CAT) เป็นการทดสอบที่ใช้หลักการของทฤษฎีการตอบสนองข้อสอบ ซึ่งสามารถประมาณค่าคุณลักษณะแฝงของบุคคล (Proficiency หรือ Theta ซึ่งใช้สัญลักษณ์ " θ ") ด้วยข้อมูลที่ได้จากการตอบข้อคำถามเป็นรายข้อ ตัวอย่างเช่น การประเมินภาวะซึมเศร้า ทฤษฎีการตอบสนองข้อสอบมีหลักการว่า การตอบข้อคำถามของผู้ทดสอบในแต่ละข้อ สามารถบ่งบอกถึงระดับความรุนแรงของภาวะซึมเศร้าได้ ข้อมูลที่ได้รับจากการตอบข้อคำถามจึงสามารถนำไปวิเคราะห์ เพื่อประมาณค่าคุณลักษณะแฝงของบุคคลนั้น (ระดับความรุนแรงของภาวะซึมเศร้า) และเลือกข้อคำถามข้อถัดไปที่มีความเหมาะสมกับคุณลักษณะแฝงของผู้ทดสอบมากที่สุด (ข้อคำถามที่มีสารสนเทศใกล้เคียงกับระดับภาวะซึมเศร้าที่ประเมินได้) (Weiss, 2011) ดังนั้น ในกระบวนการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ผู้ทดสอบแต่ละคนจึงได้รับข้อคำถามที่แตกต่างกัน ซึ่งให้ข้อมูลสารสนเทศของข้อสอบสอดคล้องกับระดับคุณลักษณะแฝงของบุคคลนั้นมากที่สุด ส่งผลให้จำนวนข้อคำถามที่ใช้กับผู้ทดสอบแต่ละคนมีจำนวนลดลง ขณะที่ประสิทธิภาพในการทดสอบเพิ่มมากขึ้น เนื่องจากความแม่นยำในการประมาณค่า หรือความเที่ยงของการทดสอบสามารถคำนวณได้โดยตรงจากฟังก์ชันสารสนเทศของมาตรวัด ขณะเดียวกันฟังก์ชันสารสนเทศของมาตรวัดนั้น มีการแปรผกผันกับค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่า (Standard Error of Estimation: SEE) การกำหนดค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่าเป็นเกณฑ์การยุติการทดสอบ จึงทำให้ได้ค่าความเที่ยงในระดับที่ต้องการ นอกจากนี้การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ยังช่วยลดความเหนื่อยล้าจากการทดสอบได้อีกด้วย (Gibbons et al., 2008) การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ จึงเป็นอีกทางเลือกหนึ่งและเป็นมิติใหม่ของการทดสอบ ที่ช่วยให้ประสิทธิภาพในการทดสอบดียิ่งขึ้น

การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์เริ่มต้นใช้ในการประเมินทักษะความสามารถต่าง ๆ เช่น ความสามารถทางด้านภาษา ความสามารถทางการคำนวณ แต่ในช่วงหลายปีที่ผ่านมาได้มีการศึกษา การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ในมิติอื่นมากขึ้น ตัวอย่างเช่น โครงการ "Patient Reported Outcomes Measurement Information System" หรือชื่อย่อว่า "PROMIS" เป็นโครงการที่ได้รับการสนับสนุนทุนดำเนินงานจากสถาบันสุขภาพแห่งชาติ (National Institutes of Health) ซึ่งเป็นหน่วยงานวิจัยในประเทศสหรัฐอเมริกาที่รับผิดชอบด้านงานวิจัยทาง

การแพทย์ PROMIS มีหน้าที่สำคัญในการสร้างระบบที่มีประสิทธิภาพในการประเมินสุขภาพของผู้ป่วยทั้งทางด้านร่างกาย จิตใจ และสังคม ทั้งในเด็กและผู้ใหญ่ พันธกิจที่สำคัญประการหนึ่ง คือ การสร้างเครื่องมือวัดที่ได้มาตรฐาน โดย PROMIS ให้ความสนใจกับการพัฒนามาตรวัดทั้งในส่วนของ การสร้างมาตรวัดในรูปแบบกระดาษ-ดินสอ และการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ปัจจุบัน PROMIS ได้สร้างคลังข้อคำถามซึ่งครอบคลุมการประเมินปัญหาทางด้านสุขภาพที่หลากหลาย เช่น การนอนหลับ ความเจ็บปวด ความวิตกกังวล ภาวะซึมเศร้า ความโกรธ (Patient-Reported Outcome Measurement Information System [PROMIS], 2013)

การวัดคุณลักษณะทางจิตวิทยาซึ่งมีการบริหารการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ มิได้จำกัดอยู่เพียงการวัดในลักษณะเอกมิติเท่านั้น โครงการการสร้างแบบวัดความผิดปกติทางบุคลิกภาพ (CAT-PD Project) เป็นอีกตัวอย่างหนึ่งที่พยายามพัฒนาวิธีการประเมินบุคลิกภาพที่มีประสิทธิภาพเพื่อวัดคุณลักษณะทางจิตวิทยาที่มีลักษณะพหุมิติ ด้วยเล็งเห็นว่าภาวะบุคลิกภาพผิดปกติเป็นภาวะที่พบได้บ่อย และส่งผลต่อความสามารถในการดำเนินชีวิตของบุคคล อีกทั้งเป็นภาวะที่มีความซับซ้อนในการรักษา หากสามารถประเมินได้ว่าผู้ป่วยมีคุณลักษณะเช่นใด จะช่วยให้สามารถดูแลรักษาได้อย่างเหมาะสมยิ่งขึ้น โครงการนี้แบ่งการศึกษาออกเป็น 5 ระยะ ได้แก่ 1) การระบุคุณลักษณะของบุคลิกภาพที่ต้องการศึกษา และการพัฒนาคลังข้อคำถามเริ่มต้น (Initial Item Bank) 2) การเก็บข้อมูลกับผู้ป่วยและชุมชน รวมทั้งการพัฒนามาตรวัดและปรับให้สมบูรณ์ 3) การปรับเทียบมาตรฐาน (Calibration) ของชุดข้อมูลสุดท้ายด้วยทฤษฎีการตอบสนองข้อสอบ และการจำลองสถานการณ์การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ 4) การพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับการประเมินบุคลิกภาพ และ 5) การตรวจสอบคุณภาพของมาตรวัดและโปรแกรมคอมพิวเตอร์ที่พัฒนาขึ้น (Simms et al., 2011)

จากการศึกษางานวิจัยที่เกี่ยวข้องกับมาตรวัดความสุขทั้งในประเทศไทยและต่างประเทศ ปรากฏว่า มาตรวัดความสุขส่วนใหญ่ยังอยู่ในรูปแบบกระดาษ-ดินสอ (Paper and Pencil Test: P&P) และมาตรวัดออนไลน์ (Online Test) ในประเทศไทย ยังไม่มีการนำแนวคิดเรื่องการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์มาใช้กับมาตรวัดทางด้านจิตใจ นอกจากนี้ การวิเคราะห์คุณภาพข้อคำถามส่วนใหญ่อยู่บนพื้นฐานทฤษฎีการทดสอบแบบดั้งเดิม ดังนั้น ผู้วิจัยจึงมีความสนใจที่จะสร้างมาตรวัดความสุขของคนไทยในรูปแบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ซึ่งมีการวิเคราะห์คุณภาพของข้อคำถามตามแนวคิดของทฤษฎีการตอบสนองข้อสอบ เพื่อเป็นอีกทางเลือกหนึ่งของการพัฒนามาตรวัดความสุขที่มีการประยุกต์เทคโนโลยีทางคอมพิวเตอร์และทฤษฎีการตอบสนองข้อสอบมาใช้ ทำให้ได้มาตรวัดความสุขรูปแบบใหม่ที่มีประสิทธิภาพ สำหรับการนำไปใช้สำรวจความสุขของคนไทย หรือเพื่อศึกษาวิจัยในด้านนี้ต่อไป นอกจากนี้ ผลการศึกษาสามารถนำไปเป็นแนวทางในการพัฒนามาตรวัดทางจิตใจด้านอื่น ๆ ซึ่งแสดงให้เห็นถึงการผสมผสานความรู้ในศาสตร์ขั้นสูงทางด้านจิตวิทยา การวัดผล และเทคโนโลยีทางคอมพิวเตอร์ได้อย่างลงตัว ทำให้เกิดการต่อยอดองค์ความรู้ทางวิชาการให้เจริญก้าวหน้ายิ่งขึ้น

วัตถุประสงค์ของการวิจัย

1. เพื่อจัดทำคลังข้อคำถามสำหรับมาตรวัดความสุขของคนไทย รูปแบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์
 - 1.1 เพื่อสร้างข้อคำถามสำหรับวัดความสุขของคนไทย
 - 1.2 เพื่อตรวจสอบคุณภาพรายข้อของข้อคำถาม
 - 1.3 เพื่อตรวจสอบความตรงเชิงโครงสร้างของข้อคำถามในคลังข้อคำถาม
 - 1.4 เพื่อจัดทำคลังข้อคำถามซึ่งประกอบด้วยข้อคำถามที่มีคุณภาพ
2. เพื่อพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย
3. เพื่อศึกษาความสัมพันธ์ของผลการประมาณค่าความสุขที่ได้จากการทดสอบด้วยโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย กับการทดสอบด้วยข้อคำถามทั้งหมดในคลังข้อคำถาม

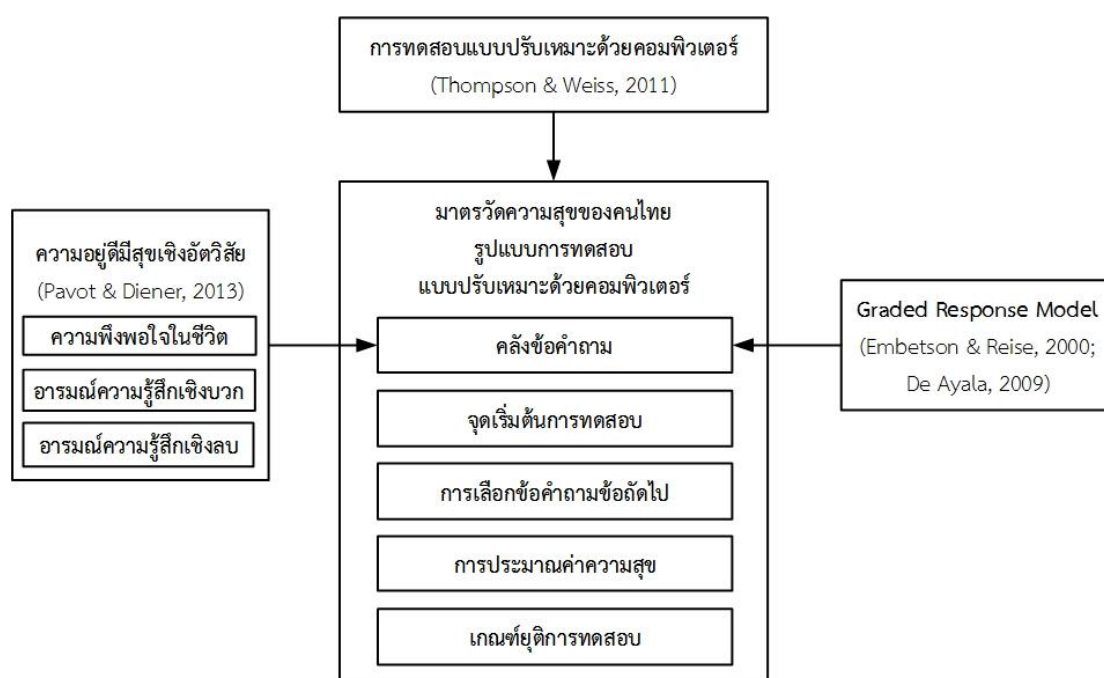
สมมติฐานการวิจัย

สมมติฐานการวิจัยจากวัตถุประสงค์ของการวิจัยข้อที่ 3 มีดังนี้
 ผลการประมาณค่าความสุขที่ได้จากโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย กับผลการประมาณค่าความสุขที่ได้จากการทดสอบด้วยข้อคำถามทั้งหมดในคลังข้อคำถามมีความสัมพันธ์กันทางบวก

กรอบแนวคิดในการวิจัย

การพัฒนามาตรวัดความสุขของคนไทย รูปแบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ เป็นการพัฒนามาตรวัดให้มีประสิทธิภาพในการทดสอบมากขึ้นโดยอาศัยหลักการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ (Thompson & Weiss, 2011) ร่วมกับทฤษฎีการตอบสนองข้อสอบ (Embretson & Reise, 2000, pp. 97-102; De Ayala, 2009, pp. 217-224) ทำให้สามารถจัดการทดสอบด้วยชุดของข้อคำถามที่มีความเฉพาะเจาะจงกับระดับความสุขของแต่ละบุคคล ส่งผลให้จำนวนข้อคำถามที่ใช้ในการทดสอบและระยะเวลาที่ใช้ในการทดสอบลดลง จึงช่วยลดความเหนื่อยล้าในการทดสอบได้เป็นอย่างดี การวิจัยครั้งนี้พัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทยขึ้น โดยมีองค์ประกอบสำคัญของการทดสอบ 5 ประการ ได้แก่ 1) คลังข้อคำถาม สำหรับบรรจุข้อคำถามที่มุ่งวัดความสุขตามทฤษฎีความอยู่ดีมีสุขเชิงอัตวิสัย (Pavot & Diener, 2013) ซึ่งผ่านการวิเคราะห์คุณภาพตามทฤษฎีการตอบสนองข้อสอบ รวมทั้งมีการประมาณค่าพารามิเตอร์ความชันและพารามิเตอร์เทรซโฮลด์ด้วย Graded Response Model สำหรับใช้ในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ 2) จุดเริ่มต้นการทดสอบ (Starting Point) เป็นการคัดเลือกข้อคำถามเริ่มต้นที่ให้ข้อมูลสารสนเทศของผู้ทดสอบสูงที่สุด ณ ตำแหน่งคุณลักษณะแฝง (Theta: θ) ที่มีค่าเท่ากับ 0.00 3) การคัดเลือกข้อคำถามข้อถัดไป (Item Selection Algorithm) เป็นการคัดเลือกข้อคำถามที่ให้ข้อมูลสารสนเทศของผู้ทดสอบสูงที่สุด (Maximum Information) ณ ตำแหน่งคุณลักษณะแฝงที่ประมาณค่าได้ภายหลังการตอบ

ข้อคำถามแต่ละครั้ง 4) การประมาณค่าความสุข (Scoring Algorithm) ใช้วิธีการประมาณค่าแบบ Expected a Posteriori (EAP) และ 5) เกณฑ์ยุติการทดสอบ (Termination Criterion) เป็นการกำหนดให้การทดสอบยุติลงเมื่อความคลาดเคลื่อนมาตรฐานจากการประมาณค่ามีค่าน้อยกว่า 0.30 หรือเมื่อข้อคำถามหมดคลัง ซึ่งสามารถเขียนเป็นกรอบแนวคิดในการวิจัยได้ดังภาพที่ 1-1



ภาพที่ 1-1 กรอบแนวคิดการประยุกต์โมเดลการตอบสนองข้อสอบในการพัฒนามาตรวัดความสุขของคนไทย รูปแบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์

ประโยชน์ที่ได้รับจากการวิจัย

1. ได้คลังข้อคำถามสำหรับวัดความสุขตามทฤษฎีความอยู่ดีมีสุขเชิงอัตวิสัย ซึ่งประกอบด้วยข้อคำถามที่มีคุณภาพ
2. ได้โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทยที่สามารถประมาณค่าความสุขได้อย่างรวดเร็ว และมีความแม่นยำ
3. โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย สามารถนำไปใช้ในการสำรวจระดับความสุขของประชาชนชาวไทย ที่มีอายุตั้งแต่ 18 ปีขึ้นไป บุคคลหรือหน่วยงานที่เกี่ยวข้องสามารถนำผลการสำรวจไปใช้ในการส่งเสริม วางแผนป้องกัน กำหนดนโยบายเพื่อให้ประชาชนมีสุขภาพจิตที่ดีหรือมีความสุขมากยิ่งขึ้น
4. โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย สามารถนำไปใช้เป็นเครื่องมือสำหรับวัดความสุขเพื่อการศึกษาวิจัยในด้านต่าง ๆ เช่น ศึกษาปัจจัยที่มีอิทธิพลต่อความสุข ศึกษาผลของวิธีการหรือโปรแกรมที่พัฒนาขึ้นเพื่อส่งเสริมความสุข

5. กระบวนการที่ใช้ในการประยุกต์ทฤษฎีการตอบสนองข้อสอบในการพัฒนามาตรวัดความสุขของคนไทย รูปแบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สามารถนำไปใช้เป็นแนวทางในการพัฒนามาตรวัดตัวแปรทางจิตวิทยาอื่น ๆ ในรูปแบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ได้ เช่น คุณภาพชีวิต เจตคติ ความเครียด ความซึมเศร้า ความวิตกกังวล

ขอบเขตของการวิจัย

การวิจัยนี้เป็นการประยุกต์โมเดลการตอบสนองข้อสอบในการพัฒนามาตรวัดความสุขของคนไทย รูปแบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ซึ่งประกอบด้วย 5 องค์ประกอบสำคัญ ได้แก่ 1) คลังข้อคำถาม สำหรับบรรจุข้อคำถามที่มุ่งวัดความสุขตามทฤษฎีความอยู่ดีมีสุขเชิงอัตวิสัย ซึ่งผ่านการวิเคราะห์คุณภาพตามทฤษฎีการตอบสนองข้อสอบ รวมทั้งมีการประมาณค่าพารามิเตอร์ความชันและพารามิเตอร์เทรซโฮลต์ด้วย Graded Response Model ซึ่งเป็นค่าที่ใช้ในทุกขั้นตอนของการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ 2) จุดเริ่มต้นการทดสอบ ซึ่งคัดเลือกข้อคำถามเริ่มต้นที่ให้ข้อมูลสารสนเทศของผู้ทดสอบสูงที่สุด ณ ตำแหน่งคุณลักษณะแฝงที่มีค่าเท่ากับ 0.00 3) การคัดเลือกข้อคำถามข้อถัดไป โดยคัดเลือกข้อคำถามที่ให้ข้อมูลสารสนเทศของผู้ทดสอบสูงที่สุด ณ ตำแหน่งคุณลักษณะแฝงที่ประมาณค่าได้ภายหลังการตอบข้อคำถามแต่ละครั้ง 4) การประมาณค่าความสุข ใช้วิธีการประมาณค่าแบบ Expected a Posteriori และ 5) เกณฑ์ยุติการทดสอบ โดยกำหนดให้การทดสอบยุติลงเมื่อความคลาดเคลื่อนมาตรฐานจากการประมาณค่ามีค่าน้อยกว่า 0.30 หรือเมื่อข้อคำถามหมดคลัง การวิจัยนี้มีขอบเขตการวิจัย ดังนี้

1. ประชากรที่ใช้ในการศึกษา เป็นประชาชนทั่วไป ที่มีสัญชาติไทย อายุ 18 ปีขึ้นไป

2. เนื้อหาที่ใช้ในการสร้างข้อคำถามเพื่อวัดความสุขของคนไทย สร้างขึ้นจากนิยามความสุขตามทฤษฎีความอยู่ดีมีสุขเชิงอัตวิสัย ซึ่งกล่าวว่า ความสุขประกอบด้วย 3 องค์ประกอบ ได้แก่ ความพึงพอใจในชีวิต อารมณ์ความรู้สึกเชิงบวก และ อารมณ์ความรู้สึกเชิงลบ คนที่มีความสุข คือ คนที่มีความพึงพอใจในชีวิตของตนเอง มีอารมณ์ความรู้สึกเชิงบวกเกิดขึ้นบ่อยครั้ง และมีอารมณ์ความรู้สึกเชิงลบเกิดขึ้นน้อยครั้งหรือแทบไม่มีเลย

3. ตัวแปรที่ใช้ในการวิจัย มีดังนี้

ตัวแปรที่ใช้ในการวิจัย คือ ผลการประมาณค่าความสุขที่ได้จากการวัดความสุขสองรูปแบบ ได้แก่ 1) การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ และ 2) การทดสอบโดยใช้ข้อคำถามทั้งหมดในคลังข้อคำถาม

นิยามศัพท์เฉพาะ

ความสุข (Happiness) หมายถึง คะแนนที่ได้จากการแปลงค่าเฉลี่ยของคุณลักษณะแฝง (Theta: θ) ซึ่งประมาณค่าได้จากมาตรวัดความสุขของคนไทย รูปแบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ในสามมาตรวัดย่อย ได้แก่ มาตรวัดย่อยความพึงพอใจในชีวิต มาตรวัดย่อยอารมณ์ความรู้สึกเชิงบวก และมาตรวัดย่อยอารมณ์ความรู้สึกเชิงลบ

ความพึงพอใจในชีวิต (Life Satisfaction) หมายถึง คะแนนที่ซึ่งได้จากการแปลงค่าคุณลักษณะแฝง (Theta: θ) ที่ประมาณค่าได้จากโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย ในส่วนของมาตรวัดย่อยความพึงพอใจในชีวิต

อารมณ์ความรู้สึกเชิงบวก (Positive Affect) หมายถึง คะแนนที่ซึ่งได้จากการแปลงค่าคุณลักษณะแฝง (Theta: θ) ที่ประมาณค่าได้จากโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย ในส่วนของมาตรวัดอารมณ์ความรู้สึกเชิงบวก

อารมณ์ความรู้สึกเชิงลบ (Negative Affect) หมายถึง คะแนนที่ซึ่งได้จากการแปลงค่าคุณลักษณะแฝง (Theta: θ) ที่ประมาณค่าได้จากโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย ในส่วนของมาตรวัดอารมณ์ความรู้สึกเชิงลบ

คุณลักษณะแฝง (Latent Trait) ใช้สัญลักษณ์แทนด้วย θ (Theta) หมายถึง คุณลักษณะของบุคคลที่ไม่สามารถสังเกตได้โดยตรง แต่มีอิทธิพลอยู่เบื้องหลังพฤติกรรม การตอบข้อคำถาม ซึ่งในงานวิจัยนี้ คุณลักษณะแฝง ได้แก่ ความสุข ความพึงพอใจในชีวิต อารมณ์ความรู้สึกเชิงบวก และอารมณ์ความรู้สึกเชิงลบ

พารามิเตอร์ความชัน (Slope Parameter: α) เป็นพารามิเตอร์หนึ่งใน Graded Response Model ที่เทียบเคียงได้กับพารามิเตอร์อำนาจจำแนก (Discrimination Parameter: a) ใน Two/Three Parameter Logistic Model ซึ่งบ่งบอกถึงขนาดความสัมพันธ์ระหว่างข้อคำถามรายข้อกับคุณลักษณะแฝงที่มุ่งวัด โดยแสดงให้เห็นว่าข้อคำถามสามารถจำแนกผู้ทดสอบกลุ่มที่มีคุณลักษณะแฝงที่มุ่งวัดอยู่ในระดับสูงกับกลุ่มที่มีคุณลักษณะแฝงที่มุ่งวัดอยู่ในระดับต่ำได้ดีเพียงใด

พารามิเตอร์เทอร์ชโฮลด์ (Threshold Parameter: β) เป็นพารามิเตอร์หนึ่งใน Graded Response Model ที่เทียบเคียงได้กับพารามิเตอร์ความยาก (Difficulty Parameter: b) ใน One/Two/Three Parameter Logistic Model ซึ่งบ่งบอกถึงค่าคุณลักษณะแฝงที่ตัวเลือกรายการคำตอบของข้อคำถามแต่ละข้อสามารถวัดได้จากช่วงของค่าคุณลักษณะแฝงทั้งหมด

โค้งลักษณะปฏิบัติการ (Operating Characteristic Curve: OCC) หมายถึง ฟังก์ชันคณิตศาสตร์ที่แสดงถึง ความน่าจะเป็นของการตอบข้อคำถาม (x) ของผู้ทดสอบที่มีระดับคุณลักษณะแฝงเท่ากับ θ ในการเลือกรายการคำตอบระหว่างช่วงเทอร์ชโฮลด์ 2 ช่วงที่อยู่ติดกัน พบได้ใน Graded Response Model ซึ่งมีการคำนวณค่าความน่าจะเป็นแบบสองขั้นตอน จำนวนโค้งลักษณะปฏิบัติการของข้อคำถามแต่ละข้อ จะมีจำนวนเท่ากับจำนวนตัวเลือกรายการคำตอบ หักลบออกหนึ่ง เช่น ข้อคำถามที่มี 5 ตัวเลือกรายการคำตอบ จะมีโค้งลักษณะปฏิบัติการจำนวน 4 โค้ง

โค้งการเลือกรายการคำตอบ (Category Response Curve: CRC) หรือ Category Boundary Curves หรือ Cumulative Probability Curves หรือ Boundary Characteristic Curves หมายถึง โค้งที่แสดงถึง ความน่าจะเป็นในการเลือกรายการคำตอบแต่ละตัวเลือกของผู้ทดสอบที่มีระดับคุณลักษณะแฝงแตกต่างกัน โค้งการเลือกรายการคำตอบพบได้ใน Graded Response Model ซึ่งเป็นโมเดลการตอบสนองข้อสอบที่มีการให้คะแนนมากกว่าสองค่า (Polytomous IRT Model)

ฟังก์ชันสารสนเทศของข้อคำถาม (Item Information Function: IIF) เป็นค่าที่แสดงถึงสารสนเทศทางการวัด (Psychometric Information) ของข้อคำถามตลอดช่วงของค่าคุณลักษณะแฝง ข้อคำถามจะมีฟังก์ชันสารสนเทศสูงที่สุดบริเวณพารามิเตอร์เทรซโฮลด์ นอกจากนี้ ข้อคำถามที่มีค่าพารามิเตอร์ความชันมาก โค้งจะมีลักษณะโด่งมาก (Peak) และมีค่าฟังก์ชันสารสนเทศสูง แต่มักจะครอบคลุมคุณลักษณะแฝงในช่วงแคบ ขณะที่โค้งที่มีค่าพารามิเตอร์ความชันน้อย โค้งจะมีลักษณะแบนเรียบ (Flat) ทำให้ข้อคำถามนั้นมีฟังก์ชันสารสนเทศต่ำ แต่ครอบคลุมคุณลักษณะแฝงในช่วงกว้าง

ฟังก์ชันสารสนเทศของมาตรวัด (Test Information Function: TIF) หมายถึง ผลรวมของฟังก์ชันสารสนเทศของข้อคำถามแต่ละข้อรวมเข้าด้วยกันทั้งฉบับ ณ ตำแหน่งคุณลักษณะแฝงเดียวกัน ซึ่งมีความสัมพันธ์โดยตรงกับความแม่นยำในการวัด (Precision)

คลังข้อคำถาม (Item Bank) หมายถึง แหล่งจัดเก็บข้อคำถามที่ใช้ในการทดสอบ ซึ่งบันทึกข้อมูลทั้งหมดที่เกี่ยวข้องกับข้อคำถามแต่ละข้อ ได้แก่ คำถาม ตัวเลือกรายการคำตอบ ค่าพารามิเตอร์ความชันของข้อคำถามแต่ละข้อ และค่าพารามิเตอร์เทรซโฮลด์ของตัวเลือกรายการคำตอบในข้อคำถามแต่ละข้อ คลังข้อคำถามในงานวิจัยนี้แบ่งเป็นสามคลังข้อคำถามย่อย ได้แก่ ความพึงพอใจในชีวิต อารมณ์ความรู้สึกเชิงบวก และอารมณ์ความรู้สึกเชิงลบ คลังข้อคำถามเป็นองค์ประกอบสำคัญอย่างหนึ่งในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ เพราะเป็นแหล่งข้อมูลสำหรับการคัดเลือกข้อคำถามเพื่อจัดการทดสอบให้เหมาะสมกับแต่ละบุคคล

โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับมาตรวัดความสุขของคนไทย (Computerized Adaptive Testing for Thai Happiness Scale: THS-CAT) หมายถึง โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่ผู้วิจัยจัดทำขึ้น ซึ่งประกอบด้วย 5 องค์ประกอบสำคัญ ได้แก่ 1) คลังข้อคำถาม สำหรับบรรจุข้อคำถามที่มุ่งวัดความสุขตามทฤษฎีความอยู่ดีมีสุขเชิงอัตวิสัย 2) จุดเริ่มต้นการทดสอบ (Starting Point) ซึ่งคัดเลือกข้อคำถามเริ่มต้นที่ให้ข้อมูลสารสนเทศของผู้สอบสูงที่สุด ณ ตำแหน่งคุณลักษณะแฝง (Theta: θ) ที่มีค่าเท่ากับ 0.00 3) การคัดเลือกข้อคำถามข้อถัดไป (Item Selection Algorithm) โดยคัดเลือกข้อคำถามที่ให้ข้อมูลสารสนเทศของผู้สอบสูงที่สุด (Maximum Information) ณ ตำแหน่งคุณลักษณะแฝงที่ประมาณค่าได้ภายหลังการตอบข้อคำถามแต่ละครั้ง 4) การประมาณค่าความสุข (Scoring Algorithm) ใช้วิธีการประมาณค่าแบบ Expected a Posteriori (EAP) และ 5) เกณฑ์ยุติการทดสอบ (Termination Criterion) โดยกำหนดให้การทดสอบยุติลงเมื่อความคลาดเคลื่อนมาตรฐานจากการประมาณค่ามีค่าน้อยกว่า 0.30 หรือเมื่อข้อคำถามหมดคลัง โดยโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์นี้ พัฒนาขึ้นในรูปแบบเว็บแอปพลิเคชันตามแนวคิดของวงจรการพัฒนา ระบบ (System Development Life Cycle: SDLC)

การทดสอบด้วยข้อคำถามทั้งหมดในคลังข้อคำถาม หมายถึง การดำเนินการทดสอบด้วยโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทยที่พัฒนาขึ้น โดยมีองค์ประกอบของการทดสอบด้านคลังข้อคำถาม จุดเริ่มต้นการทดสอบ การคัดเลือกข้อคำถามข้อถัดไป และการประมาณค่าความสุขเช่นเดียวกับโปรแกรมการทดสอบแบบปรับเหมาะ

ด้วยคอมพิวเตอร์ แต่มีเกณฑ์ยุติการทดสอบเพียงเกณฑ์เดียว คือ เมื่อข้อคำถามหมดคลัง การทดสอบรูปแบบนี้จึงใช้ผลการตอบข้อคำถามทุกข้อในคลังข้อคำถามเป็นข้อมูลในการประมาณค่าความสุข

บทที่ 2

เอกสารและงานวิจัยที่เกี่ยวข้อง

การประยุกต์โมเดลการตอบสนองข้อสอบ ในการพัฒนามาตรวัดความสุขของคนไทย รูปแบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ มีวัตถุประสงค์เพื่อสร้างคลังข้อคำถามที่มีคุณภาพ ด้วยการวิเคราะห์คุณภาพรายข้อของข้อคำถามตามทฤษฎีการตอบสนองข้อสอบและตรวจสอบความตรงเชิงโครงสร้างของข้อคำถามในคลังข้อคำถาม นอกจากนี้ ได้พัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับมาตรวัดความสุขของคนไทยตามแนวคิดของวงจรการพัฒนาระบบ และศึกษาความสัมพันธ์ของผลการประมาณค่าความสุขที่ได้จากการทดสอบด้วยโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย กับการทดสอบด้วยข้อคำถามทั้งหมดในคลังข้อคำถาม การวิจัยนี้นำเสนอทฤษฎี แนวคิด และงานวิจัยที่เกี่ยวข้อง เป็น 3 ตอน ดังนี้

ตอนที่ 1 ทฤษฎีเกี่ยวกับความสุขและงานวิจัยที่เกี่ยวข้อง

ตอนที่ 2 ทฤษฎีการตอบสนองข้อสอบและงานวิจัยที่เกี่ยวข้อง

ตอนที่ 3 การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์และงานวิจัยที่เกี่ยวข้อง

ตอนที่ 1 ทฤษฎีเกี่ยวกับความสุขและงานวิจัยที่เกี่ยวข้อง

นิยามของความสุข

ความสุขเป็นแนวคิดที่มีการให้คำจำกัดความที่แตกต่างกันจำนวนมาก ดังนี้

พจนานุกรม ฉบับราชบัณฑิตยสถาน พ.ศ. 2542 (2542, หน้า 1,201) ได้อธิบายว่า “สุข, สุข-” เป็นคำที่ได้มาจากภาษาบาลีและสันสกฤต ทำหน้าที่เป็นได้ทั้งคำนามและคำวิเศษณ์ โดย สุข หมายถึง ความสบายกายสบายใจ

กรมสุขภาพจิต กระทรวงสาธารณสุข กล่าวว่า ความสุขสามารถเทียบเท่ากับการมีสุขภาพจิตที่ดี ซึ่งหมายถึง สภาพชีวิตที่เป็นสุข อันเป็นผลมาจากการมีความสามารถในการจัดการปัญหาในการดำเนินชีวิต มีศักยภาพที่จะพัฒนาตนเองเพื่อคุณภาพชีวิตที่ดี โดยครอบคลุมถึงความดีงามภายในจิตใจ ภายใต้อิทธิพลสิ่งแวดล้อมที่เปลี่ยนแปลงไป (ประเวช ตันติพิวัฒนสกุล และเอกอนงค์ สิตลาภินันท์, 2554, หน้า 29-32)

ศูนย์สุขภาวะทางจิต คณะจิตวิทยา จุฬาลงกรณ์มหาวิทยาลัย (2556) ให้คำจำกัดความว่า ความสุข เป็นผลรวมที่เกิดจากการที่บุคคลได้ทำในสิ่งที่ตนต้องการและรับรู้ถึงความสำเร็จ สามารถสร้างสัมพันธ์ภาพที่ดีกับผู้อื่น มีจิตใจสงบ มีความคิดเชิงบวก มีความเป็นตัวของตัวเอง กระตือรือร้นในการดำเนินชีวิตซึ่งนำไปสู่การมีสุขภาพที่ดีทั้งร่างกายและจิตใจ พร้อมที่จะเรียนรู้และพัฒนาตนเองอยู่เสมอ รวมทั้งสามารถดำเนินชีวิตอย่างพอเพียงและภูมิใจในสิ่งที่ตนได้ลงมือทำ

ทีมงานเว็บไซต์ธรรมจักรได้เขียนบทความเรื่อง “สู่ความสุข” ซึ่งได้กล่าวถึง นิยามของความสุขว่า หมายถึง ความสบาย หรือความสำราญ ซึ่งมี 2 ประเภท ได้แก่ 1) ความสุขทางกาย หมายถึง ความสุขที่สัมผัสได้จากประสาทสัมผัสทั้งห้าทั้งในส่วนจากรูป เสียง กลิ่น รส และผิวหนัง

เรียกว่า "กามคุณ 5" จัดว่าเป็นฝ่ายรูป หรือความสุขที่เกิดจากเนื้อหนังมังสา อันเป็นสิ่งสกปรก และ 2) ความสุขทางใจ หมายถึง ความสุขที่สัมผัสได้จากจิต เป็นความสบายใจ ความสุขใจ ความอึ้งใจ ความพอใจ อันเกิดจากจิตใจที่สงบและเย็น จัดว่าเป็นฝ่ายนาม อันเป็นความสุขที่สะอาด ความสุขทั้งสองประเภทนี้มีส่วนสัมพันธ์กัน ไม่สามารถแยกจากกันได้อย่างชัดเจน และต่างต้องพึ่งพาอาศัยกัน และกัน (ธรรมจักร, 2556)

Coan (1974, 1977 cited in Compton & Hoffman, 2013, pp. 46-47) เสนอทฤษฎีรูปแบบของการเติมเต็มชีวิต (The Modes of Fulfillment) โดยมีแนวคิดที่ว่า ความสุขเกิดจากแนวทาง หรือวิถีของแต่ละบุคคล ในการค้นหาและเติมเต็มชีวิตของตนเอง ทฤษฎีนี้มีแนวคิดของวัฒนธรรมตะวันออกผสมผสานอยู่ และต่างจากทฤษฎีอื่นตรงที่ ไม่ได้เป็นโมเดลที่มีแนวคิดที่ว่า ความสุขเกิดจากผลรวมขององค์ประกอบต่าง ๆ แต่รูปแบบในการเติมเต็มชีวิตนั้น ไม่ว่าจะบุคคลจะมีหนึ่งรูปแบบ หรือหลายรูปแบบผสมผสานกัน ก็สามารถเป็นรูปแบบการเติมเต็มชีวิตที่สร้างความสุขได้ไม่ต่างกัน รูปแบบพื้นฐานในการเติมเต็มชีวิตของ Coan มี 5 รูปแบบ ได้แก่

1. ศักยภาพ (Efficiency) เป็นรูปแบบที่เน้นการใช้ทักษะความสามารถของตนเองเพื่อให้เกิดความเป็นเลิศในแนวทางที่ตนชื่นชอบ
2. ความสร้างสรรค์ (Creativity) เป็นรูปแบบในการเติมเต็มความสุขของผู้มีใจศิลปิน
3. ความเป็นอันหนึ่งอันเดียวภายในตน (Inner Harmony) เป็นรูปแบบที่เน้นความสำคัญทางด้านจิตใจ เช่น การค้นหาตัวตนที่แท้จริงของตน การค้นหาบุคลิกภาพที่แท้ของตน
4. ความเกี่ยวข้องสัมพันธ์ (Relatedness) เป็นรูปแบบที่เน้นการพัฒนาความสัมพันธ์ระหว่างบุคคล ความรัก ครอบครัว ความสนิทสนมกลมเกลียวกับผู้อื่น
5. ความเลื่อมใสศรัทธา (Self-Transcendence) เป็นรูปแบบที่เกี่ยวข้องกับความสัมพันธ์ของบุคคลกับพระเจ้า สิ่งศักดิ์สิทธิ์ ความเชื่อ สิ่งลึกลับ หรือจิตวิญญาณ

Ryff (1989) ได้เสนอทฤษฎีการอยู่ดีมีสุขเชิงจิตวิทยา (Psychological Well-Being) โดยอธิบายว่า ความสุขประกอบด้วย 6 องค์ประกอบ ได้แก่

1. การยอมรับตนเอง (Self-Acceptance) หมายถึง การมีเจตคติที่ดีต่อตนเอง สามารถยอมรับตนเองในมุมต่าง ๆ ได้ และรู้สึกดีกับชีวิตในอดีต
2. การเติบโตภายในตน (Personal Growth) หมายถึง ความรู้สึกถึงการพัฒนาตนเองอย่างต่อเนื่อง และเป็นไปอย่างมีประสิทธิภาพ เปิดรับประสบการณ์ใหม่ ๆ ด้วยความรู้สึกล้าหาญ
3. การมีจุดมุ่งหมายในชีวิต (Purpose in Life) หมายถึง การมีเป้าหมายในชีวิตตนเอง และเชื่อมั่นในเป้าหมายนั้นว่าเป็นสิ่งที่นำทางชีวิตตน ทำให้รู้สึกถึงความหมายในการดำเนินชีวิต
4. ความสามารถในการปรับตัวเข้ากับสิ่งแวดล้อม (Environmental Mastery) หมายถึง ความรู้สึกว่าคุณมีความสามารถ และมีศักยภาพเพียงพอที่จะจัดการกับสถานการณ์แวดล้อมที่ซับซ้อน รวมถึงความสามารถในการปรับตนเองให้เข้ากับสิ่งแวดล้อม หรือสถานการณ์ที่ตนเผชิญ
5. ความเป็นตัวของตัวเอง (Autonomy) หมายถึง รู้สึกมั่นใจในการนำชีวิตตนเอง มีค่านิยม บรรทัดฐานของตน และสามารถอดทนต่อแรงกดดันทางสังคมที่มาจากผู้อื่นได้
6. การมีสัมพันธภาพที่ดีกับผู้อื่น (Positive Relations with Others) หมายถึง การมีสัมพันธภาพที่อบอุ่น พึงพอใจ และไว้วางใจได้กับผู้อื่น สามารถที่จะแสดงความเห็นใจ และสร้าง

ความสัมพันธ์กับผู้อื่นได้

Keyes (1998) ได้เสนอทฤษฎีความอยู่ดีมีสุขเชิงสังคม ที่เรียกว่า Social Well-Being ซึ่งประกอบด้วยองค์ประกอบ 5 ประการ ได้แก่

1. การยอมรับผู้อื่นในสังคม (Social Acceptance) หมายถึง การมีเจตคติที่ดีต่อผู้อื่น รวมทั้ง มีความเข้าใจถึงความซับซ้อนของพวกเขา
2. การเชื่อมั่นในการเติบโตของสังคม (Social Actualization) หมายถึง มีความสนใจ และเชื่อมั่นว่าบุคคลต่าง ๆ ในสังคมเป็นบุคคลที่มีศักยภาพ สังคมสามารถที่จะเติบโต และดำเนินไปในทางที่ดี
3. การอุทิศตนเพื่อสังคม (Social Contribution) หมายถึง ความรู้สึกถึงควมมีคุณค่า และมีประโยชน์ต่อสังคม สามารถที่จะทำประโยชน์ให้กับผู้อื่นได้
4. ความรู้สึกถึงความสอดคล้องภายในสังคม (Social Coherence) หมายถึง การให้ความสนใจใส่ใจต่อสังคม เชื่อมั่นว่าโลกในสังคมเป็นสิ่งที่เราทำความเข้าใจได้ กลไกต่าง ๆ เกิดขึ้นอย่างมีเหตุและผล สามารถคาดการณ์ได้ และสิ่งที่เกิดขึ้นทุกอย่างมีความหมาย
5. ความรู้สึกเป็นอันหนึ่งอันเดียวกับสังคม (Social Integration) หมายถึง ความรู้สึกเป็นส่วนหนึ่งในสังคม รู้สึกมั่นคง สบาย และรู้สึกว่าได้รับการช่วยเหลือจากสังคม

Seligman (2002 cited in Compton & Hoffman, 2013, p. 45) กล่าวว่า ความสุขประกอบด้วย 3 องค์ประกอบ ได้แก่ การมีชีวิตที่พึงพอใจ (Pleasant Life), การมีชีวิตที่ดี (Good Life) และการมีชีวิตที่มีความหมาย (Meaningful Life) ชีวิตที่มีความสุขโดยสมบูรณ์ คือ ชีวิตที่มีองค์ประกอบทั้งสามนี้อย่างครบถ้วน

Deci and Ryan (2008) ผู้คิดค้นทฤษฎีอัตลิติต (Self-determination Theory) ได้เสนอว่าความสุขเป็นผลมาจากความรู้สึกที่ได้รับการเติมเต็ม จากความต้องการพื้นฐานสามประการ ได้แก่ ความรู้สึกเป็นตัวของตัวเอง (Autonomy) ความรู้สึกมีศักยภาพ (Competence) และความรู้สึกถึงสัมพันธภาพที่ดีกับผู้อื่น (Relatedness) โดยความรู้สึกเป็นตัวของตัวเองสามารถถูกเติมเต็มได้ด้วยการมีอิสระในการเลือกทำกิจกรรมต่าง ๆ ด้วยตนเองมากกว่าต้องทำตามผู้อื่น ส่งผลให้เกิดภาพพจน์ที่มั่นคงต่อตนเอง ขณะที่ความรู้สึกมีศักยภาพ เกิดขึ้นจากความพยายามทำสิ่งใดสิ่งหนึ่งจนได้รับผลสัมฤทธิ์ตามที่ต้องการ จนกระทั่งเกิดความมั่นใจในความสามารถของตนเอง ส่วนความรู้สึกถึงสัมพันธภาพที่ดีกับผู้อื่นนั้น ถูกเติมเต็มได้ด้วยการมีสัมพันธภาพที่ใกล้ชิด และราบรื่นกับผู้อื่น ซึ่งมีการวิจัยสนับสนุนว่า เมื่อมนุษย์ได้รับการตอบสนองความต้องการทั้งสามด้านนี้ สามารถส่งผลต่อความรู้สึกเป็นสุขได้

Seligman (2011 cited in Compton & Hoffman, 2013, p. 45) ได้ปรับเปลี่ยนแนวคิดเกี่ยวกับความสุขและเสนอทฤษฎีใหม่ซึ่งเชื่อว่าจะมีความครอบคลุมมุมมองทางด้านความสุขมากกว่าทฤษฎีเดิม ทฤษฎีใหม่ มีชื่อว่า ทฤษฎีความอยู่ดีมีสุข (Well-Being Theory) ซึ่งประกอบด้วยองค์ประกอบสำคัญ 5 ประการ ได้แก่ อารมณ์เชิงบวก (Positive Emotion) ความผูกพัน (Engagement) สัมพันธภาพ (Relationships) ความหมาย (Meaning) และความสำเร็จ (Accomplishment) โดยใช้ตัวย่อแทนองค์ประกอบทั้งหมดว่า PERMA

Pavot and Diener (2013) อธิบายความหมายของความสุข (Happiness) ในมุมมองของความอยู่ดีมีสุขเชิงอัตวิสัย (Subjective Well-being) โดยมีแนวคิดว่าความสุขของบุคคลเป็นกระบวนการในการตัดสิน และประเมินคุณภาพชีวิตของบุคคลตามมุมมองส่วนตัว แต่ละบุคคลอาจมีเกณฑ์ในการประเมินที่แตกต่างกัน การระบุว่าคุณภาพชีวิตของตนดีหรือไม่ จึงขึ้นอยู่กับ การตัดสินใจของบุคคลนั่นเอง

ความอยู่ดีมีสุขเชิงอัตวิสัย (Subjective Well-being) (Diener, 1984; Myers & Diener, 1995; Diener, Suh, Lucas, & Smith, 1999; Diener, 2000; Pavot & Diener, 2013) เป็นคุณลักษณะที่ประกอบด้วยสามองค์ประกอบย่อยที่แยกออกจากกัน ได้แก่

1) ความพึงพอใจในชีวิต (Life Satisfaction) เป็นการประเมินชีวิตตนเองในภาพรวม โดยไม่เจาะจงที่ด้านใดด้านหนึ่งโดยเฉพาะ ซึ่งประเมินใน 5 ส่วน ได้แก่ (Diener, Suh, Lucas, & Smith, 1999)

ก) ความต้องการเปลี่ยนแปลงชีวิตของตน (Desire to Change Life) หมายถึง ความคิด ความรู้สึก และพฤติกรรมของบุคคลที่แสดงให้เห็นว่าขาดความพอใจในสภาพชีวิตที่เป็นอยู่ โดยมีความต้องการที่จะเปลี่ยนแปลงแก้ไขชีวิตของตนเองในอดีต มีความมุ่งมั่นที่จะดำเนินชีวิตในปัจจุบันที่แตกต่างจากเดิม และมีความคาดหวังว่าชีวิตในอนาคตจะดีขึ้นกว่าที่เป็นอยู่

ข) ความพึงพอใจในชีวิตปัจจุบัน (Satisfaction with Current Life) หมายถึง การรับรู้ และการประเมินชีวิตของบุคคลในสภาพที่เป็นอยู่ตามปัจจุบัน ซึ่งเป็นทัศนคติในเชิงบวกที่มีต่อชีวิตในภาพรวมทั่วไป มากกว่าที่จะเป็นเพียงสภาวะอารมณ์ใดอารมณ์หนึ่ง หรือความรู้สึกที่มีต่อเหตุการณ์ใดเหตุการณ์หนึ่งที่เฉพาะเจาะจง

ค) ความพึงพอใจในชีวิตอดีต (Satisfaction with Past Life) หมายถึง การรับรู้ และการประเมินชีวิตของบุคคลจากประสบการณ์ที่ผ่านมาในอดีต ซึ่งเป็นทัศนคติในเชิงบวกที่มีต่อชีวิตในภาพรวมทั่วไป มากกว่าที่จะเป็นเพียงสภาวะอารมณ์ใดอารมณ์หนึ่ง หรือความรู้สึกที่มีต่อเหตุการณ์ใดเหตุการณ์หนึ่งที่เฉพาะเจาะจง

ง) ความพึงพอใจในชีวิตอนาคต (Satisfaction with Future Life) หมายถึง การรับรู้ และการประเมินชีวิตของบุคคลในสถานการณ์ข้างหน้า การคาดหวังที่จะได้พบกับประสบการณ์ที่ดีในชีวิต การเตรียมพร้อมและความมั่นใจที่จะเผชิญกับสิ่งต่าง ๆ ที่กำลังจะเกิดขึ้น ซึ่งเป็นทัศนคติในเชิงบวกที่มีต่อชีวิตในภาพรวมทั่วไป มากกว่าที่จะเป็นเพียงสภาวะอารมณ์ใดอารมณ์หนึ่ง หรือความรู้สึกที่มีต่อเหตุการณ์ใดเหตุการณ์หนึ่งที่เฉพาะเจาะจง

จ) ความคิดเห็นจากคนสำคัญที่มีต่อชีวิตของบุคคล (Significant Other's Views of One's Life) หมายถึง การรับรู้และการประเมินชีวิตของบุคคลจากภาพสะท้อนที่ได้รับจากบุคคลอื่น ๆ ซึ่งเป็นบุคคลที่มีความสำคัญ เป็นการประเมินว่าบุคคลที่มีความสำคัญในชีวิตของตนนั้น มีทัศนคติที่ดีต่อชีวิตของตน มีการชื่นชมชีวิตของตน ซึ่งทำให้เกิดความภาคภูมิใจและเกิดความพึงพอใจในชีวิตของตน

2) อารมณ์ความรู้สึกเชิงบวก (Positive Affect) เป็นลักษณะที่ใช้อธิบายถึงประสบการณ์ที่แตกต่างกันของบุคคลในการมีปฏิสัมพันธ์กับบุคคล หรือสิ่งรอบข้างด้วยอารมณ์ความรู้สึกที่ดี ซึ่งทำให้บุคคลเกิดความพึงพอใจ ประสบการณ์ทางอารมณ์ความรู้สึกเชิงบวกเกิดขึ้นได้หลากหลายรูปแบบ

และมีแนวโน้มของระดับอารมณ์ความรู้สึกเชิงบวกที่คงที่ปานกลาง แม้เวลาผ่านไป หรืออยู่ในสถานการณ์ที่แตกต่างกัน นอกจากนี้ อารมณ์ความรู้สึกเชิงบวก มีความเป็นอิสระจากอารมณ์ความรู้สึกเชิงลบอย่างชัดเจน และไม่ได้เป็นคุณลักษณะที่ตรงข้ามกัน นั่นหมายความว่า การมีอารมณ์ความรู้สึกเชิงบวกไม่ได้รับประกันว่า จะไม่มีอารมณ์ความรู้สึกเชิงลบ ดังนั้น การประเมินจึงต้องแยกจากกัน อารมณ์ความรู้สึกเชิงบวก ซึ่งเป็นตัวชี้วัดความสุขของบุคคล นิยมประเมินด้านความถี่ (Frequency) ของการเกิดอารมณ์ความรู้สึกต่าง ๆ มากกว่าความเข้มข้น (Intensity) ของอารมณ์ความรู้สึกที่เกิดขึ้น เพราะทำให้ผลการประเมินที่ได้มีความถูกต้องเที่ยงตรงมากกว่า (Diener & Emmons, 1985; Diener, Sandvik, & Pavot, 1991, pp. 119-139; Naragon & Watson, 2013, p. 707, Watson, Clark, & Tellegen, 1988)

3) อารมณ์ความรู้สึกเชิงลบ (Negative Affect) เป็นลักษณะที่ใช้อธิบายถึงประสบการณ์ที่แตกต่างกันของบุคคลในการมีปฏิสัมพันธ์กับบุคคล หรือสิ่งรอบข้างด้วยอารมณ์ความรู้สึกที่ไม่ดี ซึ่งทำให้บุคคลเกิดความรู้สึกไม่พึงพอใจ รวมถึงมีภาพพจน์ต่อตนเองในเชิงลบ ประสบการณ์ทางอารมณ์ความรู้สึกเชิงลบเกิดขึ้นได้หลากหลายรูปแบบ และเป็นอิสระจากอารมณ์ความรู้สึกเชิงบวก ในการประเมินอารมณ์ความรู้สึกเชิงลบ เพื่อเป็นตัวชี้วัดความสุขของบุคคลนิยมประเมินด้านความถี่ (Frequency) ของการเกิดอารมณ์ความรู้สึกเช่นเดียวกับอารมณ์ความรู้สึกเชิงบวก (Diener & Emmons, 1985; Diener, Sandvik, & Pavot, 1991, pp. 119-139; Naragon & Watson, 2013, p. 707)

คนที่มีความสุข คือ คนที่มีความรู้สึกพึงพอใจในชีวิต (Life Satisfaction) มีประสบการณ์ทางอารมณ์ความรู้สึกเชิงบวก (Positive Affect) บ่อยครั้ง ขณะที่ประสบการณทางอารมณ์ความรู้สึกเชิงลบ (Negative Affect) น้อยครั้งหรือแทบไม่มีเลย

Wikipedia (2013) ได้กล่าวถึงความสุขว่าเป็นสภาวะทางด้านจิตใจ หรืออารมณ์ ซึ่งแสดงออกให้เห็นด้วยอารมณ์เชิงบวก เช่น ความรู้สึกสบายใจ ความรู้สึกสนุกสนาน นักคิดและนักทฤษฎีต่างพยายามศึกษานิยามของความสุข และแหล่งที่มาของความสุขในหลากหลายแง่มุมทั้งด้านชีวภาพ จิตวิทยา ศาสนา และปรัชญา นักจิตวิทยาเชิงบวกเป็นอีกกลุ่มหนึ่งที่ใช้กระบวนการที่เป็นวิทยาศาสตร์เพื่อจะศึกษาว่าความสุขคืออะไร และสามารถสร้างได้อย่างไร

การวัดความสุข

นักคิดและนักทฤษฎี พยายามใช้วิธีการที่หลากหลายในการวัดความสุข เช่น การสังเกตพฤติกรรม (Behavioral Observations) การวัดทางกายภาพ (Physiological Measurements) และการประเมินตนเอง (Self-Reports) (Shiota & Kalat, 2012, pp. 7-13) วิธีการที่ได้รับความนิยมมากที่สุดในการวัดความสุข คือ การประเมินตนเอง ซึ่งสังเกตได้จากจำนวนงานวิจัย ตั้งแต่อดีตถึงปัจจุบัน ที่เน้นการสร้างมาตรวัดในรูปแบบการประเมินตนเอง มาตรวัดต่าง ๆ ที่สร้างขึ้นมีความแตกต่างกันไปตามแนวคิด และนิยามที่ใช้ในการศึกษาความสุข มาตรวัดบางฉบับเป็นการประเมินความสุขในภาพรวม ในขณะที่มาตรวัดบางฉบับมีแนวคิดต่อความสุขว่าประกอบด้วยองค์ประกอบย่อยต่าง ๆ จึงพัฒนามาตรวัดขึ้นในลักษณะของมาตรวัดย่อย ซึ่งวัดคุณลักษณะใดคุณลักษณะหนึ่งอย่างเฉพาะเจาะจง แต่เป็นลักษณะเชิงบวกที่ก่อให้เกิดความสุข นอกจากนี้ เมื่อพิจารณาที่จำนวนข้อคำถามจะพบว่า มาตรวัดความสุขบางฉบับสร้างขึ้นโดยมีข้อคำถามเพียง

ข้อเดียว ในขณะที่มาตรวัดความสุขบางฉบับ ประกอบด้วยข้อคำถามจำนวนมาก การศึกษาวิจัยที่เกี่ยวข้องกับความสุขได้รับความสนใจมากขึ้น การพัฒนามาตรวัดความสุขต่าง ๆ จึงเกิดขึ้นอย่างแพร่หลาย ทั้งในบริบทของสังคมไทย และต่างประเทศ รายละเอียดที่เกี่ยวข้องกับการพัฒนามาตรวัดความสุข มีดังนี้

ชูฤทธิ์ เตังไตรสรณ์ พระครุฑชิต คุณวโร และอรุณี วชิราพรทิพย์ (2551) พัฒนาแบบวัดสุขภาวะองค์กรรวมแนวพุทธ โดยมีวัตถุประสงค์เพื่อพัฒนาเครื่องมือวัดสุขภาวะองค์กรรวมแนวพุทธสำหรับคนไทย จากแนวคิดท้าวานา 4 ตามกรอบแนวคิดของ ป.อ. ปยุตโต ในการสำรวจภาวะที่สุขสมบูรณ์ทั้งทางร่างกาย จิตใจ สังคม และปัญญา มาตรวัดนี้มีลักษณะเป็นมาตรประมาณค่า (Rating Scale) จำนวน 190 ข้อ ข้อคำถามทั้งหมดผ่านการทดลองใช้ในกลุ่มตัวอย่างจำนวน 30 คน จากนั้นจึงนำข้อมูลที่ได้จากการทดลองใช้มาปรับปรุงแก้ไขมาตรวัด แล้วนำเสนอผู้เชี่ยวชาญเพื่อตรวจสอบความตรงเชิงเนื้อหา นอกจากนี้ ได้นำไปทดลองใช้ในกลุ่มผู้รู้ชัด (Known Group) 2 กลุ่ม เพื่อหาค่าอำนาจจำแนกของข้อคำถาม ตรวจสอบความตรงเชิงโครงสร้าง วิเคราะห์ค่าสหสัมพันธ์ระหว่างข้อคำถามกับมาตรวัด และวิเคราะห์ค่าความเที่ยง ข้อคำถามในแบบวัดสุขภาวะองค์กรรวมแนวพุทธ ถูกคัดเลือกจนเหลือข้อคำถามสุดท้ายจำนวน 100 ข้อ ผลการศึกษา ปรากฏว่า แบบวัดฉบับนี้ มีความตรงเชิงเนื้อหาในระดับปานกลาง มีความตรงเชิงโครงสร้างในระดับดี มีค่าอำนาจจำแนกเฉลี่ย 0.51 โดยค่าอำนาจจำแนกมีค่าอยู่ระหว่าง 0.20 ถึง 0.79 และมีค่าความเที่ยงของแบบวัดเท่ากับ 0.93 เมื่อนำแบบวัดไปทดลองใช้กับกลุ่มตัวอย่าง คือ 1) กลุ่มผู้ป่วย โสต ศอ นาสิก จำนวน 210 คน และ 2) กลุ่มผู้ติดเชื้อ HIV และผู้ป่วยเอดส์ จำนวน 121 คน ที่มารับการรักษาที่สถาบันบำราศนราดูร ปรากฏว่า แต่ละกลุ่มได้คะแนนเฉลี่ยจากแบบวัดสุขภาวะองค์กรรวมแนวพุทธเท่ากับ 3.33 และ 3.43 ตามลำดับ กล่าวคือ มีสุขภาวะองค์กรรวมแนวพุทธอยู่ในระดับปานกลาง คะแนนที่ได้จากการวัดทั้ง 4 หมวดจากกลุ่มตัวอย่าง 2 กลุ่ม ปรากฏว่าคะแนนที่ได้จากหมวดต่าง ๆ มีความสัมพันธ์กันอย่างน้อยสำคัญทางสถิติที่ระดับ .05

กรมสุขภาพจิต กระทรวงสาธารณสุขได้พัฒนาแบบวัดสุขภาพจิต ที่ชื่อว่า Thai Mental Health Indicator – 55 (TMHI – 55) จำนวน 55 ข้อ ตามนิยามความสุข ซึ่งมีองค์ประกอบ ดังนี้ (ประเวช ตันติพิวัฒนสกุล และเอกอนงค์ สีตลาภินันท์, 2554, หน้า 44-48)

1) สภาพจิตใจ (Mental State) ประกอบด้วย 3 องค์ประกอบย่อย ได้แก่ ความรู้สึกในทางที่ดี (General Well-Being Positive Affect) ความรู้สึกในทางที่ไม่ดี (General Well-Being Negative Affect) และ การรับรู้ภาวะสุขภาพและการเจ็บป่วยทางด้านจิตใจ (Perceived Ill-Health and Mental Illness)

2) สมรรถภาพของจิตใจ (Mental Capacity) ประกอบด้วย 4 องค์ประกอบย่อย ได้แก่ สัมพันธภาพระหว่างบุคคล (Interpersonal Relationship) ความสามารถที่จะบรรลุความสำเร็จตามความคาดหวัง (Expectation Achievement Congruence) ความมั่นใจในการเผชิญปัญหา (Confidence in Coping) และการควบคุมจิตใจของตนเองอย่างมีประสิทธิภาพ (Adequate Mental Mastery)

3) คุณภาพของจิตใจ (Mental Quality) ประกอบด้วย 4 องค์ประกอบย่อย ได้แก่ เมตตา กรุณา และเสียสละ (Kindness and Altruism) การนับถือตนเอง (Self-Esteem) ความศรัทธา

(Faith) และความคิดสร้างสรรค์และความกระตือรือร้นในการดำเนินชีวิต (Creativity Thinking and Enthusiasm)

4) ปัจจัยสนับสนุน (Supporting Factors) ประกอบด้วย 4 องค์ประกอบย่อย ได้แก่ การสนับสนุนทางสังคม (Social Support) การสนับสนุนจากครอบครัว (Family Support) ความปลอดภัยทางร่างกายและความมั่นคงในชีวิต (Physical Safety and Security) และการดูแลสุขภาพและบริการทางสังคม (Health and Social Care)

ผู้วิจัยได้ศึกษาและปรับลดจำนวนข้อคำถามจากแบบวัดสุขภาพจิตฉบับสมบูรณ์ เพื่อให้ได้แบบวัดสุขภาพจิตฉบับสั้นซึ่งเทียบเท่ากับแบบวัดความสุข โดยใช้ชื่อว่า Thai Mental Health Indicator – 15 (TMHI – 15) จำนวน 15 ข้อ แบบวัดความสุขฉบับสั้นคงเหลือองค์ประกอบย่อยต่าง ๆ ดังนี้ ความรู้สึกในทางที่ดี ความรู้สึกในทางที่ไม่ดี ความมั่นใจในการเผชิญปัญหา เมตตา กรุณา และเสียสละ การนับถือตนเอง และการสนับสนุนจากครอบครัว ในมุมมองของกรมสุขภาพจิตนั้นมีแนวคิดว่าคุณภาพจิตและความสุขเป็นเรื่องเดียวกัน แบบวัดสุขภาพจิตนี้จึงสามารถนำมาใช้ในการวัดความสุขได้เช่นกัน แบบวัดทั้งสองฉบับมีรูปแบบในการตอบข้อคำถามแบบมาตราประมาณค่า โดยมีทั้งข้อคำถามทางบวกและข้อคำถามทางลบ ระดับความสุขได้มาจากผลรวมของคะแนนที่ได้จากข้อคำถามในมาตรวัดทั้งหมด โดยคะแนนสูงแสดงถึงการมีสุขภาพจิตที่ดีหรือการมีความสุข

โครงการจับตาสถานการณ์ความสุขของคนทำงานในประเทศไทย โดยความร่วมมือของสำนักงานกองทุนสนับสนุนการสร้างเสริมสุขภาพ (สสส.) และสถาบันวิจัยประชากรและสังคม มหาวิทยาลัยมหิดล (ศิริพันธ์ กิตติสุขสถิต และคณะ, 2555, หน้า 23-42) ได้จัดทำคู่มือและเครื่องมือในการวัดคุณภาพชีวิตและความสุขด้วยตนเองของคนทำงานในองค์กรต่าง ๆ ในชื่อว่า “แบบสำรวจความสุขด้วยตนเอง Happinometer” ซึ่งความสุขในมุมมองของ Happinometer ประกอบด้วย 9 มิติ ซึ่งมีตัวชี้วัดรวมทั้งหมด 56 ตัวชี้วัด ดังนี้

1) สุขภาพดี (Happy Body) หมายถึง การที่บุคคลมีสุขภาพร่างกายแข็งแรง สมส่วน มีพฤติกรรมการบริโภคที่ดี/เหมาะสม มีความพึงพอใจสภาวะสุขภาพกายของตนเอง โดยมี 6 ตัวชี้วัด คือ ค่า BMI ภาวะอ้วนลงพุง ออกกำลังกาย สูดบุหรี่ยุติ ต้มเครื่องต้มแอลกอฮอล์ และความพึงพอใจต่อสุขภาพกาย

2) ผ่อนคลายดี (Happy Relax) หมายถึง การที่บุคคลสามารถบริหารเวลาในแต่ละวันเพื่อการพักผ่อนได้อย่างมีคุณภาพ พึงพอใจกับการบริหารจัดการปัญหาของตนเอง และทำชีวิตให้ง่ายสบาย ๆ โดยมี 5 ตัวชี้วัด คือ การพักผ่อนที่เพียงพอ กิจกรรมพักผ่อนหย่อนใจ ความเครียด ชีวิตเป็นไปตามที่คาดหวัง และการจัดการกับปัญหาในชีวิต

3) น้ำใจดี (Happy Heart) หมายถึง การที่บุคคลมีจิตสาธารณะ มีส่วนร่วมในการสร้างประโยชน์ให้กับส่วนรวม และมีเมตตากับคนรอบข้าง โดยมี 9 ตัวชี้วัด คือ ความรู้สึกเอื้ออาทร การให้ความช่วยเหลือแก่คนรอบข้าง การทำงานเป็นทีม ความสัมพันธ์ฉันท์พี่น้อง การสื่อสารกับเพื่อนร่วมงานในองค์กร การถ่ายทอดแลกเปลี่ยนวิธีการทำงาน การทำประโยชน์เพื่อส่วนรวม และการเข้าร่วมกิจกรรมที่เป็นประโยชน์ต่อสังคม

4) จิตวิญญาณดี (Happy Soul) หมายถึง การที่บุคคลมีความตระหนักถึงคุณธรรม และศีลธรรม รู้แพ้รู้ชนะ รู้จักให้ และมีความกตัญญูรู้คุณ โดยมีมิตินี้ มี 5 ตัวชี้วัด คือ การทำนุบำรุง ศิลปวัฒนธรรมและศาสนา การให้ทานและการปฏิบัติกิจตามศาสนา การให้อภัย การยอมรับและการขอโทษ รวมถึง การตอบแทนผู้มีพระคุณ

5) ครอบครัวดี (Happy Family) หมายถึง การที่บุคคลมีความรู้สึกผูกพัน เชื่อใจ มั่นใจ และอุ่นใจกับบุคคลในครอบครัวของตนเอง โดยมีมิตินี้ มี 3 ตัวชี้วัด คือ เวลาที่ใช้กับครอบครัว การทำกิจกรรมกับครอบครัว และการมีความสุขกับครอบครัว

6) สังคมดี (Happy Society) หมายถึง การที่บุคคลมีความสัมพันธ์ที่ดีต่อเพื่อนบ้าน ไม่ทำให้ผู้อื่นเดือดร้อน ไม่เอาเปรียบผู้คนรอบข้าง ไม่ทำให้สังคมเสื่อมถอย โดยมีมิตินี้ มี 6 ตัวชี้วัด คือ ความสัมพันธ์กับเพื่อนบ้าน การปฏิบัติตามกฎระเบียบของสังคม ความปลอดภัยในชีวิตและทรัพย์สิน การขอความช่วยเหลือจากคนในชุมชน ความสงบสุขของสังคม และการใช้ชีวิตในสังคมอย่างมีความสุข

7) ใฝ่รู้ดี (Happy Brain) หมายถึง การที่บุคคลมีความตื่นตัว กระตือรือร้นในการเรียนรู้สิ่งใหม่ ๆ เพื่อปรับตัวให้เท่าทัน และตั้งรับการเปลี่ยนแปลงได้ตลอดเวลา รวมถึงพอใจที่จะแสดง ความทันสมัยอยู่เสมอ มิตินี้ มี 3 ตัวชี้วัด คือ การแสวงหาความรู้ใหม่ ๆ การพัฒนาตนเอง และโอกาสในการพัฒนาตนเอง

8) สุขภาพเงินดี (Happy Money) หมายถึง การที่บุคคลมีวินัยในการใช้จ่ายเงิน มีความสามารถและพึงพอใจในการบริหารจัดการระบบการรับ จ่าย และออมเงินในแต่ละเดือน มิตินี้ มี 4 ตัวชี้วัด คือ การผ่อนชำระหนี้สินต่าง ๆ การชำระหนี้ เงินออม และความเพียงพอของค่าตอบแทนที่ได้รับ

9) การงานดี (Happy Work-Life) หมายถึง การที่บุคคลมีความสบายใจในที่ทำงาน มีความรัก ความผูกพัน และความภาคภูมิใจในองค์กร มีความมั่นใจในอาชีพ รายได้ และมีความพึงพอใจกับความก้าวหน้าของตนเองในองค์กร มิตินี้ มี 15 ตัวชี้วัด คือ ความสุขต่อสภาพแวดล้อมในองค์กร การได้รับการดูแลด้านสุขภาพ ความพึงพอใจต่อสวัสดิการ การได้รับการปฏิบัติอย่างถูกต้องตามกฎหมายแรงงาน ความเหมาะสมของการเลื่อนขั้น เลื่อนตำแหน่ง ความมั่นคงในอาชีพ ความเสี่ยงจากการทำงาน การลาออกจากงาน การเปลี่ยนสถานที่ทำงาน การมีส่วนร่วมแสดงความคิดเห็น ความถูกต้องของการจ่ายค่าจ้าง ความตรงเวลาของการจ่าย ค่าจ้าง และการทำงานอย่างมีความสุข

ตัวชี้วัดความสุขทั้ง 56 ตัวชี้วัด ของ Happinometer ได้มีกระบวนการตรวจสอบ ความเที่ยงตรงและน่าเชื่อถือ โดยศึกษาความตรงเชิงเนื้อหา ด้วยการทบทวนวรรณกรรมที่เกี่ยวข้อง และประชุมร่วมกับผู้ทรงคุณวุฒิทางด้านคุณภาพชีวิต ความสุข สุขภาพจิต เพื่อพิจารณาความตรงเชิงเนื้อหาที่ต้องการประเมิน จำนวน 9 ครั้ง ศึกษาความตรงตามโครงสร้าง โดยการวิเคราะห์ องค์ประกอบ (Factor Analysis) (ยกเว้นข้อคำถามที่ผู้ทรงคุณวุฒิลงความเห็นว่าเป็นต้องมีจึงไม่คำนึงถึง ค่า Factor Loading) และศึกษาค่าความเที่ยงของเครื่องมือ (Reliability) โดยมีค่า Cronbach's Alpha Coefficient จำแนกตามองค์ประกอบหลัก

แบบสำรวจนี้แบ่งเป็น 2 ชุด คือ ฉบับสมบูรณ์จำนวน 56 ข้อ และฉบับสั้นจำนวน 20 ข้อ โดยแบบวัดความสุขฉบับเต็ม สามารถวัดความสุขแบบละเอียดแยกรายมิติได้ ส่วนแบบวัดความสุขฉบับสั้น สามารถวัดระดับความสุขในภาพรวม โดยการวัดมีรูปแบบให้เลือกใช้ได้ 3 แบบ ได้แก่ การทดสอบในรูปแบบกระดาษ (Paper-Based) การทดสอบผ่านระบบออนไลน์ (Online-Based) และ การทดสอบผ่านแอปพลิเคชันของโทรศัพท์มือถือ (Mobile Apps-Based) ผลการวัดแสดงในลักษณะค่าเฉลี่ยความสุขรายมิติ ค่าเฉลี่ยความสุขรวมรายบุคคล และภาพรวมขององค์กร ค่าเฉลี่ยที่ได้สามารถนำมาเทียบกับเกณฑ์ที่พัฒนาขึ้นโดยค่าเฉลี่ยสูงแสดงถึงความสุขในระดับสูง

Diener, Emmons, Larsen, and Griffin (1985) ได้พัฒนามาตรวัด Satisfaction with Life Scale (SWLS) เพื่อประเมินความพึงพอใจต่อชีวิตของบุคคลในภาพรวม ซึ่งมีได้เจาะจงที่ด้านใดด้านหนึ่งของชีวิตโดยเฉพาะ เช่น สุขภาพ เศรษฐฐานะ แต่ให้ผู้ทดสอบสังเคราะห์ และให้นำหนักกับชีวิตด้านต่าง ๆ ตามที่ต้องการ มาตรวัดนี้ประกอบด้วยข้อคำถามจำนวน 5 ข้อ ซึ่งเลือกตอบได้ด้วยมาตรประมาณค่า 7 ระดับ ตั้งแต่รู้สึกเห็นด้วยกับข้อความที่ระบุ จนถึงไม่เห็นด้วยกับข้อความที่ระบุ โดยใช้เวลาในทดสอบ เพียงไม่กี่นาที คะแนนทั้งหมดจะอยู่ระหว่าง 7 ถึง 35 คะแนน โดยคะแนนสูงแสดงถึงมีความพึงพอใจในชีวิตมาก

Watson, Clark, and Tellegen (1988) พัฒนามาตรวัด The Positive and Negative Affect Schedule (PANAS) ซึ่งเป็นมาตรวัดอารมณ์ความรู้สึกเชิงบวกและเชิงลบ ที่ประกอบด้วยข้อคำถามจำนวน 20 ข้อ เป็นส่วนของมาตรวัดอารมณ์ความรู้สึกเชิงบวก (PA) 10 ข้อ และมาตรวัดอารมณ์ความรู้สึกเชิงลบ (NA) อีก 10 ข้อ โดยให้ผู้ทดสอบเลือกตอบข้อคำถาม ในลักษณะของความถี่ในการประสบกับอารมณ์ความรู้สึกต่าง ๆ ในช่วงเวลาที่กำหนด ในลักษณะมาตรความถี่ 5 ระดับ โดย 1 แสดงถึง ประสบกับอารมณ์ความรู้สึกตามข้อคำถามน้อยมากหรือแทบจะไม่เคยประสบเลย จนถึง 5 แสดงถึง ประสบกับอารมณ์ความรู้สึกตามข้อคำถามเป็นประจำหรือบ่อยครั้ง การศึกษาคุณสมบัติทางการวัดของ PANAS ปรากฏว่า มาตรวัดอารมณ์ความรู้สึกเชิงบวก (PA) มีค่าความเที่ยงเท่ากับ 0.89 ขณะที่มาตรวัดอารมณ์ความรู้สึกเชิงลบ (NA) มีค่าความเที่ยงเท่ากับ 0.85 เมื่อศึกษาความสัมพันธ์ระหว่างผลการทดสอบที่ได้จาก PANAS กับมาตรวัดความซึมเศร้าและมาตรวัดความวิตกกังวล ปรากฏว่า มาตรวัด PANAS มีความสัมพันธ์ทางลบกับมาตรวัดทั้งสองซึ่งสอดคล้องกับสมมติฐานที่ตั้งไว้ (Crawford & Henry, 2004)

McCullough, Emmons, and Tsang (2002) พัฒนา Gratitude Questionnaire - 6 (GQ-6) ซึ่งเป็นแบบประเมินตนเองในด้านความรู้สึกยินดี ประกอบด้วยข้อคำถามจำนวน 6 ข้อ ซึ่งสามารถเลือกตอบได้ในรูปแบบมาตรประมาณค่า 7 ระดับ โดย 1 แสดงถึง ไม่เห็นด้วยกับข้อคำถาม จนถึง 7 แสดงถึง เห็นด้วยกับข้อคำถามอย่างยิ่ง ข้อคำถาม 2 ข้อ จาก 6 ข้อ เป็นข้อคำถามเชิงลบที่สร้างขึ้นเพื่อป้องกันการตอบคำถามที่ไม่ตรงกับความเป็นจริง มาตรวัดนี้มีค่าความสอดคล้องภายในของข้อคำถามที่ดี ซึ่งมีค่าอยู่ระหว่าง 0.82 ถึง 0.87 นอกจากนี้ การศึกษาความสัมพันธ์ระหว่างผลการทดสอบที่ได้จากมาตรวัดนี้กับมาตรวัดตัวแปรทางจิตวิทยาเชิงบวกอื่น ๆ ได้แก่ การมองโลกในแง่ดี (Optimism) ความพึงพอใจในชีวิต (Life Satisfaction) ความหวัง (Hope) การมีจิตวิญญาณ และความเชื่อในศาสนา (Spirituality and Religiousness) การให้อภัย (Forgiveness) ความเห็นอก

เห็นใจและพฤติกรรมการช่วยเหลือสังคม (Empathy and Prosocial Behavior) พบความสัมพันธ์เชิงบวก ในขณะที่การศึกษาความสัมพันธ์ระหว่างผลการทดสอบที่ได้จากมาตรวัดนี้กับมาตรวัดตัวแปรทางจิตวิทยาเชิงลบอื่น ๆ ได้แก่ ความซึมเศร้า (Depression) ความวิตกกังวล (Anxiety) การให้ความสำคัญกับวัตถุนิยม (Materialism) และความอิจฉาริษยา (Envy) พบความสัมพันธ์เชิงลบ มาตรวัดนี้ใช้เวลาในการทดสอบน้อยกว่า 5 นาที อย่างไรก็ตาม มาตรวัดนี้ไม่ได้มีการกำหนดข้อจำกัดทางด้านเวลา ผู้ทดสอบสามารถใช้เวลาในการทดสอบได้นานเท่าที่ต้องการ

Thrash and Elliot (2003) สร้าง Inspiration Scale (IS) เพื่อใช้วัดแรงบันดาลใจและแรงจูงใจ มาตรวัดนี้ประกอบด้วยข้อคำถาม 4 ข้อ ผลรวมของคะแนนทั้งหมด สามารถบ่งบอกถึงแรงบันดาลใจของผู้ทดสอบ มาตรวัดนี้มีคุณสมบัติทางมาตรวัดที่ดี (Psychometric Properties) เป็นมาตรวัดที่มีโครงสร้างเป็นสององค์ประกอบ มีค่าความสอดคล้องภายในดี มีความคงที่ของผลที่ได้จากการทดสอบซ้ำ มีความไม่แปรเปลี่ยนทางการวัดตามเวลาหรือกลุ่มประชากร นอกจากนี้ มาตรวัดนี้ยังมีความตรงเชิงโครงสร้างที่ดี และมีประโยชน์ในการนำไปใช้ รวมทั้ง ใช้เวลาในการประเมินเพียง 1-2 นาที

Diener et al. (2010) ได้ร่วมกันพัฒนาเครื่องมือสำหรับวัดความอยู่ดีมีสุข (Well-being) โดยสร้างมาตรวัดที่สำคัญสองฉบับคือ มาตรวัด Flourishing Scale (FS) และมาตรวัด Scale of Positive and Negative Experience (SPANE) โดยศึกษาและเก็บรวบรวมข้อมูลในกลุ่มของนักศึกษาระดับมหาวิทยาลัยจำนวน 689 คน ที่อาศัยอยู่ในภูมิภาคที่แตกต่างกัน 6 แห่ง โดยมาตรวัด Flourishing Scale (FS) ซึ่งใช้วัดความเจริญรุ่งเรืองทางด้านจิตใจ ประกอบด้วยข้อคำถามจำนวน 8 ข้อ ซึ่งสอบถามเกี่ยวกับการรับรู้ถึงความสำเร็จของตนเองในองค์ประกอบที่สำคัญของชีวิต ได้แก่ สัมพันธภาพกับผู้อื่น ความรู้สึกมีคุณค่าในตนเอง จุดมุ่งหมายในชีวิต และการมองโลกในแง่ดี ผู้ทดสอบสามารถเลือกตอบได้ ในรูปแบบของมาตรประมาณค่า 7 ระดับ โดย 1 แสดงถึง ผู้ทดสอบไม่เห็นด้วยกับข้อคำถาม จนถึง 7 แสดงถึง ผู้ทดสอบเห็นด้วยกับข้อคำถามอย่างยิ่ง คะแนนรวมที่ได้มีค่าอยู่ระหว่าง 8 ถึง 56 คะแนน โดยคะแนนสูง แสดงให้เห็นว่าผู้ทดสอบนั้นมีแหล่งสนับสนุนทางด้านจิตใจมาก และมีจุดแข็งหลายประการ การศึกษาคุณสมบัติในการวัดของมาตรวัดนี้ ปรากฏว่า มาตรวัดนี้เป็นมาตรวัดที่มีหนึ่งองค์ประกอบ มีค่าความเที่ยงอยู่ในระดับสูง ค่าสัมประสิทธิ์ความเที่ยงเท่ากับ 0.87 การศึกษาความสัมพันธ์ของผลการทดสอบที่ได้จากมาตรวัดนี้ กับผลจากการทดสอบที่ได้จากมาตรวัดอื่นซึ่งศึกษาตัวแปรคุณลักษณะเดียวกัน พบว่า มีค่าความสัมพันธ์อยู่ระหว่าง 0.73 ถึง 0.78 ซึ่งแสดงถึงคุณสมบัติในการวัดที่ดี

ในส่วนของ Scale of Positive and Negative Experience (SPANE) ซึ่งใช้วัดความถี่ของการเกิดประสบการณ์ทางอารมณ์ความรู้สึกเชิงบวกและเชิงลบ ประกอบด้วย ข้อคำถามที่แสดงประสบการณ์ทางอารมณ์ความรู้สึกเชิงบวก 6 ข้อ และประสบการณ์ทางอารมณ์ความรู้สึกเชิงลบ 6 ข้อ ผู้ทดสอบสามารถเลือกตอบข้อคำถามได้ในลักษณะของมาตรวัดความถี่ 5 ระดับ โดย 1 แสดงถึงไม่เคยเกิดประสบการณ์ทางอารมณ์ความรู้สึกดังข้อคำถามเลย จนถึง 5 แสดงถึง มีประสบการณ์ทางอารมณ์ความรู้สึกดังข้อคำถามบ่อยมากหรือประสบอยู่เป็นประจำ การคิดคะแนนใช้วิธีรวมคะแนนจากข้อคำถามทั้งหมด ซึ่งได้เป็นค่า Affect Balance (SPANE-B) โดยนำคะแนนที่ได้จากข้อคำถามประสบการณ์ทางอารมณ์ความรู้สึกเชิงลบ หักออกจากคะแนนที่ได้จากข้อคำถามประสบการณ์ทาง

อารมณ์ความรู้สึกเชิงบวก ซึ่งจะได้ผลต่างของคะแนนอยู่ระหว่าง -24 คะแนน ซึ่งแสดงถึงความรู้สึกไม่เป็นสุข จนถึง 24 คะแนน ซึ่งแสดงให้เห็นว่า ผู้ทดสอบนั้นประสบกับประสบการณ์ทางอารมณ์ความรู้สึกเชิงบวกอยู่เสมอ และแทบจะไม่เคยประสบกับประสบการณ์ทางอารมณ์เชิงลบเลย อย่างไรก็ตาม มาตรการนี้สามารถแยกคิดคะแนนเป็นมาตรการส่วนย่อยได้อีกสองส่วน คือ คะแนน Positive Feeling (SPANE-P) ซึ่งได้จาก ผลรวมคะแนนข้อคำถามประสบการณ์ทางอารมณ์ความรู้สึกเชิงบวกทั้ง 6 ข้อ ซึ่งจะมีคะแนนอยู่ระหว่าง 6 ถึง 30 คะแนน อีกส่วนหนึ่ง คือ คะแนน Negative Feeling (SPANE-N) ซึ่งได้จาก ผลรวมคะแนนข้อคำถามประสบการณ์ทางอารมณ์ความรู้สึกเชิงลบทั้ง 6 ข้อ ซึ่งจะมีคะแนนอยู่ระหว่าง 6 ถึง 30 คะแนนเช่นกัน การศึกษาคุณสมบัติในการวัดของมาตรการนี้พบว่า มาตรการนี้มีค่าความเที่ยงของคะแนนในแต่ละมาตรการย่อยอยู่ในระดับสูง ค่าความเที่ยงของคะแนน SPANE-P SPANE-N และ SPANE-B มีค่าเท่ากับ 0.87 0.81 และ 0.89 ตามลำดับ และเมื่อศึกษาความสัมพันธ์กับมาตรการอื่น ๆ ที่วัดตัวแปรทางจิตวิทยาประเภทเดียวกัน พบว่า มาตรการที่สร้างขึ้นมีความสัมพันธ์ที่ดีกับมาตรการทางด้านอารมณ์ ความอยู่ดีมีสุข ความสุข และความพึงพอใจในชีวิต

จากการทบทวนเอกสารเกี่ยวกับนิยามของความสุข ปรากฏว่า ความสุขมีนิยามที่แตกต่างกันไปหลากหลายแนวคิด บางแนวคิดเชื่อว่าความสุขมีองค์ประกอบเดียว ในขณะที่บางแนวคิดเชื่อว่าความสุขมีหลากหลายองค์ประกอบ บางมุมมองเน้นการพิจารณาความสุขในระดับพื้นฐาน ซึ่งสามารถสังเกตหรือวัดได้จากพฤติกรรมที่แสดงออก ในขณะที่บางมุมมองเป็นการพิจารณาความสุขในระดับลึกซึ้ง มีความซับซ้อน และยากที่จะวัดหรือประเมินได้อย่างชัดเจน

มาตรการความสุขที่สร้างขึ้นต่างก็มีความแตกต่างกันไปตามนิยามความสุขที่เลือกใช้ ในต่างประเทศ นิยมสร้างมาตรการที่เน้นวัดคุณลักษณะเฉพาะที่แสดงถึงองค์ประกอบของความสุขเพียงองค์ประกอบใดองค์ประกอบหนึ่ง แล้วจึงนำผลที่ได้จากมาตรการต่าง ๆ มาประเมินความสุขในภาพรวม มาตรการความสุขส่วนใหญ่ในต่างประเทศ จึงประกอบด้วยข้อคำถามจำนวนไม่มากนัก และใช้เวลาในการทดสอบไม่นาน ในขณะที่ประเทศไทย การสร้างมาตรการความสุขส่วนใหญ่ พยายามสังเคราะห์แนวคิดของความสุขให้ครอบคลุมองค์ประกอบที่เกี่ยวข้องทั้งหมด โดยมีการปรับแนวคิดให้เหมาะสมกับบริบทของคนไทยมากยิ่งขึ้น เช่น การนำแนวคิดความสุขในมุมมองของพระพุทธศาสนา เข้ามาเชื่อมโยง การออกแบบให้เหมาะสมกับกลุ่มที่ต้องการวัดอย่างเฉพาะเจาะจง การตีความหมายความสุขให้เทียบเท่ากับภาวะสุขภาพจิต ลักษณะมาตรการความสุขในประเทศไทยส่วนใหญ่ จึงประกอบด้วยข้อคำถามจำนวนมากที่วัดครอบคลุมในสิ่งที่ผู้วิจัยมุ่งเน้น อย่างไรก็ตาม ข้อจำกัดทางด้านเวลา และความสะดวกในการทดสอบ ผู้พัฒนามาตรการหลายกลุ่มจึงได้ลดทอนจำนวนข้อคำถามจากฉบับสมบูรณ์ เพื่อให้ได้ข้อคำถามฉบับสั้นที่สะดวกในการทดสอบมากยิ่งขึ้น

ตอนที่ 2 ทฤษฎีการตอบสนองข้อสอบและงานวิจัยที่เกี่ยวข้อง

การวัดตัวแปรทางจิตวิทยา (Psychological Construct) เป็นการวัดคุณลักษณะภายในของบุคคลที่ไม่สามารถสังเกตได้โดยตรง เพราะมีลักษณะเป็นตัวแปรแฝง (Latent Variables) ที่มีอิทธิพลอยู่เบื้องหลังพฤติกรรมที่แสดงออก การวัดคุณลักษณะภายในจึงต้องวัดผ่านตัวแปรสังเกตได้

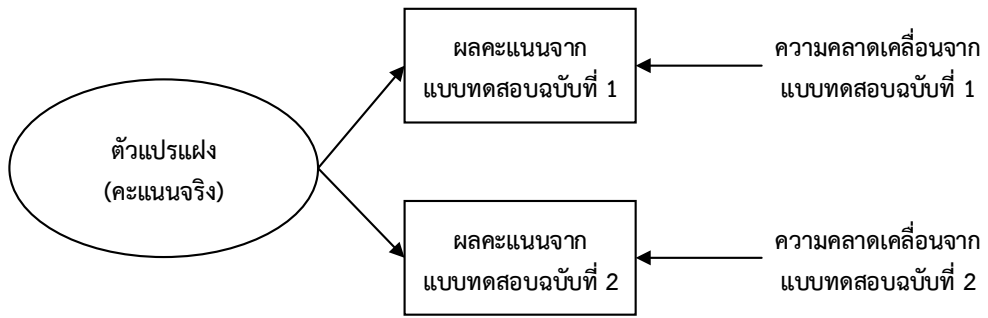
(Observable Variables) เช่น คะแนนที่ได้จากแบบทดสอบ (Test Scores) หรือการตอบข้อคำถาม (Item Responses)

อย่างไรก็ดี คะแนนที่ได้จากแบบทดสอบ และการตอบข้อคำถามนั้น เป็นเพียงตัวชี้วัดตัวหนึ่งที่บ่งบอกถึงระดับคุณลักษณะของบุคคล แต่ไม่สามารถอธิบายตัวแปรแฝงนั้นได้อย่างสมบูรณ์ การวัดคุณลักษณะทางจิตวิทยาส่วนใหญ่เป็นการวัดทางอ้อมจากพฤติกรรมที่สังเกตได้ ซึ่งผู้ทดสอบแสดงออกผ่านงาน (Tasks) หรือข้อคำถาม (Items) ที่เกี่ยวข้อง การวัดคุณสมบัติ (Properties) ต่าง ๆ ของบุคคลและข้อคำถาม ในมิติทางด้านจิตวิทยาจึงต้องวัดผ่านทางพฤติกรรม โดยนำทฤษฎีการวัดมาเป็นแนวทางในการแสดงความสัมพันธ์ระหว่างพฤติกรรมกับคุณลักษณะทางจิตวิทยาที่ต้องการวัดอย่างเป็นเหตุเป็นผล (Rationale) โดยทั่วไป ทฤษฎีการวัด (Measurement Theory) จะแสดงความสัมพันธ์ระหว่างพฤติกรรมกับคุณลักษณะทางจิตวิทยาที่ต้องการวัด ด้วยโมเดลเชิงพฤติกรรม (Model of Behavior) ทั้งทฤษฎีการทดสอบแบบดั้งเดิม (Classical Test Theory: CTT) และทฤษฎีการตอบสนองข้อสอบ (Item Response Theory: IRT) ต่างก็กล่าวถึงความสัมพันธ์ระหว่างพฤติกรรมและคุณลักษณะทางจิตวิทยาเหมือนกัน อย่างไรก็ตาม ทั้งสองทฤษฎีมีแนวทางในการอธิบายความสัมพันธ์ดังกล่าวแตกต่างกันอย่างมาก (Embretson & Reise, 2000, pp. 40-41)

หลักการพื้นฐานของทฤษฎีการตอบสนองข้อสอบ

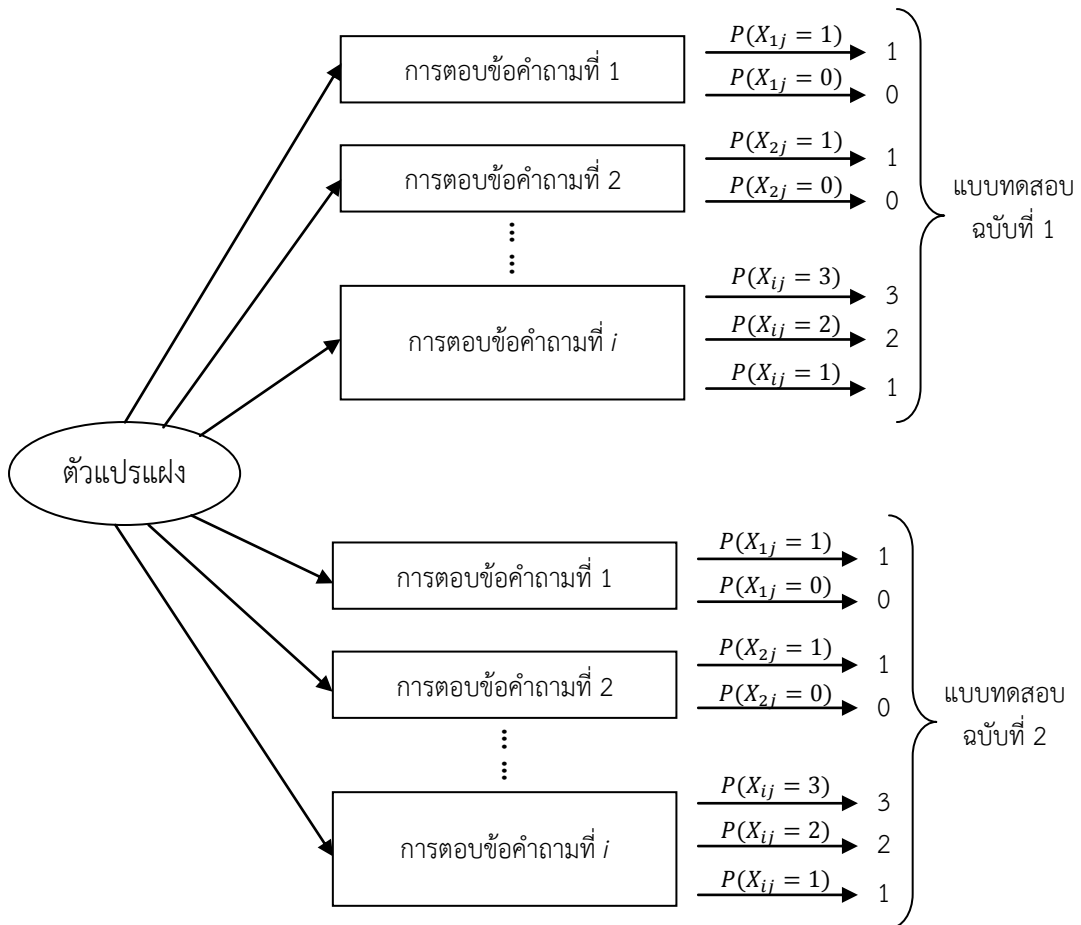
โมเดลการวัด (Measurement Model) เป็นโมเดลเชิงคณิตศาสตร์ (Mathematical Model) ซึ่งแสดงความสัมพันธ์ระหว่างกลุ่มของตัวแปรต้น (Independent Variables) เพื่อทำนายตัวแปรตาม (Dependent Variables) โมเดลดังกล่าว อาจเป็นโมเดลอย่างง่ายที่ใช้ผลรวมของตัวแปรต้นเพื่อทำนายตัวแปรตามในลักษณะของโมเดลถดถอยเชิงเส้นตรง (Linear Regression Model) หรืออาจเป็นโมเดลความสัมพันธ์ที่มีความซับซ้อนมากขึ้น เช่น การรวมตัวกันของตัวแปรต้นโดยที่ตัวแปรอาจมีปฏิสัมพันธ์กัน หรือเกิดขึ้นภายใต้ฟังก์ชันการแจกแจงรูปแบบต่าง ๆ เช่น การแจกแจงปกติสะสม (Cumulative Normal Distribution) การแจกแจงแบบโลจิสติก (Logistic Distribution)

ภาพที่ 2-1 และ ภาพที่ 2-2 แสดงให้เห็นแนวคิดของทฤษฎีการทดสอบแบบดั้งเดิมและทฤษฎีการตอบสนองข้อสอบ เพื่ออธิบายความสัมพันธ์ระหว่างตัวแปรแฝงกับพฤติกรรมที่แสดงออก โดยภาพที่ 2-1 แสดงให้เห็นแนวคิดของโมเดลการทดสอบแบบดั้งเดิม (CTT Model) ซึ่งอธิบายว่าคะแนนที่สังเกตได้ (ผลคะแนนที่ได้จากแบบทดสอบ) ประกอบด้วยคะแนนจริง (True Score) และความคลาดเคลื่อน (Error)



ภาพที่ 2-1 โมเดลการทดสอบแบบดั้งเดิม (Embretson & Reise, 2000, p. 42)

ภาพที่ 2-2 แสดงให้เห็นโมเดลการตอบสนองข้อสอบ (IRT Model) ซึ่งเป็นโมเดลเชิงคณิตศาสตร์ที่อธิบายว่า ตัวแปรสังเกตได้ คือ ผลการตอบข้อคำถามแต่ละข้อ (เช่น 0 หรือ 1) ขณะที่ตัวแปรแฝงจะเป็นตัวแปรที่มีอิทธิพลอยู่เบื้องหลังความน่าจะเป็น (Probability) ของการตอบข้อคำถามถูก หรือความน่าจะเป็นในการเลือกรายการคำตอบ (Response Options) ความน่าจะเป็นนี้แสดงในรูปของ $P(X_{is} = 1)$



ภาพที่ 2-2 โมเดลการตอบสนองข้อสอบ (Embretson & Reise, 2000, p. 42)

โมเดลการทดสอบแบบดั้งเดิมเป็นโมเดลอย่างง่ายมีตัวแปรตาม คือ ผลรวมคะแนนดิบของผู้ทดสอบ (แทนด้วย X_{Os}) และตัวแปรต้นสองตัว ได้แก่ คะแนนจริงของผู้ทดสอบซึ่งเป็นคุณลักษณะแฝงที่ต้องการวัด (แทนด้วย X_{Ts}) กับคะแนนความคลาดเคลื่อนที่เกิดจากการวัด (แทนด้วย X_{Es}) ผลรวมเชิงเส้นตรงของตัวแปรต้นสามารถทำนายตัวแปรตามได้ ดังสมการ

$$X_{Os} = X_{Ts} + X_{Es} \quad (1)$$

โดยมีข้อตกลงเบื้องต้น ของคะแนนความคลาดเคลื่อน คือ 1) ค่าคาดหวัง (Expected Value) ของความคลาดเคลื่อนในกลุ่มผู้ทดสอบมีค่าเป็นศูนย์ และ 2) คะแนนความคลาดเคลื่อนต้องไม่มีความสัมพันธ์กับคะแนนจริง หรือคะแนนความคลาดเคลื่อนอื่น ๆ นอกจากนี้ ในการแปลความหมายตัวแปรต่าง ๆ จากโมเดลการทดสอบแบบดั้งเดิมอาจมีข้อตกลงเบื้องต้นเพิ่มเติม เช่น เมื่อต้องการแปลความหมายความคลาดเคลื่อนมาตรฐานจากการวัด (Standard Error of Measurement) จะต้องมีข้อตกลงเบื้องต้นเพิ่มขึ้นว่า คะแนนความคลาดเคลื่อนของแต่ละบุคคล จะต้องมีการแจกแจงแบบปกติ และการแจกแจงของคะแนนความคลาดเคลื่อนระหว่างบุคคล จะต้องมีความเป็นเอกพันธ์

โมเดลการทดสอบแบบดั้งเดิมมีข้อจำกัดหลายประการ ดังนี้

1. คะแนนจริงเป็นค่าเฉพาะซึ่งได้มาจากแบบทดสอบหรือมาตรวัดที่ใช้วัดในแต่ละครั้ง ค่าพารามิเตอร์ข้อคำถามจึงเป็นค่าเฉพาะของแบบทดสอบหรือมาตรวัดที่ใช้วัดในครั้งนั้น ๆ หากมีชุดของข้อคำถามมากกว่าหนึ่งชุดที่มุ่งวัดคุณลักษณะเดียวกัน การอ้างอิงคะแนนจริงจากแบบทดสอบหรือมาตรวัดต่างชุดทำได้โดยใช้แนวคิดของแบบทดสอบคู่ขนาน (Parallelism) หรือการปรับเทียบแบบทดสอบ (Test Equating) เท่านั้น

2. โมเดลการทดสอบแบบดั้งเดิมแสดงให้เห็นถึงตัวแปรอิสระสองตัวที่แยกจากกัน ได้แก่ คะแนนจริงและความคลาดเคลื่อน แต่ในความเป็นจริงแล้วตัวแปรทั้งสองไม่สามารถแยกจากกันได้ โดยสมบูรณ์ โมเดลการทดสอบแบบดั้งเดิมจึงใช้เพื่ออธิบายความเป็นเหตุเป็นผลของการประมาณค่าทางสถิติของประชากรเท่านั้น จากสมการที่ 1 จะเห็นได้ว่าต้องมีการประมาณค่าความแปรปรวนของคะแนนจริง (True Variance) และความแปรปรวนของความคลาดเคลื่อน (Error Variance) แต่ในสถานการณ์จริงการวิเคราะห์คะแนนจริงแยกออกจากความคลาดเคลื่อนทำได้ยาก เพราะการวิเคราะห์ดังกล่าวอยู่ภายใต้เงื่อนไขและข้อจำกัดจำนวนมาก

3. โมเดลการทดสอบแบบดั้งเดิม (ดังสมการที่ 1) แสดงให้เห็นว่าคุณสมบัติของข้อคำถาม (Item Properties) ไม่ได้เชื่อมโยงกับพฤติกรรมโดยตรง ดังนั้น การวิเคราะห์ค่าคุณสมบัติต่าง ๆ ของข้อคำถาม เช่น ค่าความยากของข้อคำถาม (Item Difficulty) ค่าอำนาจจำแนกของข้อคำถาม (Item Discrimination) ซึ่งใช้ในการคัดเลือกข้อคำถามที่มีความเหมาะสม จึงต้องใช้หลักการอื่น นอกเหนือจากโมเดลการทดสอบแบบดั้งเดิม เช่น ความแปรปรวน (Variance) และความเที่ยง (Reliability) (Embretson & Reise, 2000, pp. 41-43)

ทฤษฎีการตอบสนองข้อสอบ มีแนวคิดที่สำคัญสองประการ คือ 1) พฤติกรรมการตอบ

ข้อคำถามของผู้ทดสอบ (Examinee) สามารถทำนาย หรืออธิบายได้ด้วยชุดขององค์ประกอบ ที่เรียกว่า คุณลักษณะ (Traits) คุณลักษณะแฝง (Latent Traits) หรือความสามารถ (Abilities) และ 2) ความสัมพันธ์ระหว่างพฤติกรรมคำตอบข้อคำถามของผู้ทดสอบ กับชุดของคุณลักษณะแฝงซึ่งมีอิทธิพลอยู่เบื้องหลังพฤติกรรมคำตอบข้อคำถามนั้น สามารถอธิบายได้ด้วยฟังก์ชันคณิตศาสตร์ที่เรียกว่า ฟังก์ชันลักษณะข้อสอบ (Item Characteristic Function) หรือ โค้งลักษณะข้อสอบ (Item Characteristic Curve: ICC) ซึ่งเมื่อระดับของคุณลักษณะเพิ่มขึ้น ความน่าจะเป็นในการตอบข้อคำถามถูก (Correct Response) จะเพิ่มขึ้นด้วย (Hambleton, Swaminathan, & Rogers, 1991, p. 7)

โค้งลักษณะข้อสอบมีลักษณะเป็นฟังก์ชันโลจิส (Logistic Function) หรือฟังก์ชันปกติสะสม (Normal Ogive Function) บางครั้งอาจเรียกว่า โมเดลโลจิส หรือโมเดลปกติสะสม ซึ่งมีหลากหลายโมเดล (ศิริชัย กาญจนวาสี, 2555, หน้า 53) โดยทั่วไป การนำโมเดลการตอบสนองข้อสอบไปใช้ทางการศึกษา หรือทางจิตวิทยา จะต้องมีการตรวจสอบข้อตกลงเบื้องต้นของโมเดลการตอบสนองข้อสอบ 2 ประการ ได้แก่ (ศิริชัย กาญจนวาสี, 2555, หน้า 75-76; Cohen, Swerdlik, & Sturman, 2013, pp. 170-171; Embretson & Reise, 2000, pp. 45-48; Hambleton, Swaminathan, & Rogers, 1991, pp. 9-12)

1. ความเป็นเอกมิติ (Unidimensionality)

ความเป็นเอกมิติ หมายถึง การที่ชุดของข้อคำถาม (หรือข้อสอบ) ในมาตรวัด (หรือแบบทดสอบ) มุ่งวัดคุณลักษณะแฝงเพียงคุณลักษณะเดียว (หรือความสามารถของบุคคลเพียงความสามารถเดียว)

ในความเป็นจริง ข้อตกลงเบื้องต้นดังกล่าว ไม่สามารถพบได้อย่างสมบูรณ์ เนื่องจากปัจจัยทางด้านการรู้คิด (Cognitive) บุคลิกภาพ (Personality) และปัจจัยของการทดสอบ เช่น แรงจูงใจ ความกังวลในการทดสอบ ความรวดเร็วในการทำแบบทดสอบ การเดาคำตอบ และทักษะทางการรู้คิด อาจส่งผลร่วมกับพฤติกรรมที่แสดงออกในการทดสอบอยู่เสมอ ผลการทดสอบที่เกิดขึ้นจึงมิได้เกิดจากคุณลักษณะที่ต้องการวัดเพียงคุณลักษณะเดียวอย่างแท้จริง อย่างไรก็ตาม ในทางปฏิบัติ การตรวจสอบความเป็นเอกมิติ ต้องการเพียงหลักฐานที่แสดงให้เห็นว่า ชุดของข้อคำถามมุ่งวัดองค์ประกอบหลัก (Dominant Factor) เพียงองค์ประกอบเดียว แบบทดสอบใดที่มีคุณลักษณะดังกล่าว จะถือว่า แบบทดสอบนั้นมีคุณลักษณะเป็นโมเดลเอกมิติ (Unidimensional Model) ขณะที่แบบทดสอบใดมีการมุ่งวัดองค์ประกอบหลักมากกว่าหนึ่งองค์ประกอบ แบบทดสอบนั้นแสดงให้เห็นถึงโมเดลพหุมิติ (Multidimensional Model) ซึ่งเป็นโมเดลที่มีความซับซ้อนมากกว่า

การตรวจสอบความเป็นเอกมิติของเครื่องมือวัด สามารถกระทำได้โดยการใช้เทคนิควิเคราะห์ทางสถิติ ได้แก่ การวิเคราะห์องค์ประกอบ (Factor Analysis) เพื่อคำนวณค่าไอเกน (Eigen Value) สำหรับศึกษาอัตราส่วนระหว่างค่าไอเกนขององค์ประกอบแรก กับองค์ประกอบถัดไป ถ้ามีอัตราส่วนสูง แสดงว่าเครื่องมือวัดนั้นมุ่งวัดคุณลักษณะหลักเพียงคุณลักษณะเดียว (Single Dominant Factor) นอกจากนี้ สามารถวิเคราะห์องค์ประกอบเชิงยืนยัน (Confirmatory Factor Analysis) เพื่อเพิ่มความมั่นใจ และช่วยยืนยันได้ว่า เครื่องมือวัดนั้นมุ่งวัดเพียงคุณลักษณะเดียว

2. ความเป็นอิสระ (Local Independence)

ความเป็นอิสระ หมายถึง เมื่อควบคุมคุณลักษณะหรือความสามารถของบุคคลให้คงที่แล้ว ผลการตอบข้อคำถามในข้อคำถามแต่ละคู่ จะต้องมีความเป็นอิสระจากกันทางสถิติ หรือกล่าวอีกนัยหนึ่งได้ว่า คุณลักษณะแฝงภายในบุคคล ต้องเป็นปัจจัยเดียวที่ทำให้การตอบข้อคำถามของผู้ทดสอบมีความแตกต่างกัน ความแตกต่างของการตอบข้อคำถามต้องไม่เกิดขึ้นจากความสัมพันธ์รูปแบบอื่น ความเป็นอิสระสามารถอธิบายในเชิงคณิตศาสตร์ ได้ดังนี้

ความเป็นอิสระ หมายถึง ความน่าจะเป็นในการตอบข้อคำถามข้อที่ i ($Prob(X_{is} = 1)$) มีความเป็นอิสระจากผลการตอบข้อคำถามข้ออื่น ๆ (i') เมื่อมีการควบคุมพารามิเตอร์บุคคล (θ_s) และพารามิเตอร์ข้อสอบ (ε_i) ดังสมการที่ 2

$$Prob(X_{is} = 1 | X_{i's} = 1, \varepsilon_k, \theta_s) = Prob(X_{is} = 1 | \varepsilon_k, \theta_s) \quad (2)$$

จากสมการที่ 2 จะสังเกตได้ว่า ความน่าจะเป็นในการตอบข้อสอบข้อที่ i อย่างมีเงื่อนไข จากผลการตอบข้อคำถามข้อที่ i' มีความน่าจะเป็นเท่ากับ ความน่าจะเป็นในการตอบข้อสอบข้อที่ i หากพิจารณาอย่างผิวเผิน แนวคิดเรื่องความเป็นอิสระของข้อคำถามอาจดูขัดแย้งกับแนวคิดเรื่องความสอดคล้องภายในของข้อคำถามตามทฤษฎีการทดสอบแบบดั้งเดิม ที่ระบุว่า ข้อคำถามที่มีความสัมพันธ์กันมาก แสดงให้เห็นว่ามาตรวัดนั้นมีความเที่ยงในระดับสูง อย่างไรก็ตาม เมื่อพิจารณาอย่างละเอียดจะพบว่า แนวคิดตามทฤษฎีทั้งสองนั้น มิได้ขัดแย้งกัน เพราะความเป็นอิสระของข้อคำถามตามทฤษฎีการตอบสนองข้อสอบนั้น เป็นความอิสระของข้อคำถาม ที่เกิดขึ้นภายหลังการควบคุมพารามิเตอร์บุคคล และพารามิเตอร์ข้อสอบที่มีอิทธิพลโดยตรงต่อการตอบข้อคำถามในโมเดลแล้ว

โมเดลการตอบสนองข้อสอบ

ทฤษฎีการตอบสนองข้อสอบมีคุณลักษณะเป็นโมเดลทางคณิตศาสตร์ การเลือกใช้โมเดลอย่างเหมาะสมจึงส่งผลสำคัญต่อคุณภาพของมาตรวัด ในเบื้องต้น การเลือกใช้โมเดลการตอบสนองข้อสอบต้องพิจารณาเลือกโมเดลที่มีความสอดคล้องกับข้อมูล เพราะโมเดลการตอบสนองข้อสอบมีให้เลือกใช้อย่างหลากหลาย แต่ละโมเดลมีความเหมาะสมกับข้อมูลที่มีลักษณะแตกต่างกันออกไป โมเดลการตอบสนองข้อสอบสามารถแบ่งตามชนิดของข้อมูลได้ 2 ประเภท ได้แก่ โมเดลการตอบสนองข้อสอบแบบตรวจให้คะแนน 2 ค่า (Binary IRT Model หรือ Dichotomous IRT Model) และโมเดลการตอบสนองข้อสอบแบบตรวจให้คะแนนมากกว่า 2 ค่า (Polytomous IRT Model) (ศิริชัยกาญจนวาสี, 2555, หน้า 51-114; De Ayala, 2009, pp. 11-38, 99-274; Embretson & Reise, 2000, pp. 65-124; Hambleton, Swaminathan, & Rogers, 1991, pp. 12-28)

1. โมเดลการตอบสนองข้อสอบแบบตรวจให้คะแนน 2 ค่า (Dichotomous IRT Models)

โมเดลการตอบสนองข้อสอบรูปแบบนี้เหมาะกับข้อมูลที่เมื่อตรวจให้คะแนน (Scored) แล้วมีคะแนนที่เป็นไปได้สองค่า (Binary Data) คือ 0 กับ 1 เช่น สำเร็จ (Success ให้ 1 คะแนน) และล้มเหลว (Failure ให้ 0 คะแนน) ลักษณะของการตรวจให้คะแนน 2 ค่า อาจได้มาจากการตอบ

ข้อคำถามหลากหลายรูปแบบ เช่น ในแบบทดสอบวัดความสามารถที่มีตัวเลือก “ถูกหรือผิด” (Right vs. Wrong) ในแบบประเมินตนเองด้านบุคลิกภาพที่มีตัวเลือกกว่า “เป็นจริงหรือไม่เป็นจริง” (True vs. Not True) ในแบบทดสอบเจตคติที่มีตัวเลือกกว่า “เห็นด้วยหรือไม่เห็นด้วย” (Agree vs. Disagree) และในมาตราประมาณค่าทางด้านพฤติกรรมที่มีตัวเลือกกว่า “ใช่หรือไม่ใช่” (Yes vs. No) การตอบข้อคำถามดังที่กล่าวมานี้ ล้วนมีการตรวจให้คะแนนแบบสองค่าทั้งสิ้น นอกจากนี้ หากพิจารณาด้านมิติทางการวัดจะเห็นได้ว่า โมเดลการตอบสนองข้อสอบแบบตรวจให้คะแนน 2 ค่า นั้น สามารถแบ่งตามมิติของการวัดตัวแปรได้ 2 ประเภท คือ โมเดลการตอบสนองข้อสอบแบบเอกมิติ (Unidimensional IRT Models) ซึ่งมุ่งวัดความแตกต่างของบุคคลที่กำหนดโดยคุณลักษณะแฝงเพียงหนึ่งคุณลักษณะ และโมเดลการตอบสนองข้อสอบแบบพหุมิติ (Multidimensional IRT Models) ซึ่งมุ่งวัดความแตกต่างของบุคคลที่กำหนดโดยคุณลักษณะมากกว่าหนึ่งคุณลักษณะ

1.1 โมเดลการตอบสนองข้อสอบแบบเอกมิติสำหรับข้อมูลที่มีลักษณะสองค่า

(Unidimensional IRT Models for Binary Data)

โมเดลการตอบสนองข้อสอบแบบเอกมิติมีแนวคิดว่าคุณลักษณะแฝงเพียงหนึ่งคุณลักษณะสามารถอธิบายความแตกต่างที่เกิดขึ้นระหว่างบุคคลได้ โมเดลนี้เหมาะสมกับข้อมูลที่มีการสร้างข้อคำถามจากคุณลักษณะแฝงเพียงหนึ่งคุณลักษณะ อย่างไรก็ตาม บางครั้งโมเดลการตอบสนองข้อสอบแบบเอกมิติอาจเหมาะสมกับข้อคำถามที่สร้างขึ้นจากองค์ประกอบมากกว่าหนึ่งองค์ประกอบ เมื่อข้อคำถามทั้งหมดเกี่ยวข้องกับผลรวมของแต่ละองค์ประกอบ อย่างไรก็ตาม โมเดลการตอบสนองข้อสอบแบบเอกมิติไม่เหมาะสมกับข้อมูลที่มีลักษณะดังนี้ 1) ข้อคำถามได้รับอิทธิพลจากคุณลักษณะแฝงมากกว่าหนึ่งคุณลักษณะ และ 2) บุคคลมีกลวิธี โครงสร้างความรู้ หรือการแปลความข้อคำถามแตกต่างกัน ข้อมูลที่มีลักษณะสองประการนี้ เหมาะที่จะวิเคราะห์ด้วยโมเดลการตอบสนองข้อสอบแบบพหุมิติ

1.1.1 โมเดลโลจิสแบบดั้งเดิม (Traditional Logistic Models)

โมเดลโลจิสแบบดั้งเดิม อยู่บนพื้นฐานการแจกแจงแบบโลจิสติก ซึ่งแสดงความน่าจะเป็นด้วยรูปแบบอย่างง่าย โดยเฉพาะอย่างยิ่งเมื่อ ω_{is} แทนผลรวมพารามิเตอร์บุคคลและพารามิเตอร์ข้อสอบในโมเดล ความน่าจะเป็นสามารถคำนวณได้จากสมการ

$$P(X_{is} = 1 | \omega_{is}) = \frac{\exp(\omega_{is})}{1 + \exp(\omega_{is})} \quad (3)$$

เมื่อ $\exp(\omega_{is})$ (สามารถเขียนได้ในรูป $e^{\omega_{is}}$) หมายถึง ลอการิทึมธรรมชาติ (มีค่าเท่ากับ 2.718) ยกกำลัง ω_{is} ค่าความน่าจะเป็นในการตอบข้อคำถามสำเร็จ (Success) สามารถคำนวณได้จากการหาค่าแอนติล็อกกาลีทึมตามสมการที่ 3

1.1.1.1 โมเดลโลจิสแบบหนึ่งพารามิเตอร์ หรือ ราส์ชโมเดล (One-Parameter Logistic Model (1PL) or Rasch Model)

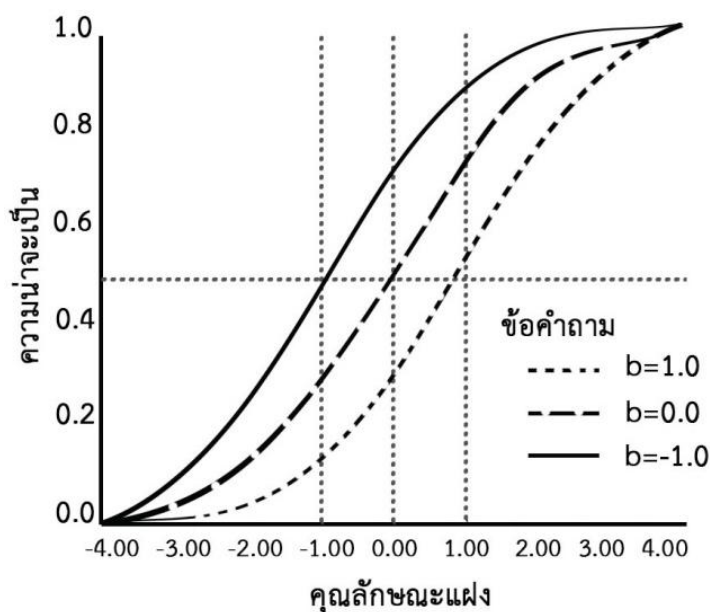
ราส์ชโมเดลคาดการณ์ความน่าจะเป็นในการตอบข้อคำถามสำเร็จ (Success) หรือตอบข้อสอบถูก ของบุคคลที่ S ต่อข้อคำถามข้อที่ i (ได้แก่ $P(X_{is} = 1)$) ดังสมการ

$$P(X_{is} = 1 | \theta_s, \beta_i) = \frac{\exp(\theta_s - \beta_i)}{1 + \exp(\theta_s - \beta_i)} \quad (4)$$

โดย

- X_{is} หมายถึง การตอบข้อคำถามของบุคคลที่ S ต่อข้อคำถามข้อที่ i (0 หรือ 1)
 θ_s หมายถึง ระดับคุณลักษณะแฝงที่ S
 β_i หมายถึง พารามิเตอร์ความยากของข้อคำถามข้อที่ i

จากสมการที่ 4 $\theta_s - \beta_i$ คือ ผลต่างของระดับของคุณลักษณะแฝง (Trait level) กับ ค่าความยากของข้อคำถาม (Item Difficulty) ซึ่งแทนที่ ω_{is} ในสมการที่ 3 ราส์ชโมเดลสามารถเรียกได้ว่าเป็นโมเดลโลจิสแบบหนึ่งพารามิเตอร์ (One-parameter Logistic (1PL) Model) เนื่องจากมีพารามิเตอร์ข้อสอบเพียงหนึ่งพารามิเตอร์ ตัวอย่างโค้งลักษณะข้อสอบของราส์ชโมเดล แสดงได้ดังภาพที่ 2-3



ภาพที่ 2-3 โค้งลักษณะข้อสอบของโมเดลโลจิสแบบหนึ่งพารามิเตอร์หรือราส์ชโมเดล (Embretson & Reise, 2000, p. 68)

จากภาพที่ 2-3 แสดงถึงโค้งลักษณะข้อสอบ (Item Characteristic Curves: ICCs) จาก ราส์ชโมเดลของข้อคำถามจำนวนสามข้อ โค้งลักษณะข้อสอบนี้ มีลักษณะหลายประการ คือ

- 1) ความน่าจะเป็นของข้อคำถามแต่ละข้อค่อย ๆ เพิ่มขึ้นตามคุณลักษณะแฝงที่เพิ่มขึ้น
- 2) ข้อคำถามแต่ละข้อมีคุณลักษณะที่แตกต่างกันเฉพาะในส่วนของพารามิเตอร์ความยาก แต่มีความชันของโค้งลักษณะข้อสอบไม่แตกต่างกัน โดยโค้งลักษณะข้อสอบมีการลู่ออกกันแต่ไม่ตัดกัน และ
- 3) จุดเปลี่ยนโค้งลักษณะข้อสอบ มีความน่าจะเป็นในการตอบข้อคำถามถูก (หรือเลือกคำตอบที่ได้ 1 คะแนน) เท่ากับ 0.50 นอกจากนี้ ภาพนี้ได้แสดงเส้นอ้างอิงที่ระดับคุณลักษณะแฝงต่าง ๆ (เส้นประ

ที่ลากจากแกนนอน) ซึ่งมีค่าเท่ากับค่าพารามิเตอร์ความยากของข้อคำถามเมื่อโคงลักษณะข้อสอบมีค่าความน่าจะเป็นในการตอบข้อคำถามถูกเท่ากับ 0.50 ความสัมพันธ์ดังกล่าวทำให้สามารถแปลความหมายระดับคุณลักษณะแฝงเชื่อมโยงกับค่าพารามิเตอร์ความยากของข้อคำถามได้ เช่น ข้อคำถามที่มีค่าพารามิเตอร์ความยากเท่ากับ 1.00 หมายความว่า บุคคลที่มีระดับคุณลักษณะแฝงเท่ากับ 1.00 มีความน่าจะเป็นที่จะตอบข้อคำถามถูกเท่ากับ 0.50

1.1.1.2 โมเดลโลจิสแบบสองพารามิเตอร์ (Two-Parameter Logistic Model (2PL))

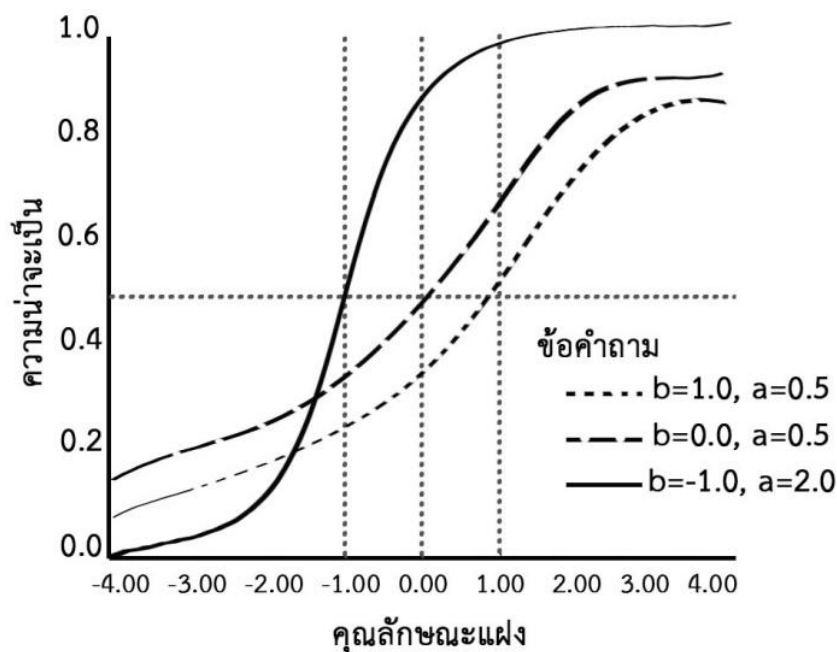
โมเดลโลจิสแบบสองพารามิเตอร์ มีพารามิเตอร์ข้อสอบเพิ่มขึ้นอีกหนึ่งพารามิเตอร์ คือ พารามิเตอร์อำนาจจำแนก (Discrimination Parameters) ในกรณีนี้ ความน่าจะเป็นของบุคคลที่ S ในการตอบข้อคำถามข้อที่ i สามารถคำนวณได้จากสมการที่ 5 ดังนี้

$$P(X_{is} = 1 | \theta_s, \beta_i, \alpha_i) = \frac{\exp[\alpha_i(\theta_s - \beta_i)]}{1 + \exp[\alpha_i(\theta_s - \beta_i)]} \quad (5)$$

โดย

X_{is}	หมายถึง	การตอบข้อคำถามของบุคคลที่ S ต่อข้อคำถามข้อที่ i (0 หรือ 1)
θ_s	หมายถึง	ระดับคุณลักษณะแฝงที่ S
β_i	หมายถึง	พารามิเตอร์ความยากของข้อคำถามข้อที่ i
α_i	หมายถึง	พารามิเตอร์อำนาจจำแนกของข้อคำถามข้อที่ i

เมื่อเปรียบเทียบสมการที่ 4 และสมการที่ 5 จะพบความแตกต่างในส่วนของพารามิเตอร์อำนาจจำแนก ซึ่งแต่ละข้อคำถามจะมีค่าอำนาจจำแนกที่แตกต่างกันไป ดังนั้น โมเดลโลจิสแบบสองพารามิเตอร์ จึงเหมาะกับข้อมูลที่ไม่สามารถอธิบายความแตกต่างระหว่างบุคคลได้ด้วยตำแหน่งหรือระดับของคุณลักษณะแฝง (พารามิเตอร์ความยาก) เพียงอย่างเดียว แต่ต้องอาศัยมุมมองอื่นในการอธิบายความแตกต่างของบุคคลเพิ่มเติมด้วย ตัวอย่างโคงลักษณะข้อสอบของโมเดลโลจิสแบบสองพารามิเตอร์ แสดงได้ดังภาพที่ 2-4



ภาพที่ 2-4 โค้งลักษณะข้อสอบของโมเดลโลจิสแบบสองพารามิเตอร์ (Embretson & Reise, 2000, p. 71)

จากภาพ 2-4 แสดงโค้งลักษณะข้อสอบของข้อคำถามสามข้อภายใต้โมเดลโลจิสแบบสองพารามิเตอร์ ซึ่งถึงแม้ระดับความยากของข้อคำถามแต่ละข้อจะมีค่าเท่ากับข้อคำถามที่แสดงในภาพที่ 2-3 แต่ในภาพนี้จะพบว่าพารามิเตอร์อำนาจจำแนกมีค่าแตกต่างกันไป จากภาพจะเห็นได้ว่าโค้งลักษณะข้อสอบมีการตัดกัน อย่างไรก็ตาม พารามิเตอร์ความยากยังคงบ่งบอกถึงตำแหน่งของระดับคุณลักษณะแฝงที่มีความน่าจะเป็นในการตอบข้อคำถามถูกเท่ากับ 0.50 เช่นเดิม

1.1.1.3 โมเดลโลจิสแบบสามพารามิเตอร์ (Three-Parameter Logistic Model (3PL))

โมเดลโลจิสแบบสามพารามิเตอร์มีลักษณะโค้งลักษณะข้อสอบที่แตกต่างไปจากสองโมเดลที่ผ่านมา โดยจุดเริ่มต้นของโค้งลักษณะข้อสอบไม่ได้เริ่มจากศูนย์ ซึ่งหมายถึงผู้ทดสอบมีโอกาสตอบข้อคำถามถูกต้องโดยไม่ต้องใช้ความสามารถ หรือคุณลักษณะแฝงใด ๆ ที่ต้องการวัดเลย แต่อาจตอบถูกต้อง เพราะผลจากการเดา (Guessing) เช่น ลักษณะการตอบข้อคำถามแบบหลายตัวเลือก (Multiple Choice) ความน่าจะเป็นในการตอบข้อคำถามถูกอาจมีค่ามากกว่าศูนย์ แม้ผู้ทดสอบจะมีความสามารถหรือมีคุณลักษณะแฝงต่ำก็ตาม โมเดลโลจิสแบบสามพารามิเตอร์จึงมีพารามิเตอร์การเดา (Guessing Parameter or Lower-Asymptote Parameter) เพิ่มขึ้นมาหนึ่งตัว ความน่าจะเป็นในการตอบข้อคำถามถูก ในโมเดลโลจิสแบบสามพารามิเตอร์สามารถคำนวณได้ดังสมการที่ 6

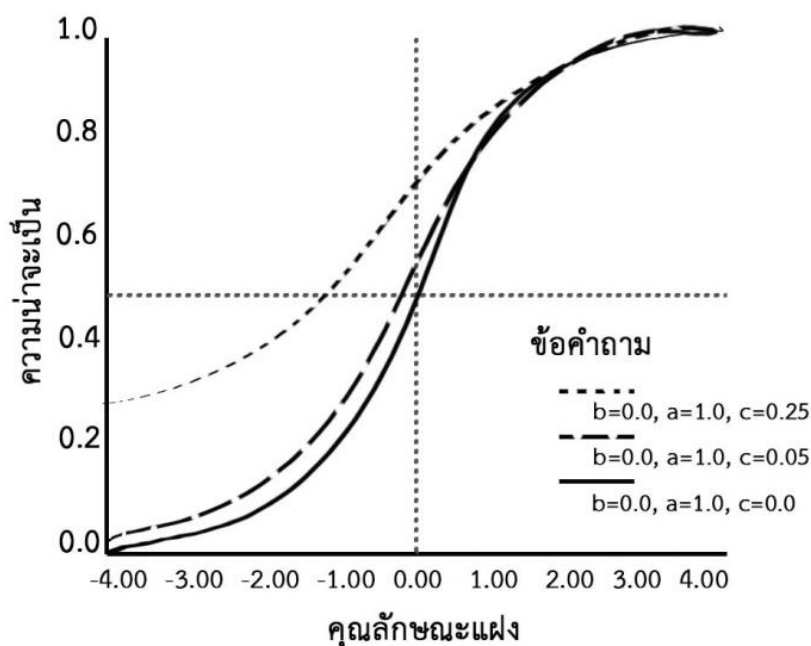
$$P(X_{is} = 1 | \theta_s, \beta_i, \alpha_i, \gamma_i) = \gamma_i + (1 - \gamma_i) \frac{\exp[\alpha_i(\theta_s - \beta_i)]}{1 + \exp[\alpha_i(\theta_s - \beta_i)]} \quad (6)$$

โดย

X_{is}	หมายถึง	การตอบข้อความของบุคคลที่ S ต่อข้อความข้อที่ i (0 หรือ 1)
θ_s	หมายถึง	ระดับคุณลักษณะแฝงที่ S
β_i	หมายถึง	พารามิเตอร์ความยากของข้อความข้อที่ i
α_i	หมายถึง	พารามิเตอร์อำนาจจำแนกของข้อความข้อที่ i
γ_i	หมายถึง	พารามิเตอร์การเดาของข้อความข้อที่ i

ข้อความที่มีตัวเลือกการตอบ 4 ตัวเลือก ปกติแล้วหากผู้ทดสอบเดาสุ่ม จะมี
ความน่าจะเป็นในการตอบข้อความถูก เท่ากับ 0.25 อย่างไรก็ตาม การประมาณค่าพารามิเตอร์การเดา
จากโมเดลโลจิสแบบสามพารามิเตอร์ จะมีค่าแตกต่างจากความน่าจะเป็นในการตอบข้อความถูก
จากการเดาสุ่ม (Random Guessing Probability) ยกตัวอย่างเช่น ถ้าผู้สอบสามารถตัดตัวเลือกอื่น ๆ
ได้ และเลือกคำตอบที่ถูกต้องที่สุดจากตัวเลือกที่เหลืออยู่ ความน่าจะเป็นในการตอบข้อความถูก
จะมีค่าสูงขึ้น ในโมเดลโลจิสแบบสามพารามิเตอร์ หากมีค่าพารามิเตอร์การเดาของข้อความแต่ละ
ข้อแตกต่างกันอาจนำไปสู่ปัญหาเกี่ยวกับการประมาณค่า ดังนั้น เพื่อหลีกเลี่ยงปัญหาดังกล่าวจึงนิยม
หาค่าพารามิเตอร์การเดาที่เป็นตัวแทนของข้อความทุกข้อ หรือสำหรับกลุ่มผู้สอบที่ใช้ข้อความ
ชุดเดียวกัน

นอกจากนี้ มีข้อสังเกตที่สำคัญอีกประการหนึ่ง คือ พารามิเตอร์ความยากในโมเดลโลจิส
แบบสามพารามิเตอร์นั้นมีความแตกต่างไปจากโมเดลอื่น ๆ โดยค่าพารามิเตอร์ความยากยังคงเกิดขึ้น
ที่จุดเปลี่ยนโค้งของโค้งลักษณะข้อสอบ แต่จุดเปลี่ยนโค้งดังกล่าวไม่ได้อยู่ที่ระดับคุณลักษณะแฝงซึ่งมี
ความน่าจะเป็นในการตอบข้อความถูก เท่ากับ 0.50 อีกต่อไป เนื่องจากจุดเปลี่ยนโค้งมีการเปลี่ยน
ตำแหน่ง ซึ่งเป็นผลมาจากโค้งที่ไม่ได้มีจุดเริ่มต้นที่ศูนย์ ค่าความน่าจะเป็นในการตอบข้อความถูก
จึงมีแนวโน้มเพิ่มขึ้นกว่าเดิม ดังภาพที่ 2-5



ภาพที่ 2-5 โค้งลักษณะข้อสอบของโมเดลโลจิสแบบสามพารามิเตอร์ (Embretson & Reise, 2000, p. 72)

ภาพที่ 2-5 แสดงโค้งลักษณะข้อสอบของข้อคำถามสามข้อ จะเห็นได้ว่าข้อคำถามทั้งสามข้อมีค่าพารามิเตอร์ความยากเท่ากับ 0.00 และมีพารามิเตอร์อำนาจจำแนกเท่ากับ 1.00 แต่ข้อคำถามทั้งสามข้อ มีพารามิเตอร์การเดาที่แตกต่างกัน (จุดเริ่มต้นโค้งแตกต่างกัน) ซึ่งแสดงให้เห็นว่าในระดับคุณลักษณะแฝงที่ต่ำที่สุด ความน่าจะเป็นในการตอบข้อคำถามถูกต้องมีค่ามากกว่าศูนย์ โดยในข้อคำถามที่มีค่า μ_i เท่ากับ 0.25 ค่าความน่าจะเป็นน้อยที่สุดในการตอบข้อคำถามถูก มีค่าเท่ากับ 0.25 นอกจากนี้จะสังเกตได้ว่าที่จุดเปลี่ยนโค้ง ค่าพารามิเตอร์ความยากไม่ได้มีค่าเท่ากับค่าความน่าจะเป็นในการตอบข้อคำถามถูก เท่ากับ 0.50 เหมือนในโมเดลโลจิสแบบหนึ่งพารามิเตอร์ และโมเดลโลจิสแบบสองพารามิเตอร์

1.1.2 โมเดลปกติสะสม (Traditional Normal Ogive Models)

โมเดลปกติสะสมเป็นโมเดลที่ประกอบด้วยพารามิเตอร์ต่าง ๆ เช่นเดียวกับโมเดลโลจิส อย่างไรก็ตาม โค้งลักษณะข้อสอบของโมเดลปกติสะสมสร้างขึ้นจากฟังก์ชันทางคณิตศาสตร์ที่แตกต่างจากโมเดลโลจิส โดยเป็นฟังก์ชันทางคณิตศาสตร์ที่มีความซับซ้อนมากกว่า เนื่องจากความน่าจะเป็นในการตอบข้อคำถามถูก ได้มาจากค่าสัดส่วนสะสมของเคส (Cases) จากการแจกแจงแบบปกติ ดังนั้น โมเดลปกติสะสมจึงสร้างโค้งลักษณะข้อสอบในลักษณะค่าสัดส่วนของเคส (Cases) ภายใต้คะแนนมาตรฐานที่มีค่าแน่นอน (Z_{is}) ดังสมการ

$$P(X_{is} = 1) = \int_{-\infty}^{Z_{is}} \frac{1}{(2\pi)^{1/2}} \exp(-t^2/2) dt \quad (7)$$

โดย

$\int_{-\infty}^{Z_{is}} dt$ หมายถึง ปริพันธ์สำหรับการหาพื้นที่ของการแจกแจงตั้งแต่ $-\infty$ ถึง Z_{is}
 π หมายถึง ค่าคงที่ซึ่งมีค่าเท่ากับ 3.14

ดังนั้น ในสมการที่ 7 จึงประกอบไปด้วยฟังก์ชันความหนาแน่นปกติ (Normal Density Function) และ การหาค่าสัดส่วนของเคส (Cases) จนถึงตำแหน่งคะแนนมาตรฐานซี (Z_{is}) โดยคะแนนมาตรฐานซี (Z_{is}) จะประกอบไปด้วยพารามิเตอร์ต่าง ๆ ของโมเดลการตอบสนองข้อสอบ

โมเดลปกติสะสมแบบสองพารามิเตอร์ และโมเดลปกติสะสมแบบสามพารามิเตอร์ถูกนำไปใช้ในหลายโอกาส ซึ่งมีโครงสร้างของพารามิเตอร์ในโมเดลเหมือนกับโมเดลโลจิสแบบสองพารามิเตอร์ และโมเดลโลจิสแบบสามพารามิเตอร์ ตามลำดับ ในขณะที่โมเดลปกติสะสมแบบหนึ่งพารามิเตอร์มีแนวคิดทางทฤษฎีและการนำไปใช้แตกต่างไปจากโมเดลโลจิสแบบหนึ่งพารามิเตอร์ ดังนั้น ในส่วนนี้จึงนำเสนอเพียงโมเดลปกติสะสมแบบสองพารามิเตอร์และโมเดลปกติสะสมแบบสามพารามิเตอร์เท่านั้น

1.1.2.1 โมเดลปกติสะสมแบบสองพารามิเตอร์ (Two-Parameter Normal Ogive Model)

โมเดลปกติสะสมแบบสองพารามิเตอร์มีโครงสร้างเช่นเดียวกับโมเดลโลจิสแบบสองพารามิเตอร์ โดย Z_{is} สามารถคำนวณได้ดังสมการที่ 8

$$Z_{is} = \alpha_i(\theta_s - \beta_i) \quad (8)$$

โดย

θ_s หมายถึง ระดับคุณลักษณะแฝงที่ S
 β_i หมายถึง พารามิเตอร์ความยากของข้อคำถามข้อที่ i
 α_i หมายถึง พารามิเตอร์อำนาจจำแนกของข้อคำถามข้อที่ i

ค่าของพารามิเตอร์ทำหน้าที่เป็นคะแนนมาตรฐาน (Z_{is}) ในการหาสัดส่วนสะสมในขอบเขตของ Z_{is} ซึ่งสามารถคำนวณความน่าจะเป็นในการตอบข้อคำถามถูกต้องจากสมการที่ 9

$$P(X_{is} = 1 | \theta_s, \beta_i, \alpha_i) = \int_{-\infty}^{\alpha_i(\theta_s - \beta_i)} \frac{1}{(2\pi)^{1/2}} \exp(-t^2/2) dt \quad (9)$$

โดย

X_{is} หมายถึง การตอบข้อคำถามของบุคคลที่ S ต่อข้อคำถามข้อที่ i (0 หรือ 1)
 θ_s หมายถึง ระดับคุณลักษณะแฝงที่ S
 β_i หมายถึง พารามิเตอร์ความยากของข้อคำถามข้อที่ i
 α_i หมายถึง พารามิเตอร์อำนาจจำแนกของข้อคำถามข้อที่ i
 $\int_{-\infty}^{Z_{is}} dt$ หมายถึง ปริพันธ์สำหรับการหาพื้นที่ของการแจกแจงตั้งแต่ $-\infty$ ถึง Z_{is}

π หมายถึง ค่าคงที่ซึ่งมีค่าเท่ากับ 3.14

ตัวอย่างเช่น ถ้า $\alpha_i = 1.00$ $\theta_s = 2.00$ และ $\beta_i = 1.00$ ดังนั้น เมื่อคำนวณค่าของ Z_{is} ตามสมการที่ 8 จะได้ค่าเท่ากับ 1.00 สัดส่วนภายใต้ค่า Z_{is} ที่มีค่าเท่ากับ 1.00 จากการแจกแจงปกติ มีค่าเท่ากับ 0.8413 ดังนั้น ความน่าจะเป็นในการตอบข้อคำถามถูกที่คำนวณได้จากสมการที่ 9 มีค่าเท่ากับ 0.8413 อีกตัวอย่างหนึ่ง คือ หากผลรวมของค่าพารามิเตอร์ทำให้ได้ค่าคะแนนมาตรฐานเท่ากับ -1.50 ความน่าจะเป็นที่คำนวณได้จากการแจกแจงปกติสะสมจะเท่ากับ 0.0668

1.1.2.2 โมเดลปกติสะสมแบบสามพารามิเตอร์ (Two-Parameter Normal Ogive Model)

โมเดลนี้มีลักษณะเช่นเดียวกับโมเดลโลจิสแบบสามพารามิเตอร์ซึ่งเพิ่มพารามิเตอร์การเดาเข้าไปในโมเดล โมเดลปกติสะสมแบบสามพารามิเตอร์สามารถคำนวณได้จากสมการที่ 10 ดังนี้

$$P(X_{is} = 1 | \theta_s, \beta_i, \alpha_i, \gamma_i) = \gamma_i + (1 - \gamma_i) \int_{-\infty}^{\alpha_{is}(\theta_s - \beta_i)} \frac{1}{(2\pi)^{1/2}} \exp(-t^2/2) dt \quad (10)$$

โดย

X_{is}	หมายถึง	การตอบข้อคำถามของบุคคลที่ S ต่อข้อคำถามข้อที่ i (0 หรือ 1)
θ_s	หมายถึง	ระดับคุณลักษณะแฝงที่ S
β_i	หมายถึง	พารามิเตอร์ความยากของข้อคำถามข้อที่ i
α_i	หมายถึง	พารามิเตอร์อำนาจจำแนกของข้อคำถามข้อที่ i
γ_i	หมายถึง	พารามิเตอร์การเดาของข้อคำถามข้อที่ i
$\int_{-\infty}^{Z_{is}} dt$	หมายถึง	ปริพันธ์สำหรับการหาพื้นที่ของการแจกแจงตั้งแต่ $-\infty$ ถึง Z_{is}
π	หมายถึง	ค่าคงที่ซึ่งมีค่าเท่ากับ 3.14

เช่นเดียวกับโมเดลโลจิสแบบสามพารามิเตอร์ ค่าพารามิเตอร์ความยากซึ่งเป็นบริเวณจุดเปลี่ยนโค้ง ไม่ได้มีค่าเท่ากับความน่าจะเป็นในการตอบข้อคำถามถูก เท่ากับ 0.50

1.2 โมเดลการตอบสนองข้อสอบแบบพหุมิติสำหรับข้อมูลที่มีลักษณะสองค่า (Multidimensional IRT Models for Binary Data)

โมเดลการตอบสนองข้อสอบแบบพหุมิติประกอบไปด้วยพารามิเตอร์ที่บ่งบอกคุณลักษณะของบุคคลตั้งแต่สองพารามิเตอร์ขึ้นไป มิติการวัดที่หลากหลายทำให้ข้อมูลการตอบข้อคำถามมีความสอดคล้องมากขึ้นเมื่อบุคคลตอบข้อคำถามที่มีความยากง่ายแตกต่างกันออกไปอย่างเป็นระบบ โมเดลการตอบสนองข้อสอบแบบพหุมิติสามารถจัดได้เป็นสองกลุ่ม คือ โมเดลเชิงสำรวจ (Exploratory Models) และโมเดลเชิงยืนยัน (Confirmatory Models) เมื่อเปรียบเทียบกับ การวิเคราะห์องค์ประกอบ (Factor Analysis) จะพบว่าโมเดลการตอบสนองข้อสอบแบบพหุมิติเชิงสำรวจ (Exploratory Multidimensional Models) เกี่ยวข้องกับการประมาณค่าพารามิเตอร์ข้อสอบและพารามิเตอร์บุคคลมากกว่าหนึ่งมิติ เพื่อให้โมเดลมีความสอดคล้องกับข้อมูล โดยไม่มีการกำหนดจำนวนขององค์ประกอบอย่างชัดเจน ซึ่งตรงข้ามกับโมเดลการตอบสนองข้อสอบแบบ

พหุมิติเชิงยืนยัน (Confirmatory Multidimensional Models) ซึ่งเป็นการประมาณค่าพารามิเตอร์ โดยกำหนดมิติขององค์ประกอบอย่างชัดเจน ตัวอย่างโมเดลการตอบสนองข้อสอบแบบพหุมิติ เช่น โมเดลการวิเคราะห์องค์ประกอบแบบไม่ใช่เส้นตรงของแมคโดนัลด์ (McDonald's Non-Linear Factor Analysis Model) โมเดลโลจิสแบบพหุมิติ (Multidimensional Logistic Models) โมเดลปกติสะสมแบบพหุมิติ (Multidimensional Normal Ogive Models) โมเดลคุณลักษณะแฝงแบบพหุองค์ประกอบ (Multicomponent Latent Trait Model; MLTM) โมเดลคุณลักษณะแฝงแบบองค์ประกอบทั่วไป (General Component Latent Trait Model) โมเดลสำหรับการเรียนรู้และการเปลี่ยนแปลง (Models for Learning and Change) โมเดลที่มีโครงสร้างระดับคุณลักษณะแฝงอย่างเจาะจง (Models With Specified Trait Level Structures) และโมเดลสำหรับการแยกกลุ่มของบุคคล (Models for Distinct Classes of Persons)

2. โมเดลการตอบสนองข้อสอบแบบตรวจให้คะแนนมากกว่า 2 ค่า (Polytomous IRT Models)

มาตรวัดทางด้านจิตวิทยาส่วนใหญ่มีลักษณะการตรวจให้คะแนนที่ไม่ได้จำกัดอยู่เพียงการให้คะแนนแบบ 0 กับ 1 ดังนั้น โมเดลการตอบสนองข้อสอบแบบตรวจให้คะแนน 2 ค่า (Dichotomous IRT Model) จึงมีข้อจำกัดสำหรับการนำไปใช้ ในกรณีที่รูปแบบการตอบสามารถตรวจให้คะแนนได้มากกว่า 2 ค่า เช่น มาตรวัดเจตคติ หรือแบบประเมินบุคลิกภาพ ส่วนใหญ่ประกอบด้วยข้อคำถามที่มีตัวเลือกคำตอบแบบเรียงลำดับหลายตัวเลือก ตัวเลือกคำตอบแบบเรียงลำดับหลายตัวเลือกสามารถให้ข้อมูลเกี่ยวกับผู้ทดสอบได้มากกว่า และมีความน่าเชื่อถือมากกว่า ตัวเลือกคำตอบแบบตรวจให้คะแนน 2 ค่า โมเดลที่เหมาะสมกับข้อมูลลักษณะนี้ คือ โมเดลการตอบสนองข้อสอบแบบตรวจให้คะแนนมากกว่า 2 ค่า (Polytomous IRT Models) ซึ่งเป็นโมเดลที่ใช้ในการอธิบายความสัมพันธ์ระหว่างระดับคุณลักษณะแฝงของผู้ทดสอบ (Examinee Trait Level) กับความน่าจะเป็นในการเลือกรายการคำตอบ โมเดลการตอบสนองข้อสอบในกลุ่มนี้มีหลากหลายโมเดล และมีผู้พัฒนาโมเดลใหม่ ๆ ขึ้นมาเป็นระยะ แต่โมเดลที่เป็นที่คุ้นเคย และรู้จักกันโดยทั่วไป มี 6 โมเดลหลัก ได้แก่ (Embretson & Reise, 2000, p. 95)

Graded Response Model (GRM) พัฒนาโดย Samejima ในปี ค.ศ. 1969 เป็นโมเดลการตอบสนองข้อสอบแบบตรวจให้คะแนนมากกว่า 2 ค่า ที่เหมาะสำหรับข้อมูลที่มีตัวเลือกรายการคำตอบแบบเรียงลำดับ (Ordered categories responses) เช่น มาตรประมาณค่าของลิเคิร์ต (Likert Rating Scale) โมเดล GRM เป็นโมเดลที่ขยายมาจากโมเดลโลจิสแบบสองพารามิเตอร์ (2PL) และมีการคำนวณค่าความน่าจะเป็นในการเลือกรายการคำตอบแบบสองขั้นตอน (Two-Step Process) ซึ่งจัดเป็นโมเดลการตอบสนองข้อสอบแบบทางอ้อม (Indirect IRT Models) โดยในแต่ละข้อคำถามอาจมีตัวเลือกรายการคำตอบเท่ากัน หรือไม่เท่ากันก็ได้ เพราะไม่ว่าจะมีจำนวนรายการคำตอบเท่ากันหรือไม่เท่ากัน ก็ไม่ส่งผลต่อความยุ่งยากในการประมาณค่าพารามิเตอร์

ในโมเดล GRM ข้อคำถามแต่ละข้อ (i) จะประกอบด้วยพารามิเตอร์ความชัน (Slope Parameter: α_i) และพารามิเตอร์เทรชโฮลด์ (Threshold Parameter: β_{ij}) ซึ่ง j แสดงถึง ช่วงเทรชโฮลด์ในข้อคำถามนั้น ๆ (ช่วงระหว่างตัวเลือกรายการคำตอบ) และมีค่าได้ตั้งแต่ 1 ถึง m_i โดย $m_i + 1 = K_i$ ซึ่ง K_i เป็นค่าที่แสดงถึงจำนวนตัวเลือกรายการคำตอบ ในข้อคำถามนั้น ๆ โมเดล

GRM ใช้วิธีการคำนวณค่าความน่าจะเป็นในการเลือกรายการคำตอบแบบ 2 ขั้นตอน โดยเริ่มต้นด้วยการคำนวณหาค่า β_{ij} ในทุกช่วงเทรชโฮลต์ ซึ่งหากข้อคำถามมีจำนวนตัวเลือกรายการคำตอบ 5 รายการ ($K = 5$) ผู้ทดสอบสามารถได้คะแนน (x) ตั้งแต่ 0 ถึง 4 คะแนน ดังนั้น m_j มีค่าเท่ากับ 4 ($j = 1..4$) ซึ่งหมายถึง ช่วงเทรชโฮลต์มี 4 ช่วง โดยโมเดล GRM จะเริ่มต้นด้วยการประมาณค่า β_{ij} ในแต่ละช่วงเทรชโฮลต์เหล่านี้ ดังตัวอย่าง

ตัวอย่างข้อคำถาม: ฉันรู้สึกสนุกสนานเมื่อได้เข้าร่วมงานคืนสู่เหย้า

	ไม่ชอบ	ไม่แน่ใจ	ชอบเล็กน้อย	ค่อนข้างชอบ	ชอบ
คะแนน (x)	0	1	2	3	4
	----- ----- ----- -----				
เทรชโฮลต์ (j)	1	2	3	4	

ขั้นตอนแรกของโมเดล GRM เป็นการประมาณค่าความน่าจะเป็นในการเลือกรายการคำตอบในแต่ละช่วงเทรชโฮลต์ ($P_{ix}^*(\theta)$) ด้วยสมการที่ 11

$$P_{ix}^*(\theta) = \frac{\exp[\alpha_i(\theta - \beta_{ij})]}{1 + \exp[\alpha_i(\theta - \beta_{ij})]} \quad (11)$$

โดย

x มีค่าเท่ากับ j ซึ่งมีค่าตั้งแต่ 1 ถึง m_j

θ หมายถึง ระดับคุณลักษณะแฝง

β_{ij} หมายถึง พารามิเตอร์ตำแหน่ง (พารามิเตอร์เทรชโฮลต์) ของข้อคำถามข้อที่ i ในตัวเลือกรายการคำตอบที่ j

α_i หมายถึง พารามิเตอร์ความชัน (พารามิเตอร์อำนาจจำแนก) ของข้อคำถามข้อที่ i

โค้งฟังก์ชัน ($P_{ix}^*(\theta)$) ที่คำนวณได้จากสมการที่ 11 จะแสดงถึงค่าความน่าจะเป็นของการตอบข้อคำถาม (x) ของผู้ทดสอบที่มีระดับคุณลักษณะแฝงเท่ากับ θ ในการเลือกรายการคำตอบระหว่างช่วงเทรชโฮลต์นั้น กับ รายการคำตอบที่อยู่ในช่วงเทรชโฮลต์ที่สูงกว่า โดยโค้ง $P_{ix}^*(\theta)$ มีชื่อเรียกว่า โค้งลักษณะปฏิบัติการ (Operating Characteristic Curves: OCC) ซึ่งจะต้องมีการประมาณค่าโค้งลักษณะปฏิบัติการในทุกช่วงเทรชโฮลต์ ดังนั้น ในข้อคำถามที่มีตัวเลือกรายการคำตอบ 5 รายการ จะต้องมีโค้งลักษณะปฏิบัติการจำนวนทั้งสิ้น 4 โค้ง และต้องประมาณค่า β_{ij} ทั้งหมด 4 ค่า (ค่าพารามิเตอร์ β_{ij} เป็นค่าที่แสดงถึง ตำแหน่งหรือระดับคุณลักษณะแฝงของ

ผู้ทดสอบ ที่มีความน่าจะเป็นในการเลือกรายการคำตอบในระดับที่สูงกว่าช่วงเทรซโฮลด์ที่กำหนด (เท่ากับ 0.50) การคำนวณค่า β_{ij} ในโมเดล GRM นั้น ช่วงแรกใช้วิธีการแบบโมเดลการตอบสนองข้อสอบที่มีการตรวจให้คะแนนแบบสองค่า โดยในแต่ละช่วงเทรซโฮลด์จะมีการประมาณค่าความน่าจะเป็นที่ผู้ทดสอบจะเลือกรายการคำตอบในช่วงเทรซโฮลด์ที่สูงกว่า เช่น มาตรฐานเจตคติซึ่งมีตัวเลือกการคำตอบ 5 ระดับ ตั้งแต่ไม่เห็นด้วยอย่างยิ่ง ถึง เห็นด้วยอย่างยิ่ง และมีการให้คะแนนแต่ละรายการคำตอบตั้งแต่ 0 ถึง 4 ในช่วงเทรซโฮลด์ต่ำที่สุด ค่า β_{ij} แสดงถึง ค่าความน่าจะเป็นที่ผู้ทดสอบจะเลือกรายการคำตอบระหว่างรายการคำตอบที่ 0 กับรายการคำตอบที่ 1, 2, 3, 4 ขณะที่ช่วงเทรซโฮลด์ต่อ ๆ ไป ค่า β_{ij} แสดงถึง ค่าความน่าจะเป็นที่ผู้ทดสอบจะเลือกรายการคำตอบระหว่างรายการคำตอบที่ 0, 1 กับรายการคำตอบที่ 2, 3, 4 ระหว่างรายการคำตอบที่ 0, 1, 2 กับรายการคำตอบที่ 3, 4 และระหว่างรายการคำตอบที่ 0, 1, 2, 3 กับรายการคำตอบที่ 4 ตามลำดับ (ในโมเดลการตอบสนองข้อสอบที่มีการตรวจให้คะแนนแบบสองค่าเป็นการคำนวณความน่าจะเป็นที่ผู้ทดสอบจะเลือกคำตอบถูกซึ่งเป็นการเลือกค่าระหว่าง 0 กับ 1) อย่างไรก็ตาม ทั้งหมดยังมีค่าพารามิเตอร์ความชัน (α_i) ร่วมกันเพียงค่าเดียว จำนวนค่า β_{ij} ที่ต้องคำนวณ ในแต่ละข้อคำถาม อาจมีจำนวนแตกต่างกันไปขึ้นอยู่กับว่า ข้อคำถามนั้นมีตัวเลือกการคำตอบจำนวนเท่าใด โดยจะมีจำนวนค่า β_{ij} ที่ต้องคำนวณ เท่ากับ m_i ซึ่งมีค่าเท่ากับ $K-1$ (โดย K หมายถึง จำนวนตัวเลือกการคำตอบในแต่ละข้อคำถาม) เมื่อทราบค่า β_{ij} และ α_i แล้วจะสามารถประมาณค่า $P_{ix}^*(\theta)$ ในแต่ละช่วงเทรซโฮลด์ได้ หลังจากนั้นจึงเริ่มการคำนวณในขั้นตอนที่ 2 ซึ่งเป็นการคำนวณค่าความน่าจะเป็นในการเลือกรายการคำตอบแต่ละรายการ (โดย $x=0..4$) ซึ่งสามารถคำนวณได้ดังสมการที่ 12

$$P_{ix}(\theta) = P_{ix}^*(\theta) - P_{i(x+1)}^*(\theta) \quad (12)$$

เมื่อ

$P_{i0}^*(\theta)$ หมายถึง ความน่าจะเป็นในการเลือกรายการคำตอบระดับต่ำที่สุด ซึ่งมีค่าเท่ากับ 1.0

$P_{i5}^*(\theta)$ หมายถึง ความน่าจะเป็นในการเลือกรายการคำตอบระดับสูงที่สุด ซึ่งมีค่าเท่ากับ 0.0

ค่าความน่าจะเป็นในการเลือกรายการคำตอบแต่ละรายการ ในมาตรฐานที่มีตัวเลือกการคำตอบ 5 ระดับ สามารถคำนวณได้ ดังนี้

$$P_{i0}(\theta) = 1.0 - P_{i1}^*(\theta)$$

$$P_{i1}(\theta) = P_{i1}^*(\theta) - P_{i2}^*(\theta)$$

$$P_{i2}(\theta) = P_{i2}^*(\theta) - P_{i3}^*(\theta)$$

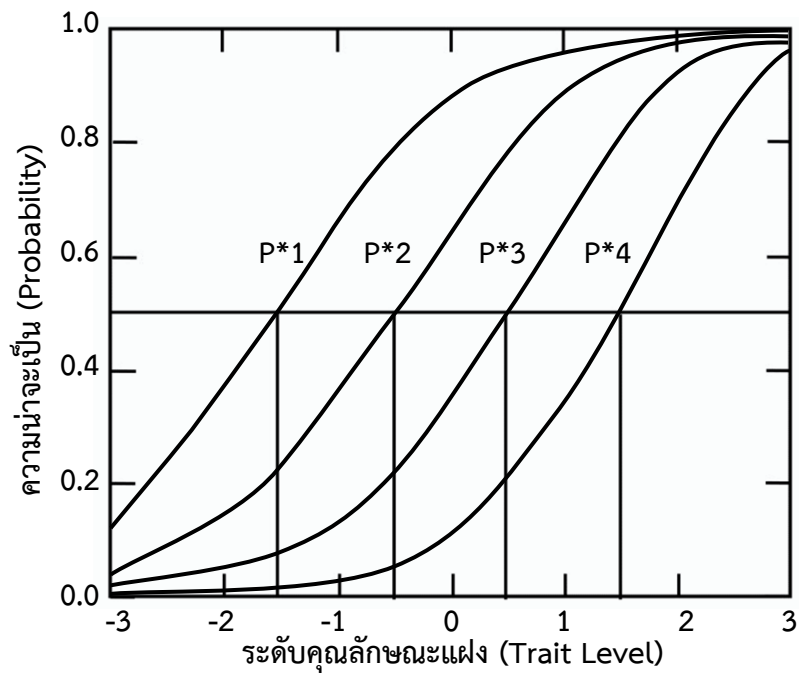
$$P_{i3}(\theta) = P_{i3}^*(\theta) - P_{i4}^*(\theta)$$

$$P_{i4}(\theta) = P_{i4}^*(\theta) - 0$$

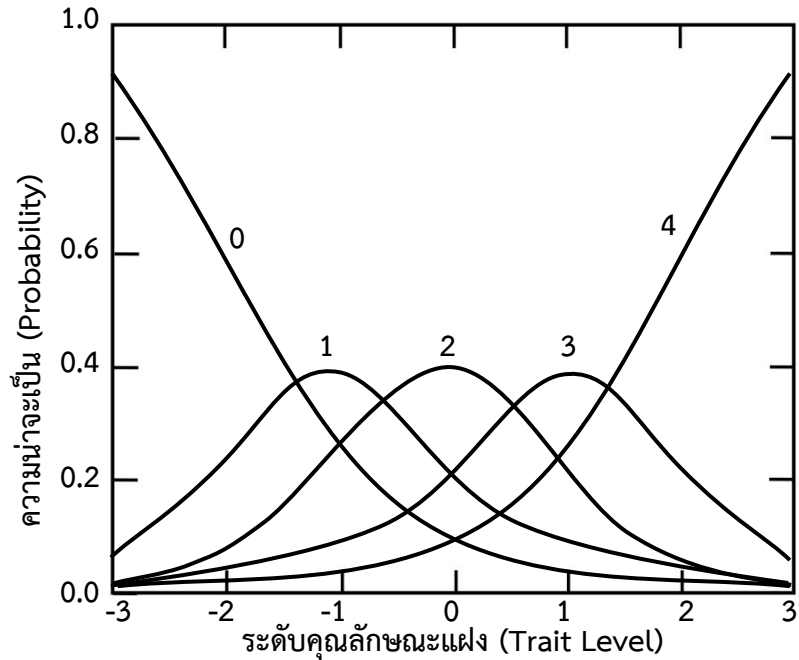
ค่าฟังก์ชันหรือโค้งที่คำนวณได้นี้ เรียกว่า โค้งการเลือกรายการคำตอบ (Category Response Curves: CRC) หรืออาจเป็นที่รู้จักในชื่ออื่น ๆ ได้แก่ Category Boundary Curves, Cumulative Probability Curves, Boundary Characteristic Curves ซึ่งโค้งดังกล่าวนี้แสดงถึงความน่าจะเป็นในการเลือกรายการคำตอบของผู้ทดสอบที่มีระดับคุณลักษณะแฝงแตกต่างกัน โค้งลักษณะปฏิบัติการและโค้งการเลือกรายการคำตอบ แสดงได้ดังภาพที่ 2-6 และ 2-7 ตามลำดับ

ในภาพที่ 2-6 แสดงให้เห็นถึงโค้งลักษณะปฏิบัติการของข้อคำถามที่มีตัวเลือกการคำตอบ 5 ตัวเลือก ซึ่งมีค่า α_i เท่ากับ 1.5 ค่า $\beta_{i1} - \beta_{i4}$ เท่ากับ -1.5, -0.5, 0.5 และ 1.5 ตามลำดับ ซึ่งจากรูปแสดงให้เห็นว่า พารามิเตอร์เทรซโฮลด์ของตัวเลือกการคำตอบเป็นสิ่งที่แสดงถึงตำแหน่งของค่าคุณลักษณะแฝงของผู้ทดสอบที่มีความน่าจะเป็นในการเลือกรายการคำตอบในช่วงเทรซโฮลด์ที่สูงกว่า เท่ากับ 0.5 ดังนั้น ค่าพารามิเตอร์ β_{ij} จึงแสดงให้เห็นถึงระดับคุณลักษณะแฝงของผู้ทดสอบในการตอบ x ที่ทำให้ค่าความน่าจะเป็นในการเลือกรายการคำตอบในช่วงเทรซโฮลด์ที่สูงกว่า รายการคำตอบที่ j มีค่าเท่ากับ 0.5 ขณะที่ภาพที่ 2-7 แสดงโค้งการเลือกรายการคำตอบ ซึ่งโค้งเหล่านี้แสดงค่าความน่าจะเป็นในการเลือกรายการคำตอบแต่ละรายการ ($x = 0...4$) ตามระดับคุณลักษณะแฝงของผู้ทดสอบ จะสังเกตได้ว่า ณ ระดับคุณลักษณะแฝงของผู้ทดสอบตำแหน่งใดตำแหน่งหนึ่ง ผลรวมของค่าความน่าจะเป็นในการเลือกรายการคำตอบทุกตัวเลือกการคำตอบจะมีค่าเท่ากับ 1.0

สิ่งที่กำหนดรูปร่าง และตำแหน่งของโค้งการเลือกรายการคำตอบและโค้งลักษณะปฏิบัติการ คือ พารามิเตอร์ข้อสอบที่ได้จากโมเดล GRM โดยปกติ หากพารามิเตอร์ความชัน (α_i) มีค่าสูง โค้งลักษณะปฏิบัติการจะมีความชันมาก (ภาพที่ 2-6) ในขณะที่โค้งการเลือกรายการคำตอบ จะมีความโด่งสูง (Peak) และมีช่วงแคบลง (Narrow) (ภาพที่ 2-7) ซึ่งลักษณะดังกล่าวแสดงให้เห็นว่าตัวเลือกการคำตอบสามารถจำแนกผู้ทดสอบที่มีระดับคุณลักษณะแฝงแตกต่างกันได้ดี นอกจากนั้น พารามิเตอร์เทรซโฮลด์ (β_{ij}) จะเป็นตัวกำหนดตำแหน่งของโค้งลักษณะปฏิบัติการ (ภาพที่ 2-6) และระบุตำแหน่งกึ่งกลางซึ่งมีความโด่งสูงสุดของโค้งการเลือกรายการคำตอบด้วย โดยโค้งการเลือกรายการคำตอบนี้จะมีค่าโด่งสูงสุด บริเวณจุดกึ่งกลางระหว่างพารามิเตอร์เทรซโฮลด์สองพารามิเตอร์ที่อยู่ติดกัน (De Ayala, 2009, pp. 217-224; Embretson & Reise, 2000, pp. 97-102)



ภาพที่ 2-6 โค้งลักษณะปฏิบัติการของข้อคำถามที่มีตัวเลือกรายการคำตอบ 5 รายการ (Embretson & Reise, 2000, p. 100)



ภาพที่ 2-7 โค้งการเลือกรายการคำตอบของข้อคำถามที่มีตัวเลือกรายการคำตอบ 5 รายการ (Embretson & Reise, 2000, p. 101)

ขนาดกลุ่มตัวอย่างที่เหมาะสมสำหรับการประมาณค่าพารามิเตอร์ในโมเดล GRM นั้น มีนักวิชาการได้เสนอไว้หลายแนวทาง หนึ่งในนั้น คือ De Ayala (2009, pp. 223-224) ได้เสนอว่าขนาดกลุ่มตัวอย่างที่ใช้ในการประมาณค่าพารามิเตอร์ในโมเดล GRM ควรมีไม่ต่ำกว่า 500 คน เพื่อให้สามารถวิเคราะห์ความสอดคล้องของโมเดล ประเมินความเป็นเอกมิติของคุณลักษณะแฝงที่ต้องการวัด ทำให้สามารถประมาณค่าพารามิเตอร์ความชันได้ไม่ยากเกินไป และเพื่อให้มั่นใจได้ว่ามีข้อมูลการตอบของผู้ทดสอบในทุกตัวเลือกการตอบ อย่างไรก็ตาม ปรากฏการณ์นี้ไม่ใช่กฎตายตัว ในทางปฏิบัติการกำหนดขนาดกลุ่มตัวอย่างที่เหมาะสมควรพิจารณาจากปัจจัยหลากหลายด้าน เช่น วัตถุประสงค์ในการนำค่าพารามิเตอร์ที่ประมาณค่าได้ไปใช้ วิธีการประมาณค่าที่เลือกใช้ ความยาวของมาตรวัด ลักษณะการตอบข้อคำถามของผู้ทดสอบ จำนวนข้อมูลที่ขาดหาย เป็นต้น

Modified Graded Response Model (M-GRM) เป็นโมเดลที่ปรับปรุงมาจากโมเดล GRM โดย Muraki (1990 cited in Embretson & Reise, 2000, pp. 102-105) เหมาะสำหรับวิเคราะห์แบบสอบถามที่มีลักษณะเป็นมาตราประมาณค่า (Rating Scale) เช่น แบบวัดทัศนคติ ข้อคำถามในมาตรวัดจะต้องมีจำนวนตัวเลือกการตอบเท่ากัน แต่สามารถมีค่าพารามิเตอร์ความชันแตกต่างกันได้ โมเดล M-GRM ใช้วิธีการคำนวณความน่าจะเป็นในการเลือกการตอบแบบสองขั้นตอนเหมือนในโมเดล GRM อย่างไรก็ตาม พารามิเตอร์เทอร์ชโฮลด์ (β_{ij}) ในโมเดล M-GRM ถูกแบ่งออกเป็นสองส่วน ได้แก่ 1) พารามิเตอร์ตำแหน่ง (Location Parameters: b_i) ของข้อคำถามแต่ละข้อ และ 2) ชูตของพารามิเตอร์เทอร์ชโฮลด์ (Threshold Parameters: c_j) ของมาตรวัดทั้งฉบับ ความแตกต่างระหว่างโมเดล GRM กับโมเดล M-GRM คือ โมเดล GRM จะมีการประมาณค่าชูตของพารามิเตอร์เทอร์ชโฮลด์ (β_{ij}) สำหรับข้อคำถามแต่ละข้อ แต่โมเดล M-GRM มีชูตของพารามิเตอร์เทอร์ชโฮลด์ (c_j) เพียงชูตเดียวที่ใช้กับมาตรวัดทั้งฉบับ แต่มีการประมาณค่าพารามิเตอร์ตำแหน่ง (b_i) หนึ่งค่าสำหรับข้อคำถามแต่ละข้อ โมเดล M-GRM ได้ชื่อว่าเป็นโมเดล GRM แบบจำกัด (Restricted Model) เพราะว่ามีจำนวนตัวเลือกการตอบ (Category Boundaries) ของข้อคำถามแต่ละข้อเท่ากันตลอดทั้งมาตรวัด ในขณะที่โมเดล GRM นั้น ข้อคำถามแต่ละข้ออาจมีจำนวนตัวเลือกการตอบของข้อคำถามแต่ละข้อแตกต่างกันได้ ดังนั้น โมเดล M-GRM จึงมีการประมาณค่าพารามิเตอร์ต่าง ๆ น้อยกว่าโมเดล GRM จุดเด่นของโมเดล M-GRM เมื่อเปรียบเทียบกับโมเดล GRM คือ มีพารามิเตอร์ตำแหน่งแยกออกมาจากชูตของพารามิเตอร์เทอร์ชโฮลด์ อย่างไรก็ตาม การเลือกใช้ขึ้นอยู่กับรูปแบบของข้อคำถามด้วย ถ้ามาตรวัดมีรูปแบบข้อคำถามหลากหลายแบบอยู่ในมาตรวัดฉบับเดียวกัน โมเดล GRM จะมีความเหมาะสมมากกว่าโมเดล M-GRM เพราะในกรณีดังกล่าวโมเดล M-GRM จะต้องแบ่งมาตรวัดออกเป็นบล็อกต่าง ๆ (Block) ตามรูปแบบของข้อคำถาม จากนั้นจึงคำนวณพารามิเตอร์ตำแหน่งและพารามิเตอร์เทอร์ชโฮลด์ในแต่ละบล็อก ซึ่งค่าพารามิเตอร์ในแต่ละบล็อกจะไม่สามารถเปรียบเทียบกับค่าพารามิเตอร์ในบล็อกอื่น ๆ การเปรียบเทียบต้องใช้วิธีการที่ซับซ้อนมากขึ้น ดังนั้น การประมาณค่าพารามิเตอร์ในมาตรวัดที่มีรูปแบบข้อคำถามที่หลากหลายจึงทำได้ยากเมื่อใช้โมเดล M-GRM แต่หากมาตรวัดมีรูปแบบของข้อคำถามแบบเดียวกันทั้งฉบับ โมเดล M-GRM จะมีความง่ายและเหมาะสมมากกว่าโมเดล GRM ซึ่งพารามิเตอร์ตำแหน่ง (Location Parameters: b_i) สามารถใช้เพื่อบอกระดับความยาก (Difficulty) หรือค่าของมาตรวัด

(Scale Value) ได้ทันที ขณะที่ชุดของพารามิเตอร์เทรชโฮลด์ (Threshold Parameters: c_j) สามารถแสดงถึงระยะห่างของมาตรวัดทางจิตวิทยา การใช้พารามิเตอร์ที่ได้จากโมเดล M-GRM ทั้งสองนี้ร่วมกัน คุณสมบัติทางการวัดดังกล่าวจะคล้ายคลึงกับวิธีการวัดของเทอร์สตัน

Partial Credit Model (PCM) เป็นโมเดลที่ถูกออกแบบมาสำหรับวิเคราะห์ข้อคำถามที่มีหลายขั้นตอน โดยมีการให้คะแนนสำหรับแต่ละขั้นตอนที่ผู้ทดสอบทำได้ ซึ่งพบได้บ่อยในแบบวัดผลสัมฤทธิ์ทางการเรียน เช่น โจทย์ปัญหาทางคณิตศาสตร์ เป็นต้น นอกจากนั้น โมเดล PCM ยังเหมาะสำหรับวิเคราะห์มาตรวัดเจตคติ หรือมาตรวัดบุคลิกภาพที่มีการประเมินความเชื่อ หรือการตอบสนองต่อข้อความโดยมีการให้คะแนนแบบหลายค่าด้วย โมเดล PCM จะแตกต่างจากโมเดล GRM และ M-GRM เพราะเป็นโมเดลแบบทางตรง (Direct Model) โมเดลนี้เป็นโมเดลที่มีการขยายต่อมาจากโมเดลโลจิสแบบหนึ่งพารามิเตอร์ (1PL Model) จึงมีคุณลักษณะมาตรฐานต่าง ๆ เช่นเดียวกับ Rasch Model เช่น การแยกจากกันของค่าพารามิเตอร์ข้อสอบกับค่าพารามิเตอร์บุคคล ในโมเดล PCM นั้นข้อคำถามแต่ละข้อในมาตรวัดจะมีความชันเท่ากัน ดังนั้น พารามิเตอร์ความชันจึงไม่ปรากฏอยู่ในโมเดลนี้ คงเหลือเพียงพารามิเตอร์ความยากแบบลำดับขั้น (Step Difficulty: δ_{ij}) ที่ใช้อธิบายระดับความยากของการผ่านจากขั้นตอนหนึ่งไปสู่อีกขั้นตอนหนึ่ง หรือจากตัวเลือกรายการคำตอบหนึ่งสู่อีกตัวเลือกหนึ่ง นอกจากนี้ ค่านี้ยังบ่งบอกถึงระดับคุณลักษณะแฝงบริเวณที่โค้งการเลือกรายการคำตอบ 2 รายการตัดกัน บางครั้งจึงนิยมเรียกว่า พารามิเตอร์จุดตัดรายการคำตอบ (Category Intersections Parameter) โดยข้อคำถามที่มีตัวเลือกรายการคำตอบ 5 รายการ จะมีพารามิเตอร์ความยากแบบลำดับขั้น 4 ค่า จุดเด่นของโมเดลนี้ คือ สามารถใช้คะแนนดิบรวมของมาตรวัดเพื่อประมาณค่าระดับคุณลักษณะแฝงได้ ดังนั้น บุคคลที่มีคะแนนดิบรวมของมาตรวัดเท่ากัน จึงมีระดับคุณลักษณะแฝงเท่ากัน อย่างไรก็ตาม โมเดลนี้นำไปใช้ในทางปฏิบัติได้ยาก เพราะในการประมาณค่าพารามิเตอร์อาจพบปัญหาว่าบริเวณจุดตัดโค้งการเลือกรายการคำตอบ ไม่ได้เรียงลำดับจากง่ายไปยาก หรือจากระดับคุณลักษณะแฝงต่ำไปสูง ทำให้เกิดปรากฏการณ์ ที่เรียกว่า การผกผัน (Reversal) ซึ่งแสดงให้เห็นว่ามีตัวเลือกรายการคำตอบอย่างน้อยหนึ่งรายการที่เป็นตัวเลือกที่ไม่สอดคล้องกับคุณลักษณะแฝงที่ต้องการวัด หรือข้อมูลที่เก็บรวบรวมได้ไม่สอดคล้องกับโมเดล (Embretson & Reise, 2000, pp. 105-110)

Generalized Partial Credit Model (G-PCM) เป็นโมเดลที่ปรับปรุงเพิ่มเติมจากโมเดล PCM โดย Muraki (1992; 1993 cited in Embretson & Reise, 2000, pp. 110-115) ซึ่งกำหนดให้ข้อคำถามแต่ละข้อในมาตรวัดมีพารามิเตอร์ความชัน (α_j) แตกต่างกันได้ โดยพารามิเตอร์ความชันในโมเดล G-PCM เป็นค่าที่บ่งบอกถึงระดับการตอบสนองต่อตัวเลือกรายการคำตอบที่แตกต่างกันไปในแต่ละข้อคำถาม ซึ่งขึ้นอยู่กับระดับคุณลักษณะแฝง หรือระดับความสามารถ (θ) ที่เปลี่ยนแปลงไป ในโมเดล G-PCM นี้ หากข้อคำถามใดมีค่าพารามิเตอร์ความชันน้อยกว่า 1.0 โค้งการเลือกรายการคำตอบ (Category Response Curves) ที่ได้จะมีลักษณะแบนราบคล้ายในโมเดล PCM แต่หากค่าพารามิเตอร์ความชันมีค่ามากกว่า 1.0 โค้งการเลือกรายการคำตอบในโมเดล G-PCM จะมีลักษณะโค้งมากกว่าโค้งการเลือกรายการคำตอบที่พบในโมเดล PCM อย่างไรก็ตาม พารามิเตอร์จุดตัดรายการคำตอบ (δ_{ij}) ในโมเดล G-PCM มีความหมายเช่นเดียวกับโมเดล PCM

Rating Scale Model (RSM) เป็นกลุ่มของโมเดลที่มีหลากหลายรูปแบบและอาจก่อให้เกิดความสับสนได้ เพราะมีลักษณะคล้ายคลึงกับโมเดลอื่น ๆ เช่น โมเดล PCM อย่างไรก็ตาม โมเดล RSM ตามแนวคิดของ Andrich (1978 cited in Embretson & Reise, 2000, pp. 115-118) มีความแตกต่างที่สำคัญ คือ ข้อคำถามในมาตรวัดซึ่งมีรูปแบบการตอบแบบเดียวกันนั้น ข้อคำถามแต่ละข้อจะมีพารามิเตอร์ตำแหน่ง (Location Parameter: λ_i) เพียงค่าเดียว ซึ่งแสดงถึงความยากหรือง่ายของข้อคำถามนั้น ๆ ขณะที่พารามิเตอร์จุดตัดรายการคำตอบ (Category Intersections Parameter: δ_j) ใช้อธิบายช่วงเทรซโฮลด์ของแต่ละรายการคำตอบ (Category Threshold) ซึ่งมี J ช่วง ($J = K - 1$ โดย K คือ จำนวนตัวเลือกรายการคำตอบ) โมเดล PCM กับโมเดล RSM แตกต่างกันที่โมเดล PCM ไม่มีแนวคิดเกี่ยวกับระดับความยากง่ายของแต่ละรายการคำตอบ (แต่ละขั้นตอน หรือแต่ละจุดตัดรายการคำตอบ) ในข้อคำถามนั้น ๆ ดังนั้น ข้อคำถามแต่ละข้อจึงมีระดับความยากของแต่ละขั้นตอน (Difficulties of the Steps) แตกต่างกันได้ ในขณะที่โมเดล RSM จะมีชุดของพารามิเตอร์จุดตัดรายการคำตอบเพียงชุดเดียวที่ใช้ตลอดทั้งมาตรวัด ยกตัวอย่างเช่น ข้อคำถามทางคณิตศาสตร์ที่ผู้ทดสอบต้องแก้โจทย์ปัญหาหลายขั้นตอน ขั้นตอนต่าง ๆ ภายในข้อคำถาม หรือระหว่างข้อคำถามอาจมีความยากง่ายต่างกันได้ บางขั้นตอนอาจง่าย ในขณะที่บางขั้นตอนอาจยาก อย่างไรก็ตามมาตรวัดเจตคติซึ่งส่วนใหญ่ทุกข้อมีการตอบข้อคำถามรูปแบบเดียวกัน เช่น 1 หมายถึง ไม่เห็นด้วย 2 หมายถึง เห็นด้วยปานกลาง และ 3 หมายถึง เห็นด้วยอย่างยิ่ง จะสังเกตได้ว่า ในมาตรวัดนั้น ๆ ข้อคำถามแต่ละข้อมีระดับความยากง่ายหรือช่วงของแต่ละรายการคำตอบไม่แตกต่างกัน ความยากง่ายในแต่ละขั้นตอน (Step Difficulties หรือ Intersections) ซึ่งพบในโมเดล PCM สามารถแบ่งออกเป็นสองส่วน ในโมเดล RSM ได้แก่ พารามิเตอร์ตำแหน่งคุณลักษณะแฝงที่ต้องการวัดของข้อคำถาม (Location Parameter: λ_i) และพารามิเตอร์จุดตัดรายการคำตอบ (Category Intersection Parameters: δ_j) ซึ่ง δ_j มีค่าเท่ากับ $\delta_j + \lambda_i$ โมเดล RSM เหมาะกับข้อคำถามที่มีการกำหนดช่วงรายการคำตอบแบบคงที่ซึ่งใช้กับข้อคำถามทุกข้อเหมือนกันหมดตลอดมาตรวัด แต่ไม่เหมาะกับมาตรวัดที่มีลักษณะการตอบข้อคำถามที่หลากหลาย

Nominal Response Model (NRM) พัฒนาโดย Bock (1972 cite in Embretson & Reise, 2000, pp. 119-122) ซึ่งเป็นโมเดลสำหรับใช้วิเคราะห์ข้อสอบหรือข้อคำถามที่รายการคำตอบไม่ได้ถูกจัดเรียงลำดับ ได้แก่ ข้อสอบแบบหลายตัวเลือก (Multiple Choices) ข้อคำถามที่ใช้วัดเจตคติ ข้อคำถามสำหรับประเมินบุคลิกลักษณะ ความตั้งใจเริ่มแรกในการพัฒนาโมเดลนี้เกิดขึ้นเพื่อให้สามารถศึกษาคุณลักษณะของตัวलगซึ่งเป็นตัวเลือกหนึ่งในข้อสอบแบบหลายตัวเลือกด้วยเส้นร่องรอย (Trace Line) การคำนวณความน่าจะเป็นในการเลือกรายการคำตอบคำนวณได้จากพารามิเตอร์ที่เกี่ยวข้อง 2 พารามิเตอร์ ได้แก่ พารามิเตอร์ความชัน (Slope Parameter: α_{ix}) ซึ่งแสดงถึงความชันของเส้นร่องรอยหรือค่าอำนาจจำแนก และพารามิเตอร์จุดตัดรายการคำตอบ (Intercept Parameter: c_{ix})

โมเดลการตอบสนองข้อสอบที่มีการตรวจให้คะแนนมากกว่า 2 ค่า ที่เป็นที่ยอมรับส่วนใหญ่พัฒนามาจากโมเดลโลจิสแบบหนึ่งพารามิเตอร์ หรือโมเดลโลจิสแบบสองพารามิเตอร์ คำศัพท์ที่ใช้ในการอธิบายรายละเอียดเฉพาะของโมเดลจึงมีชื่อเรียกแตกต่างกัน แต่พอที่จะนำมาเทียบเคียงกัน

เพื่อให้สามารถศึกษารายละเอียดเพิ่มเติมได้ง่ายขึ้น ตัวอย่างการเทียบเคียงรายละเอียดเฉพาะของ Two-Parameter Logistic Model (2PL) ซึ่งเป็นโมเดลการตอบสนองข้อสอบแบบที่มีการตรวจให้คะแนนสองค่ากับ Graded Response Model (GRM) และ Generalized Partial Credit Model (G-PCM) ซึ่งเป็นโมเดลการตอบสนองข้อสอบที่มีการตรวจให้คะแนนมากกว่า 2 ค่า แสดงได้ดังตารางที่ 2-1

ตารางที่ 2-1 การเทียบเคียงรายละเอียดเฉพาะของ Two-Parameter Logistic Model (2PL) กับ Graded Response Model (GRM) และ Generalized Partial Credit Model (G-PCM)

โมเดล	พารามิเตอร์ตัวที่ 1	พารามิเตอร์ตัวที่ 2	โค้งที่ใช้อธิบายความสัมพันธ์ระหว่างตัวแปร
2PL	พารามิเตอร์อำนาจจำแนก (Discrimination Parameter: a)	พารามิเตอร์ความยาก (Difficulty Parameter: b)	โค้งลักษณะข้อสอบ (Item Characteristic Curves: ICCs)
GRM	พารามิเตอร์ความชัน (Slope Parameter: α)	พารามิเตอร์เทรชโฮลด์ (Threshold Parameter: β)	- โค้งลักษณะปฏิบัติการ (Operating Characteristic Curves: OCCs) - โค้งตัวเลือกรายการคำตอบ (Category Response Curves: CRCs)
G-PCM	พารามิเตอร์ความชัน (Slope Parameter: α)	พารามิเตอร์ความยากแบบลำดับขั้น (Step Difficulty: δ_{ij})	โค้งตัวเลือกรายการคำตอบ (Category Response Curves: CRCs)

การเลือกใช้โมเดลการตอบสนองข้อสอบ (Embretson & Reise, 2000, pp. 123-124; Edelen & Reeve, 2007; De Ayala, 2009, pp. 162-236; Thomas, 2011) มีประเด็นที่ต้องพิจารณา 2 ประเด็นดังนี้

1. การเลือกโมเดลที่มีความสอดคล้องกับข้อมูล

1.1 พิจารณาจากรูปแบบการตรวจให้คะแนน

มาตรวัดที่ถูกพัฒนาขึ้นนั้นมีหลายรูปแบบ แต่สามารถจัดกลุ่มตามการตรวจให้คะแนนได้ 2 ประเภท คือ

1.1.1 มาตรวัดที่มีการตรวจให้คะแนนแบบ 2 ค่า (Dichotomous Item Response Theory) ให้คะแนนเป็น 0 และ 1 มาตรวัดลักษณะนี้สามารถเลือกใช้โมเดลการตอบสนองข้อสอบได้หลายรูปแบบ แต่ที่นิยมมาก คือ One-Parameter Logistic Model, Two-Parameter Logistic Model และ Three-Parameter Logistic Model

1.1.2 มาตรฐานที่มีการตรวจให้คะแนนมากกว่า 2 ค่า (Polytomous Item Response Theory) เช่น มาตรฐานค่าที่มีการตรวจให้คะแนนตั้งแต่ 0 ถึง 4 โมเดลที่นิยมในกลุ่มนี้ ได้แก่ Graded Response Model (GRM), Modified Graded-Response Model (M-GRM), Partial Credit Model (PCM), (RSM), Generalized-Partial Credit Model (G-PCM), Rating Scale Model และ Nominal Response Model (NRM)

1.2 พิจารณาจากค่าพารามิเตอร์อำนาจจำแนก โอกาสการเดาข้อสอบ และลักษณะการเรียงลำดับของตัวเลือกในแต่ละข้อคำถาม ซึ่งแบ่งโมเดลได้เป็น 2 กลุ่ม ดังนี้

1.2.1 โมเดลกลุ่มราสช์ (Rasch Family) เป็นโมเดลในกลุ่มที่มีค่าพารามิเตอร์อำนาจจำแนก หรือความชันของข้อคำถามแต่ละข้อเท่ากัน ซึ่งการวิเคราะห์ข้อมูลทำได้สะดวก แต่ในสถานการณ์จริงเป็นการยากมากที่ข้อคำถามของมาตรฐานนั้นจะมีค่าพารามิเตอร์อำนาจจำแนกของข้อคำถามแต่ละข้อเท่ากัน

1.2.2 โมเดลกลุ่มนัรราสช์ (Non-Rasch Family) เป็นโมเดลในกลุ่มที่มีค่าพารามิเตอร์อำนาจจำแนก หรือความชันของข้อคำถามแต่ละข้อในมาตรฐานแตกต่างกัน

นอกจากนี้ อาจพิจารณาว่ามาตรฐานนั้นมีพารามิเตอร์การเดาหรือไม่ พารามิเตอร์นี้แสดงถึงความน่าจะเป็นที่ผู้ทดสอบสามารถตอบข้อคำถามถูก โดยไม่ได้ใช้ความสามารถ หรือคุณลักษณะแฝงที่มุ่งวัดใด ๆ เลย แต่เกิดจากความบังเอิญเท่านั้น

มาตรฐานที่มีการให้คะแนนมากกว่าสองค่า สามารถพิจารณาเลือกใช้โมเดลการตอบสนองข้อสอบ ตามลักษณะของการตอบข้อคำถาม หากมาตรฐานมีตัวเลือกคำตอบที่หลากหลาย แต่ไม่ได้มีการเรียงลำดับตัวเลือก ก็เหมาะที่จะใช้ Nominal Model ขณะที่บริบทของการประเมินด้านสุขภาพส่วนใหญ่ มาตรฐานมีลักษณะการตอบข้อคำถามแบบเรียงลำดับจากมากไปน้อย ลักษณะเช่นนี้เหมาะกับโมเดล Generalized-Partial Credit Model (GPCM) และ โมเดล Graded Response Model (GRM) โดยโมเดลทั้งสองนี้มีความแตกต่างกันเพียงเล็กน้อย การเลือกใช้จึงขึ้นอยู่กับความพึงพอใจและความคุ้นเคยกับโปรแกรมที่ใช้ในการวิเคราะห์ข้อมูลด้วย

1.3 พิจารณาจากกราฟข้อคำถาม (Item plots)

โปรแกรม TestGraf สามารถที่จะสร้าง Non-Parametric ICC Plots เพื่อใช้ในการอธิบายความสัมพันธ์ระหว่างความน่าจะเป็นในการตอบข้อคำถาม กับคุณลักษณะแฝงที่มุ่งวัดได้ ซึ่งสามารถให้ข้อมูลที่ช่วยในการพิจารณาเลือกโมเดลที่เหมาะสมด้วย

2. การตรวจสอบความสอดคล้องระหว่างโมเดลกับข้อมูล

วิธีการพิจารณาความสอดคล้องระหว่างโมเดลการตอบสนองข้อสอบกับข้อมูลว่า ข้อมูลที่เก็บรวบรวมได้นั้น มีความสอดคล้องกับโมเดลการตอบสนองข้อสอบมากน้อยเพียงใด สามารถทำได้หลากหลายวิธี การพิจารณาความสอดคล้องควรพิจารณาใน 3 ระดับ ได้แก่ ระดับข้อคำถาม ระดับบุคคล และระดับโมเดล อย่างไรก็ตาม ไม่มีวิธีการใดที่สามารถตัดสินได้อย่างเด็ดขาดว่า โมเดลกับข้อมูลที่เก็บรวบรวมได้มีความสอดคล้อง หรือไม่สอดคล้อง ดังนั้น จึงควรใช้วิธีการตรวจสอบที่หลากหลาย และพิจารณาอย่างรอบคอบ ก่อนตัดสินใจเลือกโมเดลการตอบสนองข้อสอบที่มีความเหมาะสมกับข้อมูลมากที่สุด (Embretson & Reise, 2000, pp. 233-244; Edelen & Reeve, 2007)

2.1 ความสอดคล้องในระดับข้อคำถาม (Item Fit) การออกแบบมาตรวัดโดยใช้ข้อมูลที่ได้จากการวิเคราะห์ด้วยทฤษฎีการตอบสนองข้อสอบ ช่วยให้สามารถออกแบบมาตรวัดให้มีรูปแบบการตอบข้อคำถามเป็นแบบเดียวกันตลอดทั้งมาตรวัด หรือ ออกแบบให้ภายในมาตรวัดเดียวกัน มีรูปแบบการตอบข้อคำถามที่หลากหลายและแตกต่างกันก็ได้ เนื่องจากการวิเคราะห์และการเลือกใช้โมเดลการตอบสนองข้อสอบ สามารถพิจารณาเป็นรายข้อ โดยดูความเหมาะสมของโมเดลกับข้อคำถามแต่ละข้อหรือมาตรวัดโดยรวมก็ได้ การตรวจสอบความสอดคล้องของโมเดลในระดับข้อคำถามมีสองแนวทางหลัก ได้แก่ 1) การพิจารณาจากกราฟ ซึ่งเป็นการเปรียบเทียบโค้งการตอบสนองข้อสอบ (Item Response Curve: IRC) ที่ประมาณค่าได้ กับ โค้งการตอบสนองข้อสอบที่ได้จากข้อมูลจริง (Empirical IRC) และ 2) การใช้สถิติตรวจสอบความแตกต่างระหว่างข้อมูลที่ได้จากการประมาณค่าและข้อมูลจริงว่าแตกต่างกันเกินระดับที่ยอมรับได้หรือไม่

2.2 ความสอดคล้องในระดับบุคคล (Person Fit) เป็นการใช่วิธีการทางสถิติเพื่อตรวจสอบความสอดคล้องของโมเดลการตอบสนองข้อสอบ ว่าสามารถอธิบายคุณลักษณะแฝงที่ระดับต่าง ๆ ได้ดีเพียงใด เพราะบางครั้งโมเดลการตอบสนองข้อสอบก็อาจไม่สามารถทำนายคุณลักษณะแฝงของบางคนได้ หลักการตรวจสอบในระดับบุคคลเป็นการประเมินความคงที่ (Consistency) ของลักษณะรูปแบบ (Pattern) การตอบข้อคำถามของผู้ทดสอบ

2.3 ความสอดคล้องในระดับโมเดล (Model Fit) ดำเนินการโดยเปรียบเทียบความสอดคล้องของชุดข้อมูล กับโมเดลที่เป็นไปได้หลากหลายรูปแบบ เช่น เปรียบเทียบว่าชุดของข้อมูลนั้นมีความเหมาะสมกับโมเดลสองพารามิเตอร์ หรือโมเดลสามพารามิเตอร์มากกว่ากัน

ประเด็นสำคัญอีกประเด็นหนึ่งที่เกี่ยวข้องกับการวิเคราะห์ข้อคำถามตามทฤษฎีการตอบสนองข้อสอบ คือ ขนาดของกลุ่มตัวอย่างที่ใช้ในการวิเคราะห์ข้อมูล ในความเป็นจริงแล้วไม่มีข้อสรุปที่ชัดเจนว่า การประยุกต์ทฤษฎีการตอบสนองข้อสอบมาใช้ในการสร้างมาตรวัดนั้น จะต้องใช้ขนาดกลุ่มตัวอย่างอย่างน้อยเพียงใด แต่มีแนวทางในการพิจารณาขนาดกลุ่มตัวอย่างที่เหมาะสม ดังนี้

1. ขนาดของกลุ่มตัวอย่างควรเพิ่มขึ้นเมื่อโมเดลมีความซับซ้อนมากขึ้น
2. การประมาณค่าพารามิเตอร์จะมีความคลาดเคลื่อนลดลง เมื่อกลุ่มตัวอย่างมีขนาดเพิ่มขึ้น
3. กลุ่มตัวอย่างควรมีขนาดใหญ่เพียงพอที่ทำให้ข้อมูลมีการกระจาย อย่างน้อยที่สุดควรมีข้อมูลการตอบข้อคำถามในทุกตัวเลือกการคำตอบของมาตรวัดนั้น

ประเด็นอื่น ๆ ที่เกี่ยวข้องกับทฤษฎีการตอบสนองข้อสอบ

1. ความเที่ยง (Reliability) สารสนเทศทางการวัด (Psychometric Information) และความคลาดเคลื่อนมาตรฐานของการประมาณค่า (Standard Error of Estimation)

แนวคิดเรื่องสารสนเทศทางการวัด (Psychometric Information) เป็นแนวคิดพื้นฐานของทฤษฎีการตอบสนองข้อสอบ ซึ่งพบได้ทั้งในข้อคำถามที่มีรูปแบบการตอบแบบที่มีการตรวจให้คะแนนสองค่า (Binary Item-Response Format) และข้อคำถามที่มีรูปแบบการตอบแบบที่มีการตรวจให้คะแนนมากกว่า 2 ค่า (Polytomous Item-Response Format) โค้งลักษณะข้อสอบ (Item Characteristic Curve: ICC) หรือโค้งการตอบสนองข้อสอบ (Item Response Curve: IRC)

ซึ่งพบในโมเดลการตอบสนองข้อสอบที่มีการตรวจให้คะแนนสองค่า (Dichotomous IRT Model) และโค้งการเลือกรายการคำตอบ (Category Response Curve: CRC) ซึ่งพบในโมเดลการตอบสนองข้อสอบที่มีการตรวจให้คะแนนมากกว่า 2 ค่า (Polytomous IRT Model) ล้วนสามารถนำมาแปลงเป็นโค้งหรือฟังก์ชันสารสนเทศของข้อคำถาม (Item Information Curve: IIC) ได้ โดยโค้งสารสนเทศของข้อสอบนี้ แสดงถึงข้อมูลทางการวัด (Psychometric Information) ของข้อคำถามตลอดช่วงคุณลักษณะแฝงที่มุ่งวัด (Latent-Trait Continuum) โดยข้อคำถามที่มีรูปแบบการตอบแบบที่มีการตรวจให้คะแนนสองค่า สามารถคำนวณฟังก์ชันสารสนเทศของข้อคำถามได้ ดังสมการที่ 13

$$I(\theta) = \frac{P_i^*(\theta)^2}{P_i(\theta)(1 - P_i(\theta))} \quad (13)$$

โดย

$P_i(\theta)$ หมายถึง ความน่าจะเป็นแบบมีเงื่อนไขของข้อคำถาม ซึ่งหมายถึง โค้งการตอบสนองข้อสอบ (IRC) ซึ่งได้จากการประมาณค่าพารามิเตอร์ตามโมเดลการตอบสนองข้อสอบ

$P_i^*(\theta)$ หมายถึง อนุพันธ์ลำดับที่ 1 (First Derivative) ของโค้งการตอบสนองข้อสอบ ณ ตำแหน่งคุณลักษณะแฝงที่เฉพาะเจาะจง

ในกรณีที่ เป็นโมเดลโลจิสแบบ 3 พารามิเตอร์ (3PL Model) สามารถเขียนสมการที่ 13 ใหม่เพื่อให้ชัดเจนขึ้นสำหรับการนำไปคำนวณได้ ดังสมการที่ 14

$$I(\theta) = \left[\alpha_i^2 \frac{1 - P_i(\theta)}{P_i(\theta)} \right] \left[\frac{(P_i(\theta) - c_i)^2}{(1 - c_i)^2} \right] \quad (14)$$

ในกรณีที่ เป็นโมเดลโลจิสแบบ 2 พารามิเตอร์ (2PL Model) สามารถเขียนสมการที่ 14 ให้ง่ายขึ้นได้ดังสมการที่ 15

$$I(\theta) = \alpha_i^2 P_i(\theta)(1 - P_i(\theta)) \quad (15)$$

และในกรณีที่ เป็นโมเดลโลจิสแบบ 1 พารามิเตอร์ (1PL Model) สามารถเขียนสมการที่ 15 ให้ง่ายขึ้นได้ดังสมการที่ 16

$$I(\theta) = P_i(\theta)(1 - P_i(\theta)) \quad (16)$$

สมการเหล่านี้กำหนดค่า $D=1.7$ ซึ่งได้รวมอยู่ในพจน์เดียวกับพารามิเตอร์อำนาจจำแนก แล้ว การนำสมการเหล่านี้ไปใช้มีกฎเกณฑ์บางประการ ได้แก่ 1) ในโมเดล 1PL และ 2PL นั้น ฟังก์ชันสารสนเทศจะมีค่าสูงที่สุดบริเวณตำแหน่งพารามิเตอร์ความยาก (Difficulty Parameter) ส่งผลให้ข้อคำถามใดที่มีความเหมาะสมกับระดับความสามารถของผู้ทดสอบ ข้อคำถามนั้นจะมีสารสนเทศสูงที่สุด ในขณะที่โมเดล 3PL ค่าสารสนเทศสูงสุดจะอยู่บริเวณช่วงที่ต่ำกว่าค่าพารามิเตอร์ความยากเล็กน้อย กล่าวอีกนัยหนึ่ง คือ พารามิเตอร์การเดาทำให้สารสนเทศของข้อสอบมีค่าต่ำกว่าที่ควรจะเป็น ซึ่งทำให้ข้อคำถามในโมเดล 1PL และ 2PL มีสารสนเทศสูงกว่าข้อคำถามในโมเดล 3PL และ 2) ค่าสารสนเทศของข้อคำถามขึ้นอยู่กับพารามิเตอร์อำนาจจำแนก (Discrimination Parameter) โดยข้อคำถามที่มีค่าพารามิเตอร์อำนาจจำแนกสูง จะแสดงให้เห็นถึงสารสนเทศทางการวัดบริเวณพารามิเตอร์ความยากที่สูงด้วย ข้อคำถามที่มีพารามิเตอร์อำนาจจำแนกสูง โค้งจะมีลักษณะโด่งมาก (Peak) ทำให้ข้อคำถามนั้น มีค่าฟังก์ชันสารสนเทศมากแต่ครอบคลุมค่าคุณลักษณะแฝงในช่วงแคบ ในขณะที่ โค้งที่มีค่าพารามิเตอร์อำนาจจำแนกต่ำ โค้งจะมีลักษณะราบเรียบ (Flat) ทำให้ข้อคำถามมีฟังก์ชันสารสนเทศต่ำ แต่ครอบคลุมค่าคุณลักษณะแฝงในช่วงกว้าง

ลักษณะสำคัญประการหนึ่งของฟังก์ชันสารสนเทศของข้อคำถาม คือ หากข้อสอบหรือข้อคำถามวัดคุณลักษณะแฝงเดียวกัน โค้งเหล่านี้สามารถนำมารวมกันและคำนวณได้ว่า ข้อคำถามทั้งหมดที่นำมารวมกัน ให้ข้อมูลสารสนเทศทั้งหมดมาน้อยเพียงใด โดยคำนวณได้จากสมการที่ 17

$$TI(\theta) = \sum_{i=1}^I I(\theta) \quad (17)$$

ข้อมูลสารสนเทศที่คำนวณได้จากสมการที่ 17 เรียกว่า ฟังก์ชันสารสนเทศของมาตรวัด (Test Information Function) ซึ่งเป็นข้อมูลสำคัญที่ทำให้พิจารณาได้ว่ามาตรวัด หรือแบบทดสอบมีคุณภาพมากน้อยเพียงใด เนื่องจากฟังก์ชันสารสนเทศของมาตรวัดมีความสัมพันธ์โดยตรงกับความคลาดเคลื่อนมาตรฐานในการวัด (Standard Error of Measurement) แสดงได้ดังสมการที่ 18

$$SE(\theta) = \frac{1}{\sqrt{TI(\theta)}} \quad (18)$$

ดังนั้น เมื่อทราบค่าฟังก์ชันสารสนเทศของมาตรวัด ก็ทำให้ทราบว่ามาตรวัดหรือแบบทดสอบนั้นสามารถวัดคุณลักษณะแฝงของบุคคลได้แม่นยำ (Precision) เพียงใด ความแม่นยำในที่นี้ เทียบได้กับความเที่ยง (Reliability) ตามทฤษฎีการทดสอบแบบดั้งเดิม ฟังก์ชันสารสนเทศของมาตรวัดตามแนวคิดของทฤษฎีการตอบสนองข้อสอบมีความเป็นอิสระจากกลุ่มผู้ทดสอบ ผู้ทดสอบที่มีระดับคุณลักษณะแฝงแตกต่างกัน เมื่อวัดด้วยแบบทดสอบหรือมาตรวัดเดียวกันจะได้ค่าฟังก์ชันสารสนเทศของมาตรวัดที่แตกต่างกันด้วย (ยกเว้นกรณีที่ฟังก์ชันสารสนเทศของมาตรวัดมีลักษณะราบเรียบ) ค่าฟังก์ชันสารสนเทศของมาตรวัดจึงมีหลายค่าและค่าเหล่านี้จะแตกต่างกันไปตามระดับคุณลักษณะแฝงของผู้ทดสอบ แนวคิดนี้แตกต่างจากแนวคิดของทฤษฎีการทดสอบแบบดั้งเดิมที่มีค่า

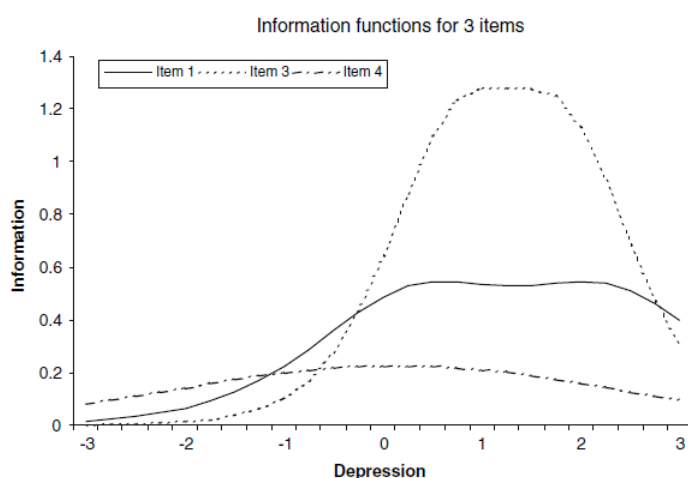
ความเที่ยงเพียงค่าเดียว ซึ่งอธิบายได้ว่าไม่ว่าผู้ทดสอบมีระดับคุณลักษณะแฝงสูงต่ำเพียงใด มาตรฐานวัดหรือแบบทดสอบก็สามารถประมาณค่าคุณลักษณะแฝงนั้นได้แม่นยำไม่แตกต่างกัน ทั้งที่ในสถานการณ์จริงปรากฏว่า มาตรฐานวัดและแบบทดสอบส่วนใหญ่ เมื่อวัดคุณลักษณะแฝงของบุคคลที่อยู่ในระดับสูงหรือต่ำมาก จะมีค่าความคลาดเคลื่อนในการวัดสูงกว่าการวัดคุณลักษณะแฝงของบุคคลที่อยู่ในระดับปานกลาง มาตรฐานวัดที่พัฒนาขึ้นส่วนใหญ่ จึงไม่สามารถวัดบุคคลที่มีคุณลักษณะแฝงแตกต่างกันได้เที่ยงตรงเท่ากัน เพราะมาตรฐานวัดจะมีความเที่ยงสูงเมื่อนำไปวัดคุณลักษณะแฝงของบุคคลที่ระดับใดระดับหนึ่ง และมีความเที่ยงลดลงเมื่อนำไปวัดคุณลักษณะแฝงของบุคคลที่อยู่ในระดับอื่น (Embretson & Reise, 2000, pp. 183-186; Reynolds & Livingston, 2012, pp. 236-242)

แนวคิดเกี่ยวกับฟังก์ชันสารสนเทศของข้อคำถามและฟังก์ชันสารสนเทศของมาตรฐานวัดสามารถนำไปใช้กับข้อคำถามที่มีรูปแบบการตรวจให้คะแนนมากกว่า 2 ค่า (Polytomous Response Format) โดยฟังก์ชันสารสนเทศของข้อคำถามแต่ละข้อจะได้มาจากผลรวมของโค้งการเลือกรายการคำตอบ (Category Response Curves: CRC) ในข้อคำถามนั้น และเมื่อรวมฟังก์ชันสารสนเทศของข้อคำถามเหล่านี้เข้าด้วยกันก็จะได้ฟังก์ชันสารสนเทศของมาตรฐานวัด ฟังก์ชันสารสนเทศของข้อคำถามสามารถคำนวณได้ดังสมการที่ 19

$$I_i(\theta) = \sum_{x=0}^m \frac{P_{ix}^*(\theta)^2}{P_{ix}(\theta)} \quad (19)$$

สมการที่ 19 เป็นสมการทั่วไปที่ใช้ในการคำนวณฟังก์ชันสารสนเทศของข้อคำถามในข้อคำถามที่มีรูปแบบการตรวจให้คะแนนมากกว่า 2 ค่า อย่างไรก็ตามวิธี รายละเอียดของการคำนวณค่อนข้างซับซ้อนและมีความแตกต่างกันไปตามโมเดลการตอบสนองข้อสอบที่เลือกใช้ (Embretson & Reise, 2000, pp. 183-186)

ฟังก์ชันสารสนเทศของข้อคำถาม (Item Information Function) แสดงได้ดังภาพที่ 2-8



ภาพที่ 2-8 ตัวอย่างสารสนเทศของข้อคำถาม 3 ข้อ จากมาตรฐานภาวะซึมเศร้า (Edelen & Reeve, 2007)

จากภาพที่ 2-8 แสดงให้เห็นว่า ข้อคำถามแต่ละข้อให้ข้อมูลสารสนเทศที่แตกต่างกัน ซึ่งในภาพรวมจะสังเกตได้ว่าคำถามข้อที่ 3 สามารถให้ข้อมูลสารสนเทศสูงที่สุด หากผู้ทดสอบมีภาวะซึมเศร้าในระดับปานกลางถึงรุนแรง อย่างไรก็ตาม ข้อคำถามอีกสองข้อจะให้ข้อมูลสารสนเทศของภาวะซึมเศร้าในระดับปานกลางถึงรุนแรงได้ต่ำกว่า แต่เมื่อต้องการวัดภาวะซึมเศร้าในระดับที่มีอาการน้อยหรือไม่รุนแรง จะเห็นได้ว่าคำถามข้อที่ 1 และข้อที่ 4 สามารถวัดภาวะซึมเศร้าในระดับดังกล่าวได้ดีกว่าคำถามข้อที่ 3 (Edelen & Reeve, 2007)

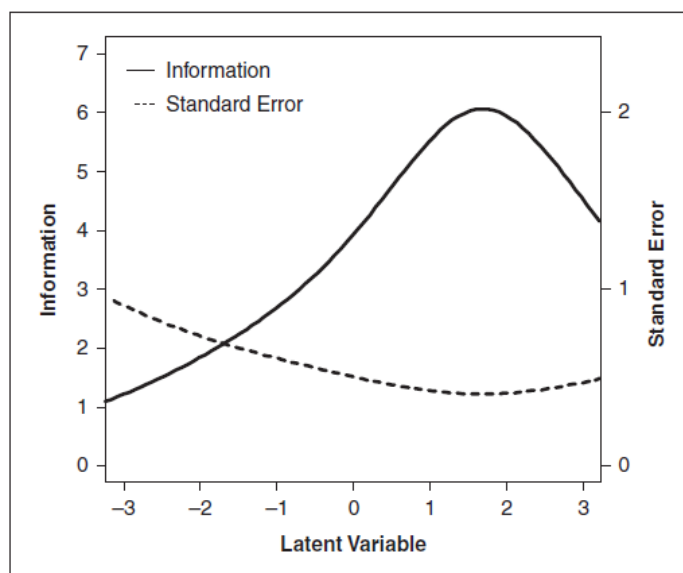
ฟังก์ชันสารสนเทศของข้อคำถามและฟังก์ชันสารสนเทศของมาตรวัดมีประโยชน์หลายประการ เช่น ในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ฟังก์ชันสารสนเทศจะช่วยให้ตัดสินใจได้ว่าข้อคำถามใดที่มีความเหมาะสมกับคุณลักษณะแฝงของบุคคล (Embretson & Reise, 2000, p. 185) นอกจากนี้ ฟังก์ชันสารสนเทศของข้อคำถามเป็นเครื่องมือสำคัญในการพัฒนามาตรวัดหรือแบบทดสอบ เพราะใช้ในการคัดเลือกข้อคำถามที่เหมาะสม มาตรวัดทางด้านจิตวิทยาส่วนใหญ่ มักประกอบด้วยข้อคำถามจำนวนมาก การพิจารณาฟังก์ชันสารสนเทศของข้อคำถามที่ได้จากข้อคำถามแต่ละข้อ จะช่วยให้ผู้วิจัยสามารถคัดเลือกเฉพาะข้อคำถามที่มีความเหมาะสมกับสิ่งที่ผู้วิจัยมุ่งวัด ทำให้ข้อคำถามในมาตรวัดมีจำนวนลดลง แต่มีประสิทธิภาพในการวัดไม่แตกต่างจากมาตรวัดที่มีข้อคำถามจำนวนมาก การคัดเลือกข้อคำถามอาจพิจารณาจากกลุ่มเป้าหมายของมาตรวัด เช่น นักวิจัยที่ต้องการศึกษากลุ่มผู้ป่วยที่มีภาวะซึมเศร้าในระดับรุนแรงอาจเลือกข้อคำถามที่มีฟังก์ชันสารสนเทศของข้อคำถามสูงในระดับคุณลักษณะแฝงที่มีค่าสูง (ภาวะซึมเศร้าในระดับรุนแรง) ในขณะที่ข้อคำถามซึ่งมีฟังก์ชันสารสนเทศของข้อคำถามสูงในระดับคุณลักษณะแฝงที่มีค่าต่ำ (ภาวะซึมเศร้าในระดับไม่รุนแรง) อาจไม่จำเป็นต้องนำไปใช้ในกรณีนี้ (Cohen, Swedlik, & Sturman, 2013)

นอกจากนี้ ฟังก์ชันสารสนเทศของข้อคำถามยังเปรียบเสมือนสัญญาณเตือน ที่ช่วยให้สามารถพิจารณาเกี่ยวกับคุณภาพของข้อคำถาม ข้อคำถามบางข้ออาจให้สารสนเทศของข้อมูลต่ำ และไม่สามารถจำแนกกลุ่มคนที่มีคุณลักษณะแฝงแตกต่างกันได้ ในกรณีนี้ ผู้วิจัยควรตั้งข้อสังเกต และพิจารณาในประเด็นต่าง ๆ ดังนี้

- 1) เนื้อหาของข้อคำถามตรงตามโครงสร้างที่ต้องการวัดหรือไม่
- 2) ถ้อยคำที่ใช้ในการเขียนข้อคำถาม สื่อความหมายได้ชัดเจนหรือไม่
- 3) ข้อคำถามมีความซับซ้อนและไม่เหมาะสมกับระดับการศึกษาของผู้ทดสอบหรือไม่
- 4) การจัดวางลำดับของข้อคำถาม ทำให้สูญเสียบริบทที่สื่อความหมายหรือไม่
- 5) ปัจจัยทางด้านวัฒนธรรม ส่งผลให้ความสามารถในการจำแนกกลุ่มคนของข้อคำถามต่ำกว่าที่ควรหรือไม่

ฟังก์ชันสารสนเทศของมาตรวัดมีประโยชน์ในการนำไปใช้ เพื่อคัดเลือกมาตรวัดที่วัดคุณลักษณะแฝงเดียวกัน โดยพิจารณาว่า มาตรวัดใดให้ค่าฟังก์ชันสารสนเทศของมาตรวัด ที่เหมาะสมกับกลุ่มเป้าหมายของการวัดมากกว่า (Embretson & Reise, 2000, p. 185) ฟังก์ชันสารสนเทศของมาตรวัด แสดงได้ดังภาพที่ 2-9 ซึ่งแสดงให้เห็นว่า มาตรวัดนี้เหมาะกับกลุ่มเป้าหมายที่มีระดับความสามารถ (ระดับคุณลักษณะแฝง) อยู่ในระดับปานกลางถึงสูง แต่ไม่เหมาะกับกลุ่มเป้าหมายที่มีระดับความสามารถ (ระดับคุณลักษณะแฝง) อยู่ในระดับต่ำ นอกจากนี้ ภาพนี้ยังแสดงให้เห็นถึง

ความคลาดเคลื่อนมาตรฐานของการประมาณค่า ซึ่งมีความสัมพันธ์โดยตรงกับฟังก์ชันสารสนเทศของมาตรวัด (Thomas, 2011)



ภาพที่ 2-9 ฟังก์ชันสารสนเทศของมาตรวัดและความคลาดเคลื่อนมาตรฐานของการประมาณค่า (Thomas, 2011)

2. การกำหนดมาตรการวัด (Scaling) และการปรับเทียบคะแนน (Equating)

ทฤษฎีการตอบสนองข้อสอบเป็นโมเดลการวัดที่พารามิเตอร์ข้อคำถามและพารามิเตอร์บุคคลมีความหมายสมบูรณ์ในตัวเองอย่างเห็นได้ชัด ในขณะที่ทฤษฎีการทดสอบแบบดั้งเดิมไม่สามารถใช้ข้อมูลจากข้อคำถามข้อใดข้อหนึ่งเพื่ออธิบายเชื่อมโยงถึงคุณลักษณะแฝงที่ต้องการวัดได้โดยตรง อย่างไรก็ตาม ทฤษฎีการทดสอบแบบดั้งเดิมมีแนวคิดที่ว่าหากผู้ทดสอบยืนยันคำตอบที่บ่งบอกถึงระดับคุณลักษณะแฝงใดซ้ำ ๆ ก็มีแนวโน้มว่าผู้ทดสอบจะมีคุณลักษณะแฝงอยู่ในระดับนั้น ผลรวมคะแนน (Total Scores) จากมาตรวัดหรือแบบทดสอบจึงมีลักษณะแบบเรียงลำดับตามระดับของคุณลักษณะแฝงที่มุ่งวัด ดังนั้น การแปลความหมายผลรวมคะแนนจึงสามารถทำได้โดยการคำนวณตำแหน่งเปอร์เซ็นต์ไทล์ (Percentile Rank) ซึ่งเป็นการอธิบายผลในระดับมาตรฐานอันดับ ในขณะที่แนวคิดของทฤษฎีการตอบสนองข้อสอบเป็นการอธิบายความสัมพันธ์ระหว่างพารามิเตอร์ข้อคำถาม (Item Parameter) กับพารามิเตอร์บุคคล (Person Parameter) ในมาตรอันตรภาคชั้น โดยค่าคุณลักษณะแฝงที่เพิ่มขึ้น เทียบเท่ากับการเพิ่มขึ้นของระดับการยืนยันคำตอบของข้อคำถาม (Item Endorsement) ด้วย

ตัวอย่างเช่น คะแนนจากมาตรวัดภาวะซึมเศร้าของสมชายและสมบัติ หากพิจารณาในบริบทของทฤษฎีการทดสอบแบบดั้งเดิม ความสัมพันธ์ของคะแนนระหว่างสมชายกับสมบัติ สามารถอธิบายได้โดยการอ้างอิงถึงปกติวิสัยที่วิเคราะห์ได้จากกลุ่มประชากร เช่น คะแนนรวมของสมบัติมี

ตำแหน่งเปอร์เซ็นต์ไทล์ที่ 99 ซึ่งสูงกว่าคะแนนรวมของสมชายซึ่งอยู่ในตำแหน่งเปอร์เซ็นต์ไทล์ที่ 7 หมายความว่า สมบัติมีภาวะซึมเศร้ามากกว่าสมชาย ความหมายของคะแนนรวมที่ได้จากแนวคิดของ ทฤษฎีการทดสอบแบบดั้งเดิมจึงจำกัดอยู่ที่การเปรียบเทียบ แต่ไม่สามารถอธิบายรายละเอียดเพิ่มเติม ต่อได้

ในขณะที่ทฤษฎีการตอบสนองข้อสอบสามารถอธิบายความแตกต่างของคุณลักษณะแฝงที่ วัดได้จากสมชายและสมบัติ ด้วยการเชื่อมโยงถึงพฤติกรรม ความคิด และความรู้สึกที่เกี่ยวข้องกับ ภาวะซึมเศร้าซึ่งเป็นโครงสร้างพื้นฐานสำหรับมาตรวัดนั้น โมเดลโลจิสติก (Logistic Model) หรือ โมเดลปกติสะสม (Normal Ogive Model) สามารถแสดงให้เห็นถึงโอกาสหรือความน่าจะเป็นใน การยืนยันคำตอบที่เกี่ยวข้องกับอาการต่าง ๆ ของภาวะซึมเศร้าซึ่งขึ้นอยู่กับคุณลักษณะแฝง เช่น หาก สมบัติมีระดับคุณลักษณะแฝง (ภาวะซึมเศร้า) เท่ากับ +2.50 เมื่อพิจารณาเนื้อหาของข้อคำถามร่วม ด้วยจะสามารถอธิบายได้ว่า สมบัติมีโอกาสที่จะรู้สึกเศร้า ร้อยละ 99 มีโอกาสที่จะรู้สึกสิ้นหวัง ร้อยละ 90 และมีโอกาสที่จะคิดฆ่าตัวตาย ร้อยละ 30 ในขณะที่ หากพบว่าสมชายมีระดับคุณลักษณะแฝง (ภาวะซึมเศร้า) เท่ากับ -0.50 เมื่อพิจารณาเนื้อหาของข้อคำถามร่วมด้วยจะสามารถอธิบายได้ว่า สมชายมีโอกาสที่จะรู้สึกเศร้า ร้อยละ 20 มีโอกาสที่จะรู้สึกสิ้นหวัง ร้อยละ 5 และมีโอกาสที่จะคิด ฆ่าตัวตาย ร้อยละ 1

นอกจากนี้ ผลการประมาณค่าระดับคุณลักษณะแฝงตามแนวคิดทฤษฎีการตอบสนอง ข้อสอบยังช่วยให้ผู้ประเมินตัดสินใจได้ว่าผู้ประเมินต้องช่วยเหลือสมบัติน้อยเพียงใด จึงจะทำให้ ระดับคุณลักษณะแฝง (ภาวะซึมเศร้า) ของเขาลดลงจนกระทั่งเหลือโอกาสที่เขาจะตัดสินใจฆ่าตัวตาย เพียงร้อยละ 5 ความหมายของค่าคุณลักษณะแฝงที่ได้นี้ไม่เพียงแต่ช่วยในการวินิจฉัยโรคทางจิตเวช แต่ยังช่วยให้ผู้ประเมินสามารถวางแผนการบำบัดรักษาที่เน้นการบรรเทาอาการต่าง ๆ ของภาวะ ซึมเศร้าได้ และทำให้สามารถจัดการความเสี่ยงต่าง ๆ ที่เกิดขึ้นได้อีกด้วย

ในบริบทของการศึกษาวิจัย การที่มาตรวัดสามารถรายงานค่าพารามิเตอร์ข้อสอบได้พร้อม กับค่าพารามิเตอร์บุคคล ทำให้ผู้วิจัยสามารถทำความเข้าใจความสัมพันธ์ระหว่างคุณลักษณะแฝงที่ ต้องการวัดกับข้อคำถามได้อย่างมีความหมายมากยิ่งขึ้น เช่น ในมาตรวัดภาวะซึมเศร้า การยืนยัน คำตอบในข้อคำถามที่เกี่ยวข้องกับการฆ่าตัวตายย่อมต้องเกิดขึ้นกับผู้ทดสอบที่มีภาวะซึมเศร้าใน ระดับรุนแรง และเป็นไปได้น้อยมากที่ผู้ทดสอบซึ่งมีภาวะซึมเศร้าอยู่ในระดับต่ำจะยืนยันคำตอบใน ข้อนี้ ในขณะที่ข้อคำถามที่เกี่ยวข้องกับการร้องไห้นั้น ไม่ว่าผู้ทดสอบจะมีระดับภาวะซึมเศร้ารุนแรง หรือไม่รุนแรงก็มีโอกาสที่จะยืนยันคำตอบได้ไม่แตกต่างกัน เป็นต้น

ตัวแปรแฝงตามโมเดลการตอบสนองข้อสอบเน้นที่ภาวะสันนิษฐาน (Construct) ไม่ได้เน้น ที่ผลรวมคะแนน (Total Scores) ดังนั้น มาตรวัดหลายฉบับที่มุ่งเน้นภาวะสันนิษฐานเดียวกันจึง สามารถนำมาปรับเทียบมาตรฐาน (Calibrate) เพื่อให้อยู่ในมาตรการวัดเดียวกันได้ ในขณะที่แนวคิด ของทฤษฎีการทดสอบแบบดั้งเดิมปรับเทียบมาตรวัดต่างฉบับกันได้ยากเพราะพารามิเตอร์ข้อสอบ และพารามิเตอร์บุคคลไม่ได้เป็นอิสระจากกัน (Thomas, 2011)

3. ความลำเอียงของข้อคำถามและมาตรวัด (Item and Test Bias)

เมื่อพารามิเตอร์ข้อสอบที่ประมาณค่าได้มีค่าแตกต่างกันตามกลุ่มประชากรที่ศึกษา พารามิเตอร์ข้อสอบที่ได้จากประชากรต่างกลุ่มกันนี้ไม่สามารถนำมาปรับเทียบเพื่อให้อยู่ในมาตร

การวัด (Scale of Measurement) เดียวกันได้ ส่งผลให้ข้อคำถามมีความลำเอียง (Item Bias) ข้อคำถามที่ปราศจากความลำเอียง คือ ข้อคำถามที่ไม่ว่าผู้ทดสอบเป็นประชากรกลุ่มใด แต่ถ้ามีระดับคุณลักษณะแฝงอยู่ในระดับเดียวกัน ก็มีค่าความน่าจะเป็นของการตอบข้อสอบถูก หรือค่าความน่าจะเป็นในการเลือกรายการคำตอบไม่ต่างกัน กล่าวอีกนัยหนึ่งได้ว่า ความแตกต่างของกลุ่มประชากรไม่ทำให้ค่าความน่าจะเป็นในการตอบข้อคำถาม หรือการเลือกรายการคำตอบเปลี่ยนแปลงไป ข้อคำถามที่มีค่าพารามิเตอร์เปลี่ยนแปลงไปตามกลุ่มประชากร เรียกว่า ข้อคำถามมีการทำหน้าที่ต่างกันของข้อสอบ (Differential Item Functioning: DIF) ขณะที่ข้อคำถามที่ค่าพารามิเตอร์ไม่เปลี่ยนแปลงไปตามกลุ่มประชากร เรียกว่า ข้อคำถามมีการไม่แปรเปลี่ยนของการวัด (Measurement Invariance)

การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ ตามแนวคิดของทฤษฎีการตอบสนองข้อสอบจะมีความถูกต้องแม่นยำ เพราะมีการจำแนกพารามิเตอร์บุคคลและพารามิเตอร์ข้อสอบออกจากกันอย่างชัดเจน ความแตกต่างในความสามารถของกลุ่มตัวอย่างแต่ละกลุ่มจึงแยกจากการตรวจสอบเรื่องความลำเอียงของแบบสอบได้ การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบเป็นกระบวนการที่พบได้ทั่วไปในบริบทที่มีความหลากหลายทางวัฒนธรรม การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบทำให้ผู้พัฒนาแบบสอบสามารถคัดเลือกข้อสอบที่ไม่เหมาะสมออกจากแบบสอบหรือมาตรวัดได้ นอกจากนี้ การที่แบบสอบหรือมาตรวัดปราศจากความลำเอียงยังแสดงให้เห็นถึงหลักการทางด้านจริยธรรมที่ผู้พัฒนาแบบสอบหรือมาตรวัดควรปฏิบัติตาม (Thomas, 2011)

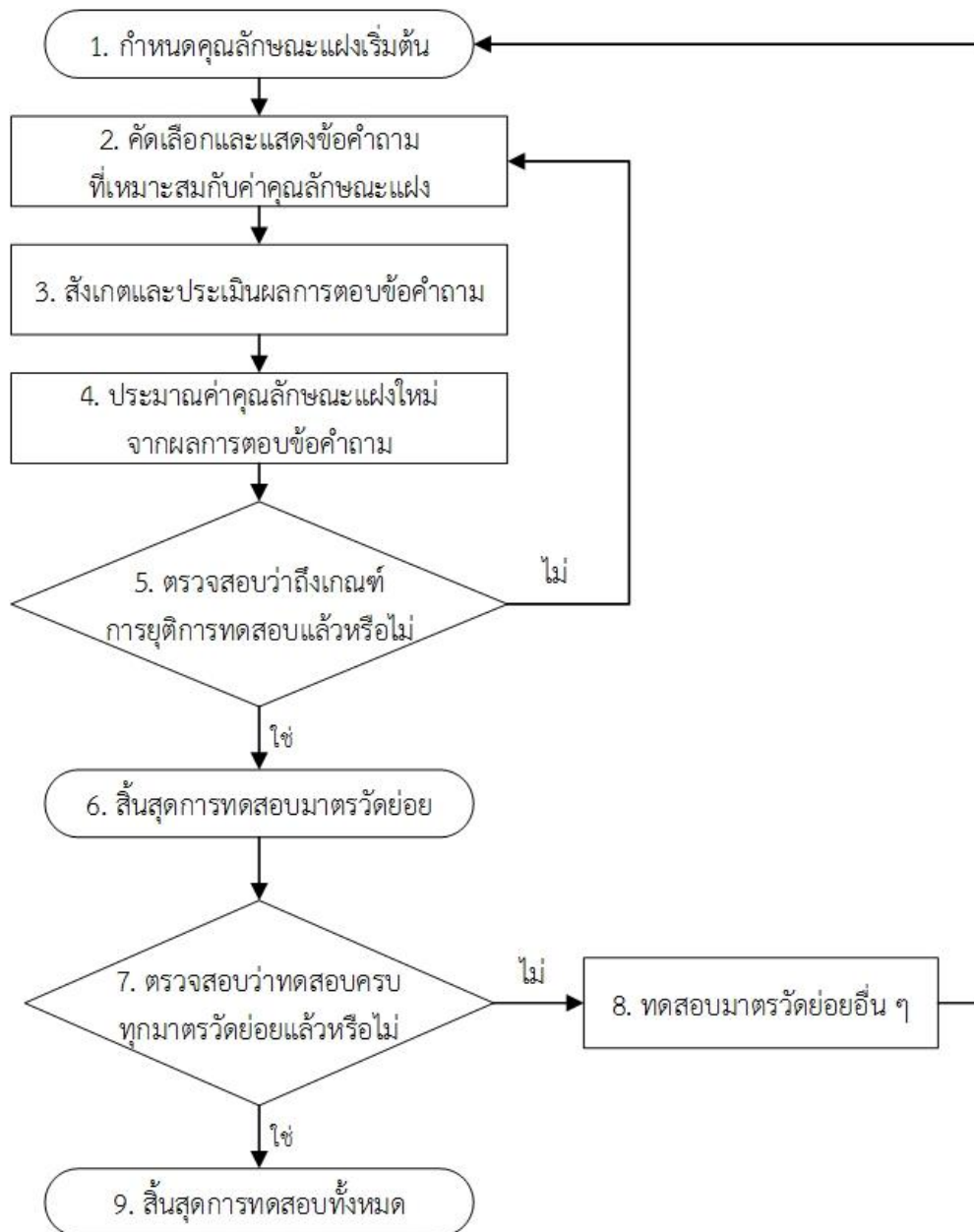
ตอนที่ 3 การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์และงานวิจัยที่เกี่ยวข้อง

หลักการการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์

การประยุกต์คอมพิวเตอร์ร่วมกับทฤษฎีการตอบสนองข้อสอบ ทำให้เกิดทางเลือกการทดสอบอีกรูปแบบหนึ่ง ซึ่งเรียกว่า การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ (Computerized Adaptive Testing: CAT) การทดสอบนี้ใช้แนวคิดของทฤษฎีการตอบสนองข้อสอบ (Item Response Theory: IRT) ทำให้สามารถประมาณค่าความสามารถหรือคุณลักษณะแฝง (Proficiency หรือ Theta ซึ่งใช้สัญลักษณ์ " θ ") โดยใช้ข้อมูลจากการตอบข้อคำถามแต่ละข้อ และเลือกข้อคำถามที่เหมาะสมข้อถัดไปสำหรับการทดสอบ โดยพิจารณาจากคุณลักษณะแฝงที่ประมาณค่าได้ใหม่ ตัวอย่างเช่น การประเมินภาวะซึมเศร้าตามแนวคิดของทฤษฎีการตอบสนองข้อสอบ มีแนวคิดว่าการตอบข้อคำถามของผู้ทดสอบแต่ละข้อสามารถบ่งบอกถึงระดับความรุนแรงของภาวะซึมเศร้าที่แตกต่างกัน เมื่อผู้ทดสอบได้ตอบข้อคำถามแรกแล้ว ข้อมูลการตอบข้อคำถามดังกล่าวสามารถนำไปใช้ในการประมาณค่าระดับความรุนแรงของภาวะซึมเศร้า รวมทั้งเป็นข้อมูลในการเลือกข้อคำถามข้อถัดไปที่วัดความซึมเศร้าในระดับความซึมเศร้าที่ประมาณค่าได้ การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ทำให้ผู้ทดสอบแต่ละคน ได้รับชุดของข้อคำถามที่แตกต่างกัน ผู้ทดสอบแต่ละคนจะได้รับเฉพาะข้อคำถามที่สอดคล้องกับคุณลักษณะแฝงของตนเท่านั้น ข้อคำถามใดที่มีความเกี่ยวข้องน้อย หรือไม่สัมพันธ์กับคุณลักษณะแฝงของผู้ทดสอบ จะไม่ถูกเลือกเข้ามาในการทดสอบ ส่งผลให้ข้อคำถามที่ใช้ในการทดสอบมีจำนวนลดลง ขณะที่ประสิทธิภาพในการทดสอบเพิ่มมากขึ้น และสามารถลดความเหนื่อยล้าในการทดสอบได้อีกด้วย (Gibbons et al., 2008)

การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ มีจุดเริ่มต้นมาจากการทดสอบแบบปรับเหมาะ (Adaptive Testing) ในรูปแบบกระดาษ-ดินสอ ซึ่งเป็นหลักการที่คิดค้นขึ้นโดย Alfred Binet ตั้งแต่ ปีค.ศ. 1905 Binet ได้สร้างแบบทดสอบความสามารถทางเชาวน์ปัญญา ที่ใช้ผลการตอบข้อสอบในแต่ละระดับอายุสมอง เป็นเกณฑ์ในการคัดเลือกระดับอายุสมองช่วงถัดไปของการทดสอบ (แบบทดสอบมีจำนวนข้อคำถามทั้งหมด 210 ข้อ ซึ่งจัดเป็นระดับอายุสมองตั้งแต่ 5-15 ปี โดยแต่ละระดับอายุสมองมีช่วงห่างกันครึ่งปี แต่ละครึ่งปีประกอบด้วยข้อคำถามจำนวน 10 ข้อ) ในการทดสอบนั้น หากผู้ทดสอบตอบได้ถูกต้องตามเกณฑ์ ผู้ทดสอบจะได้รับข้อคำถามในระดับอายุสมองที่สูงขึ้น แต่หากตอบผิด ผู้ทดสอบจะได้รับข้อคำถามในระดับอายุสมองที่ต่ำลง กระบวนการนี้จะดำเนินไปเรื่อย ๆ เพื่อค้นหาระดับอายุสมองพื้นฐาน (Basal Level) ซึ่งหมายถึง ระดับอายุสมองที่ผู้ทดสอบสามารถตอบคำถามในระดับนั้นได้ถูกต้องทุกข้อ และค้นหาระดับอายุสมองสูงสุด (Ceiling Level) ซึ่งหมายถึง ระดับอายุสมองที่ผู้ทดสอบตอบข้อคำถามผิดทุกข้อ การสิ้นสุดการทดสอบเกิดขึ้นเมื่อสามารถระบุระดับอายุสมองพื้นฐานและระดับอายุสมองสูงสุดได้แล้ว หลักการทดสอบของ Binet แตกต่างกับหลักการของการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ในปัจจุบัน เพราะการทดสอบแบบปรับเหมาะของ Binet นั้น การคัดเลือกข้อคำถามยังไม่สามารถเลือกเป็นรายชื่อได้อย่างแท้จริง แต่เป็นการเลือกตามระดับอายุสมอง การทดสอบแบบปรับเหมาะดังกล่าว จึงเป็นการทดสอบแบบปรับเหมาะแบบบางส่วน (Partially Adaptive Testing) ปัจจุบันการทดสอบแบบปรับเหมาะได้นำคอมพิวเตอร์และทฤษฎีการตอบสนองข้อสอบมาใช้ร่วมกัน จึงก่อให้เกิดกระบวนการปรับเหมาะอย่างสมบูรณ์ (Fully Adaptive Computerized Adaptive Testing) ที่สามารถตรวจให้คะแนนเป็นรายชื่อได้ทันทีหลังจากผู้ทดสอบตอบข้อคำถาม และสามารถเลือกข้อคำถามต่อไปได้จากผลการตอบข้อคำถามก่อนหน้า (Binet & Simon, 1905 cited in Weiss, 2011)

กระบวนการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ เป็นกระบวนการที่เป็นพลวัต (Dynamic) ซึ่งมีขั้นตอนการดำเนินการทดสอบดังภาพที่ 2-10



ภาพที่ 2-10 กระบวนการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ (ดัดแปลงจาก Wainer et al., 2000, p. 106)

การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ประกอบด้วยองค์ประกอบสำคัญ 5 องค์ประกอบ ดังนี้

1. คลังข้อคำถาม (Item Bank)

คือ แหล่งจัดเก็บข้อคำถาม (หรือข้อสอบ) ที่ใช้ในการทดสอบ รวมทั้งข้อมูลที่เกี่ยวข้องกับข้อคำถามนั้น ข้อคำถามต้องผ่านการตรวจสอบคุณภาพ และครอบคลุมช่วงคุณลักษณะแฝงที่ต้องการ

วัด การพัฒนาคลังข้อคำถาม มีขั้นตอนทั่วไป ดังนี้ (Thompson & Weiss, 2011; Wainer et al., 2000, pp. 37-45)

- 1.1 สร้างข้อคำถามในจำนวนมากพอที่จะครอบคลุมเนื้อหาทั้งหมด
- 1.2 ตรวจสอบคุณภาพของข้อคำถาม
- 1.3 นำข้อคำถามที่สร้างขึ้นในช่วงแรกไปทดลองใช้ ข้อคำถามที่จัดทำขึ้นเพื่อบรรจุในคลังข้อคำถามสำหรับการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ต้องผ่านการนำไปใช้ในรูปแบบกระดาษ-ดินสอก่อนเสมอ
- 1.4 เลือกข้อคำถามที่มีความเหมาะสมจากข้อมูลทางสถิติที่วิเคราะห์ได้
- 1.5 ตรวจสอบสัดส่วนของข้อคำถามในแต่ละเนื้อหา และประเมินกระบวนการทดสอบด้วยการจำลองสถานการณ์การทดสอบ (Simulation) เพื่อพิจารณาว่าข้อคำถามครอบคลุมช่วงของคุณลักษณะแฝงที่ต้องการวัดหรือไม่
- 1.6 ปรับข้อคำถามที่เหมาะสมทั้งหมดให้อยู่ในรูปแบบของคอมพิวเตอร์

นอกจากนี้ Walker, Bohnke, Cerny and Strasser (2010) ได้ศึกษาทบทวนอย่างเป็นระบบ (Systematic Review) และแนะนำเกี่ยวกับขั้นตอนการสร้างข้อสอบว่า ควรเริ่มต้นจากการทบทวนและกำหนดโครงสร้างของตัวแปรที่ต้องการวัด โดยศึกษาทฤษฎีและนิยามศัพท์ที่เกี่ยวข้องกับตัวแปรที่ต้องการวัด จากนั้นจึงกำหนดนิยามเชิงปฏิบัติการของตัวแปรนั้น เพื่อให้สามารถนำมาสร้างเป็นข้อคำถามได้ ซึ่งการสร้างข้อคำถามสามารถดำเนินการได้ 2 แนวทาง ได้แก่ 1) การเลือกข้อคำถามจากมาตรวัดที่มีอยู่แล้ว และ 2) การเขียนข้อคำถามขึ้นใหม่ตามนิยามเชิงปฏิบัติการ อย่างไรก็ตาม ไม่ว่าข้อคำถามที่ได้นั้นมาจากข้อคำถามที่สร้างขึ้นใหม่ หรือเป็นข้อคำถามที่ได้มาจากมาตรวัดที่มีอยู่แล้ว สิ่งสำคัญที่ต้องพิจารณา คือ ความเหมาะสมของข้อคำถาม ทั้งในด้านเนื้อหา รูปแบบการตอบ รูปแบบของมาตรวัด ระบบการให้คะแนน และการแปลความหมาย ซึ่งขั้นตอนนี้ต้องได้รับการตรวจสอบโดยผู้เชี่ยวชาญ

หลังจากนั้นจึงนำข้อคำถามที่ผ่านการตรวจสอบและแก้ไขเบื้องต้นแล้วมาวิเคราะห์ค่าพารามิเตอร์ข้อสอบ และข้อมูลสารสนเทศ (Calibration Study) ซึ่งการดำเนินการในส่วนนี้เกี่ยวข้องกับทฤษฎีการวัด ซึ่งอาจใช้ทฤษฎีการทดสอบแบบดั้งเดิม และ/ หรือ ทฤษฎีการตอบสนองข้อสอบก็ได้ อย่างไรก็ตาม การวิเคราะห์ข้อคำถามด้วยทฤษฎีการทดสอบแบบดั้งเดิมมีข้อจำกัด เนื่องจากค่าที่ได้จากการวิเคราะห์ข้อคำถามอาจแปรเปลี่ยนไปตามกลุ่มผู้สอบ และค่าคุณลักษณะแฝงอาจแปรเปลี่ยนไปได้ตามมาตรวัดที่ใช้ ขณะที่ทฤษฎีการตอบสนองข้อสอบมีจุดเด่น คือ ความไม่แปรเปลี่ยนของพารามิเตอร์ข้อสอบตามกลุ่มผู้สอบ และความไม่แปรเปลี่ยนของพารามิเตอร์บุคคลตามมาตรวัดที่ใช้ ดังนั้น ค่าที่วิเคราะห์ได้จึงมีความคงที่มากกว่า ด้วยเหตุผลดังกล่าวทำให้ปัจจุบัน การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์นิยมวิเคราะห์ข้อคำถามด้วยทฤษฎีการตอบสนองข้อสอบ การวิเคราะห์ค่าพารามิเตอร์ของข้อคำถาม จะต้องตรวจสอบความสอดคล้องของข้อมูลที่ได้จากมาตรวัดกับโมเดลการตอบสนองข้อสอบ ซึ่งมีทั้งโมเดลการตอบสนองข้อสอบแบบเอกมิติและพหุมิติ รวมถึงโมเดลการตอบสนองข้อสอบที่มีการตรวจให้คะแนนแบบสองค่าและแบบมากกว่าสองค่า นอกจากนี้ ต้องตรวจสอบข้อตกลงเบื้องต้นของโมเดลการตอบสนองข้อสอบด้วย (Walker, Bohnke, Cerny, & Strasser, 2010)

การคัดเลือกข้อคำถามที่เหมาะสมสำหรับคลังข้อคำถามนั้น มีแนวทางที่หลากหลาย สามารถสรุปแนวทางต่าง ๆ ได้ดังนี้

1) มิติของมาตรวัด (Dimensionality) ตรวจสอบได้โดยการวิเคราะห์องค์ประกอบเชิงสำรวจ (Exploratory Factor Analysis: EFA) ซึ่งอาจกำหนดค่าน้ำหนักองค์ประกอบ (Factor Loadings) ที่ต้องการเป็นเกณฑ์ในการคัดเลือกข้อคำถาม เช่น หากค่าน้ำหนักองค์ประกอบมีค่าน้อยกว่า 0.30 (Olinos et al., 2012) หรือ น้อยกว่า 0.40 (Pilkonis et al., 2011) ให้คัดข้อคำถามนั้นออกจากคลังข้อคำถาม

2) การวิเคราะห์องค์ประกอบเชิงยืนยัน (Confirmatory Factor Analysis: CFA) โดยพิจารณาคัดเลือกข้อคำถามที่มีค่าน้ำหนักองค์ประกอบน้อยกว่า 0.40 ออกจากคลังข้อคำถาม (Fliege et al., 2005; Walter, Becker, & Bjorner, 2007; Kocalevant et al., 2009)

3) ค่าสหสัมพันธ์เศษเหลือ (Residual Correlations) ซึ่งได้จากการวิเคราะห์องค์ประกอบเชิงยืนยัน โดยสามารถกำหนดเกณฑ์ค่าสหสัมพันธ์เศษเหลือที่ยอมรับได้ตามความเหมาะสม โดยทั่วไปนิยมกำหนดเกณฑ์ในการพิจารณาคัดข้อคำถามออกจากคลังข้อคำถามเมื่อข้อคำถามนั้นมีค่าสหสัมพันธ์เศษเหลือมากกว่า 0.20 (Pilkonis et al., 2011) หรือมากกว่า 0.25 (Fliege et al., 2005; Walter, Becker, & Bjorner, 2007; Kocalevant et al., 2009)

4) การแจกแจงความถี่ (Frequency Distribution) ของผลคะแนนรวมของผู้ทดสอบ โดยเลือกข้อคำถามที่มีความเหมาะสมกับการแจกแจงความถี่ของผลคะแนนรวมของผู้ทดสอบ หากข้อคำถามใดมีค่าพารามิเตอร์ตำแหน่งอยู่ในช่วงที่มีการแจกแจงความถี่ต่ำ จะพิจารณาคัดออกจากคลังข้อคำถาม (Pilkonis et al., 2011)

5) ค่าสหสัมพันธ์ปรับแก้ระหว่างข้อคำถามกับมาตรวัด (Adjusted item-total correlations) ข้อคำถามที่มีค่าสหสัมพันธ์ปรับแก้ระหว่างข้อคำถามกับมาตรวัดต่ำกว่า 0.4 จะถูกคัดออกจากคลังข้อคำถาม (Pilkonis et al., 2011)

6) โมนोटอนิกซิตี (Monotonicity) ซึ่งเป็นสิ่งที่แสดงให้เห็นว่าความน่าจะเป็นในการยืนยันคำตอบที่เพิ่มขึ้น เมื่อระดับคุณลักษณะแฝงสูงขึ้น ซึ่งสามารถตรวจสอบได้โดยใช้สถิติไม่พารามิเตอร์ (Nonparametric) ข้อคำถามใดที่ไม่พบลักษณะโมนोटอนิกซิตีจะถูกคัดออกจากคลังข้อคำถาม (Pilkonis et al., 2011)

7) ความสามารถของมาตรวัด (Scalability) ตรวจสอบได้โดยคำนวณค่า Loevinger H coefficients ซึ่งหากมีค่าต่ำกว่า 0.30 ให้พิจารณาคัดข้อคำถามนั้นออกจากคลังข้อคำถาม (Pilkonis et al., 2011)

คลังข้อคำถามเป็นปัจจัยสำคัญที่ทำให้การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ดำเนินไปได้อย่างมีประสิทธิภาพ ซึ่งโดยทั่วไปควรมีคลังข้อคำถามขนาดใหญ่ซึ่งประกอบด้วยข้อคำถามจำนวนมากที่ผ่านการวิเคราะห์ด้วยทฤษฎีทดสอบมาแล้ว ในคลังข้อคำถามอาจมีรูปแบบการตอบข้อคำถามที่หลากหลายอยู่ในคลังข้อคำถามเดียวกันได้ เช่น ข้อคำถามแบบหลายตัวเลือก (Multiple Choice Questions) แบบถูกผิด (True/ False Questions) ซึ่งเป็นรูปแบบที่มีการให้คะแนนแบบสองค่า หรือข้อคำถามที่ใช้มาตราประมาณค่า (Graded Response Format) ในทางทฤษฎีข้อคำถามในคลังข้อคำถามควรเป็นข้อคำถามที่มีความสามารถในการจำแนกสูง และข้อคำถาม

เหล่านั้นควรครอบคลุมทุกระดับความสามารถ หรือระดับคุณลักษณะแฝงที่ต้องการวัด ซึ่งจะทำให้ การทดสอบดำเนินไปได้ด้วยดี โดยทั่วไป คลังข้อคำถามควรประกอบด้วยข้อคำถามประมาณ 100 ข้อ (ซึ่งหากรูปแบบการตอบข้อคำถามเป็นแบบที่มีการตรวจให้คะแนนมากกว่า 2 ค่า จำนวนข้อคำถามใน คลังข้อคำถามสามารถมีจำนวนน้อยกว่านี้ได้) อย่างไรก็ตาม ในบริบททางด้านการวัดบุคลิกภาพและเจตคติ นั้น การสร้างคลังข้อคำถามให้มีข้อคำถามจำนวนมากเป็นสิ่งที่เป็นไปได้ยาก ส่งผลให้ที่ผ่านมา การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ถูกนำมาใช้ในการวัดผลทางการศึกษามากกว่า (Embretson & Reise, 2000, pp. 264-266)

อย่างไรก็ดี ปัจจุบันมีการนำโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์มาใช้ในการประเมินทางด้านบุคลิกภาพ เจตคติ และตัวแปรทางด้านจิตวิทยาอื่น ๆ มากยิ่งขึ้น ซึ่งงานวิจัยแต่ละฉบับต่างก็มีจำนวนข้อคำถามในคลังข้อคำถามที่แตกต่างกันไป มาตรฐานฉบับนี้มีข้อคำถามในคลังข้อคำถามไม่มากนัก แต่ก็สามารถดำเนินการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ได้เป็นอย่างดี จำนวนข้อคำถามในคลังข้อคำถามของมาตรฐานวัดต่าง ๆ ที่ใช้กับโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ แสดงได้ ดังตารางที่ 2-2

ตารางที่ 2-2 จำนวนข้อคำถามในคลังข้อคำถามของมาตรฐานวัดต่าง ๆ ที่ใช้กับโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์

มาตรวัด	จำนวนข้อคำถาม (ข้อ)
1 Depression-CAT (Fliege et al., 2005)	64
2 Anxiety-CAT (Becker et al., 2008)	50
3 Mood and Anxiety Spectrum Scales (MASS; Gibbons et al., 2008)	616
4 Stress- CAT (Kocalevent et al., 2009)	69
5 CES-D Scale (Smits, Cuijpers, & Straten, 2011)	20
6 Questionnaires of Children with Psychosocial Problem (Vogels, Jacobusse, & Reijneveld, 2011)	190
7 Anhedonic Depression CAT (Smits, Zitman, Cuijpers, Hollander-Gijsman, & Carlier, 2012)	22
8 Short Form Health Survey (SF-36; Zheng, Chang, & Chang, 2013)	36
9 Transmissible Liability Index (TLI; Kirisci et al., 2012)	65

2. การเริ่มต้นการทดสอบ (Starting the First Item)

การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ใช้ผลการตอบข้อคำถามในข้อคำถามก่อนหน้าเป็นข้อมูลสำคัญในการดำเนินการทดสอบ แต่เมื่อเริ่มต้นการทดสอบยังไม่มีข้อมูลการตอบ

ข้อคำถามก่อนหน้า ผู้ทดสอบจึงต้องเป็นผู้พิจารณาว่าจะใช้เกณฑ์ใดในการเริ่มต้นการทดสอบแนวทางในการกำหนดเกณฑ์เริ่มต้นการทดสอบมีหลากหลายแนวทาง ดังนี้

2.1 หากไม่มีข้อมูลอื่นประกอบการพิจารณา นิยมใช้ค่าเฉลี่ยคุณลักษณะแฝงของประชากร (Average Theta) ซึ่งมีค่าเท่ากับ 0.00 เป็นจุดเริ่มต้นในการเลือกข้อคำถามแรกให้เหมาะกับระดับคุณลักษณะแฝงดังกล่าว (Becker et al., 2008; Fliege et al., 2005; Kirisci et al., 2012; Kocalevent et al., 2009; Smits, Cuijpers, & Straten, 2011)

2.2 ใช้วิธีการสุ่มเลือกข้อคำถามด้วยโปรแกรมคอมพิวเตอร์ (Vogels, Jacobusse, & Reijneveld, 2011)

2.3 หากมีผลการทดสอบของมาตรวัดนั้น ๆ ในครั้งก่อนหน้า ผู้ทดสอบอาจเริ่มต้นด้วยค่าคุณลักษณะแฝงของผู้ทดสอบที่ได้จากครั้งก่อน (Thompson & Weiss, 2011; Wainer et al., 2000, pp. 107-108)

2.4 พิจารณาจากข้อมูลประกอบอื่น ๆ เช่น ผลการทดสอบจากมาตรวัดอื่น เกรดเฉลี่ยแรงจูงใจในการทดสอบ เศรษฐฐานะทางสังคม (Thompson & Weiss, 2011; Wainer et al., 2000, pp. 107-108)

3. การคัดเลือกข้อคำถามข้อถัดไป (Selecting the Next Item)

วิธีการในการเลือกข้อคำถามข้อถัดไปในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่นิยมใช้อย่างแพร่หลาย มีสองแนวทาง ได้แก่ 1) การเลือกข้อคำถามที่ให้ข้อมูลสารสนเทศสูงสุด (Maximum Information) และ 2) การเลือกข้อคำถามที่ทำให้เกิดความแม่นยำในการประมาณค่าสูงสุด (Maximum Expected Precision) หรือบางครั้งเรียกว่า การเลือกข้อคำถามที่ทำให้ส่วนเบี่ยงเบนมาตรฐานที่คาดหวังมีค่าน้อยที่สุด (Minimum Expected Posterior Standard Deviation) วิธีการในการเลือกข้อคำถามทั้งสองวิธีนี้สามารถปรับได้เมื่อพิจารณาร่วมกับประเด็นอื่น ๆ เช่น การควบคุมอัตราการใช้ข้อสอบ (Exposure Control) และการควบคุมสัดส่วนเนื้อหาของข้อคำถาม (Content Balancing Strategies)

3.1 การเลือกข้อคำถามที่ให้ข้อมูลสารสนเทศสูงสุดแบบไม่มีการควบคุม (Unconstrained Maximum Information Selection) เป็นวิธีการที่เลือกข้อคำถามข้อที่ j ซึ่งให้สารสนเทศของข้อมูลสูงสุด ณ ตำแหน่งคุณลักษณะแฝงที่ θ_i สารสนเทศของข้อคำถามสามารถคำนวณได้จากสมการที่ 20 (Wainer et al., 2000)

$$I_j(\hat{\theta}_i) = \frac{[P'_j(\hat{\theta}_i)]^2}{\{P_j(\hat{\theta}_i)[1 - P_j(\hat{\theta}_i)]\}} \quad (20)$$

โดย

$\hat{\theta}_i$ หมายถึง คุณลักษณะแฝงซึ่งประมาณค่าล่วงหน้าสำหรับผู้ทดสอบที่ i ซึ่งทดสอบด้วยข้อคำถามมาแล้ว n ข้อ

$P_j(\hat{\theta}_i)$ หมายถึง ความน่าจะเป็นที่ผู้ทดสอบซึ่งมีค่าคุณลักษณะแฝงเท่ากับ $\hat{\theta}_i$ จะตอบข้อคำถามที่ j ซึ่งได้จากคลังข้อคำถามถูก

$P'_j(\hat{\theta}_i)$ หมายถึง อนุพันธ์ของ $P_j(\theta)$ ซึ่งสัมพันธ์กับค่า θ ที่ถูกประมาณค่า ณ ตำแหน่ง $\hat{\theta}$

3.2 การเลือกข้อคำถามที่ทำให้การประมาณค่ามีความแม่นยำสูงสุดตามวิธีการของเบส์ (The Bayesian Counterpart of Maximum Information Item Selection Maximizes Posterior Precision) ซึ่งนิยมเรียกในอีกชื่อหนึ่งว่า Owen's Bayesian Method เป็นวิธีการที่นำเสนอโดย Owen (1969 cited in Wainer et al., 2000) ซึ่งในแต่ละขั้นตอนของการทดสอบจะได้ข้อมูลการแจกแจงความถี่ชุดหลัง (Posterior Distribution) ของคุณลักษณะแฝง (θ) ภายหลังจากการดำเนินการทดสอบด้วยข้อคำถามทั้งสิ้น n ข้อ แทนด้วย $p(\theta|s_n)$ วิธีการนี้ใช้คัดเลือกข้อคำถามโดยค้นหาข้อคำถามที่ให้ข้อมูลสารสนเทศสูงที่สุดซึ่งทำให้การประมาณค่ามีความแม่นยำมากที่สุด ด้วยสมการที่ 21

$$E_x \left[\text{Var}^{-1}(\theta|s_n, j, x_{n+1}) \right] = \sum_{x=0}^1 \text{Var}^{-1}(\theta|s_n, (j, x)) \cdot \int P(x_{n+1} = x|\theta, a_j, b_j, c_j) p(\theta|x_n) d\theta \quad (21)$$

โดย

$\text{Var}^{-1}(\theta|s_n, j, x_{n+1})$ หมายถึง ความแม่นยำในการประมาณค่าภายหลังจากเลือกตอบ x (ซึ่งแทนด้วย $(n+1)$) ในคำถามข้อที่ j

s_n หมายถึง สารสนเทศทั้งหมดที่ได้จากผู้ทดสอบ ซึ่งผ่านการทดสอบ n ข้อ

x_n หมายถึง ชุดของคำตอบของผู้ทดสอบในข้อก่อนหน้าทั้งหมด ซึ่งมีค่าเท่ากับ x_1 ถึง x_n

วิธีการคัดเลือกข้อคำถามเบื้องต้น เป็นเพียงหลักการโดยทั่วไปเท่านั้น การพัฒนาวิธีการคัดเลือกข้อคำถามยังคงเกิดขึ้นอย่างต่อเนื่อง แต่ละวิธีการต่างมีแนวทางที่แตกต่างกันออกไป นอกเหนือจากหลักการคัดเลือกข้อคำถามข้อถัดไปแล้ว ผู้ทดสอบต้องพิจารณาประเด็นที่เกี่ยวข้องในทางปฏิบัติ ที่สำคัญ 2 ประเด็น ได้แก่ การควบคุมอัตราการใช้ข้อสอบ (Item Exposure Control) ซึ่งเป็นการควบคุมไม่ให้ข้อคำถามข้อใดข้อหนึ่ง ถูกเลือกใช้มากกว่าข้อคำถามอื่น ๆ และประเด็นด้านการควบคุมสัดส่วนเนื้อหาของข้อคำถาม (Content Balancing Strategies) ซึ่งเป็นการควบคุมความสมดุลของการวัดตามองค์ประกอบที่กำหนดไว้ และเป็นการระมัดระวังไม่ให้ผู้ทดสอบได้รับการทดสอบในเนื้อหาใดด้านหนึ่งเพียงด้านเดียว การจำกัดการใช้ข้อคำถามทั้งสองรูปแบบอาจส่งผลกระทบต่อประสิทธิภาพในการทดสอบได้ เพราะเป็นการเข้าไปแทรกแซงกระบวนการคัดเลือกข้อคำถาม อย่างไรก็ตาม อย่างไรก็ดี กระบวนการเหล่านี้มีความจำเป็นในบริบทการทดสอบที่มีการนำผล

การทดสอบไปใช้ในการตัดสินใจที่สำคัญ (High Stake Test) ซึ่งเป็นการทดสอบที่พบได้บ่อยในทางคลินิก (Thompson & Weiss, 2011; Wainer et al., 2000, pp. 108-111)

งานวิจัยที่เกี่ยวข้องกับการนำโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์มาใช้ในการพัฒนามาตรวัดทางด้านจิตวิทยานั้น มีวิธีการเลือกข้อสอบข้อถัดไปที่หลากหลาย แตกต่างกันไป วิธีการที่ได้รับความนิยมค่อนข้างมาก คือ วิธีการเลือกข้อคำถามที่ให้สารสนเทศสูงสุด ณ ตำแหน่งคุณลักษณะแฝงของผู้ทดสอบที่ประมาณค่าได้ (Becker et al., 2008; Fliege et al., 2005; Kocalevent et al., 2009; Smits, Cuijpers, & Straten, 2011; Smits, Zitman, Cuijpers, Hollander-Gijsman, & Carlier, 2012; Zheng, Chang, & Chang, 2012)

4. การประมาณค่าคุณลักษณะแฝง (Proficiency Estimation/ Scoring)

การประมาณค่าคุณลักษณะแฝง สามารถประมาณค่าได้โดยอาศัยทฤษฎีการทดสอบแบบดั้งเดิม และทฤษฎีการตอบสนองข้อสอบ แต่วิธีการของทฤษฎีการตอบสนองข้อสอบได้รับความนิยม และสามารถประมาณค่าคุณลักษณะแฝงในแต่ละตำแหน่งได้มีประสิทธิภาพมากกว่า วิธีการที่นิยมใช้ส่วนใหญ่มี 3 วิธี ได้แก่ (De Ayala, 2009, pp. 75-79; Bock & Mislevy, 1982; Embretson & Reise, 2000, pp. 158-186; Keller, 2000; Thompson, 2009; Thompson & Weiss, 2011; Wainer et al., 2000, pp. 111-112)

4.1 Maximum Likelihood (ML) เป็นกระบวนการค้นหาค่าคุณลักษณะแฝง (θ) ที่ทำให้มีค่าควรจะเป็น (Likelihood) ของรูปแบบการตอบข้อคำถาม (Item Response Pattern) ของผู้ทดสอบมากที่สุด กล่าวอีกนัยหนึ่ง คือ หากรูปแบบการตอบข้อคำถามของผู้ทดสอบเป็น 0 (ตอบผิด หรือไม่ยืนยันคำตอบ) และ 1 (ตอบถูก หรือยืนยันคำตอบ) ซึ่งทราบค่าพารามิเตอร์ข้อสอบแล้ว วิธีการนี้ คือ การค้นหาว่าผู้ทดสอบที่มีรูปแบบการตอบข้อคำถามเช่นนี้ ควรจะมีคุณลักษณะแฝงอยู่ในระดับใด ซึ่งค่าคุณลักษณะแฝงดังกล่าวเป็นได้ทั้งค่าบวกและลบอนันต์ (Positive to Negative Infinity) ค่าควรจะเป็น (Likelihood Value) นี้สามารถคำนวณได้จากรูปแบบการตอบข้อคำถามของผู้ทดสอบ การแสดงค่าดังกล่าวด้วยกราฟฟังก์ชันจะช่วยให้เห็นภาพได้ชัดเจนยิ่งขึ้น กราฟที่ได้ เรียกว่า ฟังก์ชันควรจะเป็น (Likelihood Function) ซึ่งแกนนอน แสดงถึง ค่าคุณลักษณะแฝง (θ) ขณะที่แกนตั้ง แสดงถึง ค่าควรจะเป็นของรูปแบบการตอบข้อคำถามของผู้ทดสอบ เมื่อมีการคำนวณค่าควรจะเป็น (Likelihood Value) ที่เป็นไปได้ทั้งหมดตลอดช่วงของค่าคุณลักษณะแฝง (θ) แล้ว ผู้วิจัยสามารถเลือกระดับคุณลักษณะแฝงที่ทำให้ค่าควรจะเป็นมีค่ามากที่สุด กล่าวอีกนัยหนึ่ง คือ วิธีการประมาณค่าแบบ ML เป็นการหาผลรวมของฟังก์ชันควรจะเป็นของการตอบข้อคำถามแต่ละข้อ ซึ่งเปลี่ยนแปลงไปตามระดับคุณลักษณะแฝง จากนั้นจึงค้นหาค่าที่มีฟังก์ชันควรจะเป็นสูงที่สุด (ได้แก่ ค่าฐานนิยม) และใช้ค่าดังกล่าวเป็นค่าคุณลักษณะแฝง (θ) การประมาณค่าคุณลักษณะแฝงด้วยวิธี ML มีข้อจำกัด คือ หากผู้ทดสอบตอบข้อคำถามถูกทั้งหมด (ยืนยันคำตอบทุกข้อ) หรือตอบข้อคำถามผิดทั้งหมด (ไม่ยืนยันคำตอบทุกข้อ) จะไม่สามารถประมาณค่าคุณลักษณะแฝงได้ ส่งผลให้เกิดข้อจำกัดด้านการนำไปใช้ในบางกรณี นอกจากนี้ วิธีการประมาณค่าแบบ ML จะประมาณค่าคุณลักษณะแฝงได้ดี เมื่อกลุ่มตัวอย่างมีขนาดใหญ่ ในกรณีที่เป็นการประมาณค่าคุณลักษณะแฝงกลุ่มตัวอย่างดังกล่าว หมายถึง จำนวนข้อคำถาม ดังนั้น วิธีการนี้จึงเหมาะที่จะใช้เมื่อมาตรวัดหรือแบบทดสอบมีข้อคำถามจำนวนมาก (เช่น 50 ข้อขึ้นไป)

4.2 Maximum a Posteriori (MAP) เป็นวิธีการประมาณค่าแบบเบส์ (Bayesian Estimation Method) ซึ่งใช้ข้อมูลเกี่ยวกับค่าพารามิเตอร์ที่ทราบมาก่อน (Prior Information) ร่วมกับฟังก์ชันควรจะเป็นแบบล็อก (Log-Likelihood Function) เพื่อประมาณค่าคุณลักษณะแฝงจากค่าสูงสุดของการแจกแจงภายหลัง (Posterior Distribution) การประมาณค่าวิธีนี้ควรเข้าใจแนวคิดที่สำคัญ ดังนี้

ก) การแจกแจงก่อน (Prior Distribution) คือ การแจกแจงความน่าจะเป็นเชิงสมมติฐาน (Hypothetical Probability Distribution) ของผู้ทดสอบที่ได้มาจากการคัดเลือกแบบสุ่ม (Random Sample) ซึ่งการแจกแจงก่อนที่นิยมใช้มากสำหรับการประมาณค่าคุณลักษณะแฝงตามแนวคิดของทฤษฎีการตอบสนองข้อสอบ คือ การแจกแจงปรกติมาตรฐาน (Standard Normal Distribution) ซึ่งหมายถึง ผู้วิจัยเชื่อว่าผู้ทดสอบเป็นกลุ่มตัวอย่างที่ได้จากการแจกแจงซึ่งมีค่าเฉลี่ยเท่ากับ 0 และมีความแปรปรวน เท่ากับ 1.0 ในทางทฤษฎีนั้นการแจกแจงก่อนอาจเป็นรูปแบบใดก็ได้ อย่างไรก็ตาม การแจกแจงปรกติสะสมก็เป็นตัวเลือกหนึ่งที่มีความสมเหตุสมผล

ข) ฟังก์ชันควรจะเป็นแบบล็อก (Log-Likelihood Function) โดยปกติแล้วโค้งการตอบสนองข้อสอบ (Item Response Curves: IRCs) เป็นการประมาณค่าการแจกแจงความถี่สัมพัทธ์ ซึ่งหากมีการคำนวณการแจกแจงความถี่สัมพัทธ์ก่อนที่มีข้อมูลจริง จะเรียกว่า ความน่าจะเป็น (Probability) และเมื่อมีการคำนวณการแจกแจงความถี่สัมพัทธ์ภายหลังจากมีข้อมูลจริง จะเรียกว่า ค่าควรจะเป็น (Likelihood) ดังนั้น เมื่อมีการตอบข้อคำถาม ผู้วิจัยจึงสามารถใช้โค้งการตอบสนองข้อสอบในการคำนวณค่าควรจะเป็นตลอดช่วงของค่าคุณลักษณะแฝง ซึ่งค่าควรจะเป็น (Likelihood Values) นั้นสามารถคำนวณได้จากผลคูณของโค้งการตอบสนองข้อสอบ (IRC) ซึ่งหากโค้งการตอบสนองข้อสอบมีค่าระหว่าง 0 กับ 1 จะถูกนำมาคูณกัน และทำให้ค่าควรจะเป็นแบบมีเงื่อนไข (Conditional Likelihood) มีค่าลดลงอย่างรวดเร็ว ซึ่งเป็นผลมาจากการคำนวณที่ขาดความแม่นยำ ดังนั้น ผู้วิจัยส่วนมากจึงนิยมใช้ค่าควรจะเป็นแบบล็อกเพื่อแก้ปัญหาดังกล่าว ค่าควรจะเป็นแบบล็อกเป็นการจัดกระทำลอการิทึมธรรมชาติ (Natural Logarithm) กับโค้งการตอบสนองข้อสอบ ค่าลอการิทึมของตัวเลขระหว่าง 0 ถึง 1.0 จะมีค่าเป็นลบ และหากค่าเป็นลบมากจะทำให้ความน่าจะเป็นมีค่าต่ำ ฟังก์ชันควรจะเป็นแบบล็อกคำนวณได้จากผลรวมของโค้งการตอบสนองข้อสอบ ซึ่งแตกต่างจากฟังก์ชันควรจะเป็นทั่วไปซึ่งคำนวณจากผลคูณ จากนั้นจึงค้นหาระดับคุณลักษณะแฝงที่ทำให้ค่าควรจะเป็นแบบล็อกมีค่าสูงที่สุด โดยปกติแล้วหากข้อคำถามในแบบทดสอบหรือมาตรวัดมีจำนวนไม่มาก ผู้วิจัยสามารถนำฟังก์ชันควรจะเป็นแบบล็อกมารวมกันแล้วค้นหาระดับคุณลักษณะแฝงได้ไม่ยาก แต่ในความเป็นจริง การทดสอบอาจดำเนินการทดสอบกับผู้ทดสอบเป็นพันคน ด้วยข้อคำถามมากกว่า 50 ข้อ ทำให้นักวิจัยกลุ่มหนึ่งพยายามหาวิธีการที่จะค้นหาค่าคุณลักษณะแฝงได้สะดวกยิ่งขึ้น วิธีการที่เป็นที่นิยม คือ วิธีการของนิวตันและราฟส์สัน (An Iterative Newton-Raphson Procedure) ซึ่งเป็นการค้นหาค่าฐานนิยม (Mode) จากฟังก์ชันควรจะเป็นแบบล็อกของผู้ทดสอบแต่ละคน

ค) การแจกแจงหลัง (Posterior Distribution) เป็นผลคูณระหว่างฟังก์ชันควรจะเป็น (Likelihood Function) กับฟังก์ชันการแจกแจงก่อน (Prior Distribution Function)

วิธีการประมาณค่าแบบ MAP เป็นการค้นหาค่าคุณลักษณะแฝง (θ) ที่มีการแจกแจงหลังสูงสุด ซึ่งมีค่าเท่ากับค่าฐานนิยม (Mode) ทำให้บางครั้งวิธีการนี้มีชื่อเรียกว่า Bayes Modal Estimation

4.3 Expected a Posterior (EAP) วิธีการประมาณค่าแบบ EAP สามารถใช้ประมาณค่าคุณลักษณะแฝง เมื่อรูปแบบการตอบข้อคำถามเป็นไปในทิศทางใดทิศทางหนึ่งเพียงรูปแบบเดียว เช่น ตอบข้อคำถามถูกทั้งหมด หรือตอบข้อคำถามผิดทั้งหมด (ยืนยันคำตอบทั้งหมดหรือไม่ยืนยันคำตอบทั้งหมด) ซึ่งเป็นข้อดีที่เหนือกว่าวิธีการแบบ ML นอกจากนี้ เมื่อเปรียบเทียบวิธีการประมาณค่าแบบ MAP และ EAP พบว่าทั้งสองวิธีมีความแตกต่างกันหลายประการ ดังนี้

ก) วิธีการประมาณค่าแบบ MAP ใช้กระบวนการทำซ้ำ (Iterative Method) แต่วิธีการประมาณค่าแบบ EAP ไม่ใช้กระบวนการทำซ้ำ (Noniterative Method) แต่ใช้วิธีการหาพื้นที่แบบตัวเลข (Numerical Quadrature Methods) แทนซึ่งทำให้การประมาณค่ารวดเร็วกว่า ข) วิธีการประมาณค่าแบบ MAP ใช้การแจกแจงก่อน (Prior Distribution) แบบค่าต่อเนื่อง ในขณะที่วิธีการประมาณค่าแบบ EAP ใช้การแจกแจงก่อน (Prior Distribution) แบบไม่ต่อเนื่อง ค) วิธีการประมาณค่าแบบ MAP นั้นถึงแม้จะสามารถประมาณค่าคุณลักษณะแฝงได้ในรูปแบบการตอบข้อคำถามทุกรูปแบบ แต่พบปัญหาเกี่ยวกับการถดถอยเข้าสู่ค่าเฉลี่ยของการแจกแจงก่อน (Regression Toward the Prior's Mean) มากกว่าวิธีการประมาณค่าแบบ EAP ง) ค่าเฉลี่ยของความคลาดเคลื่อนกำลังสองในการประมาณค่าพารามิเตอร์ด้วยวิธีการประมาณค่าแบบ EAP มีค่าน้อยกว่า MAP และ จ) วิธีการนำไปใช้ของทั้งสองวิธีมีความแตกต่างกัน ซึ่งวิธีการทางคณิตศาสตร์ที่ใช้ในการประมาณค่าด้วยวิธี EAP ทำได้ง่ายกว่าวิธี MAP

การประมาณค่าคุณลักษณะแฝง (θ) ด้วยวิธีการแบบ EAP สามารถคำนวณได้จากสมการที่ 22

$$\hat{\theta}_i = \frac{\sum_{r=1}^R X_r L(X_r) A(X_r)}{\sum_{r=1}^R L(X_r) A(X_r)} \quad (22)$$

และค่าความคลาดเคลื่อนมาตรฐานภายหลัง (Posterior Standard Distribution) สามารถคำนวณได้จากสมการที่ 23

$$PSD(\hat{\theta}) = \sqrt{\frac{\sum_{r=1}^R (X_r - \hat{\theta}_i)^2 L(X_r) A(X_r)}{\sum_{r=1}^R L(X_r) A(X_r)}} \quad (23)$$

โดย

X_r	หมายถึง	จุดกึ่งกลางของพื้นที่โค้งการแจกแจงความน่าจะเป็น (Quadrature Node or Point)
$A(X_r)$	หมายถึง	ค่าฟังก์ชันการแจกแจงของคุณลักษณะแฝงบริเวณจุดกึ่งกลางของพื้นที่ (Quadrature Weight)
R	หมายถึง	จำนวนจุดกึ่งกลางของพื้นที่โค้งการแจกแจงความน่าจะเป็น
$L(X_r)$	หมายถึง	ฟังก์ชันควรจะเป็น (Likelihood Function) ณ ตำแหน่ง X_r ที่มีรูปแบบการตอบข้อคำถามเป็น $\underline{x} = x_1, \dots, x_L$ ซึ่งสามารถคำนวณได้จากสมการที่ 24

$$L(X_r) = \prod_{j=1}^L p_j(X_r)^{x_{ij}} (1-p_j(X_r))^{(1-x_{ij})} \quad (24)$$

วิธีการประมาณค่าคุณลักษณะแฝงแบบ EAP สามารถกำหนดจำนวนจุดกึ่งกลางของพื้นที่โค้งการแจกแจงความน่าจะเป็น (Quadrature Node or Point) ตามความเหมาะสม ซึ่งการเพิ่มจำนวนจุดกึ่งกลางของพื้นที่จาก 10 จุด เป็น 20 จุด จะทำให้การประมาณค่าคุณลักษณะแฝงมีความแม่นยำเพิ่มขึ้น โดยไม่ต้องคำนึงว่ากลุ่มตัวอย่างมีขนาดเท่าใด หรือ การแจกแจงก่อนจะเหมาะสมหรือไม่ (Seong, 1990 cited in De Ayala, 2009, p. 79) นอกจากนี้ จากการศึกษาวิจัยพบว่าหากใช้จุดกึ่งกลางของพื้นที่โค้งการแจกแจงความน่าจะเป็นเพียง 10 จุดทำให้ผลความคลาดเคลื่อนจากการประมาณค่าด้วยวิธีการแบบ EAP มีค่าต่ำกว่าที่ควรจะเป็น (Underestimated) ทำให้ประมาณค่าคุณลักษณะแฝงได้แม่นยำลดลง ซึ่งโดยทั่วไปแล้ว ควรใช้ประมาณ 80 จุด สำหรับการประมาณค่าคุณลักษณะแฝงในช่วง -4 ถึง +4 อย่างไรก็ตาม ในกรณีที่ไม่ได้ให้ความสำคัญกับความแม่นยำของความคลาดเคลื่อนมาตรฐาน (Standard Error) ก็อาจไม่จำเป็นต้องกำหนดจำนวนจุดกึ่งกลางของพื้นที่โค้งการแจกแจงความน่าจะเป็นมากกว่า 20 จุด เพราะการใช้จุดกึ่งกลางของพื้นที่โค้งการแจกแจงความน่าจะเป็น 10 จุด ก็สามารถประมาณค่าคุณลักษณะแฝงได้ดีสำหรับการแจกแจงแบบปกติ (De Ayala, 2009, p. 79)

มาตรวัดต่าง ๆ ที่นำโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์มาใช้ ต่างก็เลือกใช้วิธีการประมาณค่าคุณลักษณะแฝงที่แตกต่างกัน แสดงตัวอย่างได้ดังตารางที่ 2-3

ตารางที่ 2-3 วิธีการประมาณค่าคุณลักษณะแฝงของมาตรวัดต่าง ๆ ที่ใช้กับโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์

มาตรวัด	วิธีการประมาณค่าคุณลักษณะแฝง
1 Depression-CAT (Fliege et al., 2005)	EAP
2 Anxiety-CAT (Becker et al., 2008)	EAP

ตารางที่ 2-3 (ต่อ)

มาตรวัด	วิธีการประมาณค่า คุณลักษณะแฝง
3 Stress- CAT (Kocalevent et al., 2009)	EAP
4 CES-D Scale (Smits, Cuijpers, & Straten, 2011)	MAP
5 Questionnaires of Children with Psychosocial Problem (Vogels, Jacobusse, & Reijneveld, 2011)	Bayesian Approach with a Right-Skewed Log Normal Prior
6 Anhedonic Depression CAT (Smits, Zitman, Cuijpers, Hollander-Gijsman, & Carlier, 2012)	MAP
7 Short Form Health Survey (SF-36; Zheng, Chang, & Chang, 2013)	ML
8 Transmissible Liability Index (TLI; Kirisci et al., 2012)	EAP

5. เกณฑ์การยุติการทดสอบ (Termination criterion)

เกณฑ์ในการพิจารณาการยุติการทดสอบมีหลากหลายเกณฑ์ให้เลือกใช้ ดังนี้ (Thompson & Weiss, 2011; Wainer et al., 2000, p. 113)

ก) ยุติการทดสอบเมื่อข้อคำถามในคลังข้อคำถามหมด

ข) ยุติการทดสอบเมื่อการประมาณค่าคุณลักษณะแฝงมีการเปลี่ยนแปลงเพียงเล็กน้อยหรือไม่เปลี่ยนแปลงเลย

ค) ยุติการทดสอบเมื่อค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่า (Standard Error of Estimation: SE) ถึงเกณฑ์ที่ยอมรับได้ หรือความแม่นยำในการประมาณค่า (Precision/Accuracy) ถึงเกณฑ์ที่กำหนด

ง) ยุติการทดสอบเมื่อทดสอบครบจำนวนข้อคำถามตามที่กำหนด (Fixed Length)

มาตรวัดต่าง ๆ ที่นำโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์มาใช้ ต่างก็มีเกณฑ์ในการพิจารณาการยุติการทดสอบที่แตกต่างกัน แสดงตัวอย่างได้ดังตารางที่ 2-4

ตารางที่ 2-4 เกณฑ์การพิจารณาการสิ้นสุดการทดสอบของมาตรวัดต่าง ๆ ที่ใช้กับโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์

มาตรวัด	เกณฑ์ในการพิจารณา
1 Depression-CAT (Fliege et al., 2005)	$SE \leq 0.32$
2 Anxiety-CAT (Becker et al., 2008)	$SE \leq 0.32$
3 Mood and Anxiety Spectrum Scales	$SE \leq 0.30$

(MASS; Gibbons et al., 2008)

ตารางที่ 2-4 (ต่อ)

	มาตรวัด	เกณฑ์ในการพิจารณา
4	Stress- CAT (Kocalevent et al., 2009)	SE \leq 0.32
5	CES-D Scale (Smits, Cuijpers, & Straten, 2011)	SE \leq 0.30, 0.40, 0.50, 0.60, 0.70, 0.80
6	Questionnaires of Children with Psychosocial Problem (Vogels, Jacobusse, & Reijneveld, 2011)	ครบจำนวนข้อคำถามที่กำหนด และมีความแม่นยำในการประมาณค่ามากกว่าเท่ากับ 95%
7	Anhedonic Depression CAT (Smits, Zitman, Cuijpers, Hollander-Gijsman, & Carlier, 2012)	SE \leq 0.20, 0.30, 0.40, 0.50, 0.60
8	Short form Health Survey (SF-36; Zheng, Chang, & Chang, 2013)	ครบจำนวนข้อคำถามที่กำหนด
9	Transmissible Liability Index (TLI; Kirisci et al., 2012)	SE \leq 0.30 และเมื่อการประมาณค่าคุณลักษณะแฝงมีการเปลี่ยนแปลงเพียงเล็กน้อย

การประยุกต์การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ในการสร้างมาตรวัดตัวแปรทางจิตวิทยา

การศึกษาวิจัยด้านการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์นั้น เดิมทีมีขอบเขตอยู่เพียงเรื่องการวัดผลทางการศึกษาซึ่งวัดความสามารถด้านต่าง ๆ เช่น การวัดความสามารถทางคณิตศาสตร์ การวัดความสามารถด้านภาษา แต่ในช่วงหลายปีที่ผ่านมา ได้มีการประยุกต์ใช้การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ กับตัวแปรทางจิตวิทยาอื่น ๆ มากยิ่งขึ้น

Fliege et al. (2005) ได้ประยุกต์ทฤษฎีการตอบสนองข้อสอบ ร่วมกับ การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ในการประเมินภาวะซึมเศร้า โดยศึกษาจากข้อมูลการตอบแบบสอบถาม 11 ฉบับ ของผู้ป่วยโรคทางกายที่มีสาเหตุมาจากจิตใจ (Psychosomatic Disorder) จำนวน 3,270 คน คลังข้อคำถามเริ่มต้นมีข้อคำถามจำนวน 144 ข้อ ข้อคำถามทั้งหมดได้รับการตรวจสอบความสอดคล้องของข้อคำถามกับภาวะซึมเศร้าโดยผู้เชี่ยวชาญซึ่งมีการประเมินอย่างเป็นอิสระจากกัน จากนั้นข้อคำถามทุกข้อจะถูกวิเคราะห์เพื่อตรวจสอบความเป็นเอกมิติ (Unidimensionality) ความเป็นอิสระ (Local Independence) และค่าอำนาจจำแนกรายข้อ (Item Discrimination) จนกระทั่งคงเหลือข้อคำถามจำนวน 64 ข้อ สำหรับการปรับเทียบมาตรฐาน (Calibration) การประมาณค่าพารามิเตอร์ใช้โมเดล Generalized Partial Credit Model (GPCM) การประมาณค่าคุณลักษณะแฝง (ความซึมเศร้า) ใช้วิธี Expected a Posteriori (EAP) ผลการวิจัยจากการจำลองสถานการณ์การทดสอบ ปรากฏว่าค่าคุณลักษณะแฝง (ความซึมเศร้า) มีค่าอยู่ระหว่างค่าเฉลี่ยบวกลบ

สองเท่าของส่วนเบี่ยงเบนมาตรฐาน (ร้อยละ 98 ของกลุ่มตัวอย่างทั้งหมด) จำนวนข้อคำถามเฉลี่ยที่ใช้ในการทดสอบเท่ากับ 6 ข้อ กำหนดเกณฑ์ยุติการทดสอบด้วยเมื่อค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่ามีค่าเท่ากับ 0.32 หรือน้อยกว่า ซึ่งทำให้การทดสอบมีความเที่ยงเท่ากับ 0.90 หรือมากกว่า ผลการทดสอบที่ได้จากการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์กับการทดสอบด้วยข้อคำถามทั้งหมดในคลัง มีความสัมพันธ์อยู่ในระดับสูง ($r = 0.95$) นอกจากนี้ ผลการประมาณค่าที่ได้จากการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์กับผลการทดสอบที่ได้จากมาตรวัด Beck Depression Inventory ($r = 0.79$) และมาตรวัด Center of Epidemiological Studies-Depression Scale ฉบับ 8 ข้อ ($r = 0.76$) มีความสัมพันธ์ในระดับสูง จึงสรุปได้ว่า มาตรวัดความซึมเศร้ารูปแบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์มีความแม่นยำในการประมาณค่าความซึมเศร้าสูง และช่วยลดความเหนื่อยล้าในการทดสอบได้เป็นอย่างดี

Rose, Bjorner, Becker, Fries, and Ware (2006) ได้ทำวิจัยซึ่งเป็นส่วนหนึ่งในโครงการ Patient-Reported Outcomes Measurement Information System (PROMIS) ที่มีวัตถุประสงค์เพื่อพัฒนาแบบประเมินทางด้านสุขภาพให้มีความเที่ยงตรง สามารถลดความเหนื่อยล้าในการทดสอบ และมุ่งหวังที่จะนำผลการประเมินมาใช้ในตรวจสอบผลลัพธ์ด้านสุขภาพ การวิจัยนี้ นำทฤษฎีการตอบสนองข้อสอบมาเพื่อสร้างและประเมินคุณภาพของคลังข้อคำถามที่สร้างขึ้นเพื่อวัดความสามารถในการทำหน้าที่ต่าง ๆ ของร่างกาย ข้อมูลได้มาจากกลุ่มตัวอย่าง 7 กลุ่ม จำนวนทั้งสิ้น 17,726 คน ซึ่งตอบข้อคำถามจากแบบสอบถาม 9 ฉบับ รวมทั้งสิ้น 136 ข้อ วิเคราะห์ข้อคำถามด้วย Generalized Partial Credit Model โดยมีการศึกษาค่าปกติ ในประชากรชาวสหรัฐอเมริกา ซึ่งมีค่าเฉลี่ยเท่ากับ 50 และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 10 ผลการวิเคราะห์คุณภาพของข้อคำถามในคลังข้อคำถามพบว่า ข้อคำถามทั้งสิ้น 70 ข้อ จาก 136 ข้อ ผ่านเกณฑ์การพิจารณาคุณภาพตามทฤษฎีการตอบสนองข้อสอบ หลังจากนั้นใช้วิธีการจำลองสถานการณ์การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์เพื่อประเมินคุณภาพของการทดสอบ ปรากฏว่า การทดสอบใช้จำนวนข้อคำถามเฉลี่ย 10 ข้อ ค่าความคลาดเคลื่อนมาตรฐานน้อยกว่า 2.2 โดยได้ค่าคะแนนตั้งแต่ 20 ถึง 50 คะแนน และมีความเที่ยงมากกว่า 0.95 ซึ่งนับได้ว่ามีความเที่ยงหรือความแม่นยำในการประมาณค่าสูง คุณลักษณะเช่นนี้เป็นคุณสมบัติที่ไม่สามารถพบได้ในมาตรวัดที่มีการกำหนดจำนวนข้อคำถามในการทดสอบคงที่

Walter et al. (2007) ได้สร้างคลังข้อคำถาม และประยุกต์โปรแกรมการทดสอบปรับเหมาะด้วยคอมพิวเตอร์เพื่อวัดความวิตกกังวล ฉบับภาษาเยอรมัน ตามแนวคิดของทฤษฎีการตอบสนองข้อสอบ กลุ่มตัวอย่างเป็นผู้ป่วยจิตเวช จำนวน 2,348 คน กระบวนการวิจัยเริ่มต้นด้วยการสร้างคลังข้อคำถาม โดยคัดเลือกข้อคำถามจากมาตรวัดมาตรฐานที่ใช้อยู่ทั่วไปจำนวน 13 ฉบับ ซึ่งรวบรวมข้อคำถามได้ทั้งสิ้น 81 ข้อ จากนั้นจึงตรวจสอบและวิเคราะห์ข้อคำถามด้วยทฤษฎีการตอบสนองข้อสอบ โดยใช้ Generalized Partial Credit Model ปรากฏว่ามีข้อคำถามที่เหมาะสมคงเหลือในคลังข้อคำถามจำนวน 50 ข้อ การจำลองสถานการณ์ (Simulation) การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์กับกลุ่มตัวอย่างอีกกลุ่มหนึ่ง จำนวน 1,582 คน พบว่าการวัดความวิตกกังวลโดยการกำหนดให้มีความคลาดเคลื่อนมาตรฐานของการวัดน้อยกว่า 0.32 ต้องใช้ข้อคำถามในการทดสอบเฉลี่ย 6-8 ข้อ คะแนนที่ได้จากการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์กับคะแนนที่ประมาณค่าได้จากการทดสอบโดยใช้ข้อคำถามทั้งหมดในคลังและคะแนนที่ได้จากมาตรวัด State

Trait Anxiety Inventory มีความสัมพันธ์กันในระดับสูง ($r = 0.97, 0.93$ ตามลำดับ) เมื่อนำไปทดลองใช้ในโรงพยาบาล พบว่า มาตรการวัดความวิตกกังวลรูปแบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์นั้นสามารถคัดกรองผู้ที่มีความวิตกกังวลได้มีประสิทธิภาพไม่แตกต่างจากมาตรการวัดเดิมที่ใช้อยู่ ซึ่งแสดงให้เห็นถึงประโยชน์ของมาตรการวัดความวิตกกังวลรูปแบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ได้เป็นอย่างดี

Gibbons et al. (2008) ได้ประยุกต์ทฤษฎีการตอบสนองข้อสอบ กับการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับสร้างมาตรวัดตัวแปรทางจิตวิทยา เพื่อลดความเหนื่อยล้าจากการประเมินทางคลินิก และการศึกษาวิจัย โดยเก็บข้อมูลจากกลุ่มตัวอย่างจำนวน 800 คน ซึ่งเป็นผู้ป่วยนอกที่มารับบริการในเรื่องปัญหาทางอารมณ์และความวิตกกังวล โดยให้ผู้ป่วยเหล่านี้ ตอบคำถามจากมาตรวัด Mood and Anxiety Spectrum Scales (MASS) 2 ครั้ง แต่แต่ละครั้งให้ตอบคำถามจำนวนทั้งสิ้น 616 ข้อ จากข้อคำถามทั้งหมด 626 ข้อ โดยศึกษาวิจัย 2 รูปแบบ ได้แก่ 1) ศึกษาด้วยวิธีการจำลองสถานการณ์แบบโพสท์ฮอก (Post Hoc Simulation) ซึ่งเป็นการจำลองสถานการณ์จากข้อมูลจริง และ 2) การนำมาตรวัดที่พัฒนาขึ้นในรูปแบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ไปใช้ในสถานการณ์จริง ผลการศึกษา ปรากฏว่า ข้อมูลมีลักษณะเป็นสององค์ประกอบ ซึ่งประกอบด้วยองค์ประกอบหลักหนึ่งองค์ประกอบ และองค์ประกอบย่อยอีก 4 องค์ประกอบ ได้แก่ อารมณ์ (Mood) ความตื่นตระหนก (Panic-Agoraphobia) ความย้ำคิดย้ำทำ (Obsessive-Compulsive) และความกลัวสังคม (Social Phobia) การศึกษาประสิทธิภาพในการทดสอบของมาตรวัดรูปแบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ด้วยวิธีการจำลองสถานการณ์ และการนำไปใช้ในสถานการณ์จริง ปรากฏว่า ข้อคำถามที่ใช้ในการทดสอบมีจำนวนลดลงร้อยละ 95 เมื่อเทียบกับการทดสอบรูปแบบเดิม (ข้อคำถามลดลงเฉลี่ย 585 ข้อ) โดยใช้ข้อคำถามในการทดสอบเฉลี่ย 24 ข้อ เมื่อศึกษาด้วยวิธีการจำลองสถานการณ์ และเฉลี่ย 30 ข้อ เมื่อศึกษาด้วยวิธีการนำไปใช้ในสถานการณ์จริง ผลการประมาณค่าคุณลักษณะแฝงที่ได้จากมาตรวัด MASS รูปแบบเดิม และรูปแบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ มีความสัมพันธ์กันในระดับสูง ($r = 0.93$) นอกจากนี้ ได้ศึกษาเปรียบเทียบผู้ป่วยซึมเศร้าที่มี และไม่มีภาวะอารมณ์แปรปรวนร่วมด้วย ปรากฏว่า มาตรวัด MASS รูปแบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ มีความตรงในการจำแนกผู้ป่วยได้ดีกว่า ดังนั้น การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ จึงใช้ข้อคำถามในการทดสอบน้อยกว่าการทดสอบรูปแบบเดิม ขณะที่มีความตรงในการทดสอบเพิ่มขึ้น

Kocalevent et al. (2009) ได้พัฒนาและตรวจสอบประสิทธิภาพของมาตรวัดความเครียด รูปแบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ซึ่งจำแนกความเครียดออกเป็นสององค์ประกอบหลัก ได้แก่ 1) การเผชิญกับความเครียด และ 2) ปฏิกริยาตอบสนองต่อความเครียด วิเคราะห์ข้อคำถามด้วยทฤษฎีการตอบสนองข้อสอบ โดยใช้ Generalized Partial Credit Model การประเมินประสิทธิภาพทางการทดสอบจัดทำขึ้นทั้งในรูปแบบการจำลองสถานการณ์ และการนำไปใช้จริง โดยศึกษาจากผู้ป่วยโรคทางกายที่มีสาเหตุมาจากจิตใจ จำนวนทั้งสิ้น 1,092 คน (กลุ่มตัวอย่างที่ 1) และใช้กลุ่มตัวอย่างอีก 200 คน (กลุ่มตัวอย่างที่ 2) สำหรับการจำลองชุดข้อมูล จากนั้น จึงนำมาตรวัดความเครียด รูปแบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ไปใช้ในกลุ่มผู้ป่วยใน จำนวน 116 คน (กลุ่มตัวอย่างที่ 3) เพื่อตรวจสอบความตรงของเครื่องมือ ผลการศึกษาทำ

ให้ได้คลังข้อคำถาม สำหรับวัดการเผชิญกับความเครียดจำนวน 38 ข้อ และคลังข้อคำถามสำหรับวัด ปฏิกริยาตอบสนองต่อความเครียดจำนวน 31 ข้อ ในการจำลองสถานการณ์การทดสอบปรับเหมาะ ด้วยคอมพิวเตอร์ครั้งแรก ปรากฏว่า การวัดปฏิกริยาตอบสนองต่อความเครียดใช้ข้อคำถามในการทดสอบประมาณ 7.0 ± 2.3 ข้อ และการวัดการเผชิญกับความเครียดใช้ข้อคำถามในการทดสอบประมาณ 11.6 ± 1.7 ข้อ โดยกำหนดค่าความคลาดเคลื่อนมาตรฐานจากการวัดให้มีค่าน้อยกว่า 0.32 ซึ่งทำให้มีค่าความเที่ยงมากกว่า 0.90 การจำลองสถานการณ์การทดสอบครั้งที่สอง ได้วิเคราะห์ข้อมูล จากกลุ่มตัวอย่างที่ 1 อีกครั้ง พบว่า ในการวัดปฏิกริยาตอบสนองต่อความเครียดใช้ข้อคำถามในการทดสอบประมาณ 5.6 ± 2.1 ข้อ และการวัดการเผชิญกับความเครียดใช้ข้อคำถามในการทดสอบประมาณ 10.0 ± 4.9 ข้อ การศึกษาความตรง ปรากฏว่า มาตรฐานวัดความเครียด รูปแบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ มีความสัมพันธ์กับเครื่องมืออื่น ๆ ในระดับสูง ดังนั้น จึงสรุปได้ว่า มาตรฐานวัดความเครียด รูปแบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ เป็นการทดสอบที่สั้น กระชับ แต่มีความเที่ยงหรือความแม่นยำในการประมาณค่าสูง และช่วยลดความเหนื่อยล้าในการทดสอบได้เป็นอย่างดี

Lai et al. (2011) ศึกษาความตรง ความเที่ยงและประโยชน์ทางคลินิกของคลังข้อคำถาม Parent-Report Pediatric Perceived Cognitive Function (pedsPCF) ซึ่งสร้างขึ้นเพื่อวัดความ เข้าใจในเรื่องต่าง ๆ ของเด็ก ด้วยการสอบถามผ่านทางผู้ปกครอง การศึกษาครั้งนี้ ศึกษาใน ผู้ปกครองของเด็กอายุ 7-17 ปี ซึ่งผู้ปกครองแต่ละคนต้องตอบข้อคำถามจากคลังข้อคำถาม pedsPCF จำนวน 45 ข้อ การศึกษาคุณสมบัติทางการวัดใช้แนวคิดของทฤษฎีการตอบสนองข้อสอบ ใช้เส้นโค้ง ROC (Receiver Operating Characteristic Curves) และการวิเคราะห์ความสามารถในการ จำแนกเป็นตัวทำนายปัญหาด้านพฤติกรรมของเด็ก การศึกษาความสัมพันธ์ระหว่างมาตรฐานวัด pedsPCF ฉบับเดิม และมาตรฐานวัดรูปแบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ปรากฏว่า มาตรฐานวัดทั้งสองมีความสัมพันธ์กันในระดับสูง ($r = 0.98$) จึงสรุปได้ว่า คลังข้อคำถาม pedsPCF มี คุณสมบัติในการวัดที่ดี และมีความพร้อมสำหรับการนำไปใช้ในการทดสอบแบบปรับเหมาะด้วย คอมพิวเตอร์ ในกลุ่มประชากรชาวสหรัฐอเมริกา

Simms et al. (2011) กล่าวว่า ภาวะบุคลิกภาพผิดปกติ เป็นภาวะที่มีความสำคัญ และ ส่งผลต่อพยาธิสภาพทางจิตใจอื่น ๆ เป็นจำนวนมาก ปัญหาสำคัญที่พบ คือ ความยากลำบากใน การประเมินภาวะดังกล่าว จึงได้จัดทำโครงการพัฒนาแบบประเมินภาวะบุคลิกภาพผิดปกติ รูปแบบ การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ (CAT-PD Project) โดยมีวัตถุประสงค์เพื่อพัฒนาโมเดล ภาวะบุคลิกภาพผิดปกติที่มีประสิทธิภาพ และสร้างเครื่องมือที่มีคุณภาพสำหรับวัดคุณลักษณะ ดังกล่าว ซึ่งจากการศึกษาทบทวนคาดว่าข้อคำถามที่ใช้ในการวัดภาวะบุคลิกภาพผิดปกติ นั้น อาจมี จำนวนมาก จึงได้เลือกใช้กระบวนการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์มาใช้เป็นวิธีการใน การดำเนินการทดสอบ เนื่องจากวิธีการนี้ มีข้อดีเหนือกว่าการทดสอบรูปแบบกระดาษ-ดินสอ และ การสัมภาษณ์ โครงการวิจัยนี้แบ่งออกเป็น 5 ระยะ โดยผลการศึกษาที่ได้จากระยะที่ 1 แสดงให้เห็น ว่าลักษณะบุคลิกภาพที่ได้จากกลุ่มตัวอย่างทั้ง 59 ลักษณะ สามารถจัดเป็นกลุ่มใหญ่ได้ทั้งหมด 5 กลุ่ม ได้แก่ อารมณ์ทางลบ (Negative emotionality) อารมณ์ทางบวก (Positive emotionality)

การต่อต้าน (Antagonism) การยับยั้งซึ่งใจ (Constraint) และความแปลกประหลาด (Oddity) ลักษณะดังกล่าวจะนำไปใช้ในการพัฒนามาตรวัดรูปแบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ต่อไป

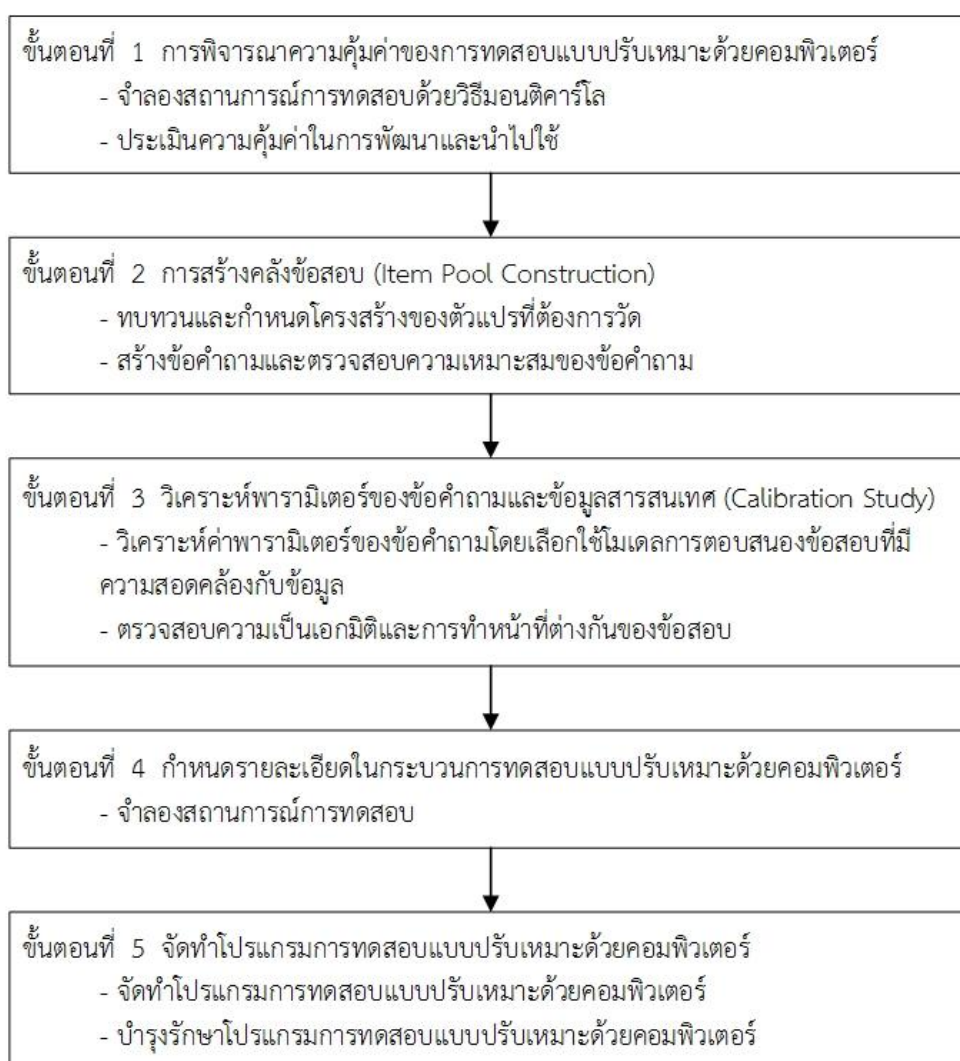
Smits, Cuijpers, and Straten (2011) ได้พัฒนามาตรวัด Center of Epidemiological Studies-Depression Scale (CES-D) รูปแบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ซึ่งดำเนินการศึกษาวิจัยในกลุ่มตัวอย่างที่เป็นวัยรุ่น จำนวนทั้งสิ้น 1,392 คน จากผลการศึกษา ปรากฏว่า ข้อคำถามในมาตรวัด CES-D มีคุณลักษณะเหมาะสมกับการพัฒนาเป็นรูปแบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ การตรวจสอบประสิทธิภาพในการทดสอบด้วยวิธีการจำลองสถานการณ์ จากข้อมูลจริงที่มีอยู่ พบว่า การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สามารถวัดภาวะซึมเศร้าได้ โดยใช้จำนวนข้อคำถามน้อยกว่าการทดสอบรูปแบบเดิม ขณะที่ความตรงที่ได้จากการวัดไม่ต่างจากการวัดรูปแบบเดิม สรุปได้ว่ามาตรวัด CES-D รูปแบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ช่วยเพิ่มประสิทธิภาพในการวัดของมาตรวัด CES-D ได้เป็นอย่างดี

Smits, Zitman, Cuijpers, Hollander-Gijsman, and Carlier (2012) ได้ศึกษาประสิทธิภาพในการทดสอบของมาตรวัด Mood and Anxiety Symptom Questionnaire - Anhedonic Depression Scale (MASQ-AD) รูปแบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ซึ่งเป็นมาตรวัดที่ใช้วัดภาวะซึมเศร้า โดยศึกษาด้วยวิธีการจำลองสถานการณ์ ในกลุ่มตัวอย่างที่เป็นผู้ป่วยจำนวน 3,597 คน จากการวิเคราะห์ข้อคำถาม ปรากฏว่า ข้อคำถามมีคุณภาพและสามารถบรรจุไว้ในคลังข้อคำถาม ของการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ได้ โดยในการจำลองสถานการณ์ ได้กำหนดเกณฑ์การยุติการทดสอบจากค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่าซึ่งกำหนดไว้หลายระดับ ผลการทดสอบ ปรากฏว่า การวัดภาวะซึมเศร้าด้วยการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ใช้ข้อคำถามในการทดสอบลดลง เมื่อเทียบกับวิธีการวัดรูปแบบเดิม ขณะที่ความตรงในการทดสอบไม่แตกต่างจากเดิม

จากการทบทวนงานวิจัยที่เกี่ยวข้องซึ่งการพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดทางด้านจิตใจ ปรากฏว่า การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์มีประสิทธิภาพในการทดสอบที่ดี เนื่องจากกระบวนการทดสอบสามารถกำหนดค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่าเป็นเกณฑ์ในการยุติการทดสอบได้ ซึ่งค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่ามีความสัมพันธ์โดยตรงกับความแม่นยำในการทดสอบ หากค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่ามีค่าต่ำ การทดสอบจะมีความแม่นยำสูง ดังนั้น ผู้วิจัยจึงสามารถกำหนดได้ว่าต้องการให้มีความแม่นยำในการทดสอบมากน้อยเพียงใด นอกจากนี้ โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สามารถประมาณค่าคุณลักษณะแฝงได้โดยอาศัยทฤษฎีการตอบสนองข้อสอบ ซึ่งเป็นการประมาณค่าคุณลักษณะแฝงที่ไม่อาศัยผลรวมคะแนนที่ได้จากข้อคำถามในมาตรวัดทั้งหมด ผู้ทดสอบแต่ละคนจึงไม่ต้องทดสอบด้วยข้อคำถามชุดเดียวกันและไม่ต้องทดสอบด้วยข้อคำถามทั้งหมดในคลัง แต่ผู้ทดสอบจะได้รับข้อคำถามต่างชุดที่มีความเหมาะสมกับคุณลักษณะแฝงของแต่ละบุคคล ส่งผลให้ข้อคำถามที่ใช้ในการทดสอบมีจำนวนลดลงจึงช่วยลดความเหนื่อยล้าในการทดสอบได้เป็นอย่างดี การศึกษาความสัมพันธ์ระหว่างผลการทดสอบที่ได้จากการทดสอบด้วยโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ กับผล

การทดสอบที่ได้จากมาตรวัด รูปแบบกระดาษ-ดินสอที่มีความเกี่ยวข้องกัน ปรากฏว่า ส่วนใหญ่มีความสัมพันธ์กันในระดับสูง แสดงให้เห็นว่านอกจากโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์จะมีประสิทธิภาพในการทดสอบที่ดีแล้วยังมีความตรงในการวัดที่ดีอีกด้วย

การพัฒนามาตรวัดตัวแปรทางจิตวิทยาในรูปแบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ เป็นโครงการใหญ่ที่ต้องอาศัยการลงทุนทั้งความคิด ค่าใช้จ่ายและเวลา วิธีการศึกษาที่พบ จึงนิยมใช้การจำลองสถานการณ์การทดสอบในช่วงต้นก่อน เพราะช่วยประหยัดค่าใช้จ่าย และสามารถประเมินความคุ้มค่า ก่อนการนำไปใช้ทดสอบในสถานการณ์จริงได้ ขั้นตอนในการพัฒนามาตรวัดตัวแปรทางจิตวิทยาในรูปแบบการทดสอบปรับเหมาะด้วยคอมพิวเตอร์นั้น สามารถสรุปได้ดังภาพที่ 2-11 (Walker, Böhnke, Cerny, & Strasser, 2010; Thompson & Weiss, 2011)



ภาพที่ 2-11 ขั้นตอนในการพัฒนามาตรวัดตัวแปรทางจิตวิทยาในรูปแบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ (ดัดแปลงจาก Walker, Böhnke, Cerny, & Strasser, 2010; Thompson & Weiss, 2011)

จากภาพที่ 2-11 มีรายละเอียด ดังนี้

1. การพิจารณาความคุ้มค่าของการทดสอบปรับเหมาะด้วยคอมพิวเตอร์

โดยทั่วไปสามารถพิจารณาถึงจุดเด่นและจุดอ่อนของการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ได้จากการทบทวนงานวิจัยที่เกี่ยวข้อง แต่ข้อมูลดังกล่าวอาจไม่ชัดเจนเพียงพอที่จะตัดสินใจได้อย่างเด็ดขาด ดังนั้น จึงควรพิจารณาประเด็นที่เกี่ยวข้องอื่น ๆ ร่วมด้วย ดังนี้

- 1.1 ความเชี่ยวชาญในศาสตร์ด้านการวัด ของนักวิจัยหรือบุคลากรในองค์กร
- 1.2 แหล่งความรู้หรือผู้เชี่ยวชาญ สำหรับการให้คำปรึกษาและแนะนำ
- 1.3 ความรู้และความสามารถในการพัฒนาคลังข้อคำถาม
- 1.4 เครื่องมือที่มีความจำเป็น เช่น ซอฟต์แวร์ เครื่องคอมพิวเตอร์
- 1.5 ประโยชน์ของการนำไปใช้ เช่น สามารถลดจำนวนข้อคำถามและระยะเวลาที่ใช้ในการทดสอบ

อย่างไรก็ดี การตัดสินใจดังกล่าวอาจทำได้ยาก สำหรับผู้ที่ไม่มีประสบการณ์ในการพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์มาก่อน ดังนั้น ในขั้นตอนนี้อาจใช้การจำลองสถานการณ์การทดสอบ ในรูปแบบมอนติคาร์โล (Monte Carlo simulation) ประกอบการพิจารณาได้ โดยการจำลองสถานการณ์การทดสอบสามารถพิจารณาจำนวนข้อคำถามที่เหมาะสมในการทดสอบ ความเที่ยงหรือความแม่นยำในการประมาณค่าที่ควรได้จากการทดสอบ ขนาดของคลังข้อคำถามที่สามารถทำให้การประมาณค่าคุณลักษณะแฝงได้อย่างแม่นยำตามที่ต้องการ รวมถึงแนวทางการควบคุมอัตราการใช้ข้อคำถาม และการควบคุมสัดส่วนเนื้อหาของข้อคำถาม การจำลองสถานการณ์นี้ดำเนินการโดยสร้างสถานการณ์การทดสอบหลากหลายรูปแบบ และเปรียบเทียบประสิทธิภาพการทดสอบแต่ละรูปแบบ ซึ่งช่วยให้ผู้วิจัยสามารถตัดสินใจได้ง่ายขึ้น การจำลองสถานการณ์อาจทำได้โดยการพัฒนาซอฟต์แวร์ขึ้นมาใหม่ หรือเลือกใช้ซอฟต์แวร์ที่จัดจำหน่ายทั่วไปก็ได้

ซอฟต์แวร์ที่จำเป็นต้องใช้ในการจำลองสถานการณ์การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์แบบวิธีมอนติคาร์โล มีสองส่วน คือ 1) ซอฟต์แวร์ที่ใช้สำหรับจำลองข้อมูล เช่น โปรแกรม WINGEN (Han, 2007) โปรแกรม PARDSIM (Yoes, 1997) ซึ่งจำลองข้อมูลบนพื้นฐานทฤษฎีการตอบสนองข้อสอบ และ 2) ซอฟต์แวร์ที่ใช้สำหรับจำลองสถานการณ์การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ เช่น โปรแกรม FireStar (Choi, 2009) โปรแกรม CATSim (Weiss & Guyer, 2012)

2. การสร้างคลังข้อคำถาม (Item Bank Construction)

2.1 ทบทวนและกำหนดโครงสร้างของตัวแปรที่ต้องการวัด โดยศึกษาทฤษฎี และนิยามศัพท์ที่เกี่ยวข้องกับตัวแปรที่ต้องการวัด จากนั้นจึงกำหนดนิยามเชิงปฏิบัติการของตัวแปรนั้น เพื่อให้สามารถนำมาสร้างเป็นข้อคำถามได้

2.2 สร้างข้อคำถาม และตรวจสอบความเหมาะสมของข้อคำถาม การสร้างข้อคำถามในคลังนั้น สามารถดำเนินการได้ 2 แนวทาง ได้แก่

- 2.2.1 การเลือกข้อคำถามจากมาตรวัดที่มีอยู่แล้ว
- 2.2.2 การเขียนข้อคำถามขึ้นใหม่ตามนิยามเชิงปฏิบัติการ

อย่างไรก็ดี ไม่ว่าข้อความนั้นเป็นข้อความที่สร้างขึ้นใหม่หรือเป็นข้อความที่ได้มาจากมาตรวัดที่มีอยู่เดิม สิ่งสำคัญที่ต้องพิจารณา คือ ความเหมาะสมของข้อความ ทั้งในด้านเนื้อหา รูปแบบการตอบ รูปแบบของมาตรวัด ระบบการให้คะแนน และการแปลความหมาย ในส่วนนี้จึงเป็นส่วนที่ต้องตรวจสอบโดยผู้เชี่ยวชาญ

3. การเปรียบเทียบมาตรฐาน (Calibration Study)

การดำเนินการในส่วนนี้เกี่ยวข้องกับทฤษฎีการวัด ซึ่งทั้งทฤษฎีการทดสอบแบบดั้งเดิมและทฤษฎีการตอบสนองข้อสอบสามารถนำมาใช้ในวิเคราะห์ข้อความได้ การวิเคราะห์ข้อความด้วยทฤษฎีการทดสอบแบบดั้งเดิมมีข้อจำกัด เนื่องจากค่าที่ได้จากการวิเคราะห์ข้อความ สามารถแปรเปลี่ยนได้ตามกลุ่มผู้สอบ และค่าคุณลักษณะแฝงสามารถแปรเปลี่ยนได้ตามมาตรวัดที่ใช้ ขณะที่ทฤษฎีการตอบสนองข้อสอบมีจุดเด่น คือ ความไม่แปรเปลี่ยนตามกลุ่มผู้สอบและมาตรวัดที่ใช้ ดังนั้นค่าที่วิเคราะห์ได้จึงมีความคงที่มากกว่า ปัจจุบันการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ จึงนิยมวิเคราะห์ข้อความด้วยทฤษฎีการตอบสนองข้อสอบ

การวิเคราะห์ค่าพารามิเตอร์ข้อสอบ จะต้องตรวจสอบความสอดคล้องของข้อมูลที่ได้จากมาตรวัดกับโมเดลการตอบสนองข้อสอบ ซึ่งมีทั้งรูปแบบโมเดลการตอบสนองข้อสอบแบบเอกมิติและพหุมิติ รวมถึงรูปแบบการตรวจให้คะแนนที่แตกต่างกันซึ่งมีทั้งการตรวจให้คะแนนแบบสองค่าและการตรวจให้คะแนนแบบมากกว่าสองค่า นอกจากนี้ยังมีประเด็นอื่น ๆ ที่ต้องพิจารณาด้วย เช่น ข้อตกลงเบื้องต้นของโมเดลการตอบสนองข้อสอบ

4. กำหนดรายละเอียดในกระบวนการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์

การกำหนดรายละเอียดในกระบวนการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สามารถพิจารณาคัดเลือกเกณฑ์ที่เหมาะสมจากการจำลองสถานการณ์การทดสอบ การจำลองสถานการณ์การทดสอบ แบ่งเป็นสองรูปแบบ ได้แก่ การจำลองสถานการณ์การทดสอบแบบโพสต์ฮอก (Post-hoc Simulations) ซึ่งเป็นการจำลองสถานการณ์จากข้อมูลจริง และการจำลองสถานการณ์การทดสอบแบบไฮบริด (Hybrid Simulations) ซึ่งเป็นการจำลองสถานการณ์โดยใช้ทั้งข้อมูลจริงและข้อมูลที่สร้างขึ้นจากคอมพิวเตอร์เพื่อให้การดำเนินการทดสอบเป็นไปด้วยความราบรื่น การจำลองสถานการณ์การทดสอบทั้งสองรูปแบบ จัดทำขึ้นเพื่อใช้ในการเปรียบเทียบและประเมินวิธีการต่าง ๆ ที่ใช้ในการดำเนินการทดสอบปรับเหมาะด้วยคอมพิวเตอร์ เพื่อค้นหารูปแบบการดำเนินการทดสอบที่เหมาะสมที่สุด โดยมีประเด็นที่ต้องพิจารณา คือ เกณฑ์การเริ่มต้นการทดสอบ วิธีการคัดเลือกข้อความข้อถัดไป วิธีการประมาณค่าคุณลักษณะ เกณฑ์การยุติการทดสอบ การควบคุมอัตราการใช้ข้อความและการควบคุมสัดส่วนเนื้อหาของข้อความ การดำเนินการในขั้นตอนนี้สามารถใช้โปรแกรม CATSim (Weiss & Guyer, 2012) ในการจำลองสถานการณ์การทดสอบได้

5. จัดทำโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์

ขั้นตอนนี้เป็นการจัดทำโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ตามรูปแบบและวิธีการที่ได้ศึกษาไว้จาก 4 ขั้นตอนก่อนหน้า รวมถึง เป็นขั้นตอนการตรวจสอบประสิทธิภาพการทดสอบ ซึ่งหากเป็นไปได้ควรนำไปศึกษาในสถานการณ์จริง นอกจากนี้ การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์นั้น จะต้องมีการดูแลและปรับปรุงโปรแกรมเป็นระยะ อาจมีการเพิ่มเติมปรับเปลี่ยนข้อความต่าง ๆ เข้าไปในคลังข้อความด้วย

การทบทวนเอกสารและงานวิจัยที่เกี่ยวข้องกับทฤษฎีของความสุข ทฤษฎีการตอบสนองข้อสอบ และการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ปรากฏว่า มาตรฐานความสุขที่สร้างขึ้นทั้งในประเทศไทยและต่างประเทศ ส่วนใหญ่แตกต่างกันที่นิยามหรือเนื้อหาที่ใช้ในการวัดความสุข และรูปแบบการบริหารการทดสอบ (เช่น การทดสอบด้วยกระดาษ-ดินสอ การทดสอบในลักษณะการทดสอบออนไลน์ การทดสอบในลักษณะของแอปพลิเคชันในโทรศัพท์มือถือ) อย่างไรก็ตาม มาตรฐานเหล่านี้ ล้วนสร้างขึ้นจากพื้นฐานทฤษฎีการทดสอบแบบดั้งเดิม ที่ใช้วิธีการรวมคะแนนจากข้อคำถามทุกข้อในมาตรวัดนั้น ๆ เพื่ออธิบายคุณลักษณะที่ต้องการวัด ทำให้ผู้ทดสอบต้องตอบข้อคำถามทุกข้อ ทั้งที่ข้อคำถามบางข้ออาจไม่จำเป็นสำหรับผู้ทดสอบ นอกจากนี้ ค่าทางสถิติของข้อคำถาม (ค่าพารามิเตอร์ของข้อคำถาม) มีความแปรเปลี่ยนไปตามกลุ่มประชากร ขณะที่ผลการทดสอบที่ได้ (ค่าพารามิเตอร์ของบุคคล) มีความแปรเปลี่ยนไปตามมาตรวัดที่ใช้ แนวโน้มในการพัฒนามาตรวัดความสุข จึงไม่ควรจำกัดอยู่เพียงการศึกษากรอบแนวคิดที่นำมาใช้ในการสร้างมาตรวัด แต่ควรขยายขอบเขต และพัฒนามาตรวัดในเชิงทฤษฎีการวัด เช่น การขยายแนวคิดในการวิเคราะห์คุณภาพของข้อคำถามจากทฤษฎีการทดสอบแบบดั้งเดิม เป็นทฤษฎีการตอบสนองข้อสอบเนื่องจากสามารถช่วยแก้ปัญหาข้อจำกัดต่าง ๆ ที่พบ นอกจากนี้ ควรมีการนำความเจริญก้าวหน้าทางเทคโนโลยีที่เกิดขึ้นตามยุคสมัยมาพัฒนารูปแบบการบริหารการทดสอบที่ไม่จำกัดอยู่เพียงรูปแบบดั้งเดิมซึ่งเป็นรูปแบบกระดาษ-ดินสอ แต่ควรปรับรูปแบบการบริหารการทดสอบให้เหมาะสมและมีประสิทธิภาพในการทดสอบเพิ่มมากขึ้น เช่น การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ที่ช่วยให้การทดสอบมีความแม่นยำสูง ขณะที่สามารถลดความเหนื่อยล้าจากการทดสอบได้เป็นอย่างดี

บทที่ 3

วิธีดำเนินการวิจัย

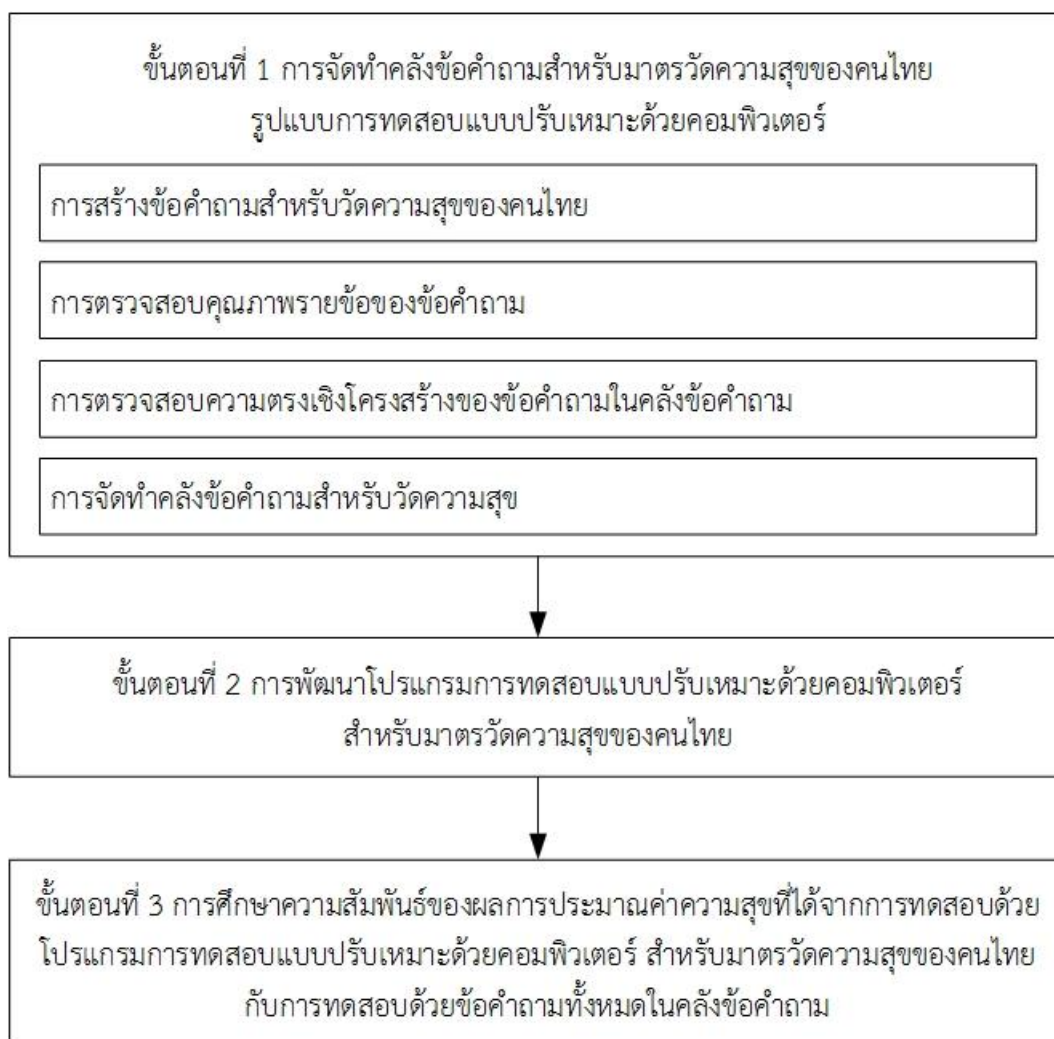
การประยุกต์โมเดลการตอบสนองข้อสอบ ในการพัฒนามาตรวัดความสุขของคนไทย รูปแบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ มีวัตถุประสงค์เพื่อสร้างคลังข้อคำถามที่มีคุณภาพ ด้วยการวิเคราะห์คุณภาพรายข้อของข้อคำถามตามทฤษฎีการตอบสนองข้อสอบและตรวจสอบความตรงเชิงโครงสร้างของข้อคำถามในคลังข้อคำถาม นอกจากนี้ ได้พัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับมาตรวัดความสุขของคนไทยตามแนวคิดของวงจรการพัฒนาระบบ และศึกษาความสัมพันธ์ของผลการประมาณค่าความสุขที่ได้จากการทดสอบด้วยโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย กับการทดสอบด้วยข้อคำถามทั้งหมดในคลังข้อคำถาม การวิจัยนี้แบ่งวิธีดำเนินการวิจัยเป็น 3 ขั้นตอนหลัก ดังนี้

ขั้นตอนที่ 1 การจัดทำคลังข้อคำถามสำหรับมาตรวัดความสุขของคนไทย รูปแบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์

ขั้นตอนที่ 2 การพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย

ขั้นตอนที่ 3 การศึกษาความสัมพันธ์ของผลการประมาณค่าความสุขที่ได้จากการทดสอบด้วยโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย กับการทดสอบด้วยข้อคำถามทั้งหมดในคลังข้อคำถาม

ขั้นตอนหลักของการวิจัยสามารถสรุปได้ ดังภาพที่ 3-1



ภาพที่ 3-1 สรุปขั้นตอนหลักของการดำเนินการวิจัย

ขั้นตอนที่ 1 การจัดทำคลังข้อคำถามสำหรับมาตรวัดความสุขของคนไทย รูปแบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์

การสร้างข้อคำถามสำหรับวัดความสุขของคนไทย

กระบวนการสร้างข้อคำถาม เริ่มต้นด้วยการกำหนดจุดมุ่งหมายของการสร้างมาตรวัด และโครงสร้างคุณลักษณะของความสุข กำหนดนิยามเชิงปฏิบัติการ กำหนดรูปแบบของมาตรวัด จากนั้นจึงเขียนข้อคำถามสำหรับบรรจุในคลังข้อคำถามเริ่มต้น (Initial Item Bank) ซึ่งการกำหนดนิยามเชิงปฏิบัติการของความสุข ใช้แนวคิดของทฤษฎีความอยู่ดีมีสุขเชิงอัตวิสัย (Pavot & Diener, 2013) ที่ระบุว่า ความสุขเป็นกระบวนการในการตัดสินใจและประเมินคุณภาพชีวิตของบุคคลตามมุมมองส่วนตัว แต่ละคนอาจมีเกณฑ์ในการประเมินที่แตกต่างกัน การระบุถึงคุณภาพชีวิตของแต่ละบุคคล จึงขึ้นอยู่กับ การตัดสินใจของบุคคลนั่นเอง ความสุขประกอบด้วยสามองค์ประกอบย่อยที่แยกออกจากกัน ได้แก่ ความพึงพอใจในชีวิต อารมณ์ความรู้สึกเชิงบวก และอารมณ์ความรู้สึกเชิงลบ คนที่มีความสุข

คือ คนที่มีความพึงพอใจในชีวิตของตน เปรียบเทียบกับสิ่งต่าง ๆ ด้วยอารมณ์ความรู้สึกเชิงบวกบ่อยครั้ง ขณะที่เผชิญกับสิ่งต่าง ๆ ด้วยอารมณ์เชิงลบน้อยครั้ง หรือแทบไม่มีเลย (Diener, 1984; Myers & Diener, 1995; Diener, Suh, Lucas, & Smith, 1999; Diener, 2000; Pavot & Diener, 2013) การสร้างข้อคำถามในแต่ละองค์ประกอบย่อย มีแนวทางดังนี้

1. ความพึงพอใจในชีวิต

การวัดความพึงพอใจในชีวิต เป็นการประเมินชีวิตตนเองในภาพรวม ไม่เจาะจงที่ด้านใดด้านหนึ่งโดยเฉพาะ ซึ่งมุมมองของความพึงพอใจในชีวิต แบ่งออกเป็น 5 มุมมอง (Diener, Suh, Lucas, & Smith, 1999) ได้แก่

- 1.1 ความต้องการเปลี่ยนแปลงชีวิตของตน
- 1.2 ความพึงพอใจในชีวิตปัจจุบัน
- 1.3 ความพึงพอใจในชีวิตอดีต
- 1.4 ความพึงพอใจในชีวิตอนาคต
- 1.5 ความคิดเห็นจากคนสำคัญที่มีต่อชีวิตของตัวเอง

นิยามดังกล่าวได้นำมาใช้เป็นแนวทางในการวิเคราะห์เนื้อหา และสร้างข้อคำถามที่เกี่ยวข้อง โดยจัดทำในลักษณะแบบรายงานตนเอง (Self-Report) ซึ่งมีตัวเลือกรายการคำตอบแบบมาตรฐานค่า 5 ระดับ (Rating Scale) เพื่อใช้สำหรับการประเมินชีวิตโดยรวมของผู้ทดสอบว่า ข้อความที่ระบุตรงกับความรู้สึกของผู้ทดสอบมากน้อยเพียงใด โดยมีเกณฑ์การให้คะแนนในแต่ละตัวเลือก ดังนี้

ไม่เห็นด้วย/เห็นด้วยน้อยที่สุด	ให้คะแนน	1
เห็นด้วยน้อย	ให้คะแนน	2
เห็นด้วยปานกลาง	ให้คะแนน	3
เห็นด้วยมาก	ให้คะแนน	4
เห็นด้วยอย่างยิ่ง	ให้คะแนน	5

2. อารมณ์ความรู้สึกเชิงบวก และอารมณ์ความรู้สึกเชิงลบ

อารมณ์ความรู้สึกเชิงบวก หมายถึง ประสบการณ์ที่แตกต่างกันของบุคคลในการมีปฏิสัมพันธ์กับบุคคล หรือสิ่งรอบข้าง ด้วยอารมณ์ความรู้สึกที่ดี ซึ่งทำให้บุคคลเกิดความพึงพอใจ ในขณะที่อารมณ์ความรู้สึกเชิงลบ หมายถึง ประสบการณ์ที่แตกต่างกันของบุคคลในการมีปฏิสัมพันธ์กับบุคคล หรือสิ่งรอบข้างด้วยอารมณ์ความรู้สึกที่ไม่ดี ซึ่งทำให้บุคคลเกิดความรู้สึกไม่พึงพอใจ และมีภาพพจน์ต่อตนเองในเชิงลบ

การสร้างมาตรวัดองค์ประกอบด้านอารมณ์ความรู้สึกเชิงบวก และอารมณ์ความรู้สึกเชิงลบ เพื่อสะท้อนถึงความสุข เน้นที่ความถี่ของการเกิดอารมณ์ความรู้สึก มากกว่ามุ่งเน้นที่ความเข้มข้น (Intensity) ของอารมณ์ความรู้สึกว่า มีมากน้อยเพียงใด เพราะความเข้มข้นของอารมณ์ความรู้สึกสามารถเปลี่ยนแปลงได้ง่าย และมีความไม่คงที่ ดังนั้น ความเข้มข้นของระดับอารมณ์ความรู้สึกที่ใช้ในการวัด จึงเน้นเพียงให้อารมณ์ความรู้สึกดังกล่าวมีความเข้มข้นเพียงพอ ที่ผู้ทดสอบจะรับรู้ได้ว่ามีอารมณ์ดังกล่าวเกิดขึ้น ข้อคำถามที่ใช้ในมาตรวัดจึงเน้นการวัดความถี่ของอารมณ์ความรู้สึก เพื่อประเมินว่า บุคคลมีประสบการณ์ทางอารมณ์ความรู้สึกต่าง ๆ เกิดขึ้น บ่อยครั้งเพียงใด

การประเมินที่เน้นความถี่ของการเกิดอารมณ์ความรู้สึก จะทำให้ผลที่วัดได้ มีความถูกต้อง (Accuracy) และความเที่ยงตรง (Validity) มากกว่าการวัดที่เน้นความเข้มข้นของการเกิดอารมณ์ความรู้สึก นอกจากนี้ อารมณ์ความรู้สึกเชิงบวกและอารมณ์ความรู้สึกเชิงลบนั้น มีความเป็นอิสระจากกันอย่างชัดเจน รวมทั้ง ไม่ได้เป็นคุณลักษณะที่ตรงข้ามกัน นั้นหมายความว่า การมีอารมณ์ความรู้สึกเชิงบวก ไม่ได้รับประกันว่าจะไม่มีอารมณ์ความรู้สึกเชิงลบ และการมีอารมณ์ความรู้สึกเชิงลบก็ไม่ได้บ่งชี้ว่า จะไม่มีอารมณ์ความรู้สึกเชิงบวก ดังนั้น การประเมินอารมณ์ความรู้สึกทั้งสอง จึงต้องแยกออกจากกัน (Diener & Emmons, 1985; Watson, Clark, & Tellegen, 1988; Diener, Sandvik, & Pavot, 1991, pp. 119-139; Naragon & Watson, 2013, p. 707)

อารมณ์ความรู้สึกที่เกิดขึ้นมีได้หลากหลายแบบ ในต่างประเทศมีคลังคำ และการแบ่งหมวดหมู่ของอารมณ์ความรู้สึกตามเกณฑ์ต่าง ๆ ซึ่งสามารถยึดเป็นแนวทางในการสร้างข้อคำถามได้โดยง่าย แต่เมื่อนำมาแปลเป็นภาษาไทยปรากฏว่า ถ้อยคำในภาษาต่างประเทศสามารถแปลเป็นภาษาไทยได้หลากหลายความหมาย ทำให้บางครั้งมีความซ้ำซ้อนกันของคำที่แสดงอารมณ์ ในขณะที่คลังคำแสดงอารมณ์ความรู้สึกในภาษาไทย ยังมีบางส่วนที่ไม่ครอบคลุมอารมณ์ความรู้สึกที่เป็นไปได้ทั้งหมด ดังนั้น การสร้างมาตรวัดจึงใช้วิธีการรวบรวมถ้อยคำแสดงอารมณ์ความรู้สึกทั้งในประเทศและต่างประเทศ เพื่อเป็นแนวทางในการสร้างมาตรวัด และให้ครอบคลุมอารมณ์ความรู้สึกที่เป็นไปได้ทั้งหมด (นรารธรรม พันธุมธา, 2549, หน้า 31-74; Watson, Clark, & Tellegen, 1988; Diener et al., 2009; Wikipedia, 2014)

ข้อคำถามเกี่ยวกับประสบการณ์ของผู้ทดสอบทางด้านอารมณ์ความรู้สึกเชิงบวกและอารมณ์ความรู้สึกเชิงลบ สร้างขึ้นในลักษณะมาตราประมาณค่า 5 ระดับ ซึ่งวัดระดับความถี่ในการเกิดประสบการณ์ทางอารมณ์ความรู้สึกเชิงบวก และอารมณ์ความรู้สึกเชิงลบ ในช่วง 1 เดือนที่ผ่านมา

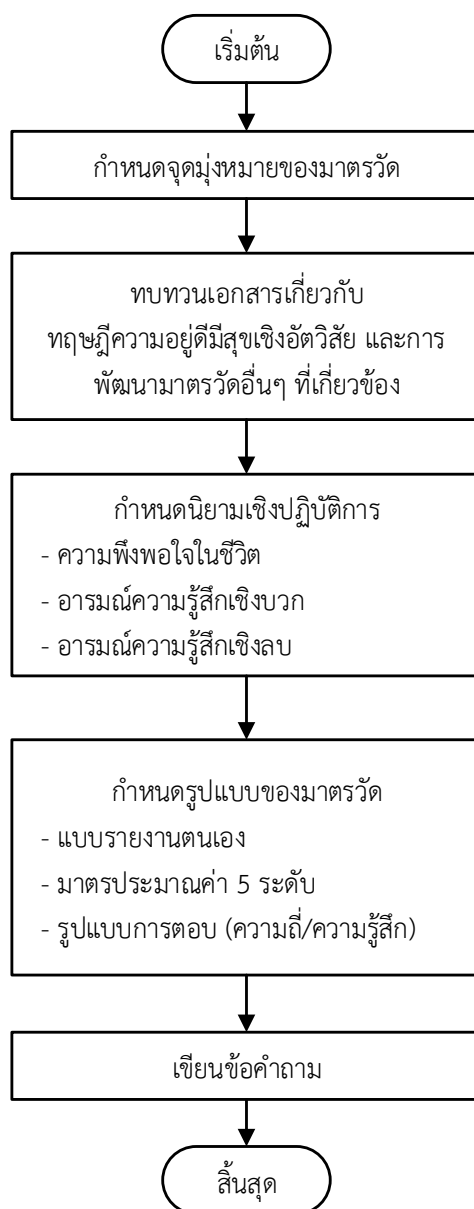
โดยมีเกณฑ์การให้คะแนนขององค์ประกอบด้านอารมณ์ความรู้สึกเชิงบวก ดังนี้

ไม่เคยหรือแทบจะไม่เคยเกิดอารมณ์ความรู้สึกดังกล่าว	ให้คะแนน	1
เกิดอารมณ์ความรู้สึกดังกล่าวน้อยครั้ง	ให้คะแนน	2
เกิดอารมณ์ความรู้สึกดังกล่าวบางครั้ง	ให้คะแนน	3
เกิดอารมณ์ความรู้สึกดังกล่าวบ่อยครั้ง	ให้คะแนน	4
เกิดอารมณ์ความรู้สึกดังกล่าวเป็นประจำ	ให้คะแนน	5

และมีเกณฑ์การให้คะแนนขององค์ประกอบด้านอารมณ์ความรู้สึกเชิงลบ ดังนี้

ไม่เคยหรือแทบจะไม่เคยเกิดอารมณ์ความรู้สึกดังกล่าว	ให้คะแนน	5
เกิดอารมณ์ความรู้สึกดังกล่าวน้อยครั้ง	ให้คะแนน	4
เกิดอารมณ์ความรู้สึกดังกล่าวบางครั้ง	ให้คะแนน	3
เกิดอารมณ์ความรู้สึกดังกล่าวบ่อยครั้ง	ให้คะแนน	2
เกิดอารมณ์ความรู้สึกดังกล่าวเป็นประจำ	ให้คะแนน	1

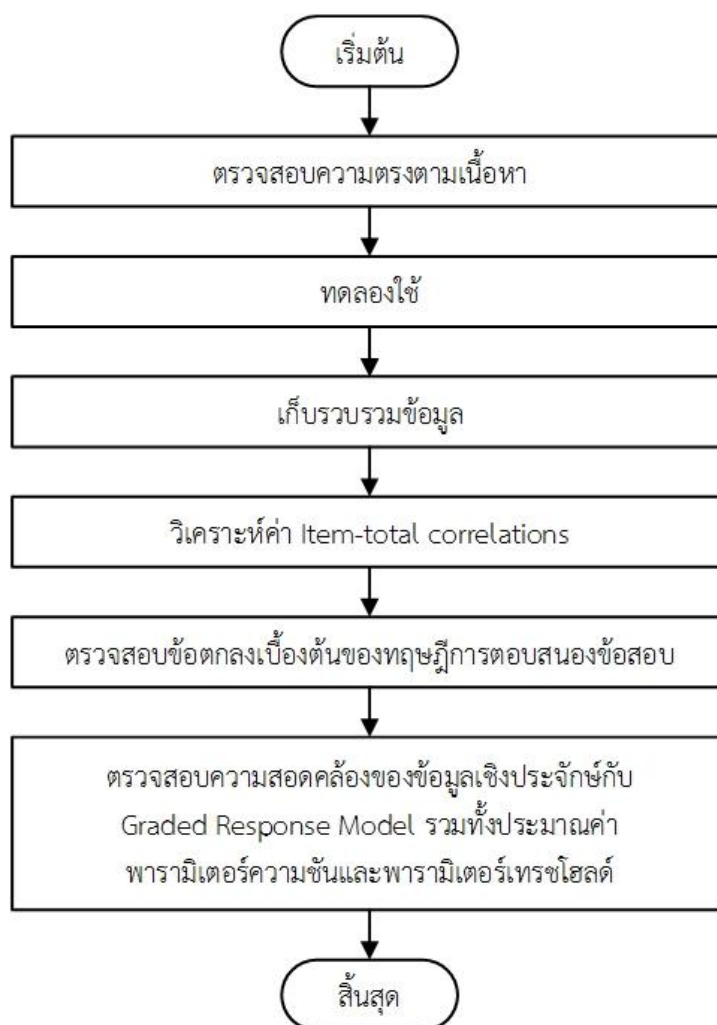
ขั้นตอนการสร้างข้อคำถาม สามารถสรุปเป็นแผนภาพได้ดังภาพที่ 3-2



ภาพที่ 3-2 ขั้นตอนการสร้างข้อคำถามสำหรับวัดความสุขของคนไทย

การตรวจสอบคุณภาพรายข้อของข้อคำถาม

การตรวจสอบคุณภาพรายข้อของข้อคำถามประกอบด้วยขั้นตอนย่อยหลายขั้นตอนซึ่งสามารถสรุปได้ ดังภาพที่ 3-3



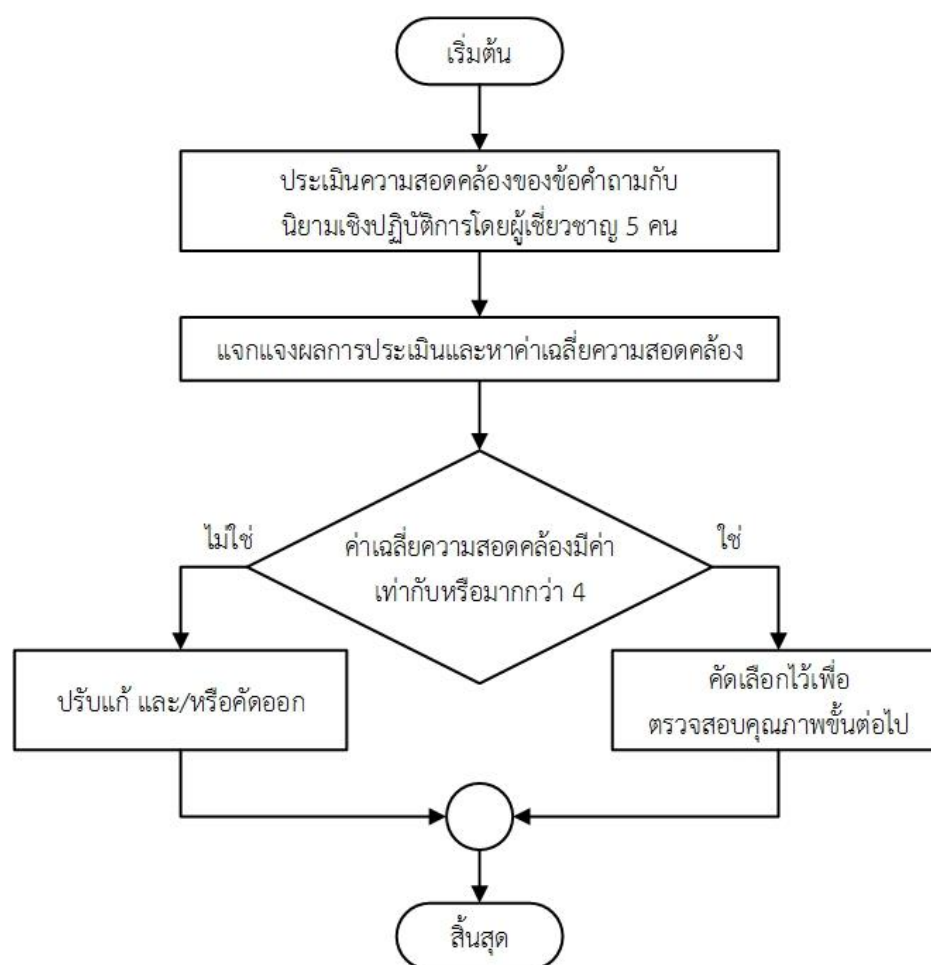
ภาพที่ 3-3 ขั้นตอนการตรวจสอบคุณภาพรายข้อของข้อคำถาม

จากภาพที่ 3-3 การตรวจสอบคุณภาพรายข้อของข้อคำถาม มีขั้นตอนย่อย ดังนี้

1) การตรวจสอบความตรงเชิงเนื้อหาของข้อคำถามทั้งหมด ด้วยวิธีการประเมินความสอดคล้องของข้อคำถามกับนิยามเชิงปฏิบัติการ โดยให้ผู้เชี่ยวชาญ 5 คน ซึ่งประกอบด้วยนักจิตวิทยาที่มีประสบการณ์การทำงานตั้งแต่ 10 ปีขึ้นไป จำนวน 3 คน และผู้เชี่ยวชาญทางการวัดผลทางการศึกษา/ จิตวิทยา จำนวน 2 คน (รายชื่อผู้เชี่ยวชาญ แสดงในภาคผนวก ก) กำหนดระดับการแสดงความคิดเห็นในลักษณะมาตราประมาณค่า 5 ระดับ (Wiersma & Jurs, 1990, pp. 272-274) ซึ่งแต่ละระดับมีความหมาย ดังนี้

- 1 หมายถึง ข้อคำถามไม่สอดคล้องกับนิยามเชิงปฏิบัติการ
- 2 หมายถึง ข้อคำถามสอดคล้องกับนิยามเชิงปฏิบัติการน้อย
- 3 หมายถึง ข้อคำถามสอดคล้องกับนิยามเชิงปฏิบัติการปานกลาง
- 4 หมายถึง ข้อคำถามสอดคล้องกับนิยามเชิงปฏิบัติการมาก
- 5 หมายถึง ข้อคำถามสอดคล้องกับนิยามเชิงปฏิบัติการมากที่สุด

จากนั้น จึงรวบรวมความคิดเห็นของผู้เชี่ยวชาญมาแจกแจงตามระดับความคิดเห็น 5 ระดับ และคำนวณค่าเฉลี่ยความสอดคล้องของข้อคำถามกับนิยามเชิงปฏิบัติการเป็นรายข้อ โดยข้อคำถามที่มีค่าเฉลี่ยความสอดคล้องต่ำกว่า 4.00 จะถูกปรับปรุงแก้ไขตามคำแนะนำของผู้เชี่ยวชาญหรือคัดออก ส่วนข้อคำถามที่มีค่าเฉลี่ยความสอดคล้องตั้งแต่ 4.00 ขึ้นไป แสดงว่า ข้อคำถามนั้นมีความสอดคล้องกับนิยามเชิงปฏิบัติการ จึงคัดเลือกไว้เพื่อการตรวจสอบคุณภาพของข้อคำถามต่อไป (เสรี ชัดเข้ม, 2544, หน้า 166) การตรวจสอบความตรงเชิงเนื้อหา สามารถสรุปได้ดังภาพที่ 3-4

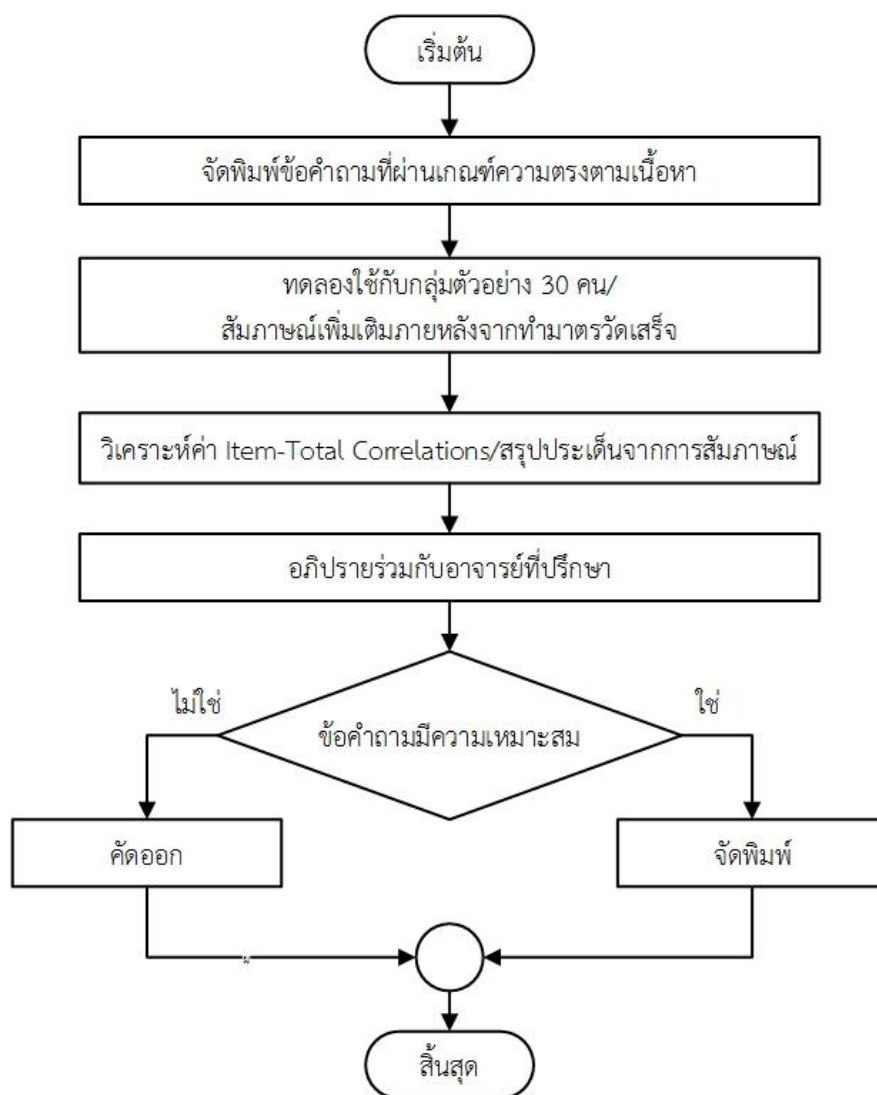


ภาพที่ 3-4 ขั้นตอนการตรวจสอบความตรงเชิงเนื้อหา

2) การทดลองใช้ (Tryout) ได้จัดพิมพ์ข้อคำถามที่ผ่านเกณฑ์การตรวจสอบความตรงเชิงเนื้อหาโดยผู้เชี่ยวชาญ ในรูปแบบกระดาษ-ดินสอ จากนั้น นำไปทดลองใช้กับกลุ่มตัวอย่าง (Tryout Sample) ที่มีลักษณะใกล้เคียงกับประชากรที่ต้องการศึกษาวิจัย จำนวน 30 คน โดยเลือกกลุ่มตัวอย่างแบบตามสะดวก (Convenience Sampling) (Gravetter & Forzano, 2012, pp. 151-152) กลุ่มตัวอย่างต้องมีสัญชาติไทย อายุ 18 ปีขึ้นไป ไม่จำกัดเพศ ไม่จำกัดการศึกษา ไม่จำกัด

สถานภาพสมรส ใช้ภาษาไทยในการติดต่อสื่อสารในชีวิตประจำวัน สามารถอ่านและเขียนภาษาไทยได้ และยินดีที่จะรับการทดสอบ

การทดลองใช้ มีจุดมุ่งหมายเพื่อประเมินความเหมาะสมของข้อคำถาม ในประเด็นด้านความเข้าใจของผู้ทดสอบ ภาษาที่ใช้ รูปแบบการจัดพิมพ์ และข้อคิดเห็นอื่น ๆ เมื่อผู้ทดสอบทำมาตรวัดทั้งหมดเสร็จสิ้น ผู้วิจัยดำเนินการสัมภาษณ์เพิ่มเติมถึงปัญหา อุปสรรค และข้อเสนอแนะอื่น ๆ จากผู้ทดสอบ จากนั้น จึงนำข้อมูลที่ได้จากการสัมภาษณ์ทั้งหมดมาสรุปประเด็น และนำข้อมูลที่ได้จากการตอบมาตรวัดมาวิเคราะห์ค่า Item-Total Correlations โดยใช้โปรแกรม LERTAP Version 5.10.4 ในการวิเคราะห์ข้อมูล จากนั้นจึงนำผลการวิเคราะห์ข้อมูลที่ได้จากกลุ่มทดลองใช้ ไปอภิปรายร่วมกับอาจารย์ที่ปรึกษา เพื่อปรับปรุงข้อคำถาม มาตรวัด รูปแบบการจัดพิมพ์ และรายละเอียดส่วนอื่น แล้วจึงจัดพิมพ์ข้อคำถามทั้งหมด เพื่อนำไปเก็บข้อมูลกับกลุ่มตัวอย่าง ขั้นตอนการทดลองใช้สามารถสรุปได้ ดังแผนภาพที่ 3-5



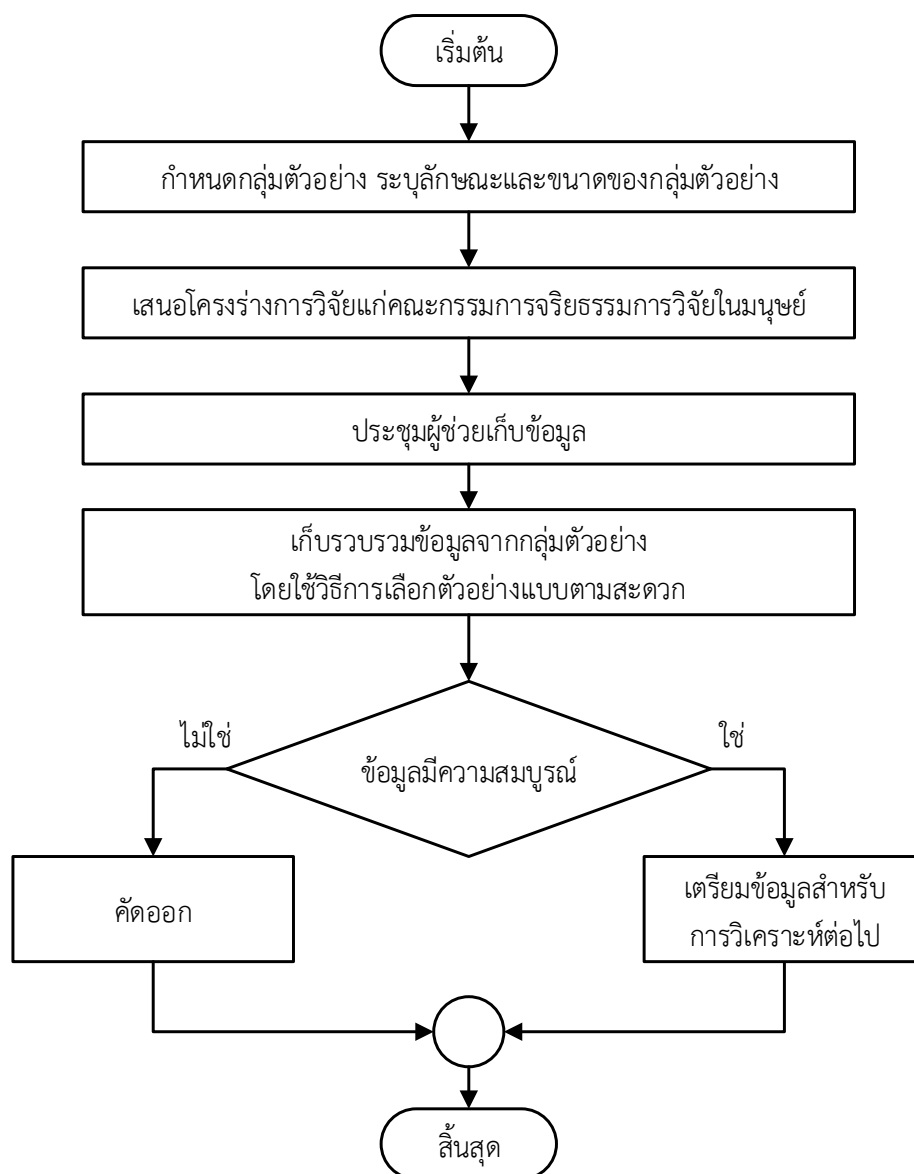
ภาพที่ 3-5 ขั้นตอนการทดลองใช้

3) การเก็บรวบรวมข้อมูล ข้อคำถามที่ผ่านการตรวจสอบความตรงเชิงเนื้อหาและการทดลองใช้ จะถูกจัดพิมพ์สำหรับเก็บข้อมูลกับกลุ่มตัวอย่าง (Calibration Sample) ซึ่งมีลักษณะคล้ายประชากรที่มุ่งนำมาตรวจวัดไปใช้ โดยแนวคิดของทฤษฎีการตอบสนองข้อสอบ กลุ่มตัวอย่างที่เป็นตัวแทนที่ดีของประชากร (Representative Sample) จะต้องมีระดับของคุณลักษณะแฝงที่หลากหลาย (Heterogeneous Sample) ครอบคลุมช่วงของคุณลักษณะแฝงที่ต้องการศึกษา เช่น กลุ่มตัวอย่างมีค่าคุณลักษณะแฝงอยู่ในช่วง -3 ถึง +3 (Embretson & Reise, 2000, pp. 269-270) โดยเกณฑ์การคัดเลือกกลุ่มตัวอย่าง คือ ต้องมีสัญชาติไทย อายุ 18 ปีขึ้นไป ไม่จำกัดเพศ ไม่จำกัดการศึกษา ไม่จำกัดสถานภาพสมรส ใช้ภาษาไทยในการติดต่อสื่อสารในชีวิตประจำวัน สามารถอ่านและเขียนภาษาไทยได้

การพิจารณาขนาดของกลุ่มตัวอย่างขึ้นอยู่กับจำนวนพารามิเตอร์ที่ใช้ในโมเดลและลักษณะของโมเดลการตอบสนองข้อสอบที่เลือกใช้ ซึ่งการเลือกโมเดลต้องพิจารณาจากความสัมพันธ์ระหว่างข้อมูลกับโมเดล Reeve and Fayers (2005) ได้ให้คำแนะนำว่า กลุ่มตัวอย่างที่ใช้ในการวิเคราะห์คุณภาพของข้อคำถามในขั้นตอนของการสร้างคลังข้อคำถาม ควรมีจำนวนไม่ต่ำกว่า 500 คน เป็นกลุ่มตัวอย่างที่มีการแจกแจงของคุณลักษณะแฝงที่ต้องการศึกษาอย่างหลากหลาย และครอบคลุมทุกค่าของคุณลักษณะแฝงนั้น ดังนั้น ในการวิจัยนี้จึงกำหนดขนาดกลุ่มตัวอย่าง จำนวน 1,000 คน เพื่อเพิ่มโอกาสให้ข้อมูลมีการแจกแจงครอบคลุมค่าคุณลักษณะแฝง (ความสุข) ตั้งแต่ -3 ถึง +3 และเพื่อให้มั่นใจว่ามีข้อมูลเพียงพอสำหรับการวิเคราะห์ทางสถิติ ในกรณีที่มีข้อมูลไม่สมบูรณ์

ผู้วิจัยได้จัดทำโครงร่างการวิจัยเพื่อขออนุมัติการดำเนินการวิจัย จากคณะกรรมการจริยธรรมการวิจัยในคน มหาวิทยาลัยบูรพา เพื่อเป็นการพิทักษ์สิทธิ์ของผู้เข้าร่วมการวิจัย ผู้ทดสอบมีสิทธิ์อย่างเต็มที่ในการตัดสินใจเข้าร่วมหรือไม่เข้าร่วมการวิจัย และสามารถถอนตัวจากการวิจัยได้ทุกเมื่อตามความต้องการ โดยการตัดสินใจดังกล่าวจะไม่ส่งผลใด ๆ ต่อผู้ทดสอบ ภายหลังจากผ่านความเห็นชอบทางด้านจริยธรรมการวิจัยจากคณะกรรมการจริยธรรมการวิจัยในคน ผู้วิจัยใช้วิธีการเลือกตัวอย่างแบบตามสะดวก (Convenience Sampling) (Gravetter & Forzano, 2012, pp. 151-152) โดยคัดเลือกผู้ทดสอบที่มีความพร้อมและความยินดีที่จะเข้าร่วมการวิจัยเท่านั้น โดยมีผู้ช่วยเก็บข้อมูล 5 คน ที่ผ่านการประชุมทำความเข้าใจถึงวัตถุประสงค์การวิจัย และรายละเอียดที่เกี่ยวข้องกับกลุ่มตัวอย่างที่ต้องการศึกษา เป็นผู้ช่วยเก็บข้อมูลกับกลุ่มเป้าหมาย โดยให้ความหลากหลายทางด้านเพศ อายุ และระดับการศึกษา การเก็บรวบรวมข้อมูลใช้ระยะเวลา 2 เดือน ตั้งแต่เดือนกรกฎาคม 2557 ถึง สิงหาคม 2557 เมื่อได้มาตรวัดกลับมา ผู้วิจัยได้ตรวจสอบความสมบูรณ์ของข้อมูล และคัดเลือกมาตรวัดที่มีความสมบูรณ์สำหรับการนำไปวิเคราะห์ต่อไป

ขั้นตอนการเก็บรวบรวมข้อมูล สามารถสรุปได้ดังภาพที่ 3-6



ภาพที่ 3-6 ขั้นตอนการเก็บรวบรวมข้อมูล

4) การวิเคราะห์ค่า Item-Total Correlations เมื่อเก็บรวบรวมข้อมูลครบถ้วนแล้ว จึงนำข้อมูลที่ได้มาวิเคราะห์ค่า Item-Total Correlations โดยใช้โปรแกรม LERTAP Version 5.10.4 จากนั้นจึงคัดเลือกข้อคำถามที่มีค่า Item-Total Correlations มากกว่า 0.30 เพื่อนำไปวิเคราะห์ด้วยทฤษฎีการตอบสนองข้อสอบในขั้นต่อไป (Reynolds & Livingston, 2012, p. 227)

5) การตรวจสอบข้อตกลงเบื้องต้นของทฤษฎีการตอบสนองข้อสอบ ดำเนินการดังนี้

5.1 การตรวจสอบความถี่ของการตอบข้อคำถามในแต่ละตัวเลือกการคำตอบ เนื่องจากการวิเคราะห์ข้อมูลด้วยทฤษฎีการตอบสนองข้อสอบนั้น ต้องมั่นใจว่าทุกตัวเลือกการคำตอบมีผู้ทดสอบเลือกตอบ ในเบื้องต้นจึงต้องตรวจสอบความถี่ของการตอบข้อคำถามในแต่ละตัวเลือกการคำตอบ (De Ayala, 2009, p. 223)

5.2 การตรวจสอบความเป็นเอกมิติของมาตรวัด (Unidimensionality) ซึ่งตรวจสอบในแต่ละมาตรวัดย่อย ได้แก่ มาตรวัดย่อยความพึงพอใจในชีวิต มาตรวัดย่อยอารมณ์ความรู้สึกเชิงบวก และมาตรวัดย่อยอารมณ์ความรู้สึกเชิงลบ โดยการวิเคราะห์องค์ประกอบเชิงยืนยัน (Confirmatory Factor Analysis: CFA) ด้วยโปรแกรม LISREL version 8.8

5.3 การตรวจสอบความเป็นอิสระของข้อคำถาม (Local Independence) โดยการพิจารณาค่าสหสัมพันธ์เศษเหลือ (Residual Correlations) ที่ได้จากการวิเคราะห์องค์ประกอบเชิงยืนยันในการตรวจสอบความเป็นเอกมิติ โดยตัดข้อคำถามที่มีค่าสหสัมพันธ์เศษเหลือมากกว่า 0.25 ออกจากคลังข้อคำถาม (Fliege et al., 2005; Walter, Becker, & Bjorner, 2007, Kocalevant et al., 2009)

6) การตรวจสอบความสอดคล้องของข้อมูลเชิงประจักษ์ที่ได้จากการตอบข้อคำถามกับ Graded-Response Model รวมทั้งการประมาณค่าพารามิเตอร์ความชันและพารามิเตอร์เทรซโฮลด์ ด้วยโปรแกรม Xcalibre version 4.2.2.1 โดยพิจารณาดัชนีความสอดคล้องของข้อคำถามกับโมเดลการตอบสนองข้อสอบเป็นรายข้อ แล้วจึงตัดข้อคำถามที่มีปัญหาเรื่องความสอดคล้องของข้อคำถามกับโมเดลการตอบสนองข้อสอบออกจากคลังข้อคำถาม (Embretson & Reise, 2000, pp. 233-244; Edelen & Reeve, 2007) ขั้นตอนที่ 4-6 สามารถสรุปได้ดังภาพที่ 3-7

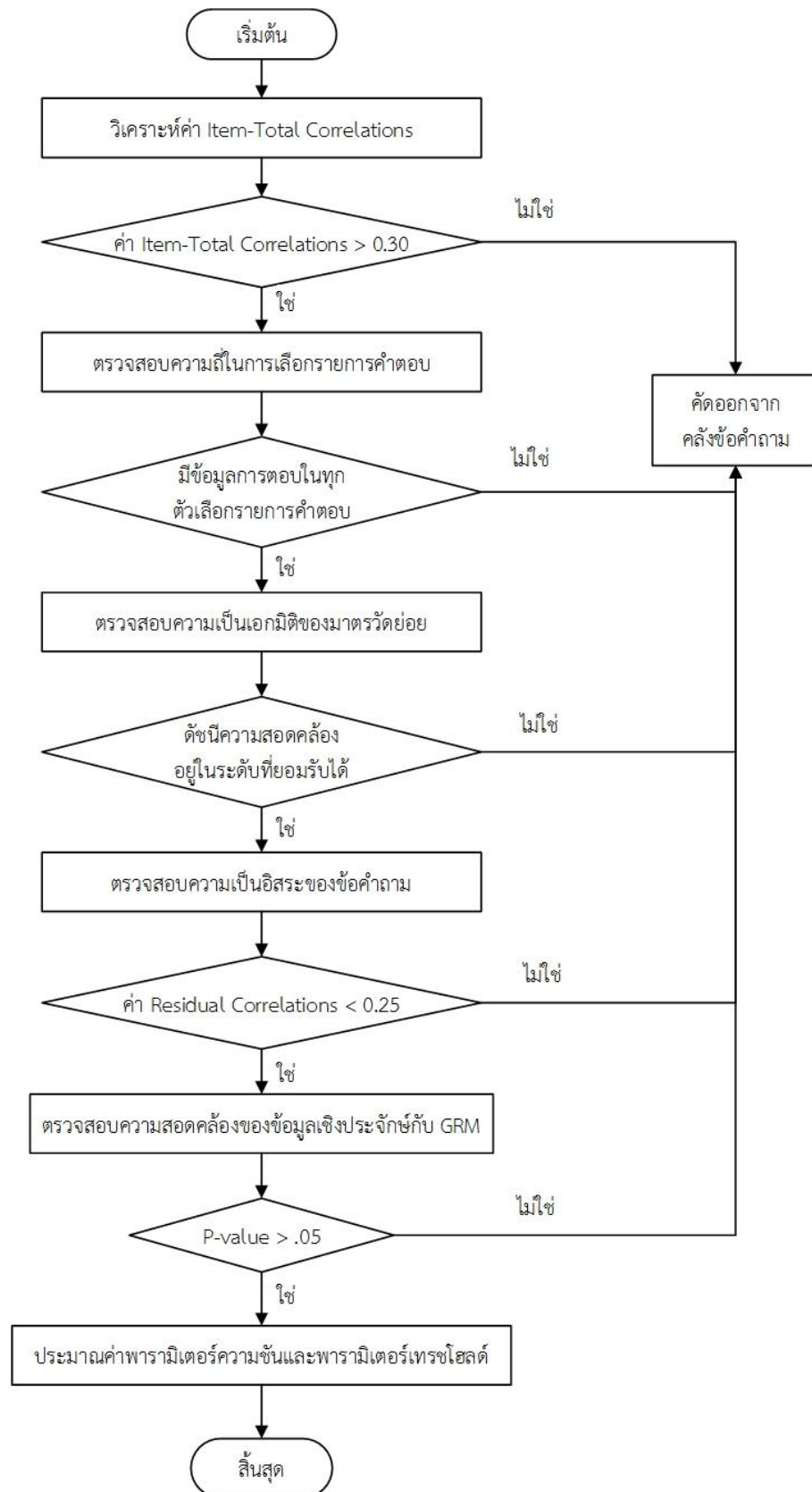
การตรวจสอบความตรงเชิงโครงสร้างของข้อคำถามในคลังข้อคำถาม

นำข้อคำถามในแต่ละมาตรวัดย่อย ที่ผ่านเกณฑ์การวิเคราะห์คุณภาพรายข้อ มาวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสอง (Second Order Confirmatory Factor Analysis) ด้วยโปรแกรม LISREL Version 8.8 เพื่อตรวจสอบว่า ข้อคำถามสำหรับวัดความสุขที่พัฒนาขึ้นประกอบด้วยองค์ประกอบ 3 ด้าน ได้แก่ ความพึงพอใจในชีวิต อารมณ์ความรู้สึกเชิงบวก และอารมณ์ความรู้สึกเชิงลบ ตามทฤษฎีความอยู่ดีมีสุขเชิงอัตวิสัย (Pavot & Diener, 2013)

การจัดทำคลังข้อคำถามสำหรับวัดความสุข

การจัดทำคลังข้อคำถาม ดำเนินการโดยรวบรวมข้อคำถามที่ผ่านเกณฑ์การตรวจสอบคุณภาพทุกข้อมาบันทึกไว้ในรูปแบบไฟล์ เพื่อเตรียมพร้อมสำหรับการพัฒนามาตรวัดความสุขของคนไทยรูปแบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ในขั้นตอนต่อไป ข้อมูลที่บันทึกประกอบด้วยคำถาม ตัวเลือกรายการคำตอบ ค่าพารามิเตอร์ความชัน และค่าพารามิเตอร์เทรซโฮลด์

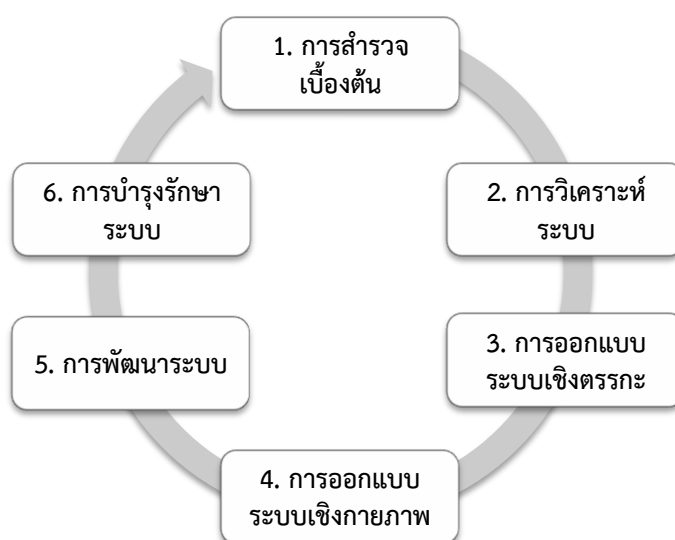
ในการจัดทำคลังข้อคำถาม ผู้วิจัยได้จัดพิมพ์ข้อคำถามในลักษณะมาตรวัดรูปแบบกระดาษ-ดินสอขึ้น สำหรับเป็นอีกทางเลือกหนึ่งในการนำไปใช้ (รายละเอียดการจัดทำมาตรวัดความสุขของคนไทย รูปแบบกระดาษ-ดินสอ แสดงในภาคผนวก ข)



ภาพที่ 3-7 ขั้นตอนการวิเคราะห์ข้อมูล

ขั้นตอนที่ 2 การพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย

การพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย ดำเนินการตามแนวคิดของวงจรการพัฒนาระบบ (System Development Life Cycle: SDLC) ซึ่งประกอบด้วย 6 ขั้นตอนหลัก ดังภาพที่ 3-8 (อรยา ปรีชาพานิช, 2557, หน้า 41-46)



ภาพที่ 3-8 วงจรการพัฒนาระบบ (System Development Life Cycle: SDLC)

การพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย มีขั้นตอนการพัฒนาตามแนวคิดของวงจรการพัฒนาระบบ ซึ่งมีรายละเอียด ดังนี้

1. ขั้นตอนการสำรวจเบื้องต้น (Preliminary Investigation Phase)

ผู้วิจัยทบทวนเอกสารที่เกี่ยวข้องกับการพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ศึกษาความเป็นไปได้ในการพัฒนาโปรแกรม จัดทำแผนการดำเนินงานซึ่งระบุรายละเอียดและระยะเวลาการดำเนินงาน รวมทั้งทรัพยากรต่าง ๆ ที่จำเป็นต้องใช้ในแต่ละขั้นตอน

2. ขั้นตอนการวิเคราะห์ระบบ (System Analysis Phase)

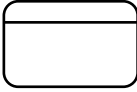



การวิเคราะห์ระบบของโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ เริ่มต้นจากการรวบรวมความต้องการใช้งานของผู้ใช้งาน (User Requirement) แล้วนำมาสรุปเป็นความต้องการของระบบ (System Requirement) จากนั้นจึงสร้างแบบจำลองกระบวนการ (Process Model) เพื่อให้เห็นกระบวนการทำงานอย่างเป็นลำดับขั้นตอน และแบบจำลองข้อมูล (Data Model) เพื่อใช้อธิบายโครงสร้างและลักษณะของข้อมูล

2.1 แบบจำลองกระบวนการ (Process Model) เป็นแบบจำลองที่ทำให้เห็นกระบวนการทำงานของระบบอย่างเป็นลำดับขั้นตอน แสดงให้เห็นถึงการนำเข้าและส่งออกข้อมูล

โดยสร้างแบบจำลองกระบวนการในรูปแบบแผนภาพกระแสข้อมูล (Data Flow Diagram: DFD) ซึ่งประกอบด้วย 4 ส่วนหลัก ได้แก่ (อริยา ปรีชาพานิช, 2557, หน้า 99-129)

- 1) กระบวนการทำงาน (Process) ประกอบด้วย
 - ก. การจัดการคลังข้อความ ประกอบด้วย มาตรการวัดย่อย และข้อความ
 - ข. การจัดการประวัติผู้ทดสอบ
 - ค. การจัดการการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ประกอบด้วย การคัดเลือกข้อความเริ่มต้น การประมาณค่าคุณลักษณะแฝงของผู้ทดสอบ การคัดเลือกข้อความข้อถัดไป และการยุติการทดสอบ
 - ง. การรายงานผลการทดสอบ
 - 2) เอนทิตีภายนอก (External Entity) ประกอบด้วย ผู้ทดสอบ และผู้ดูแลระบบ
 - 3) กระแสข้อมูล (Data Flow)
 - 4) แหล่งจัดเก็บข้อมูล (Data Store) ประกอบด้วยแฟ้มมาตรการวัดย่อย แฟ้มข้อความ แฟ้มประวัติผู้ทดสอบ และแฟ้มข้อมูลการทดสอบ
- สัญลักษณ์ที่ใช้ในการสร้างแผนภาพกระแสข้อมูล ใช้ชุดสัญลักษณ์ที่กำหนดโดย Gane and Sarson (อริยา ปรีชาพานิช, 2557, หน้า 100-106) ดังตารางที่ 3-1

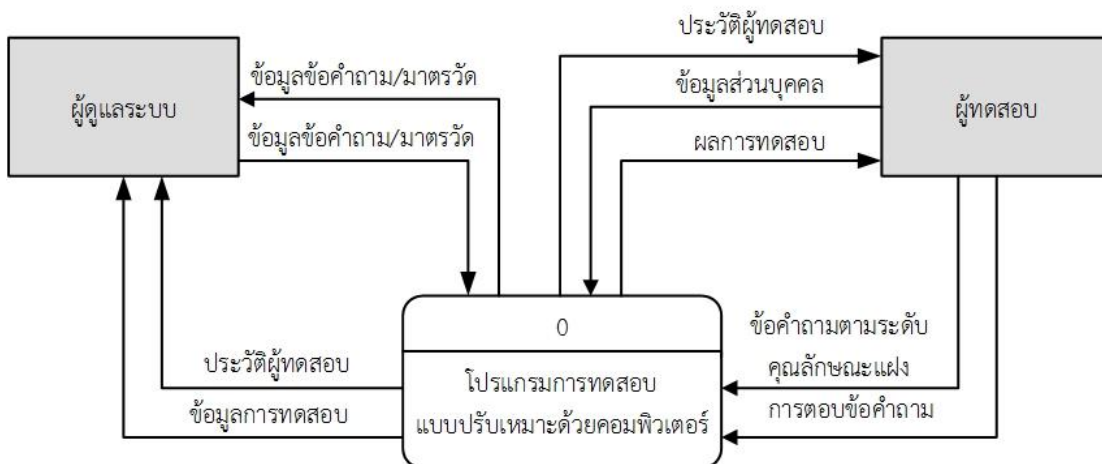
ตารางที่ 3-1 ความหมายและสัญลักษณ์ที่ใช้ในการเขียนแผนภาพกระแสข้อมูล

ภาพสัญลักษณ์	ความหมาย
	กระบวนการทำงาน (Process)
	เอนทิตีภายนอก (External Entity) หมายถึง แหล่งข้อมูล ซึ่งอาจเป็นบุคคล ระบบสารสนเทศ หรือเครื่องมือ/อุปกรณ์ที่สัมพันธ์กับกระบวนการทำงาน
	กระแสข้อมูล (Data Flow) แสดงถึง ทิศทางการไหลของข้อมูลนำเข้า และข้อมูล/สารสนเทศที่เป็นผลลัพธ์
	แหล่งจัดเก็บข้อมูล (Data Store) แสดงถึง แหล่งจัดเก็บข้อมูลทั้งในส่วนของข้อมูลนำเข้า และข้อมูล/สารสนเทศที่เป็นผลลัพธ์

โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรการวัดความสุขของคนไทย ประกอบด้วย แผนภาพกระแสข้อมูล ดังนี้

2.1.1 แผนภาพบริบท (Context Diagram) เป็นแผนภาพแสดงภาพรวมของระบบทั้งหมด ซึ่งแสดงถึงขอบเขตของระบบว่ามีปฏิสัมพันธ์กับสิ่งแวดล้อมของระบบอย่างไร โดยไม่แสดง

รายละเอียดกระบวนการทำงานภายในระบบ และไม่แสดงแหล่งจัดเก็บข้อมูลต่าง ๆ ของระบบ แสดงได้ดังภาพที่ 3-9



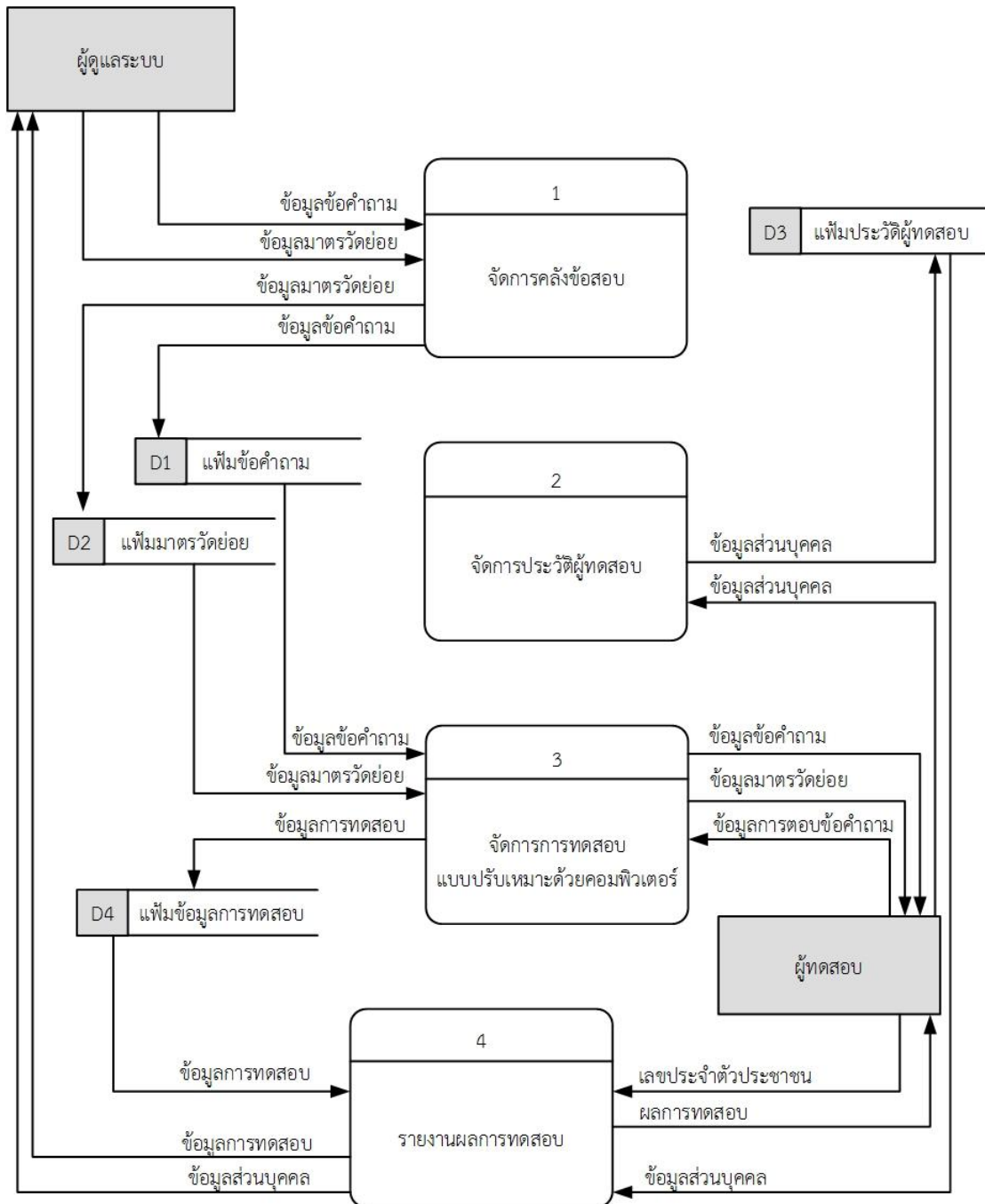
ภาพที่ 3-9 แผนภาพบริบทของโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับมาตรวัดความสุขของคนไทย

จากภาพที่ 3-9 แสดงการรับส่งข้อมูลระหว่างโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่พัฒนาขึ้น กับเอนทิตีภายนอกซึ่งเป็นผู้ที่มีความเกี่ยวข้องกับโปรแกรม ได้แก่

ผู้ดูแลระบบ หมายถึง บุคคลที่รับผิดชอบเกี่ยวกับการจัดการคลังข้อคำถาม และการจัดการประวัติการทดสอบ ซึ่งสามารถเพิ่ม ลบ และแก้ไขข้อมูลของคลังข้อคำถาม ทั้งในส่วนข้อคำถามและมาตรวัดย่อย รวมทั้ง สามารถค้นหาข้อมูลการทดสอบของผู้ทดสอบโดยละเอียด

ผู้ทดสอบ หมายถึง บุคคลทั่วไปที่มีความสนใจเข้ารับการทดสอบด้วยโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย ซึ่งกลุ่มเป้าหมายคือ ประชาชนชาวไทย ที่มีอายุตั้งแต่ 18 ปีขึ้นไป

2.1.2 แผนภาพกระแสข้อมูล ระดับที่ 0 (Data Flow Diagram Level 0) เป็นแผนภาพในระดับถัดมาจากแผนภาพบริบท ซึ่งแสดงรายละเอียดของกระบวนการทำงานหลักของระบบเกี่ยวกับการประสานงานรับ-ส่งข้อมูล และสารสนเทศกับเอนทิตีภายนอก นอกจากนี้ยังแสดงถึงการเชื่อมโยงการทำงานระหว่างกระบวนการทำงานต่าง ๆ รวมทั้งการจัดเก็บและการค้นหาข้อมูลจากแหล่งจัดเก็บข้อมูลต่าง ๆ ดังภาพที่ 3-10



ภาพที่ 3-10 แผนภาพกระแสข้อมูล ระดับที่ 0 ของโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย

จากภาพที่ 3-10 แสดงกระบวนการหลักของโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย ดังนี้

กระบวนการหลักที่ 1 คือ จัดการคลังข้อคำถาม เป็นกระบวนการในการจัดเก็บและบำรุงรักษาคลังข้อคำถาม ซึ่งสามารถเพิ่ม ลบ และแก้ไข ข้อมูลในคลังข้อคำถาม ทั้งในส่วนของข้อคำถาม และมาตรวัดย่อยได้ ข้อคำถามต้องมีลักษณะตัวเลือกการตอบ เป็นมาตรฐานค่า 5 ระดับ ที่ทราบค่าพารามิเตอร์ความชัน (α) และค่าพารามิเตอร์เทรซโฮลด์ (β) จากการวิเคราะห์คุณภาพของข้อคำถามด้วย Graded Response Model ตามแนวคิดของทฤษฎีการตอบสนองข้อสอบ

กระบวนการหลักที่ 2 คือ จัดการประวัติผู้ทดสอบ เป็นกระบวนการที่จัดเก็บข้อมูลประวัติทั่วไปของบุคคล ได้แก่ ชื่อ สกุล เพศ อายุ และเลขประจำตัวประชาชน

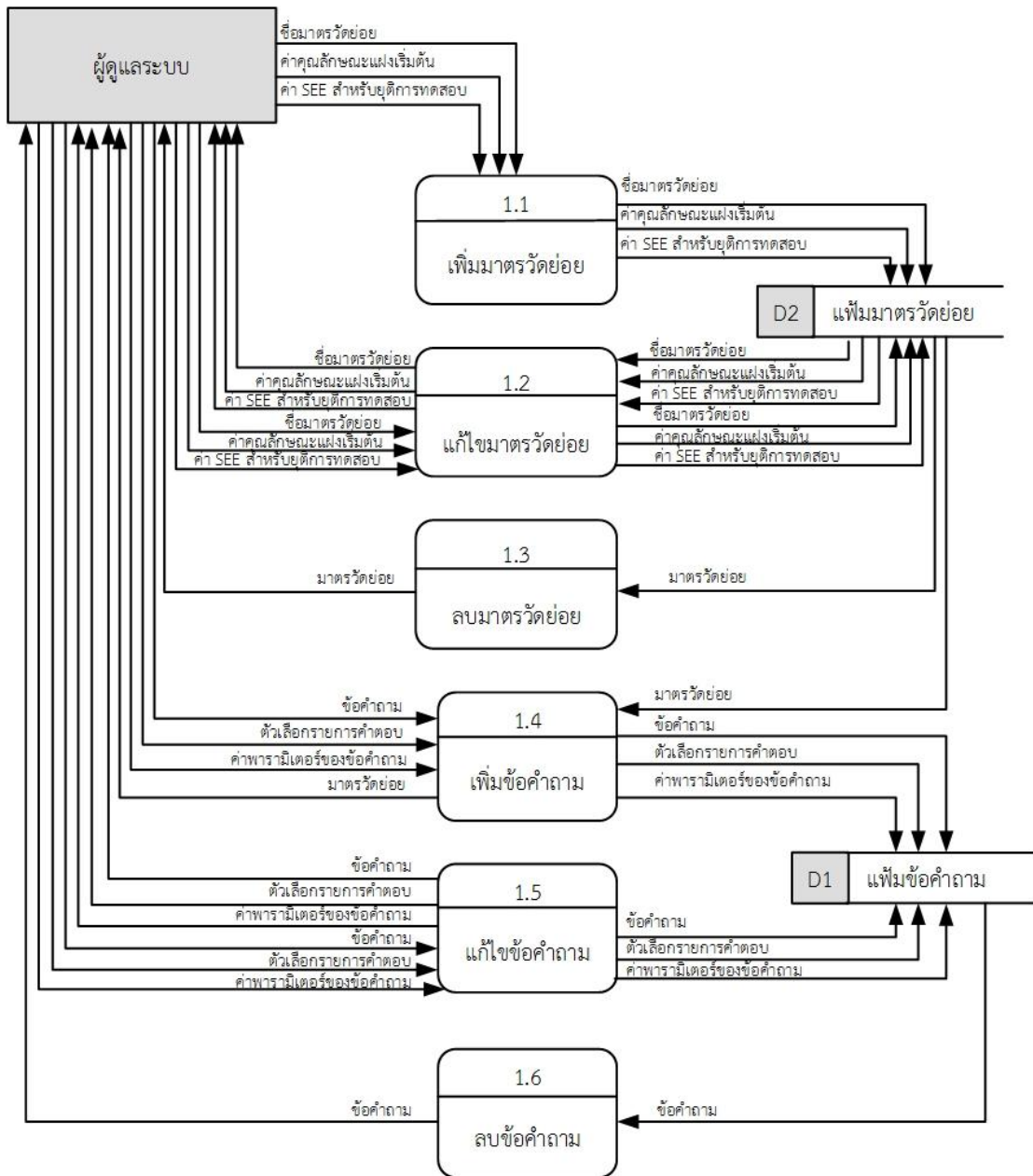
กระบวนการหลักที่ 3 คือ จัดการการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ เป็นการบริหารการทดสอบซึ่งประกอบด้วยกระบวนการย่อย ได้แก่ คัดเลือกข้อคำถามเริ่มต้น ประเมินค่าคุณลักษณะแฝงของบุคคล คัดเลือกข้อคำถามข้อถัดไป และยุติการทดสอบ

กระบวนการหลักที่ 4 คือ รายงานผลการทดสอบ เป็นกระบวนการที่นำข้อมูลผลการทดสอบของผู้ทดสอบ และประวัติผู้ทดสอบ มาจัดทำรายงานผลการทดสอบให้ผู้ใช้งาน และผู้ดูแลระบบทราบ

2.1.3 แผนภาพกระแสข้อมูล ระดับที่ 1 (Data Flow Diagram Level 1) เป็นแผนภาพย่อยที่ใช้อธิบายรายละเอียดของกระบวนการทำงานจากแผนภาพระดับที่ 0 โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย มีกระบวนการหลัก 2 กระบวนการที่มีกระบวนการย่อยระดับล่างอีกชั้น คือ กระบวนการจัดการคลังข้อคำถาม และกระบวนการจัดการการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ดังนี้

ก. กระบวนการจัดการคลังข้อคำถาม

กระบวนการจัดการคลังข้อคำถามของโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ แบ่งออกเป็นสองส่วน โดยส่วนที่หนึ่ง คือ ส่วนของมาตรวัดย่อย ประกอบด้วย 3 กระบวนการย่อย ได้แก่ 1) เพิ่มมาตรวัดย่อย 2) แก้ไขมาตรวัดย่อย และ 3) ลบมาตรวัดย่อย อีกส่วนหนึ่ง คือ ส่วนของข้อคำถาม ประกอบด้วย 3 กระบวนการย่อย ได้แก่ 1) เพิ่มข้อคำถาม 2) แก้ไขข้อคำถาม และ 3) ลบข้อคำถาม แสดงได้ดังภาพที่ 3-11



ภาพที่ 3-11 แผนภาพกระแสข้อมูล ระดับที่ 1 ของกระบวนการจัดการคลังข้อความ

จากภาพที่ 3-11 แสดงการรับ-ส่งข้อมูลของกระบวนการจัดการคลังข้อความ ซึ่งอธิบายความหมายของแต่ละกระบวนการย่อยได้ ดังนี้

กระบวนการย่อยที่ 1.1 คือ เพิ่มมาตรฐานวัดย่อย เป็นกระบวนการที่ใช้เพิ่มมาตรฐานวัดย่อยใหม่ลงในคลังข้อความ ซึ่งต้องดำเนินการก่อนเพิ่มข้อความของมาตรฐานวัดย่อยใหม่ และข้อความควรผ่านการตรวจสอบคุณภาพตามทฤษฎีการตอบสนองข้อสอบเรียบร้อยแล้ว การเพิ่มมาตรฐานวัดย่อยเป็นการระบุชื่อมาตรฐานวัดย่อย ค่าคุณลักษณะแง่เริ่มต้น และค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่า (SEE) สำหรับยุติการทดสอบ

กระบวนการย่อยที่ 1.2 คือ แก้ไขมาตรวัดย่อย เป็นกระบวนการที่ใช้แก้ไขมาตรวัดย่อยที่มีอยู่ในคลังข้อคำถาม ซึ่งเป็นการแก้ไขชื่อ ค่าคุณลักษณะแฝงเริ่มต้น และค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่า (SEE) สำหรับยุติการทดสอบของมาตรวัดย่อยที่เลือก

กระบวนการย่อยที่ 1.3 คือ ลบมาตรวัดย่อย เป็นกระบวนการลบมาตรวัดย่อยที่ไม่ต้องการออกจากคลังข้อคำถาม

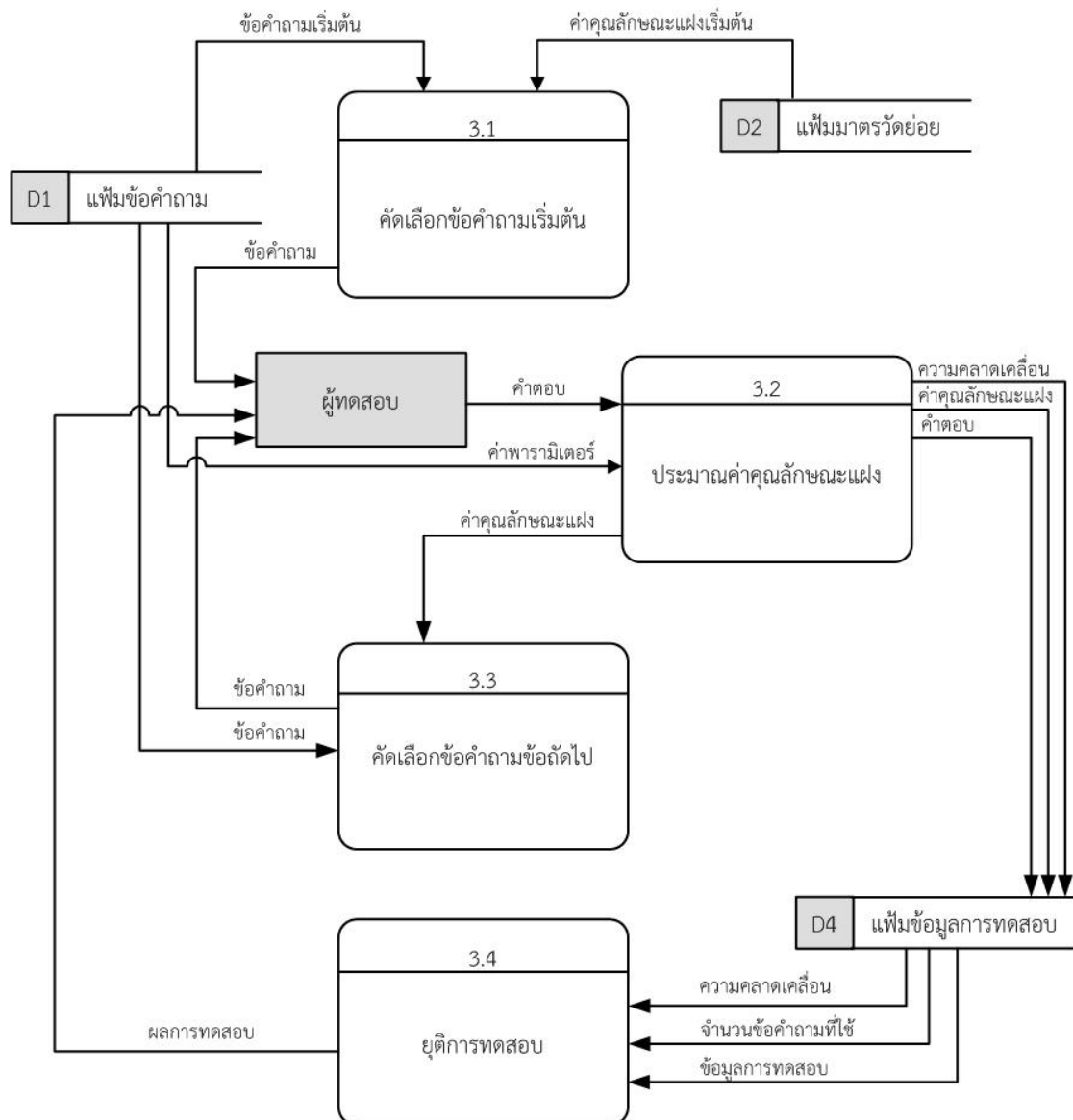
กระบวนการย่อยที่ 1.4 คือ เพิ่มข้อคำถาม เป็นกระบวนการที่ใช้เพิ่มข้อคำถามใหม่ ลงในคลังข้อคำถาม ข้อคำถามประกอบด้วยรายละเอียดย่อย ได้แก่ คำถาม ตัวเลือกรายการคำตอบซึ่งต้องมีลักษณะเป็นมาตรประมาณค่า 5 ระดับ และค่าพารามิเตอร์ของข้อคำถามที่ได้จากการวิเคราะห์ด้วย Graded Response Model ตามแนวคิดของทฤษฎีการตอบสนองข้อสอบ

กระบวนการย่อยที่ 1.5 คือ แก้ไขข้อคำถาม เป็นกระบวนการที่ใช้แก้ไขข้อคำถามที่มีอยู่ในคลังข้อคำถาม ซึ่งสามารถแก้ไขคำถาม ตัวเลือกรายการคำตอบ และค่าพารามิเตอร์ของข้อคำถามที่เกี่ยวข้องได้

กระบวนการย่อยที่ 1.6 คือ ลบข้อคำถาม เป็นกระบวนการลบข้อคำถามที่ไม่ต้องการออกจากคลังข้อคำถาม เมื่อเลือกการลบข้อคำถาม รายละเอียดที่เกี่ยวข้องกับข้อคำถามนั้น ๆ ทั้งหมดจะหายไปจากคลัง ได้แก่ คำถาม ตัวเลือกรายการคำตอบ และค่าพารามิเตอร์ของข้อคำถาม

ข. กระบวนการจัดการการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์

กระบวนการย่อยในส่วนนี้มี 4 กระบวนการย่อย ได้แก่ 1) คัดเลือกข้อคำถามเริ่มต้น 2) ประมาณค่าคุณลักษณะแฝงของผู้ทดสอบ 3) คัดเลือกข้อคำถามข้อถัดไป และ 4) ยุติการทดสอบ แสดงได้ดังภาพที่ 3-12



ภาพที่ 3-12 แผนภาพกระแสข้อมูล ระดับที่ 1 ของกระบวนการจัดการการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์

จากภาพที่ 3-12 แสดงการรับ-ส่งข้อมูลของกระบวนการย่อยในกระบวนการจัดการการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ซึ่งอธิบายความหมายของแต่ละกระบวนการย่อยได้ดังนี้

กระบวนการย่อยที่ 1.1 คือ คัดเลือกข้อคำถามเริ่มต้น เป็นการคัดเลือกข้อคำถามจากคลังข้อคำถามของมาตรวัดย่อย ซึ่งให้ค่าสารสนเทศสูงที่สุด (Maximum Information) ณ ตำแหน่งค่าคุณลักษณะแฟงที่กำหนดไว้ โดยกำหนดค่าคุณลักษณะแฟงเริ่มต้นของผู้ทดสอบ ในทั้งสามมาตรวัดย่อย ให้มีค่าคุณลักษณะแฟงเริ่มต้นเท่ากับ 0.00 โปรแกรมจะเลือกข้อคำถามที่ให้ค่าสารสนเทศข้อสอบสูงที่สุด ณ ตำแหน่งค่าคุณลักษณะแฟงดังกล่าว

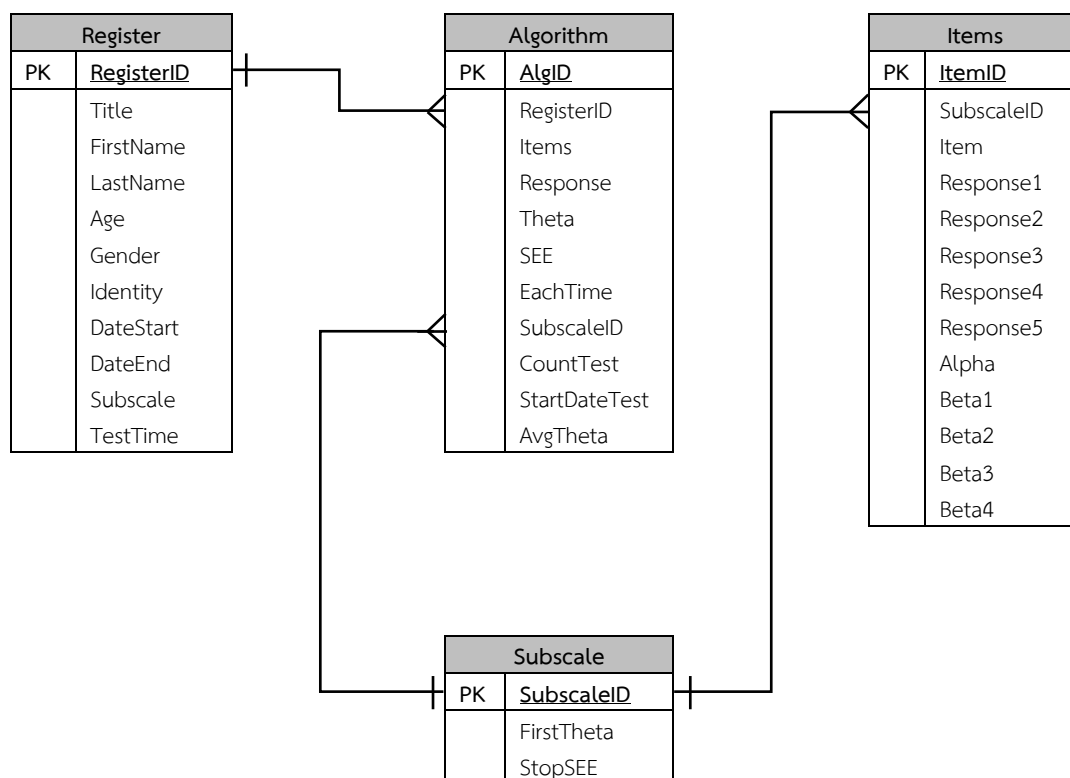
กระบวนการย่อยที่ 1.2 คือ ประมาณค่าคุณลักษณะแฝงของผู้ทดสอบ เป็นกระบวนการที่ใช้ผลการตอบข้อคำถามข้อที่ผ่านมาของผู้ทดสอบ เป็นข้อมูลในการประมาณค่าคุณลักษณะแฝงใหม่ โดยใช้วิธีการ Expected a Posteriori (EAP)

กระบวนการย่อยที่ 1.3 คือ คัดเลือกข้อคำถามข้อถัดไป เป็นกระบวนการที่นำผลการประมาณค่าคุณลักษณะแฝงที่ได้จากกระบวนการก่อนหน้ามาเป็นข้อมูลสำหรับการเลือกข้อคำถามข้อถัดไปซึ่งมีค่าสารสนเทศสูงที่สุด (Maximum Information) ณ ตำแหน่งค่าคุณลักษณะแฝงที่ประมาณค่าได้

กระบวนการย่อยที่ 1.4 คือ ยุติการทดสอบ เป็นกระบวนการที่ทำให้การทดสอบยุติลง โดยการตรวจสอบข้อมูลการทดสอบ ซึ่งโปรแกรมจะยุติการทดสอบเมื่อความคลาดเคลื่อนจากการประมาณค่าคุณลักษณะแฝงมีค่าต่ำกว่า 0.30 หรือข้อคำถามหมดคลัง

เมื่อกระบวนการทดสอบยุติลง โปรแกรมจะมีการประมาณค่าคุณลักษณะแฝงสุดท้ายในลักษณะค่า Theta ซึ่งเทียบเท่ากับคะแนนซี (Z-score) ที่มีค่าเฉลี่ยเท่ากับ 0 และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 1 จากนั้นจึงแปลงค่าคะแนนเป็นค่าคะแนนที (T-score) ซึ่งมีค่าเฉลี่ยเท่ากับ 50 และมีส่วนเบี่ยงเบนมาตรฐานเท่ากับ 10 ค่าคะแนนที่จะถูกนำไปใช้ในการรายงานผลคะแนนในแต่ละมาตรวัดย่อย ในขณะที่ผลการวัดความสุขโดยรวมสามารถคำนวณได้จากค่าเฉลี่ยของคะแนนมาตรฐานที่ได้จากสามมาตรวัดย่อย ซึ่งจะมีการรายงานคะแนนเมื่อผู้ทดสอบทำครบทั้งสามมาตรวัดเท่านั้น การแปลความหมายคะแนนที่ได้ในแต่ละมาตรวัดย่อยและคะแนนความสุขโดยรวม ทำได้โดยแปลความหมายคะแนนที่ได้โดยแบ่งออกเป็น 9 ช่วงตามคะแนนสเตนไนน์ ทั้งในส่วนของมาตรวัดย่อยและมาตรวัดความสุขโดยรวม

2.2 แบบจำลองข้อมูล (Data Model) ใช้สำหรับอธิบายโครงสร้าง และคุณลักษณะของข้อมูล รวมทั้งความสัมพันธ์ของข้อมูลในโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทยที่พัฒนาขึ้น โดยเป็นแบบจำลองข้อมูลเชิงแนวคิด (Conceptual Data Model) ซึ่งใช้วิธีการนำเสนอด้วยการสร้างแผนภาพ E-R (Entity-Relationship Diagram) ดังภาพที่ 3-13



ภาพที่ 3-13 แผนภาพ E-R (Entity Relationship Diagram) ของโปรแกรมการทดสอบแบบปรับ
 เหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย

จากภาพที่ 3-13 แสดงให้เห็นเอนทิตี (Entity) ที่จำเป็นสำหรับโปรแกรมการทดสอบแบบ
 ปรับเหมาะด้วยคอมพิวเตอร์สำหรับมาตรวัดความสุขของคนไทย ทั้งหมด 4 เอนทิตี ได้แก่

- 1) การลงทะเบียน (Register)
- 2) การดำเนินการทดสอบ (Algorithm)
- 3) ข้อคำถาม (Items)
- 4) มาตรวัดย่อย (Subscale)

ในแต่ละเอนทิตีประกอบด้วยแอตทริบิวต์ (Attribute) และคีย์หลัก (Primary Key)
 รายละเอียดดังตารางที่ 3-2

ตารางที่ 3-2 เอนทิตี แอตทริบิวต์ และคีย์หลักของโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์

เอนทิตี	แอตทริบิวต์	คำอธิบาย	คีย์หลัก
Register	RegisterID	รหัสผู้ทดสอบ	RegisterID
	Title	คำนำหน้าชื่อ	
	FirstName	ชื่อ	
	LastName	นามสกุล	
	Age	อายุ	
	Gender	เพศ	
	Identity	เลขประจำตัวประชาชน	
	DateStart	วันเวลาที่เริ่มทดสอบ	
	DateEnd	วันเวลาที่สิ้นสุดการทดสอบ	
	Subscale	มาตรวัดย่อยที่ทดสอบ	
	TestTime	ระยะเวลาที่ใช้ทดสอบ	
Algorithm	AlgID	รหัสการทดสอบ	AlgID
	RegisterID	รหัสผู้ทดสอบ	
	Items	ข้อคำถาม	
	Response	คำตอบ	
	Theta	ค่าคุณลักษณะแฝง	
	SEE	ความคลาดเคลื่อนมาตรฐานของการประมาณค่า	
	EachTime	วันเวลาที่ทดสอบในข้อคำถามนั้น	
	SubscaleID	รหัสมาตรวัดย่อย	
	CountTest	จำนวนมาตรวัดย่อยที่ทดสอบ	
	StartDateTest	วันเวลาที่ทดสอบในข้อคำถามแรก	
	AvgTheta	ค่าคุณลักษณะแฝงรวมเฉลี่ย (กรณีทดสอบครบทั้งสามมาตรวัดย่อย)	
Items	ItemID	รหัสข้อคำถาม	ItemID
	SubscaleID	รหัสมาตรวัดย่อย	
	Item	คำถาม	
	Response1	ตัวเลือกคำตอบที่ 1	
	Response 2	ตัวเลือกคำตอบที่ 2	
	Response 3	ตัวเลือกคำตอบที่ 3	
	Response 4	ตัวเลือกคำตอบที่ 4	
	Response 5	ตัวเลือกคำตอบที่ 5	

ตารางที่ 3-2 (ต่อ)

เอนทิตี	แอตทริบิวต์	คำอธิบาย	คีย์หลัก
	Alpha	ค่าพารามิเตอร์ความชัน	
	Beta1	ค่าพารามิเตอร์เทรซโฮลด์ที่ 1	
	Beta2	ค่าพารามิเตอร์เทรซโฮลด์ที่ 2	
	Beta3	ค่าพารามิเตอร์เทรซโฮลด์ที่ 3	
	Beta4	ค่าพารามิเตอร์เทรซโฮลด์ที่ 4	
Subscale	Subscale ID	รหัสมาตรวัดย่อย	Subscale ID
	Subscale Name	ชื่อมาตรวัดย่อย	
	FirstTheta	ค่าคุณลักษณะแฝงเริ่มต้น	
	StopSEE	ค่าความคลาดเคลื่อนมาตรฐานของ การประมาณค่าที่ใช้ยุติการทดสอบ	

3. ขั้นตอนการออกแบบระบบเชิงตรรกะ (Logical Design Phase)

ขั้นตอนการออกแบบระบบเชิงตรรกะ ของโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ เป็นการกำหนดรายละเอียดองค์ประกอบต่าง ๆ ของระบบให้สอดคล้องกับความต้องการของระบบ ซึ่งประกอบด้วย 4 กิจกรรมย่อย ดังนี้

3.1 การออกแบบผลลัพธ์ของระบบ (System Output Design)

โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย ประกอบด้วยผลลัพธ์ 2 ประเภท ได้แก่

1) ผลลัพธ์ภายนอก (External Output) เป็นส่วนที่แสดงรายงานผลการทดสอบให้ผู้ทดสอบทราบ ซึ่งแสดงผลลัพธ์ในรูปแบบหน้าจอ (Screen Output) ขณะเดียวกันผู้ทดสอบสามารถสั่งพิมพ์ผลลัพธ์ในรูปแบบกระดาษได้ด้วย (Paper-Based Output) โดยมีการออกแบบโครงสร้างหน้าจอรายงานผลการทดสอบสำหรับผู้ทดสอบ ดังภาพที่ 3-14

ผลการทดสอบ	
ชื่อ-สกุล	<input type="text"/>
วันที่ทดสอบ	<input type="text"/>
เพศ	<input type="text"/> อายุ (ปี) <input type="text"/>
มาตรวัดย่อย	
คะแนน	<input type="text"/>
ระดับ	<input type="text"/>
มาตรวัดย่อย	
คะแนน	<input type="text"/>
ระดับ	<input type="text"/>
มาตรวัดย่อย	
คะแนน	<input type="text"/>
ระดับ	<input type="text"/>
ผลรวม	
คะแนน	<input type="text"/>
ระดับ	<input type="text"/>

ภาพที่ 3-14 โครงสร้างหน้าจอรายงานผลการทดสอบสำหรับผู้ทดสอบ

2) ผลลัพธ์ภายใน (Internal Output) เป็นส่วนที่แสดงรายงานผลการทดสอบโดยละเอียด (Detailed Output) สำหรับผู้ดูแลระบบ โดยมีการออกแบบโครงสร้างหน้าจอรายงานผลการทดสอบสำหรับผู้ดูแลระบบ ดังภาพที่ 3-15

ประวัติการทดสอบ				
ลำดับที่	ชื่อ-สกุล	เพศ	อายุ	
<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>

ภาพที่ 3-15 โครงสร้างหน้าจอรายงานผลการทดสอบสำหรับผู้ดูแลระบบ หน้าจอที่ 1

จากภาพที่ 3-15 แสดงโครงสร้างหน้าจอรายงานผลการทดสอบสำหรับผู้ดูแลระบบ หน้าจอที่ 1 ซึ่งเป็นหน้าจอแรกสุดสำหรับการแสดงประวัติการทดสอบของผู้ทดสอบทุกคนโดยรวม ประกอบด้วยลำดับที่ของผู้ทดสอบ ชื่อ-สกุล เพศ อายุ และออกแบบให้สามารถคลิกเลือกปุ่ม “เพิ่มเติม” เพื่อเข้าสู่หน้าจอที่ 2 ซึ่งแสดงรายละเอียดผลการทดสอบของแต่ละบุคคล ดังภาพที่ 3-16

ประวัติการทดสอบ					
ครั้งที่	วัน/เวลาที่ทดสอบ	ระยะเวลาที่ใช้	จำนวนข้อที่ใช้	หมายเหตุ	
				รายละเอียด	ลบ
				รายละเอียด	ลบ
				รายละเอียด	ลบ
				รายละเอียด	ลบ

ภาพที่ 3-16 โครงสร้างหน้าจอรายงานผลการทดสอบสำหรับผู้ดูแลระบบ หน้าจอที่ 2

จากภาพที่ 3-16 แสดงโครงสร้างหน้าจอรายงานผลการทดสอบสำหรับผู้ดูแลระบบ หน้าจอที่ 2 ซึ่งเป็นหน้าจอแสดงรายละเอียดผลการทดสอบของแต่ละบุคคล เรียงตามครั้งที่ทดสอบ ประกอบด้วยหมายเลขครั้งที่ทดสอบ วัน/เวลาที่ทดสอบ ระยะเวลาที่ใช้ จำนวนข้อที่ใช้ และออกแบบให้สามารถคลิกเลือกปุ่ม “รายละเอียด” เพื่อเข้าสู่หน้าจอที่ 3 ซึ่งแสดงรายงานผลการทดสอบในแต่ละครั้งที่ทดสอบโดยละเอียด หรือ สามารถคลิกเลือกปุ่ม “ลบ” เพื่อลบข้อมูลผลการทดสอบที่ไม่ต้องการได้ ดังภาพที่ 3-17

ประวัติการทดสอบ				
ข้อที่	คำถาม	คำตอบ	ค่า Theta	ค่า SEE

ภาพที่ 3-17 โครงสร้างหน้าจอรายงานผลการทดสอบสำหรับผู้ดูแลระบบ หน้าจอที่ 3

จากภาพที่ 3-17 แสดงโครงสร้างหน้าจอรายงานผลการทดสอบสำหรับผู้ดูแลระบบ หน้าจอที่ 3 ซึ่งเป็นหน้าจอแสดงรายละเอียดการทดสอบในแต่ละครั้งที่ทดสอบ ประกอบด้วย หมายเลขข้อคำถามที่ทดสอบ คำถาม คำตอบ ค่า Theta และค่า SEE

3.2 การออกแบบส่วนนำเข้าสู่ข้อมูล (Input Design)

การออกแบบการนำเข้าสู่ข้อมูลของโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ เป็นการนำเข้าสู่ข้อมูลผ่านฟอร์ม (Form) ด้วยการกรอกข้อมูลในช่องว่างที่ได้กำหนดไว้ โดยฟอร์มสำหรับการกรอกข้อมูลของโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย แบ่งออกเป็น 3 ส่วน ได้แก่ 1) ฟอร์มข้อมูลผู้ทดสอบ 2) ฟอร์มมาตรวัดย่อย และ 3) ฟอร์มข้อคำถาม การออกแบบหน้าจอนำเข้าสู่ข้อมูลผ่านฟอร์ม แสดงได้ ดังภาพที่ 3-18 ภาพที่ 3-19 และภาพที่ 3-20

ภาพที่ 3-18 โครงสร้างหน้าจอฟอร์มข้อมูลผู้ทดสอบ

จากภาพที่ 3-18 แสดงโครงสร้างหน้าจอฟอร์มข้อมูลผู้ทดสอบ ซึ่งจะปรากฏขึ้นก่อนการเริ่มการทดสอบ โดยประกอบด้วย ตัวเลือกมาตรวัดย่อยที่ต้องการทดสอบ คำนำหน้าชื่อ ชื่อ นามสกุล เลขประจำตัวประชาชน อายุ เพศ รหัสแคปTCHA และปุ่มเลือกรูปแบบการทดสอบสองรูปแบบ คือ การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์และการทดสอบด้วยข้อคำถามทุกข้อในคลังข้อคำถาม

มาตรวัดย่อย	
ชื่อมาตรวัด	<input type="text"/>
ค่า Theta เริ่มต้น	<input type="text"/>
ค่า SEE สำหรับปฏิบัติการทดสอบ	<input type="text"/>

ภาพที่ 3-19 โครงสร้างหน้าจอฟอร์มมาตรวัดย่อย

จากภาพที่ 3-19 แสดงโครงสร้างหน้าจอฟอร์มมาตรวัดย่อย ซึ่งประกอบด้วย ชื่อมาตรวัด ค่า Theta เริ่มต้น และค่า SEE สำหรับปฏิบัติการทดสอบ

ข้อคำถาม	
มาตรวัดย่อย	<input type="text"/>
คำถาม	<input type="text"/>
คำตอบที่ 1	<input type="text"/>
คำตอบที่ 2	<input type="text"/>
คำตอบที่ 3	<input type="text"/>
คำตอบที่ 4	<input type="text"/>
คำตอบที่ 5	<input type="text"/>
Alpha	<input type="text"/>
Beta 1	<input type="text"/>
Beta 2	<input type="text"/>
Beta 3	<input type="text"/>
Beta 4	<input type="text"/>

ภาพที่ 3-20 โครงสร้างหน้าจอฟอร์มข้อคำถาม

จากภาพที่ 3-20 แสดงโครงสร้างหน้าจอฟอร์มข้อคำถาม ประกอบด้วย ชื่อมาตรวัด คำถาม ตัวเลือกคำตอบที่ 1-5 ค่าความชัน (หรือ Alpha) และค่าเทรซโฮลด์ที่ 1-4 (หรือค่า Beta)

3.3 การออกแบบในส่วนของกระบวนการทำงาน (Process Design)

การออกแบบในส่วนของกระบวนการทำงานของโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย ประกอบด้วย 4 ขั้นตอนหลัก ได้แก่

- 1) การคัดเลือกข้อความเริ่มต้น ใช้ข้อมูลค่าคุณลักษณะแฝงที่ได้จากแฟ้มข้อมูลมาตรวัดย่อย เป็นเกณฑ์ในการเลือกข้อความที่ให้ค่าสารสนเทศข้อสอบสูงที่สุด ณ ตำแหน่งคุณลักษณะแฝงที่กำหนด (Maximum Information)
- 2) การประมาณค่าคุณลักษณะแฝง ใช้ผลการตอบข้อความของผู้ทดสอบสำหรับการประมาณค่าคุณลักษณะแฝงใหม่ด้วยวิธี Expect a Posteriori (EAP)
- 3) การคัดเลือกข้อความข้อถัดไป ใช้ข้อมูลคุณลักษณะแฝงที่ประมาณค่าใหม่ เป็นเกณฑ์ในการเลือกข้อความที่ให้ค่าสารสนเทศข้อสอบสูงที่สุด ณ ตำแหน่งคุณลักษณะแฝงที่กำหนด
- 4) การยุติการทดสอบ ใช้ผลความคลาดเคลื่อนจากการประมาณค่าคุณลักษณะแฝงซึ่งกำหนดไว้ให้มีค่าน้อยกว่า 0.30 เป็นเกณฑ์ในการยุติการทดสอบ หรือยุติการทดสอบเมื่อข้อความหมดคลัง

ในส่วนของการทดสอบด้วยข้อความทั้งหมดในคลังข้อความ การทดสอบจะยุติลงเมื่อข้อความหมดคลังแล้วเท่านั้น ส่วนขั้นตอนอื่น ๆ ใช้วิธีการเดียวกับการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์

3.4 การออกแบบส่วนต่อประสานกับผู้ใช้ (User Interface Design)

โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทยออกแบบส่วนต่อประสานกับผู้ใช้แบบกราฟิก (Graphical User Interface: GUI) เพื่อให้โปรแกรมดูทันสมัย น่าใช้งาน และไม่เคร่งเครียดหรือเป็นทางการมากเกินไป

4. ขั้นตอนการออกแบบระบบเชิงกายภาพ (Physical Design Phase)

การออกแบบโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย สร้างขึ้นในรูปแบบเว็บแอปพลิเคชัน (Web Application) ซึ่งใช้งานผ่านเครือข่ายอินเทอร์เน็ต (Internet) การออกแบบฐานข้อมูลใช้ซอฟต์แวร์ (Software) MySQL เป็นระบบในการจัดการฐานข้อมูล เพื่อให้สามารถเพิ่ม ลบ และแก้ไขข้อมูลได้อย่างเป็นระบบ ในส่วนของความปลอดภัยของข้อมูล ใช้วิธีการกำหนดบัญชีผู้ใช้ (User Account) และรหัสผ่าน (Password) สำหรับการเข้าใช้งานระบบ ซึ่งจะมีเพียงผู้ดูแลระบบที่เกี่ยวข้องเท่านั้น จึงจะสามารถเข้าไปจัดการกับระบบต่าง ๆ ได้

5. ขั้นตอนการพัฒนาระบบ (System Implementation Phase)

5.1 การเขียนโปรแกรม (Programming) และการทดสอบระบบ (System Testing)

โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทยพัฒนาขึ้นในรูปแบบเว็บแอปพลิเคชัน โดยใช้ภาษา PHP ในการเขียนโปรแกรม เมื่อการเขียนโปรแกรมเสร็จสิ้น ผู้วิจัยและโปรแกรมเมอร์ได้ทดสอบระบบ โดยใช้เทคนิคการทดสอบแบบกล่องดำ (Black Box Testing) ซึ่งเป็นการทดสอบที่เน้นผลลัพธ์ที่เกิดขึ้น (Output) จากกระบวนการประมวลผลโปรแกรม (Process) โดยไม่เน้นรูปแบบการเขียนโปรแกรมของโปรแกรมเมอร์ เพื่อตรวจสอบความถูกต้อง และค้นหาข้อบกพร่องที่อาจเกิดขึ้น จากนั้น จึงปรับปรุงแก้ไขโปรแกรมจนมี

ความสมบูรณ์มากขึ้น นำโปรแกรมที่ผ่านการตรวจสอบโดยผู้วิจัยและโปรแกรมเมอร์ เสนอเพื่อขอความคิดเห็นจากอาจารย์ที่ปรึกษา แล้วจึงนำข้อคิดเห็นที่ได้มาปรับปรุงแก้ไข จนกระทั่งโปรแกรมมีความสมบูรณ์มากขึ้น

5.2 การจัดทำคู่มือการใช้งาน (User's Manual)

เมื่อทดสอบและปรับปรุงจุดบกพร่องของโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทยเรียบร้อยแล้ว จึงจัดทำคู่มือการใช้งาน ซึ่งอธิบายถึงวิธีการใช้งานอย่างละเอียด โดยแสดงตัวอย่างหน้าจอการใช้งานพร้อมทั้งผลลัพธ์ที่เกิดขึ้นจากระบบ ประกอบคำอธิบาย สำหรับผู้ใช้งาน และผู้ดูแลระบบที่เกี่ยวข้อง เพื่อให้ผู้ใช้สามารถใช้อย่างเป็นแนวทางในการใช้งานโปรแกรมได้อย่างถูกต้องและรวดเร็ว

5.3 การประเมินผลระบบ (System Evaluation)

การประเมินผลโปรแกรม แบ่งเป็นสองส่วน คือ การประเมินผลโดยผู้เชี่ยวชาญ และการประเมินผลโดยผู้ใช้งาน ซึ่งมีรายละเอียด ดังนี้

5.3.1 การประเมินผลโดยผู้เชี่ยวชาญ

เมื่อพัฒนา ทดสอบ แก้ไข และจัดทำคู่มือการใช้งานโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทยเรียบร้อยแล้ว ผู้วิจัยนำโปรแกรมและคู่มือการใช้งานที่พัฒนาขึ้น เสนอให้ผู้เชี่ยวชาญพิจารณา โดยกำหนดเกณฑ์ว่าผู้เชี่ยวชาญต้องมีความรู้ทางด้านทฤษฎี หรือ ด้านการพัฒนาโปรแกรมคอมพิวเตอร์ มีประสบการณ์ทางการวัดผล หรือ ด้านการพัฒนาโปรแกรมคอมพิวเตอร์ไม่น้อยกว่า 10 ปี และมีวุฒิการศึกษาไม่ต่ำกว่าระดับปริญญาโท ผู้เชี่ยวชาญประกอบด้วย

- 1) ดร.ปิยะทิพย์ ตินวร อาจารย์ประจำวิทยาลัยวิทยาการวิจัยและวิทยาการปัญญา มหาวิทยาลัยบูรพา จังหวัดชลบุรี
- 2) ดร.โสฬส สุขานนท์สวัสดิ์ นักประมวลผลข้อมูล 6 ฝ่ายวางแผนและบริหารทรัพยากรสารสนเทศ การไฟฟ้านครหลวง
- 3) ดร.ทิพย์ ขำอยู่ อาจารย์ประจำคณะครุศาสตร์ มหาวิทยาลัยราชภัฏสวนดุสิต กรุงเทพมหานคร

ผู้เชี่ยวชาญประเมินความเหมาะสมของโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย ใน 4 ด้าน ได้แก่ 1) ด้านความสะดวกในการนำไปใช้ 2) ด้านความถูกต้องในการใช้งาน 3) ด้านลักษณะทั่วไปของโปรแกรม และ 4) ด้านความชัดเจนของคู่มือการใช้งาน โดยใช้แบบประเมินผลที่ผู้วิจัยพัฒนาขึ้น ซึ่งมีลักษณะเป็นมาตราประมาณค่า 5 ระดับ โดยมีเกณฑ์การให้คะแนน ดังนี้

เหมาะสมมากที่สุด	ให้คะแนน	5
เหมาะสมมาก	ให้คะแนน	4
เหมาะสมปานกลาง	ให้คะแนน	3
เหมาะสมน้อย	ให้คะแนน	2
เหมาะสมน้อยที่สุด	ให้คะแนน	1

และมีเกณฑ์การแปลความหมายค่าเฉลี่ย ดังนี้

ค่าเฉลี่ยตั้งแต่ 4.51 ถึง 5.00 หมายถึง โปรแกรมมีความเหมาะสมในระดับมากที่สุด
 ค่าเฉลี่ยตั้งแต่ 3.51 ถึง 4.50 หมายถึง โปรแกรมมีความเหมาะสมในระดับมาก
 ค่าเฉลี่ยตั้งแต่ 2.51 ถึง 3.50 หมายถึง โปรแกรมมีความเหมาะสมในระดับปานกลาง
 ค่าเฉลี่ยตั้งแต่ 1.51 ถึง 2.50 หมายถึง โปรแกรมมีความเหมาะสมในระดับน้อย
 ค่าเฉลี่ยตั้งแต่ 1.00 ถึง 1.50 หมายถึง โปรแกรมมีความเหมาะสมในระดับน้อยที่สุด

5.3.2 การประเมินผลโดยผู้ใช้งาน

การประเมินผลโดยผู้ใช้งาน เป็นการประเมินความคิดเห็นในด้านความสะดวกในการนำไปใช้ และด้านลักษณะทั่วไปของโปรแกรม โดยกลุ่มตัวอย่างที่ทดลองใช้ เป็นคนไทยอายุตั้งแต่ 18 ปีขึ้นไป ซึ่งสามารถใช้คอมพิวเตอร์เบื้องต้นได้ จำนวน 30 คน คัดเลือกด้วยวิธีการเลือกตัวอย่างตามสะดวก การประเมินผลดำเนินการภายหลังจากผู้ใช้งานได้ทดลองใช้โปรแกรมผ่านทางเว็บไซต์ www.thscat.com/Test เรียบร้อยแล้ว โดยใช้แบบประเมินผลที่ผู้วิจัยพัฒนาขึ้น ในลักษณะมาตราประมาณค่า 5 ระดับ โดยมีเกณฑ์การให้คะแนน ดังนี้

เหมาะสมมากที่สุด	ให้คะแนน	5
เหมาะสมมาก	ให้คะแนน	4
เหมาะสมปานกลาง	ให้คะแนน	3
เหมาะสมน้อย	ให้คะแนน	2
เหมาะสมน้อยที่สุด	ให้คะแนน	1

และมีเกณฑ์การแปลความหมายค่าเฉลี่ย ดังนี้

ค่าเฉลี่ยตั้งแต่ 4.51 ถึง 5.00 หมายถึง โปรแกรมมีความเหมาะสมในระดับมากที่สุด
 ค่าเฉลี่ยตั้งแต่ 3.51 ถึง 4.50 หมายถึง โปรแกรมมีความเหมาะสมในระดับมาก
 ค่าเฉลี่ยตั้งแต่ 2.51 ถึง 3.50 หมายถึง โปรแกรมมีความเหมาะสมในระดับปานกลาง
 ค่าเฉลี่ยตั้งแต่ 1.51 ถึง 2.50 หมายถึง โปรแกรมมีความเหมาะสมในระดับน้อย
 ค่าเฉลี่ยตั้งแต่ 1.00 ถึง 1.50 หมายถึง โปรแกรมมีความเหมาะสมในระดับน้อยที่สุด

6. ขั้นตอนการบำรุงรักษาระบบ (System Maintenance Phase)

การบำรุงรักษาระบบ เป็นการติดตามผลการใช้งานและให้ความช่วยเหลือแก่ผู้ใช้งาน เพื่อให้สามารถใช้งานได้อย่างต่อเนื่อง และมีประสิทธิภาพ โดยดำเนินการพร้อมกับการประเมินผลระบบ จนสิ้นสุดการวิจัย

โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทยที่พัฒนาขึ้น จะมีการประมาณค่าคุณลักษณะแฝง (ความสุข) ในแต่ละมาตรวัดย่อย หากผู้ทดสอบดำเนินการทดสอบครบทั้งสามมาตรวัดย่อย จะมีการคำนวณค่าเฉลี่ยคุณลักษณะแฝงของทั้งสามมาตรวัด เพื่อเป็นตัวชี้วัดความสุขโดยรวมของผู้ทดสอบ ค่าคุณลักษณะแฝงที่ประมาณค่าได้ในแต่ละมาตรวัดย่อย และค่าเฉลี่ยคุณลักษณะแฝงของมาตรวัดความสุขโดยรวม สามารถแปลความหมายได้ โดยการเปรียบเทียบกับเกณฑ์การแปลความหมายที่แบ่งออกเป็น 9 ช่วงตามคะแนนสเตโน

ขั้นตอนที่ 3 การศึกษาความสัมพันธ์ของผลการประมาณค่าความสุขที่ได้จากการทดสอบด้วยโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย กับการทดสอบด้วยข้อคำถามทั้งหมดในคลังข้อคำถาม

การประมาณค่าความสุขจากการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์เป็นการประมาณค่าความสุขที่ใช้ผลการตอบข้อคำถามเพียงจำนวนหนึ่งจากข้อคำถามทั้งหมดที่มีอยู่ในคลังข้อคำถามเป็นข้อมูลในการประมาณค่า ข้อคำถามที่ใช้เป็นข้อคำถามที่ผ่านกระบวนการเลือกและตรวจสอบด้วยกลไกของโปรแกรมแล้วว่าจะมีความเหมาะสมกับผู้ทดสอบ การดำเนินการทดสอบจะดำเนินการอย่างต่อเนื่องจนกระทั่งถึงเกณฑ์ยุติการทดสอบ โดยทั่วไปโปรแกรมจะยุติการทดสอบด้วยเกณฑ์ความคลาดเคลื่อนมาตรฐานของการประมาณค่าที่ยอมรับได้ก่อนที่ข้อคำถามจะหมดคลัง อย่างไรก็ตาม บางแนวคิดเชื่อว่าหากข้อคำถามที่ใช้ในการทดสอบมีจำนวนมากขึ้นอาจทำให้ได้ข้อมูลการทดสอบที่ละเอียดและแม่นยำเพิ่มขึ้น การศึกษาความสัมพันธ์ของผลการประมาณค่าความสุขที่ได้จากการทดสอบทั้งสองรูปแบบนี้จึงเป็นการตรวจสอบว่าการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ซึ่งใช้ข้อคำถามในการประมาณค่าความสุขน้อยกว่าการทดสอบด้วยข้อคำถามทั้งหมดในคลังข้อคำถามสามารถประมาณค่าความสุขได้ใกล้เคียงกับการทดสอบด้วยข้อคำถามทั้งหมดในคลังข้อคำถามหรือไม่ ด้วยการศึกษผลการประมาณค่าที่ได้จากการทดสอบทั้งสองรูปแบบ

ประชากรและกลุ่มตัวอย่าง

ประชากรในการวิจัย เป็นประชาชนที่มีสัญชาติไทย อายุตั้งแต่ 18 ปีขึ้นไป ใช้ภาษาไทยในการติดต่อสื่อสาร สามารถใช้คอมพิวเตอร์เบื้องต้นได้ กลุ่มตัวอย่างคัดเลือกมาจากประชากรด้วยวิธีการเลือกตัวอย่างแบบตามสะดวก จำนวน 30 คน

เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการวิจัยประกอบด้วยโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทยซึ่งพัฒนาขึ้นในรูปแบบเว็บแอปพลิเคชัน สามารถใช้งานผ่านทางเว็บไซต์ <http://www.thscat.com/Test/> โดยมีรูปแบบการทดสอบสองรูปแบบ ได้แก่

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

การเก็บรวบรวมข้อมูล

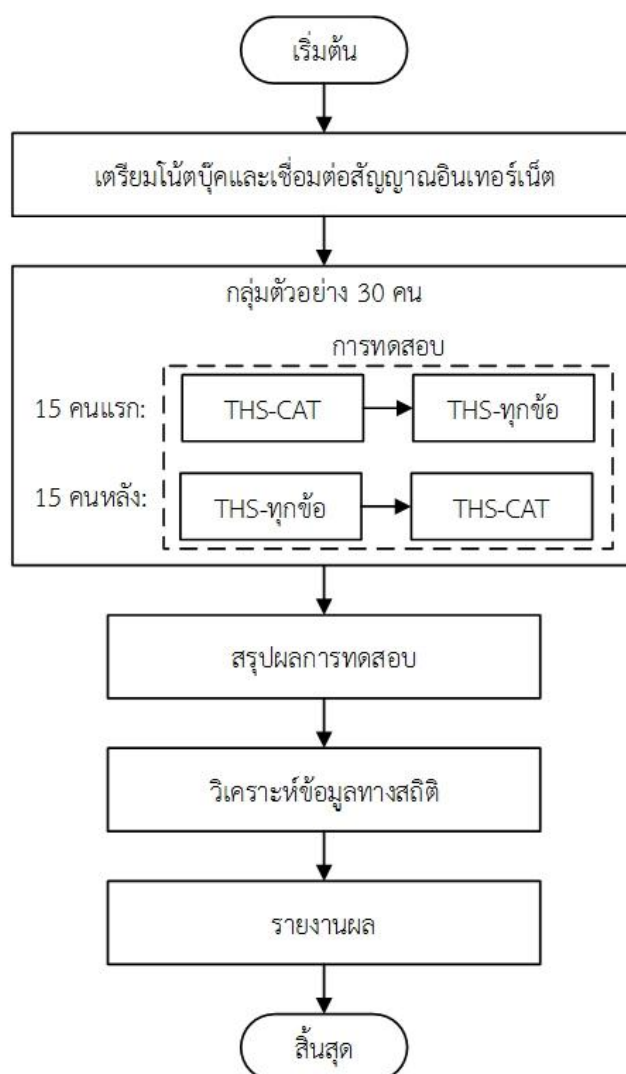
ดำเนินการเก็บรวบรวมข้อมูลด้วยตัวเองตามสถานที่ต่าง ๆ เพื่อให้ได้กลุ่มตัวอย่างที่มีความหลากหลายทางด้านเพศ อายุ และระดับการศึกษา การเก็บข้อมูลใช้โน้ตบุ๊กส่วนตัวของผู้วิจัย และเชื่อมต่ออินเทอร์เน็ตผ่านเครือข่ายโทรศัพท์มือถือที่ความเร็ว 3G-4G ทำให้สามารถเก็บข้อมูลได้ทุกที่ทุกเวลา การเก็บข้อมูลดำเนินการแบบรายบุคคล แต่แต่ละบุคคลจะได้รับการทดสอบทั้งสองรูปแบบโดยกลุ่มตัวอย่าง 15 คนแรกจะได้รับการทดสอบด้วย THS-CAT ก่อนแล้วตามด้วย THS-ทุกข้อ และกลุ่มตัวอย่าง 15 คนหลังจะได้รับการทดสอบด้วย THS-ทุกข้อ ก่อนแล้วตามด้วย THS-CAT ทั้งนี้เพื่อจัดการปัญหาเรื่องลำดับของการทดสอบ (Order Effect) ที่อาจส่งผลกระทบต่อคะแนนการทดสอบได้

การวิเคราะห์ข้อมูล

1) วิเคราะห์ข้อมูลการทดสอบทั่วไปด้วยสถิติบรรยาย (Descriptive Statistic) ได้แก่ ความถี่ ร้อยละ ค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐาน

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

การวิเคราะห์ข้อมูลทั้งสองส่วนดำเนินการด้วยโปรแกรมสำเร็จรูป SPSS การวิจัยในขั้นตอนนี้ แสดงได้ดังภาพที่ 3-21



ภาพที่ 3-21 ขั้นตอนการศึกษาความสัมพันธ์ของผลการประมาณค่าความสุขที่ได้จากการทดสอบด้วยโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์กับการทดสอบด้วยข้อคำถามทั้งหมดในคลังข้อคำถาม

บทที่ 4 ผลการวิจัย

การประยุกต์โมเดลการตอบสนองข้อสอบ ในการพัฒนามาตรวัดความสุขของคนไทย รูปแบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ มีวัตถุประสงค์เพื่อสร้างคลังข้อคำถามที่มีคุณภาพ ด้วยการวิเคราะห์คุณภาพรายข้อของข้อคำถามตามทฤษฎีการตอบสนองข้อสอบและตรวจสอบความตรงเชิงโครงสร้างของข้อคำถามในคลังข้อคำถาม นอกจากนี้ ได้พัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับมาตรวัดความสุขของคนไทยตามแนวคิดของวงจรการพัฒนาระบบ และศึกษาความสัมพันธ์ของผลการประมาณค่าความสุขที่ได้จากการทดสอบด้วยโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย กับการทดสอบด้วยข้อคำถามทั้งหมดในคลังข้อคำถาม ผลการศึกษาวิจัยครั้งนี้แบ่งเป็น 3 ตอน ดังนี้

ตอนที่ 1 ผลการสร้างคลังข้อคำถามสำหรับมาตรวัดความสุขของคนไทย รูปแบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์

ตอนที่ 2 ผลการพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย

ตอนที่ 3 ผลการศึกษาความสัมพันธ์ของผลการประมาณค่าความสุขที่ได้จากการทดสอบด้วยโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย กับการทดสอบด้วยข้อคำถามทั้งหมดในคลังข้อคำถาม

สัญลักษณ์และอักษรย่อที่ใช้ในการเสนอผลการวิเคราะห์ข้อมูล มีดังนี้

χ^2	หมายถึง	ค่าสถิติไค-สแควร์
df	หมายถึง	องศาอิสระ
CFI	หมายถึง	ค่าดัชนีวัดระดับความสอดคล้องเปรียบเทียบ
GFI	หมายถึง	ค่าดัชนีวัดระดับความกลมกลืน
AGFI	หมายถึง	ค่าดัชนีวัดระดับความกลมกลืนที่ปรับแก้
RMSEA	หมายถึง	ค่าความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์
SRMR	หมายถึง	ค่าดัชนีรากของค่าเฉลี่ยกำลังสองของส่วนที่เหลือมาตรฐาน
α	หมายถึง	ค่าพารามิเตอร์ความชัน
β	หมายถึง	ค่าพารามิเตอร์เทรซโฮลด์
$p - value$	หมายถึง	ค่าความน่าจะเป็น
SE	หมายถึง	ค่าความคลาดเคลื่อนมาตรฐานของค่าเฉลี่ย
t	หมายถึง	ค่าสถิติทดสอบที

r	หมายถึง	ค่าสัมประสิทธิ์สหสัมพันธ์เพียร์สัน
R^2	หมายถึง	ค่าสัมประสิทธิ์การทำนาย
Z-score	หมายถึง	คะแนนซี
T-score	หมายถึง	คะแนนที
Stanine	หมายถึง	คะแนนสเตโนน
Min	หมายถึง	คะแนนต่ำสุด
Max	หมายถึง	คะแนนสูงสุด
M	หมายถึง	ค่าเฉลี่ย
SD	หมายถึง	ส่วนเบี่ยงเบนมาตรฐาน
THS-CAT	หมายถึง	มาตรวัดความสุขของคนไทยรูปแบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์
THS-ทุกข้อ	หมายถึง	การทดสอบด้วยข้อคำถามทั้งหมดในคลังข้อคำถามของมาตรวัดความสุขของคนไทย

ตอนที่ 1 ผลการสร้างคลังข้อคำถามสำหรับมาตรวัดความสุขของคนไทย รูปแบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์

การสร้างคลังข้อคำถาม เป็นองค์ประกอบสำคัญ สำหรับการพัฒนามาตรวัดความสุข รูปแบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ผู้วิจัยนำเสนอผลการวิจัยเป็น 4 ประเด็น ดังนี้

1. ผลการสร้างข้อคำถามสำหรับวัดความสุขของคนไทย
2. ผลการตรวจสอบคุณภาพรายข้อของข้อคำถาม
3. ผลการตรวจสอบความตรงเชิงโครงสร้างของข้อคำถามในคลังข้อคำถาม
4. ผลการจัดทำคลังข้อคำถามสำหรับวัดความสุข

1. ผลการสร้างข้อคำถามสำหรับวัดความสุขของคนไทย

ข้อคำถาม สำหรับมาตรวัดความสุขของคนไทย สร้างขึ้นจากทฤษฎีความอยู่ดีมีสุขเชิงอัตวิสัย มีลักษณะการตอบแบบมาตราประมาณค่า 5 ระดับ จำนวนข้อคำถามที่สร้างขึ้นในแต่ละมาตรวัดย่อย แสดงได้ดังตารางที่ 4-1

ตารางที่ 4-1 จำนวนข้อคำถามที่สร้างขึ้นตามทฤษฎีความอยู่ดีมีสุขเชิงอัตวิสัย

มาตรวัดย่อย	จำนวนข้อคำถาม (ข้อ)
ความพึงพอใจในชีวิต	30
อารมณ์ความรู้สึกเชิงบวก	67
อารมณ์ความรู้สึกเชิงลบ	94
รวม	191

2. ผลการตรวจสอบคุณภาพรายข้อของข้อคำถาม

การตรวจสอบคุณภาพรายข้อของข้อคำถามประกอบการตรวจสอบหลายขั้นตอน ได้แก่ 1) การตรวจสอบความตรงเชิงเนื้อหา 2) การประเมินความเหมาะสมของข้อคำถามด้วยการทดลองใช้ 3) การวิเคราะห์ค่า Item-Total Correlations 4) การตรวจสอบข้อตกลงเบื้องต้นของทฤษฎีการตอบสนองข้อสอบ และ 5) การตรวจสอบความสอดคล้องของข้อมูลเชิงประจักษ์กับ Graded Response Model รวมทั้งการประมาณค่าพารามิเตอร์ของข้อคำถาม (พารามิเตอร์ความชันและพารามิเตอร์เทอร์ซโฮลด์)

2.1 ผลการตรวจสอบความตรงเชิงเนื้อหา ด้วยการพิจารณาความสอดคล้องระหว่างข้อคำถามกับนิยามเชิงปฏิบัติการโดยผู้เชี่ยวชาญ ปรากฏผลการคัดเลือกข้อคำถามที่มีความเหมาะสม ดังตารางที่ 4-2

ตารางที่ 4-2 ผลการคัดเลือกข้อคำถามจากการประเมินความตรงเชิงเนื้อหาโดยผู้เชี่ยวชาญ

มาตรวัดย่อย	จำนวนข้อคำถาม (ข้อ)			
	เริ่มต้น	คัดออก	ปรับปรุง	คงเหลือ
ความพึงพอใจในชีวิต	30	0	13	30
อารมณ์ความรู้สึกเชิงบวก	67	3	5	64
อารมณ์ความรู้สึกเชิงลบ	94	0	1	94
รวม	191	3	19	188

จากตารางที่ 4-2 ข้อคำถามจำนวน 3 ข้อในมาตรวัดย่อยอารมณ์ความรู้สึกเชิงบวก มีค่าเฉลี่ยความสอดคล้องระหว่างข้อคำถามกับนิยามเชิงปฏิบัติการต่ำกว่า 4.00 จึงถูกคัดออกจากคลังข้อคำถาม ในขณะที่ ข้อคำถามจำนวน 19 ข้อ ในทั้งสามมาตรวัดย่อย ได้รับการปรับปรุงตามคำแนะนำของผู้เชี่ยวชาญเพื่อให้ข้อคำถามมีความเหมาะสมยิ่งขึ้น และคงไว้ในคลังข้อคำถาม (รายละเอียดค่าเฉลี่ยความสอดคล้องของข้อคำถามกับนิยามเชิงปฏิบัติการ แสดงในภาคผนวก ค)

2.2 ผลการประเมินความเหมาะสมของข้อคำถามด้วยการทดลองใช้ ได้นำข้อคำถามที่ผ่านการตรวจสอบความตรงเชิงเนื้อหาโดยผู้เชี่ยวชาญไปทดลองใช้ (188 ข้อ) ซึ่งผลการคัดเลือกข้อคำถามโดยพิจารณาผลที่ได้จากการทดลองใช้ แสดงได้ดังตารางที่ 4-3

ตารางที่ 4-3 ผลการคัดเลือกข้อคำถามจากการทดลองใช้

มาตรวัดย่อย	จำนวนข้อคำถาม (ข้อ)			
	เริ่มต้น	คัดออก	ปรับปรุง	คงเหลือ
ความพึงพอใจในชีวิต	30	1	2	29
อารมณ์ความรู้สึกเชิงบวก	64	1	0	63
อารมณ์ความรู้สึกเชิงลบ	94	4	2	90
รวม	188	6	4	182

จากตารางที่ 4-3 ข้อคำถามจากมาตรวัดย่อยความพึงพอใจในชีวิต 1 ข้อ ถูกคัดออกจากคลังข้อคำถาม เนื่องจากมีค่า Item-Total Correlations ของข้อคำถามติดลบ ขณะที่ข้อคำถามจากมาตรวัดย่อยอารมณ์ความรู้สึกเชิงบวก 1 ข้อ และมาตรวัดย่อยอารมณ์ความรู้สึกเชิงลบ 4 ข้อ ถูกคัดออกจากคลังข้อคำถาม เนื่องจากข้อคำถามดังกล่าวสื่อความหมายได้ไม่ชัดเจน และมีความซ้ำซ้อนกับข้อคำถามอื่น นอกจากนี้ มีการปรับปรุงภาษาที่ใช้ในข้อคำถามจำนวน 4 ข้อ เพื่อให้เข้าใจง่ายขึ้น จึงคงเหลือข้อคำถาม ในคลังข้อคำถามสำหรับการนำไปเก็บข้อมูลกับกลุ่มตัวอย่าง จำนวนทั้งสิ้น 182 ข้อ (รายละเอียดค่า Item-Total Correlations รายข้อและข้อคิดเห็นที่ได้จากการทดลองใช้ แสดงในภาคผนวก ค)

2.3 ผลการวิเคราะห์ค่า Item-Total Correlations ผู้วิจัยนำข้อคำถาม 182 ข้อ ซึ่งผ่านการตรวจสอบความตรงเชิงเนื้อหาโดยผู้เชี่ยวชาญและผ่านการทดลองใช้แล้ว ไปจัดพิมพ์ในรูปแบบกระดาษ-ดินสอเพื่อเก็บข้อมูลกับกลุ่มตัวอย่างจำนวนทั้งสิ้น 1,000 คน ปรากฏว่า ได้ข้อมูลการตอบข้อคำถามที่มีความสมบูรณ์เพียงพอสำหรับการนำไปวิเคราะห์คุณภาพของข้อคำถามและมาตรวัด จำนวนทั้งสิ้น 934 ฉบับ คิดเป็นร้อยละ 93.40 ลักษณะของกลุ่มตัวอย่าง แสดงได้ดังตารางที่ 4-4

ตารางที่ 4-4 ลักษณะของกลุ่มตัวอย่าง

ลักษณะของกลุ่มตัวอย่าง	จำนวน (คน)	ร้อยละ
เพศ		
ชาย	285	30.5
หญิง	649	69.5
อายุ		
18-25 ปี	186	19.9
26-40 ปี	525	56.2
41-60 ปี	212	22.7
61 ปีขึ้นไป	11	1.2

ตารางที่ 4-4 (ต่อ)

ลักษณะของกลุ่มตัวอย่าง	จำนวน (คน)	ร้อยละ
สถานภาพสมรส		
โสด	446	47.8
แต่งงานและอยู่ร่วมกัน	372	39.8
แต่งงาน แต่ไม่ได้อยู่ด้วยกัน	32	3.4
อยู่ด้วยกัน โดยไม่ได้แต่งงาน	41	4.4
หม้าย	13	1.4
หย่า/ แยกทาง/ เลิกกัน	30	3.2
การศึกษาสูงสุด		
ประถมศึกษา	24	2.6
มัธยมศึกษาตอนต้น	52	5.6
มัธยมศึกษาตอนปลาย/ ปวช.	157	16.8
อนุปริญญา หรือเทียบเท่า	82	8.8
ปริญญาตรี	519	55.6
สูงกว่าปริญญาตรี	100	10.7
อาชีพหลัก		
ว่างงาน	8	0.9
รับจ้างทั่วไป	51	5.5
พนักงานบริษัทเอกชน	328	35.1
รับราชการ/ พนักงานของรัฐ	364	39.0
ธุรกิจส่วนตัว/ ค้าขาย	81	8.7
นักเรียน/ นักศึกษา	60	6.4
อื่น ๆ	42	4.5
ฐานะทางเศรษฐกิจ		
พอมีพอใช้	573	61.3
มีเหลือเก็บ	152	16.3
มีหนี้สิน	209	22.4
ความรับผิดชอบครอบครัว		
เป็นหลักของครอบครัว	281	30.1
ต้องช่วยเหลือครอบครัวบางส่วน	496	53.1
ไม่มีภาระดูแลครอบครัว	157	16.8
รวม	934	100.0

จากตารางที่ 4-4 ปรากฏว่า กลุ่มตัวอย่างส่วนใหญ่ เป็นเพศหญิง (ร้อยละ 69.5) อายุระหว่าง 26-40 ปี (ร้อยละ 56.2) สถานภาพโสด (ร้อยละ 47.8) การศึกษาสูงสุดระดับปริญญาตรี (ร้อยละ 55.6) อาชีพรับราชการ/พนักงานของรัฐ (ร้อยละ 39.0) ฐานะทางเศรษฐกิจพอมีพอใช้ (ร้อยละ 61.3) และมีความรับผิดชอบต้องช่วยเหลือครอบครัวบางส่วน (ร้อยละ 53.1)

การวิเคราะห์ค่า Item-Total Correlations ของข้อคำถาม ในแต่ละมาตรวัดย่อย และมาตรวัดความสุขโดยรวม จากกลุ่มตัวอย่าง 934 คน ปรากฏผลดังตารางที่ 4-5

ตารางที่ 4-5 ค่า Item-Total Correlations ของข้อคำถาม ในแต่ละมาตรวัดย่อย และมาตรวัดความสุขโดยรวม

มาตรวัดย่อย	จำนวนข้อ	ค่า Item-Total Correlations	
		ต่ำสุด	สูงสุด
ความพึงพอใจในชีวิต	29	0.14	0.64
อารมณ์ความรู้สึกเชิงบวก	63	0.06	0.58
อารมณ์ความรู้สึกเชิงลบ	90	0.08	0.59
ผลรวม	182	0.06	0.64

จากตารางที่ 4-5 แสดงให้เห็นว่า ในแต่ละมาตรวัดย่อย มีข้อคำถามบางข้อที่มีค่า Item-Total Correlations เท่ากับ 0.30 หรือต่ำกว่า (รายละเอียดค่า Item-Total Correlations รายข้อ แสดงในภาคผนวก ค) ข้อคำถามดังกล่าวจึงถูกคัดออกจากคลังข้อคำถาม ทำให้มีจำนวนข้อคำถามคงเหลือในคลังข้อคำถามของแต่ละมาตรวัดย่อย ดังตารางที่ 4-6

ตารางที่ 4-6 ผลการคัดเลือกข้อคำถาม จากค่า Item-Total Correlations

มาตรวัดย่อย	จำนวนข้อคำถาม (ข้อ)		
	เริ่มต้น	คัดออก	คงเหลือ
ความพึงพอใจในชีวิต	29	4	25
อารมณ์ความรู้สึกเชิงบวก	63	22	41
อารมณ์ความรู้สึกเชิงลบ	90	34	56
รวม	182	60	122

จากตารางที่ 4-6 ปรากฏว่าข้อคำถามที่มีค่า Item-Total Correlations ต่ำกว่าเกณฑ์ที่กำหนดไว้ (≤ 0.30) ในมาตรวัดย่อยความพึงพอใจในชีวิต มีจำนวน 4 ข้อ มาตรวัดย่อยอารมณ์ความรู้สึกเชิงบวก มีจำนวน 22 ข้อ และมาตรวัดอารมณ์ความรู้สึกเชิงลบ มีจำนวน 34 ข้อ รวมข้อ

คำถามที่ถูกคัดออกจากคลังข้อคำถาม จำนวน 60 ข้อ และคงเหลือข้อคำถามในคลังข้อคำถาม จำนวนทั้งสิ้น 122 ข้อ

2.4 ผลการตรวจสอบข้อตกลงเบื้องต้นของทฤษฎีการตอบสนองข้อสอบ แบ่งออกเป็นสองส่วน ดังนี้

1) ผลการตรวจสอบความถี่ของการเลือกรายการคำตอบ ปรากฏว่าไม่มีตัวเลือก รายการคำตอบใดที่มีความถี่เป็นศูนย์ จึงสามารถนำข้อคำถามทั้งหมดไปวิเคราะห์ในขั้นต่อไปได้ (รายละเอียดความถี่ของการเลือกรายการคำตอบ แสดงในภาคผนวก ค)

2) ผลการตรวจสอบข้อตกลงเบื้องต้น ในด้านความเป็นเอกมิติ ด้วยการวิเคราะห์องค์ประกอบเชิงยืนยัน ปรากฏว่า ในแต่ละมาตรวัดย่อยมีเพียงหนึ่งองค์ประกอบ โดยค่าดัชนีความสอดคล้องของข้อมูลเชิงประจักษ์กับโมเดลองค์ประกอบเชิงยืนยัน ในการตรวจสอบความเป็นเอกมิติ แสดงได้ดังตารางที่ 4-7

ตารางที่ 4-7 ดัชนีความสอดคล้องของข้อมูลเชิงประจักษ์กับโมเดลองค์ประกอบเชิงยืนยัน ในการตรวจสอบความเป็นเอกมิติ

ดัชนีความ สอดคล้อง	มาตรวัดย่อย		
	ความพึงพอใจในชีวิต	อารมณ์ความรู้สึกเชิงบวก	อารมณ์ความรู้สึกเชิงลบ
χ^2	186.69 ($p=0.167$)	523.973 ($p=0.879$)	375.751 ($p=1.00$)
df	169	563	533
CFI	1.00	1.00	1.00
GFI	0.99	0.98	0.99
AGFI	0.97	0.95	0.96
RMSEA	0.01	0.00	0.00
SRMR	0.03	0.02	0.02

จากตารางที่ 4-7 แสดงให้เห็นว่าข้อมูลเชิงประจักษ์มีความสอดคล้องกับโมเดลองค์ประกอบเชิงยืนยันในทุกมาตรวัดย่อย ซึ่งพิจารณาได้จากค่า χ^2 ที่ไม่พบความแตกต่างระหว่างโมเดลกับข้อมูลเชิงประจักษ์ นอกจากนี้ ค่าดัชนีความสอดคล้องทุกค่าอยู่ในระดับดี แสดงให้เห็นจากค่า CFI GFI และ AGFI มีค่ามากกว่า 0.95 ขณะที่ค่า RMSEA และ SRMR มีค่าน้อยกว่า 0.05 (Schumacker & Lomax, 2004, p. 82) ซึ่งเป็นข้อมูลยืนยันว่ามาตรวัดย่อยแต่ละมาตรวัด มีหนึ่งองค์ประกอบ และสะท้อนถึงความเป็นเอกมิติของแต่ละมาตรวัดย่อย

การตรวจสอบความเป็นอิสระของข้อคำถามจากการพิจารณาค่าสหสัมพันธ์เศษเหลือที่ได้จากการวิเคราะห์องค์ประกอบเชิงยืนยัน ปรากฏผลดังตารางที่ 4-8

ตารางที่ 4-8 ค่าสหสัมพันธ์พิเศษเหลือจากการวิเคราะห์องค์ประกอบเชิงยืนยันมาตรวัดความสุขของคนไทย

มาตรวัดย่อย	ค่าสหสัมพันธ์พิเศษเหลือ		
	ค่าต่ำสุด	ค่ากลาง	ค่าสูงสุด
ความพึงพอใจในชีวิต	-0.107	0.000	0.082
อารมณ์ความรู้สึกเชิงบวก	-0.080	0.002	0.093
อารมณ์ความรู้สึกเชิงลบ	-0.062	0.000	0.066

จากตารางที่ 4-8 มาตรวัดความพึงพอใจในชีวิต มีค่าสหสัมพันธ์พิเศษเหลืออยู่ระหว่าง -0.107 ถึง 0.082 มาตรวัดอารมณ์ความรู้สึกเชิงบวก มีค่าสหสัมพันธ์พิเศษเหลืออยู่ระหว่าง -0.080 ถึง 0.093 และมาตรวัดอารมณ์ความรู้สึกเชิงลบ มีค่าสหสัมพันธ์พิเศษเหลืออยู่ระหว่าง -0.062 ถึง 0.066 ซึ่งค่าสหสัมพันธ์พิเศษเหลือของข้อคำถามทุกข้อในสามมาตรวัดย่อย มีค่าต่ำกว่า 0.25 (Fliege et al., 2005; Walter, Becker, & Bjorner, 2007, Kocalevant et al., 2009) แสดงให้เห็นถึงความเป็นอิสระของข้อคำถาม จากผลดังกล่าวเป็นข้อมูลสนับสนุนว่าข้อคำถามทุกข้อผ่านเกณฑ์ข้อตกลงเบื้องต้นของทฤษฎีการตอบสนองข้อสอบ ซึ่งทำให้มั่นใจได้ว่าข้อคำถามดังกล่าว สามารถนำไปวิเคราะห์ค่าพารามิเตอร์ของข้อคำถามตามทฤษฎีการตอบสนองข้อสอบได้ในขั้นต่อไป

2.5 ผลการตรวจสอบความสอดคล้องของข้อมูลเชิงประจักษ์กับ Graded Response Model และการประมาณค่าพารามิเตอร์ของข้อคำถาม (พารามิเตอร์ความชันและพารามิเตอร์เทอร์โซลด์) ผู้วิจัยคัดเลือกข้อคำถามเข้าสู่คลังข้อคำถามสุดท้าย (Final Item Bank) โดยพิจารณาจากความสอดคล้องของข้อมูลเชิงประจักษ์ที่ได้จากการตอบข้อคำถามกับ Graded Response Model ซึ่งปรากฏผล ดังตารางที่ 4-9

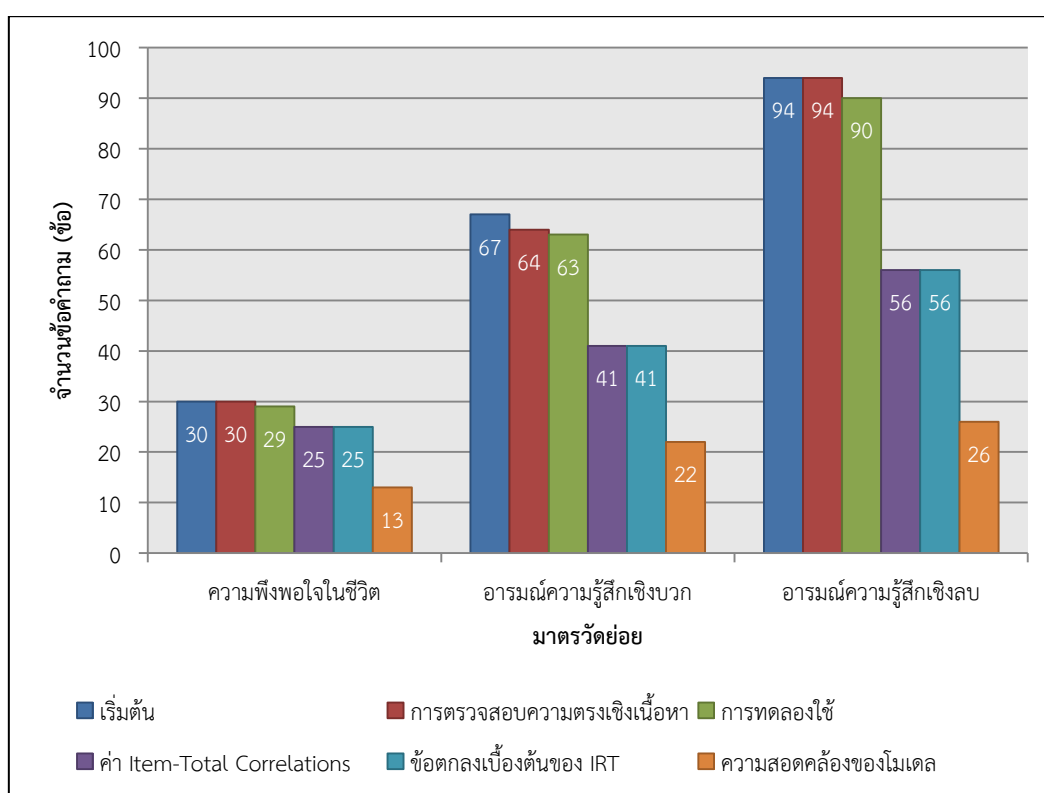
ตารางที่ 4-9 ผลการคัดเลือกข้อคำถาม ตามความสอดคล้องของข้อมูลเชิงประจักษ์ที่ได้จากการตอบข้อคำถามกับ Graded Response Model

มาตรวัดย่อย	จำนวนข้อคำถาม (ข้อ)		
	เริ่มต้น	คัดออก	คงเหลือ
ความพึงพอใจในชีวิต	25	12	13
อารมณ์ความรู้สึกเชิงบวก	41	19	22
อารมณ์ความรู้สึกเชิงลบ	56	30	26
รวม	122	61	61

จากตารางที่ 4-9 ปรากฏว่าข้อคำถามในมาตรวัดย่อยความพึงพอใจในชีวิต จำนวน 12 ข้อ มาตรวัดย่อยอารมณ์ความรู้สึกเชิงบวก จำนวน 19 ข้อ และมาตรวัดย่อยอารมณ์ความรู้สึกเชิงลบ

จำนวน 30 ข้อ มีผลการวิเคราะห์ที่แสดงให้เห็นว่าข้อมูลเชิงประจักษ์ที่ได้จากการตอบข้อคำถามไม่สอดคล้องกับ Graded Response Model จึงถูกคัดออกจากคลังข้อคำถาม ส่งผลให้คลังข้อคำถามสุดท้ายสำหรับมาตรวัดความสุขของคนไทย คงเหลือข้อคำถามจำนวนทั้งสิ้น 61 ข้อ ซึ่งเป็นข้อคำถามในมาตรวัดย่อยความพึงพอใจในชีวิต 13 ข้อ มาตรวัดอารมณ์ความรู้สึกเชิงบวก 22 ข้อ และมาตรวัดอารมณ์ความรู้สึกเชิงลบ 26 ข้อ (รายละเอียดค่าความสอดคล้องของข้อมูลเชิงประจักษ์ที่ได้จากการตอบข้อคำถามกับ Graded Response Model ค่าพารามิเตอร์ความชัน และค่าพารามิเตอร์เทอร์ชโฮลด์ แสดงในภาคผนวก ค)

ผลการคัดเลือกข้อคำถามตามขั้นตอนต่าง ๆ สรุปได้ดังภาพที่ 4-1



ภาพที่ 4-1 จำนวนข้อคำถามที่ผ่านการคัดเลือกตามเกณฑ์ต่าง ๆ ของการสร้างคลังข้อคำถาม

3. ผลการตรวจสอบความตรงเชิงโครงสร้างของข้อคำถามในคลังข้อคำถาม

การตรวจสอบความตรงเชิงโครงสร้างของข้อคำถามในคลังข้อคำถามสำหรับมาตรวัดความสุขของคนไทย ใช้วิธีการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสอง (Second Order Factor Analysis) เพื่อตรวจสอบว่าความสุขประกอบด้วย 3 องค์ประกอบ ได้แก่ ความพึงพอใจในชีวิต อารมณ์ความรู้สึกเชิงบวก และอารมณ์ความรู้สึกเชิงลบ ตามแนวคิดความอยู่ดีมีสุขเชิงอัตวิสัย ซึ่งมีข้อคำถามจำนวนทั้งสิ้น 61 ข้อ ผลการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสอง ชั้นที่ 1 แสดงได้ดังตารางที่ 4-10 และชั้นที่ 2 แสดงได้ดังตารางที่ 4-11

ตารางที่ 4-10 ผลการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองของข้อคำถามในคลังข้อคำถาม
สำหรับวัดความสุขของคนไทย ข้อคำถาม 61 ข้อ (ขั้นที่ 1)

ข้อ ที่	องค์ประกอบ									R ²
	ความพึงพอใจในชีวิต			อารมณ์ความรู้สึกเชิงบวก			อารมณ์ความรู้สึกเชิงลบ			
	น้ำหนัก องค์ประกอบ	SE	t	น้ำหนัก องค์ประกอบ	SE	t	น้ำหนัก องค์ประกอบ	SE	t	
1	0.50*	-	-	-	-	-	-	-	-	0.16
2	0.53*	0.05	9.77	-	-	-	-	-	-	0.30
3	0.48*	0.05	9.17	-	-	-	-	-	-	0.26
4	0.57*	0.06	10.23	-	-	-	-	-	-	0.46
5	0.50*	0.06	8.21	-	-	-	-	-	-	0.25
6	0.36*	0.05	7.22	-	-	-	-	-	-	0.15
7	0.55*	0.06	9.72	-	-	-	-	-	-	0.37
8	0.72*	0.07	9.68	-	-	-	-	-	-	0.46
9	0.50*	0.05	9.44	-	-	-	-	-	-	0.31
10	0.51*	0.05	9.73	-	-	-	-	-	-	0.27
11	0.61*	0.06	9.67	-	-	-	-	-	-	0.39
12	0.62*	0.06	10.18	-	-	-	-	-	-	0.43
13	0.53*	0.05	9.82	-	-	-	-	-	-	0.34
14	-	-	-	0.42*	-	-	-	-	-	0.21
15	-	-	-	0.51*	0.04	12.69	-	-	-	0.37
16	-	-	-	0.49*	0.04	12.15	-	-	-	0.26
17	-	-	-	0.54*	0.04	12.49	-	-	-	0.37
18	-	-	-	0.54*	0.04	12.00	-	-	-	0.40
19	-	-	-	0.37*	0.04	9.53	-	-	-	0.13
20	-	-	-	0.38*	0.04	9.34	-	-	-	0.15
21	-	-	-	0.61*	0.05	12.61	-	-	-	0.48
22	-	-	-	0.64*	0.05	12.70	-	-	-	0.43
23	-	-	-	0.54*	0.04	12.24	-	-	-	0.44
24	-	-	-	0.57*	0.05	12.58	-	-	-	0.42
25	-	-	-	0.52*	0.05	11.38	-	-	-	0.27
26	-	-	-	0.52*	0.04	12.28	-	-	-	0.37
27	-	-	-	0.54*	0.04	12.86	-	-	-	0.41
28	-	-	-	0.47*	0.04	11.24	-	-	-	0.26
29	-	-	-	0.52*	0.04	12.33	-	-	-	0.36

ตารางที่ 4-10 (ต่อ)

ข้อ ที่	องค์ประกอบ									R ²
	ความพึงพอใจในชีวิต			อารมณ์ความรู้สึกเชิงบวก			อารมณ์ความรู้สึกเชิงลบ			
	น้ำหนัก องค์ประกอบ	SE	t	น้ำหนัก องค์ประกอบ	SE	t	น้ำหนัก องค์ประกอบ	SE	t	
30	-	-	-	0.52*	0.04	11.87	-	-	-	0.35
31	-	-	-	0.60*	0.05	12.43	-	-	-	0.49
32	-	-	-	0.63*	0.05	12.82	-	-	-	0.54
33	-	-	-	0.63*	0.05	12.32	-	-	-	0.46
34	-	-	-	0.51*	0.04	12.18	-	-	-	0.31
35	-	-	-	0.56*	0.05	12.25	-	-	-	0.37
36	-	-	-	-	-	-	0.66*	-	-	0.37
37	-	-	-	-	-	-	0.59*	0.04	15.60	0.36
38	-	-	-	-	-	-	0.51*	0.04	14.19	0.37
39	-	-	-	-	-	-	0.63*	0.04	14.59	0.29
40	-	-	-	-	-	-	0.47*	0.04	12.83	0.23
41	-	-	-	-	-	-	0.63*	0.03	17.98	0.36
42	-	-	-	-	-	-	0.73*	0.04	17.38	0.49
43	-	-	-	-	-	-	0.48*	0.04	13.64	0.27
44	-	-	-	-	-	-	0.73*	0.04	18.27	0.47
45	-	-	-	-	-	-	0.43*	0.03	13.78	0.25
46	-	-	-	-	-	-	0.73*	0.04	18.61	0.58
47	-	-	-	-	-	-	0.40*	0.03	11.63	0.17
48	-	-	-	-	-	-	0.65*	0.04	17.54	0.48
49	-	-	-	-	-	-	0.61*	0.04	15.63	0.36
50	-	-	-	-	-	-	0.52*	0.04	14.07	0.28
51	-	-	-	-	-	-	0.73*	0.04	18.28	0.55
52	-	-	-	-	-	-	0.43*	0.03	13.01	0.27
53	-	-	-	-	-	-	0.55*	0.03	16.02	0.40
54	-	-	-	-	-	-	0.57*	0.04	15.63	0.40
55	-	-	-	-	-	-	0.63*	0.04	17.06	0.48
56	-	-	-	-	-	-	0.65*	0.04	16.39	0.46
57	-	-	-	-	-	-	0.56*	0.04	14.14	0.26
58	-	-	-	-	-	-	0.68*	0.04	16.21	0.39
59	-	-	-	-	-	-	0.50*	0.04	13.84	0.32

ตารางที่ 4-10 (ต่อ)

ข้อ ที่	องค์ประกอบ									R^2	
	ความพึงพอใจในชีวิต			อารมณ์ความรู้สึกเชิงบวก			อารมณ์ความรู้สึกเชิงลบ				
	น้ำหนัก องค์ประกอบ	SE	t	น้ำหนัก องค์ประกอบ	SE	t	น้ำหนัก องค์ประกอบ	SE	t		
60	-	-	-	-	-	-	0.68*	0.04	17.23	0.45	
61	-	-	-	-	-	-	0.48*	0.03	15.01	0.32	
Chi-Square = 1067.36 ($p = 0.47$)			$df = 1065$			GFI= 0.96			AGFI= 0.94		CFI= 1.00
Standardized RMR= 0.04			RMSEA= 0.00								

* $p < .01$

ตารางที่ 4-11 ผลการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองของข้อคำถามในคลังข้อคำถาม สำหรับวัดความสุขของคนไทย ข้อคำถาม 61 ข้อ (ขั้นที่ 2)

องค์ประกอบ	ความสุข			R^2		
	น้ำหนัก องค์ประกอบ	SE	t			
ความพึงพอใจในชีวิต (13 ข้อ)	0.89*	0.09	10.09	0.82		
อารมณ์ความรู้สึกเชิงบวก (23 ข้อ)	0.53*	0.05	10.07	0.26		
อารมณ์ความรู้สึกเชิงลบ (26 ข้อ)	0.64*	0.05	13.17	0.40		
Chi-Square = 1067.36 ($p = 0.47$)		$df = 1065$		GFI= 0.96	AGFI= 0.94	CFI= 1.00
Standardized RMR= 0.04		RMSEA= 0.00				

* $p < .01$

จากตารางที่ 4-10 และ 4-11 แสดง ผลการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสองของ ข้อคำถามในคลังข้อคำถาม สำหรับวัดความสุขของคนไทย ข้อคำถาม 61 ข้อ ปรากฏว่าโมเดลมี ความสอดคล้องกับข้อมูลเชิงประจักษ์ ซึ่งพิจารณาได้จากค่า Chi-Square มีค่าเท่ากับ 1067.36; $p = 0.47$ ที่องศาอิสระเท่ากับ 1065 แสดงให้เห็นจากค่า CFI, GFI และ AGFI มีค่ามากกว่า 0.90 (Hooper, Coughlan, & Mullen, 2008) ขณะที่ค่า RMSEA และ Standardized RMR มีค่าน้อย กว่า 0.05 (Schumacker & Lomax, 2004, p. 82) แสดงว่าโมเดลมีความสอดคล้องกับข้อมูล เชิงประจักษ์อยู่ในเกณฑ์ดี

เมื่อพิจารณาค่าน้ำหนักองค์ประกอบแต่ละข้อ ในแต่ละองค์ประกอบ ปรากฏว่า ด้าน ความพึงพอใจในชีวิต ข้อคำถาม 13 ข้อ มีค่าน้ำหนักองค์ประกอบ ตั้งแต่ 0.36 - 0.72 ด้านอารมณ์ ความรู้สึกเชิงบวก ข้อคำถาม 22 ข้อ มีค่าน้ำหนักองค์ประกอบ ตั้งแต่ 0.37 - 0.64 ด้านอารมณ์

ความรู้สึกเชิงลบ ข้อคำถาม 26 ข้อ มีค่าน้ำหนักองค์ประกอบ ตั้งแต่ 0.40 – 0.73 ค่าน้ำหนักองค์ประกอบมีนัยสำคัญทางสถิติที่ระดับ .01 ทุกค่า แสดงว่าข้อคำถามทั้ง 61 ข้อ สามารถวัดได้ตรงตามองค์ประกอบทั้ง 3 องค์ประกอบ

ส่วนค่าน้ำหนักองค์ประกอบแต่ละองค์ประกอบของมาตรวัดมีค่าตั้งแต่ 0.53 – 0.89 จำแนกได้ดังนี้ ด้านความพึงพอใจในชีวิตมีน้ำหนักองค์ประกอบ 0.89 ด้านอารมณ์ความรู้สึกเชิงบวกมีค่าน้ำหนักองค์ประกอบ 0.53 และด้านอารมณ์ความรู้สึกเชิงลบ มีค่าน้ำหนักองค์ประกอบ 0.64 ค่าน้ำหนักองค์ประกอบมีนัยสำคัญทางสถิติที่ระดับ .01 ทุกค่า แสดงว่าโครงสร้างของข้อคำถามในคลังข้อคำถามสำหรับวัดความสุขของคนไทยที่ผู้วิจัยพัฒนาขึ้นประกอบด้วย 3 องค์ประกอบตามทฤษฎีความอยู่ดีมีสุขเชิงอัตวิสัย

4. ผลการจัดทำคลังข้อคำถามสำหรับวัดความสุข

ผู้วิจัยได้รวบรวมข้อคำถามที่ผ่านเกณฑ์การตรวจสอบคุณภาพทุกขั้นตอนมาบันทึกไว้ในรูปแบบไฟล์ เพื่อเตรียมพร้อมสำหรับการพัฒนามาตรวัดความสุขของคนไทยรูปแบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ โดยมีข้อคำถามในคลังข้อคำถามสุดท้าย จำนวนทั้งสิ้น 61 ข้อ ข้อคำถามค่าพารามิเตอร์ความชัน และค่าพารามิเตอร์เทรซโฮลด์ของข้อคำถามที่ได้จากการประมาณค่าด้วย Graded Response Model แสดงได้ดังตารางที่ 4-12

ตารางที่ 4-12 ข้อคำถาม ค่าพารามิเตอร์ความชัน และค่าพารามิเตอร์เทรซโฮลด์ของข้อคำถามในคลังข้อคำถามสุดท้าย

มาตรวัดย่อย/ข้อคำถาม	ค่าพารามิเตอร์ของข้อคำถาม				
	α	β_1	β_2	β_3	β_4
ความพึงพอใจในชีวิต					
1. ถ้าฉันย้อนกลับไปปรับเปลี่ยนตัวเองได้ชีวิตฉันจะดีกว่านี้ (-)	0.4608	3.1790	1.1236	-0.4985	-2.0964
2. ฉันพึงพอใจในสภาพชีวิตที่เป็นอยู่	0.9930	-2.6308	-1.6769	0.0272	1.6477
3. ฉันดำเนินชีวิตได้อย่างที่ต้องการ	0.9696	-2.7491	-1.4157	0.3147	2.1288
4. ฉันมีชีวิตที่มีความสุขสบายตามสมควร	1.2509	-3.0465	-1.8004	-0.2792	1.5843
5. ฉันรู้สึกว่าวันเวลาที่ผ่านไปในแต่ละวันช่างยาวนานเสียเหลือเกิน (-)	0.4423	2.5466	-0.3232	-2.4575	-4.7265
6. ที่ผ่านมา พ่อแม่ หรือคนรอบข้างคอยช่วยเหลือให้กำลังใจฉัน	0.6144	-4.5804	-3.1354	-1.3170	0.5049
7. ฉันรู้สึกภาคภูมิใจในชีวิตที่ผ่านมาของฉัน	1.0960	-3.1246	-2.0540	-0.3592	1.2150
8. ชีวิตฉันที่ผ่านมามีแต่เหตุการณ์ที่เลวร้าย (-)	0.6134	0.8920	-0.8484	-2.3757	-3.6989

ตารางที่ 4-12 (ต่อ)

มาตรวัดย่อย/ข้อคำถาม	ค่าพารามิเตอร์ของข้อคำถาม				
	α	β_1	β_2	β_3	β_4
9. ฉันมีความพร้อมสำหรับการใช้ชีวิตที่ดีในอนาคตข้างหน้า	0.8669	-3.9104	-2.5097	-0.8572	0.8722
10. ฉันกังวลว่าจะไม่มีคนดูแลหรือให้ความช่วยเหลือในอนาคต (-)	0.5332	1.8571	-0.3810	-2.7770	-4.3432
11. ฉันไม่แน่ใจว่าจะใช้ชีวิตต่อไปได้อย่างมีความสุข (-)	0.6923	0.9315	-0.8414	-2.5130	-4.3132
12. ฉันประสบความสำเร็จในชีวิตอย่างที่ครอบครัวคาดหวัง	1.0261	-2.7082	-1.3901	0.3027	1.8422
13. ผู้ที่ฉันนับถือเห็นว่าชีวิตของฉันล้มเหลว (-)	0.6721	0.0107	-1.6428	-3.0400	-4.8119
อารมณ์ความรู้สึกเชิงบวก					
1. เมตตากรุณา	0.7372	-1.8332	0.0159	1.5004	2.7544
2. เข้าอกเข้าใจ	0.7780	-1.3283	0.4642	2.0174	3.2679
3. ซื่อซม	0.9314	-1.6519	-0.1095	1.1302	2.2536
4. ซาบซึ้งใจ	0.8085	-0.5866	0.9719	2.4623	3.4357
5. ดีใจ	0.8110	-1.6893	0.1544	1.8313	3.0799
6. เด็ดเดี่ยว	0.9522	0.0778	1.2794	2.3226	3.2633
7. ตื่นตาตื่นใจ	1.1122	-0.2318	1.0488	2.4484	3.1629
8. เบิกบานใจ	1.1859	-0.9805	0.5987	1.8998	2.7024
9. ผาสุก	1.2637	-0.1622	0.9852	1.9434	2.5769
10. พอใจ	1.2213	-0.5504	0.8075	1.8642	2.6317
11. มั่นคง	1.1558	-0.3911	0.8782	1.8598	2.6942
12. มีเสน่ห์	0.8881	-1.188	0.5096	2.1702	3.2611
13. มุ่งมั่น	1.1487	-0.7480	0.6259	1.9056	2.8973
14. ยินดี	0.9670	-1.3823	0.1101	1.3513	2.4580
15. โล่งใจ	0.9013	-1.1348	0.5611	2.0593	2.9031
16. ไว้วางใจ	1.0291	-0.9847	0.6735	2.0106	2.9221
17. สนุกสนม	1.2858	0.0874	1.1119	1.9289	2.6477
18. สนุกสนาน	0.9833	-1.2952	0.2836	1.4363	2.4382
19. สบายใจ	0.9963	0.3934	1.8148	2.8994	3.6813
20. สุขใจ	0.9243	0.4782	1.4642	2.4475	3.2523
21. ให้เกียรติ	1.0203	-1.2241	0.2412	1.4351	2.2371

ตารางที่ 4-12 (ต่อ)

มาตรวัดย่อย/ข้อคำถาม	ค่าพารามิเตอร์ของข้อคำถาม				
	α	β_1	β_2	β_3	β_4
22. อ่อนน้อม	0.8247	-1.1739	0.7284	2.3189	4.6966
อารมณ์ความรู้สึกเชิงลบ					
1. กังวล	1.0522	3.2024	2.3472	1.1251	-0.3375
2. แค้นใจ	0.8822	2.7522	1.5020	0.2691	-1.4965
3. ถูกเหยียดหยาม	1.2665	2.6771	1.8263	0.6823	-0.6555
4. โดดเดี่ยว	1.1564	3.0356	2.1413	0.7988	-0.7964
5. ต่อต้าน	1.2638	2.4796	1.5235	0.3098	-1.3028
6. ตึงเครียด	1.1236	2.9222	1.6834	0.4234	-1.2192
7. น้อยใจ	0.9749	2.5605	1.4701	0.2967	-1.4263
8. บ้าคลั่ง	0.9147	2.8233	1.5406	0.2076	-1.4622
9. เปื่อหน้าย	1.5338	2.1567	1.4474	0.4828	-0.7853
10. ประหม่า	1.4390	2.3923	1.7384	0.7331	-0.3685
11. ผิดหวัง	1.3786	2.4055	1.6572	0.6972	-0.4258
12. พลุ่พล่าน	1.4113	2.4193	1.7544	0.7158	-0.4217
13. ไม่มีความสุข	1.3934	2.3359	1.5641	0.5786	-0.8109
14. รำคาญ	1.1082	3.2397	2.2646	1.2795	-0.1362
15. รู้สึกผิด	1.0264	3.3295	2.3427	1.0951	-0.3996
16. เศร้าใจ	1.1004	2.3486	1.4075	0.1828	-1.3107
17. สยดสยอง	1.5270	2.4652	1.7775	0.8960	-0.4251
18. สลดใจ	1.1433	2.4226	1.5988	0.3447	-1.1949
19. สะเทือนใจ	1.4940	2.4579	1.8022	0.8721	-0.4877
20. สิ้นหวัง	1.3026	2.7328	1.9003	0.6999	-0.7103
21. เสียใจ	1.1763	3.0919	2.4004	1.3302	-0.0167
22. เหน็ดเหนื่อย	1.0479	3.2898	2.8355	1.9095	0.9151
23. อัดอั้นตันใจ	0.9175	3.4190	2.6112	1.8723	0.7383
24. อับอาย	1.0301	2.6314	2.0672	1.0023	-0.1957
25. อ่างว้าง	1.0012	2.7425	1.7186	0.2781	-1.2823
26. อิจฉา	1.2356	2.4221	1.7464	0.5597	-0.8145

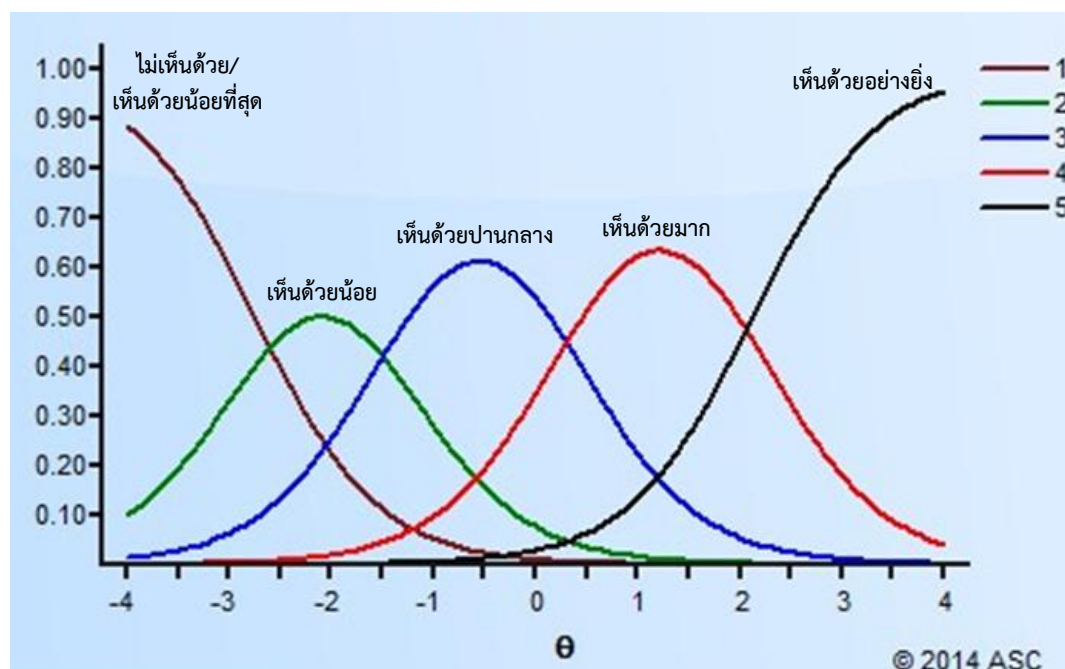
หมายเหตุ α

คือ ค่าพารามิเตอร์ความชัน

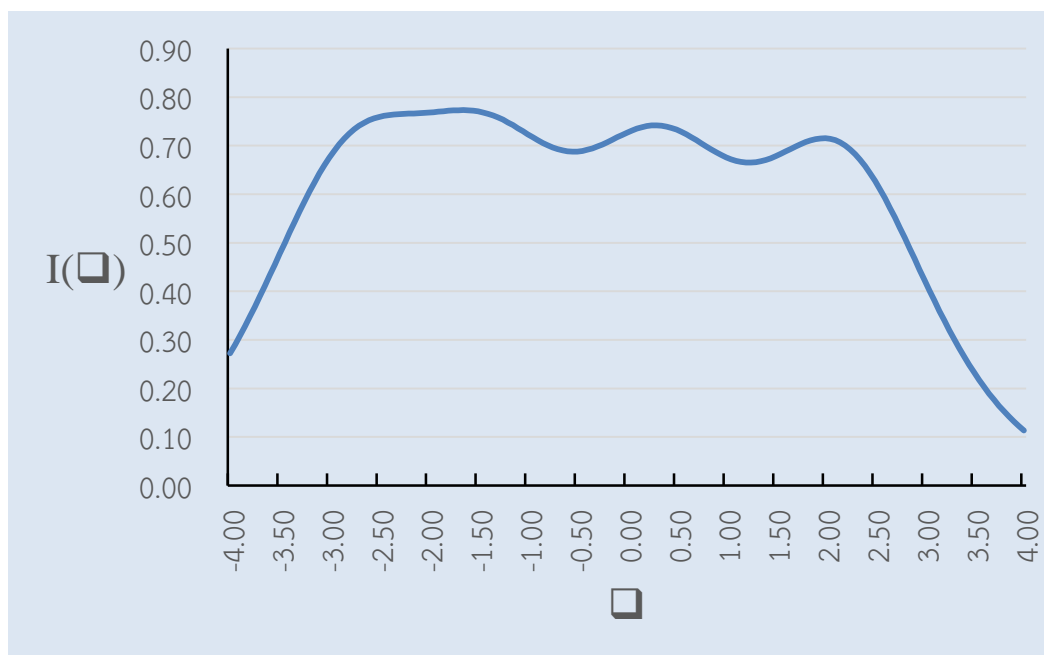
 $\beta_1, \beta_2, \beta_3, \beta_4$

คือ ค่าพารามิเตอร์เทรซโฮลด์ที่ 1 2 3 และ 4

ข้อมูลที่ได้จากการวิเคราะห์ด้วยทฤษฎีการตอบสนองข้อสอบสามารถแสดงในรูปของโค้งการเลือกรายการคำตอบ (Category Response Curve: CRC) ซึ่งทำให้เห็นภาพรวมของข้อคำถามได้ชัดเจนยิ่งขึ้น นอกจากนี้ โค้งการเลือกรายการคำตอบสามารถนำไปคำนวณเพื่อหาฟังก์ชันสารสนเทศของข้อสอบและแสดงในรูปของโค้งสารสนเทศของข้อสอบได้ (Item Information Curve: IIC) ตัวอย่างโค้งการเลือกรายการคำตอบและโค้งสารสนเทศของข้อสอบ จากข้อคำถามที่ 3 ในมาตรวัดย่อยความพึงพอใจในชีวิต ที่ว่า “ฉันดำเนินชีวิตได้อย่างที่ต้องการ” แสดงได้ดังภาพที่ 4-2 และภาพที่ 4-3 ตามลำดับ



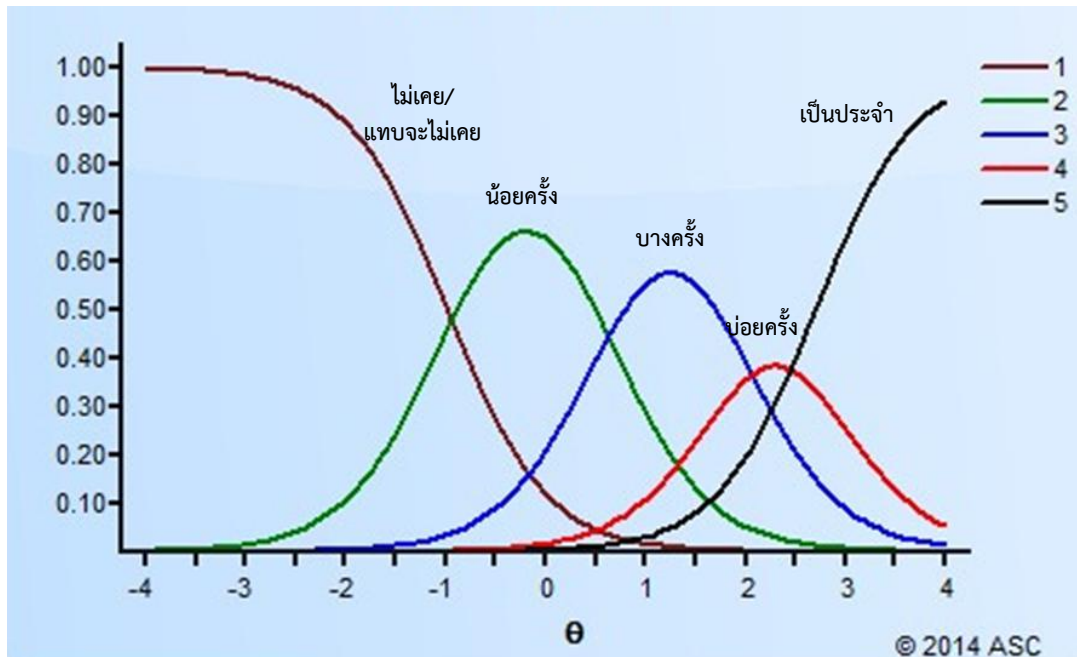
ภาพที่ 4-2 โค้งการเลือกรายการคำตอบจากข้อคำถามที่ 3 ในมาตรวัดย่อยความพึงพอใจในชีวิต ที่ว่า “ฉันดำเนินชีวิตได้อย่างที่ต้องการ”



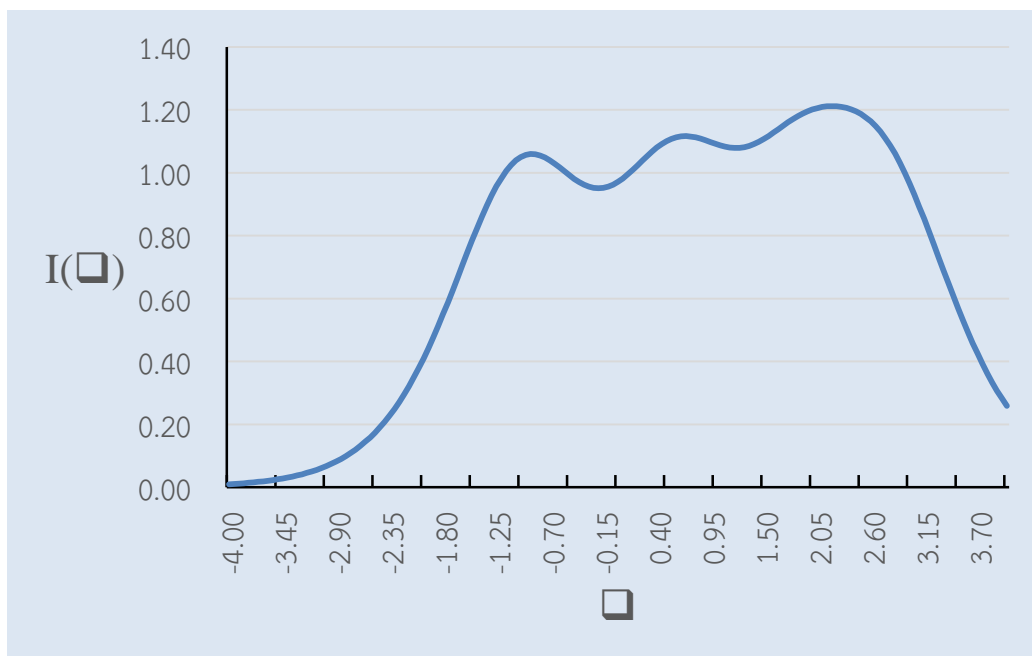
ภาพที่ 4-3 โค้งสารสนเทศของข้อสอบจากข้อคำถามที่ 3 ในมาตรวัดย่อยความพึงพอใจในชีวิตที่ว่า “ฉันดำเนินชีวิตได้อย่างที่ต้องการ”

จากภาพที่ 4-2 แสดงโค้งการเลือกรายการคำตอบของข้อคำถาม “ฉันดำเนินชีวิตได้อย่างที่ต้องการ” ซึ่งจะเห็นได้ว่าตัวเลือกการคำตอบแต่ละตัวเลือกสามารถให้ข้อมูลความน่าจะเป็นในการเลือกรายการคำตอบซึ่งสัมพันธ์กับระดับความสุขที่แตกต่างกัน เช่น ผู้ทดสอบที่มีค่าคุณลักษณะแฝง (ความสุข) เท่ากับ -1.00 มีโอกาสที่จะเลือกตอบว่า “ไม่เห็นด้วย/เห็นด้วยน้อยที่สุด” ประมาณร้อยละ 5 มีโอกาสที่จะเลือกตอบว่า “เห็นด้วยน้อย” ประมาณร้อยละ 28 มีโอกาสที่จะเลือกตอบว่า “เห็นด้วยปานกลาง” ประมาณร้อยละ 56 มีโอกาสที่จะเลือกตอบว่า “เห็นด้วยมาก” ประมาณร้อยละ 10 และมีโอกาสที่จะเลือกตอบว่า “เห็นด้วยอย่างยิ่ง” ประมาณร้อยละ 1 เมื่อพิจารณาภาพที่ 4-3 ปรากฏว่าข้อคำถาม “ฉันดำเนินชีวิตได้อย่างที่ต้องการ” สามารถให้ข้อมูลสารสนเทศของข้อสอบได้ดีในช่วงคุณลักษณะแฝง (ความสุข) ตั้งแต่ -3.00 ถึง 2.50 ในขณะที่ข้อคำถามข้อนี้จะให้ข้อมูลสารสนเทศของข้อสอบได้ลดลงเมื่อผู้ทดสอบมีคุณลักษณะแฝง (ความสุข) ต่ำกว่า -3.00 และมากกว่า 2.50

ตัวอย่างโค้งการเลือกรายการคำตอบและโค้งสารสนเทศของข้อสอบ จากข้อคำถามข้อที่ 8 ในมาตรวัดย่อยอารมณ์ความรู้สึกเชิงบวก ที่ว่า “ในระยะเวลา 1 เดือนที่ผ่านมา ฉันรู้สึกเบิกบานใจ” แสดงได้ดังภาพที่ 4-4 และภาพที่ 4-5 ตามลำดับ



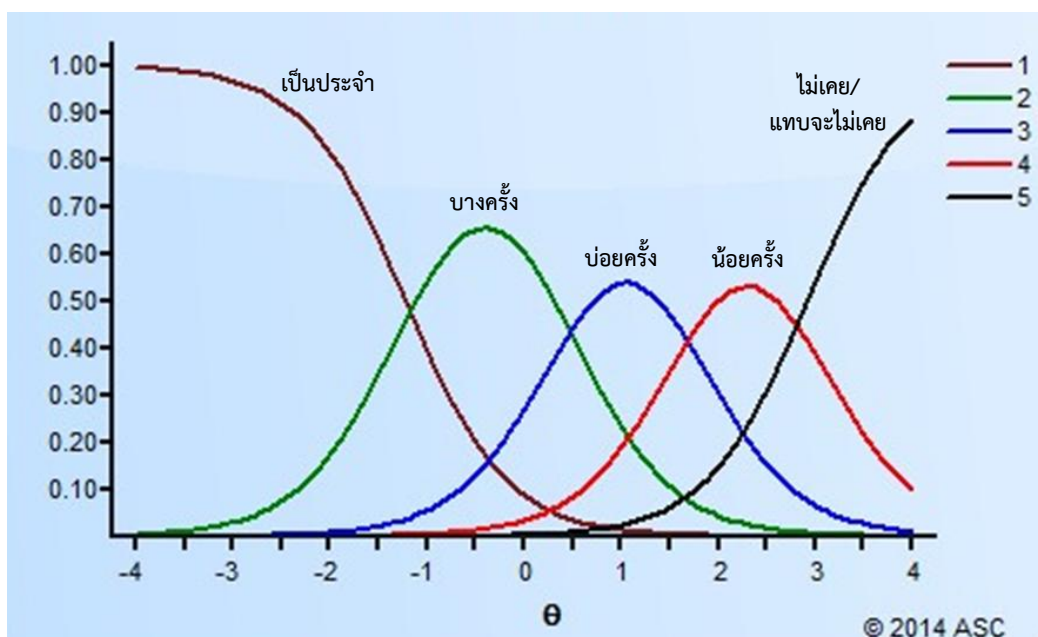
ภาพที่ 4-4 โค้งการเลือกรายการคำตอบจากข้อคำถามที่ 8 ในมาตรวัดย่อยอารมณ์ความรู้สึกเชิงบวก ที่ว่า “ในระยะเวลา 1 เดือนที่ผ่านมา ฉันรู้สึกเบิกบานใจ”



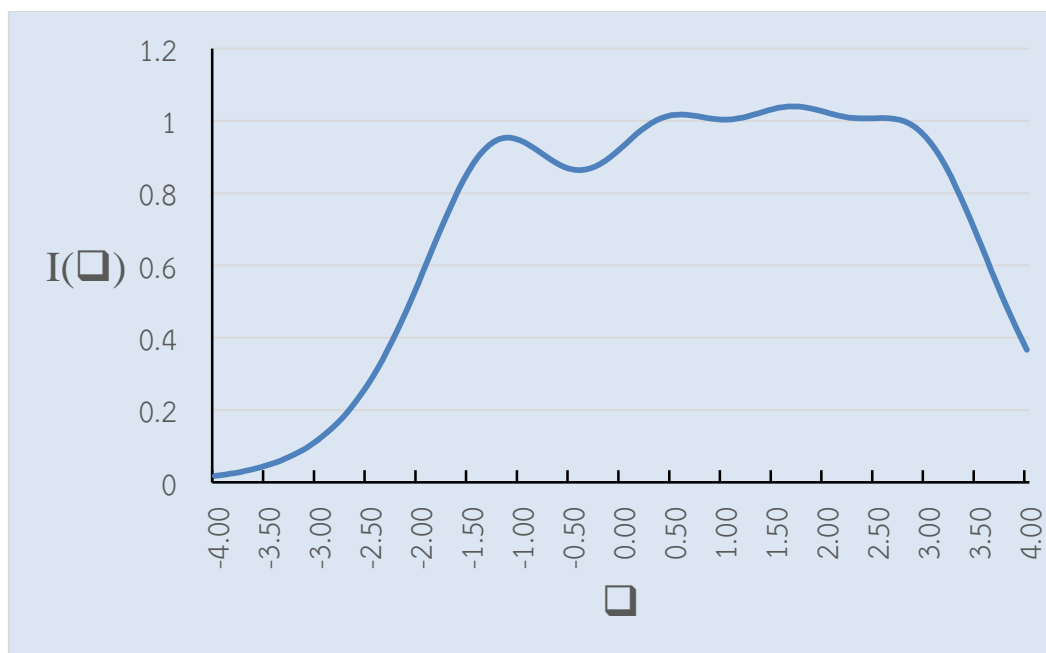
ภาพที่ 4-5 โค้งการสรสนเทศของข้อสอบจากข้อคำถามที่ 8 ในมาตรวัดย่อยอารมณ์ความรู้สึกเชิงบวก ที่ว่า “ในระยะเวลา 1 เดือนที่ผ่านมา ฉันรู้สึกเบิกบานใจ”

จากภาพที่ 4-4 แสดง โค้งการเลือกรายการคำตอบของข้อความที่ว่า “ในระยะเวลา 1 เดือนที่ผ่านมา ฉันรู้สึกเบิกบานใจ” ซึ่งจะเห็นได้ว่าตัวเลือกการคำตอบแต่ละตัวเลือกสามารถให้ข้อมูลความน่าจะเป็นในการเลือกรายการคำตอบซึ่งสัมพันธ์กับระดับความสุขที่แตกต่างกัน เช่น ผู้ทดสอบที่มีค่าคุณลักษณะแฝง (ความสุข) เท่ากับ -1.00 มีโอกาสที่จะเลือกตอบว่า “ไม่เคย/แทบจะไม่มีเคย” หรือ “น้อยครั้ง” ประมาณร้อยละ 48 มีโอกาสที่จะเลือกตอบว่า “บางครั้ง” ประมาณร้อยละ 3 มีโอกาสที่จะเลือกตอบว่า “บ่อยครั้ง” ประมาณร้อยละ 1 และมีโอกาสที่จะเลือกตอบว่า “เป็นประจำ” ประมาณร้อยละ 0 เมื่อพิจารณาภาพที่ 4-5 ปรากฏว่าข้อความที่ว่า “ในระยะเวลา 1 เดือนที่ผ่านมา ฉันรู้สึกเบิกบานใจ” สามารถให้ข้อมูลสารสนเทศของข้อสอบได้ดีในช่วงคุณลักษณะแฝง (ความสุข) ตั้งแต่ -1.00 ถึง 2.50 ในขณะที่ข้อความข้อนี้จะให้ข้อมูลสารสนเทศของข้อสอบได้ลดลงเมื่อผู้ทดสอบมีค่าคุณลักษณะแฝง (ความสุข) ต่ำกว่า -1.00 และมากกว่า 2.50

ตัวอย่างโค้งการเลือกรายการคำตอบและโค้งสารสนเทศของข้อสอบ จากข้อความข้อที่ 6 ในมาตรวัดย่อยอารมณ์ความรู้สึกเชิงลบ ที่ว่า “ในระยะเวลา 1 เดือนที่ผ่านมา ฉันรู้สึกตึงเครียด” แสดงได้ดังภาพที่ 4-6 และภาพที่ 4-7 ตามลำดับ



ภาพที่ 4-6 โค้งการเลือกรายการคำตอบจากข้อความข้อที่ 6 ในมาตรวัดย่อยอารมณ์ความรู้สึกเชิงลบ ที่ว่า “ในระยะเวลา 1 เดือนที่ผ่านมา ฉันรู้สึกตึงเครียด”



ภาพที่ 4-7 โค้งการแจกแจงของข้อสอบจากข้อคำถามที่ 6 ในมาตรวัดย่อยอารมณ์ความรู้สึกเชิงลบ ที่ว่า “ในระยะเวลา 1 เดือนที่ผ่านมา ฉันรู้สึกตึงเครียด”

จากภาพที่ 4-6 แสดง โค้งการเลือกรายการคำตอบของข้อคำถามที่ว่า “ในระยะเวลา 1 เดือนที่ผ่านมา ฉันรู้สึกตึงเครียด” ซึ่งจะเห็นได้ว่าตัวเลือกรายการคำตอบแต่ละตัวเลือกสามารถให้ข้อมูลความน่าจะเป็นในการเลือกรายการคำตอบซึ่งสัมพันธ์กับระดับความสุขที่แตกต่างกัน เช่น ผู้ทดสอบที่มีค่าคุณลักษณะแฝง (ความสุข) เท่ากับ -1.00 มีโอกาสที่จะเลือกตอบว่า “เป็นประจำ” ประมาณร้อยละ 41 มีโอกาสที่จะเลือกตอบว่า “บ่อยครั้ง” ประมาณร้อยละ 52 มีโอกาสที่จะเลือกตอบว่า “บางครั้ง” ประมาณร้อยละ 6 มีโอกาสที่จะเลือกตอบว่า “น้อยครั้ง” ประมาณร้อยละ 1 และมีโอกาสที่จะเลือกตอบว่า “ไม่เคย/แทบจะไม่เคย” ประมาณร้อยละ 0 เมื่อพิจารณาภาพที่ 4-7 ปรากฏว่าข้อคำถามที่ว่า “ในระยะเวลา 1 เดือนที่ผ่านมา ฉันรู้สึกตึงเครียด” สามารถให้ข้อมูลสารสนเทศของข้อสอบได้ดีในช่วงคุณลักษณะแฝง (ความสุข) ตั้งแต่ -1.50 ถึง 3.00 ในขณะที่ข้อคำถามข้อนี้จะให้ข้อมูลสารสนเทศของข้อสอบได้ลดลงเมื่อผู้ทดสอบมีคุณลักษณะแฝง (ความสุข) ต่ำกว่า -1.50 และมากกว่า 3.00

คลังข้อคำถามที่สร้างขึ้นนี้ นอกจากสามารถนำไปใช้ในการพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับมาตรวัดความสุขของคนไทยแล้ว ผู้วิจัยได้นำข้อคำถามมาจัดพิมพ์เป็นมาตรวัดความสุขของคนไทย รูปแบบกระดาษ-ดินสอ พร้อมทั้งตรวจสอบความเที่ยงทั้งฉบับและสร้างปกติวิสัยสำหรับการแปลผลคะแนน เพื่อเป็นอีกทางเลือกหนึ่งสำหรับการนำไปใช้ด้วย (รายละเอียดของผลการพัฒนามาตรวัดความสุขของคนไทยรูปแบบกระดาษ-ดินสอ มาตรวัดความสุขของคนไทยรูปแบบกระดาษ-ดินสอ และคู่มือการใช้ แสดงในภาคผนวก ข)

ตอนที่ 2 ผลการพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย

การพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย พัฒนาขึ้นในรูปแบบเว็บแอปพลิเคชัน และทดสอบผ่านระบบอินเทอร์เน็ต คลังข้อคำถามที่ใช้สำหรับการทดสอบ ได้มาจากการพัฒนาคลังข้อคำถามในขั้นตอนที่ 1 ผู้วิจัยนำเสนอผลการพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ใน 2 ประเด็นหลัก ดังนี้

1. ผลการพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย

1.1 ผลการกำหนดรูปแบบโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย

1.2 ผลการจัดทำคู่มือการใช้งานโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย

2. ผลการประเมินความเหมาะสมของโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย

2.1 ผลการประเมินโดยผู้เชี่ยวชาญ

2.2 ผลการประเมินโดยผู้ใช้งาน

1. ผลการพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย

1.1 ผลการกำหนดรูปแบบโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย

1.1.1 หน้าจอหลักของโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ แสดงได้ดังภาพที่ 4-8 และ 4-9



ภาพที่ 4-8 หน้าจอหลักของโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย สำหรับผู้ใช้งานทั่วไป



ภาพที่ 4-9 หน้าจอหลักของโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย สำหรับผู้ดูแลระบบ

จากภาพที่ 4-8 แสดงหน้าจอหลักของโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย สำหรับผู้ใช้งานทั่วไป ซึ่งประกอบด้วย 4 เมนู ดังนี้

1) หน้าหลัก เป็นเมนูที่แสดงหน้าจอหลักของโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับมาตรวัดความสุขของคนไทย ซึ่งแบ่งเนื้อหาเป็นสามส่วนย่อย ได้แก่ THS-CAT คืออะไร คำแนะนำในการทดสอบ (ปุ่มดาวน์โหลดคู่มือการใช้โปรแกรมแสดงอยู่ในส่วนนี้) และการเริ่มการทดสอบ

- 2) ผลการทดสอบ เป็นเมนูสำหรับการค้นหาผลการทดสอบที่ผ่านมาของผู้ใช้งาน
- 3) ติดต่อเรา เป็นเมนูสำหรับแสดงข้อมูลที่เกี่ยวข้องกับผู้วิจัยและสถาบันของผู้วิจัย
- 4) เข้าสู่ระบบ เป็นเมนูสำหรับการเข้าสู่ระบบของผู้ดูแลระบบ

จากภาพที่ 4-9 แสดงหน้าจอหลักของโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย สำหรับผู้ดูแลระบบ ซึ่งประกอบด้วย 5 เมนู ดังนี้

- 1) หน้าหลัก เป็นเมนูที่แสดงหน้าจอหลักของโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์
- 2) ผลการทดสอบ เป็นเมนูสำหรับการค้นหาผลการทดสอบที่ผ่านมาของผู้ใช้งาน
- 3) การจัดการ เป็นเมนูสำหรับการจัดการคลังข้อคำถามและการค้นหาประวัติการทดสอบโดยละเอียด
- 4) ติดต่อเรา เป็นเมนูสำหรับแสดงข้อมูลที่เกี่ยวข้องกับผู้วิจัย
- 5) ออกจากระบบ เป็นเมนูสำหรับการออกจากระบบการจัดการโดยผู้ดูแลระบบ เพื่อกลับเข้าสู่ระบบของผู้ใช้งานทั่วไป

1.1.2 รายละเอียดของโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ แบ่งออกเป็น 3 ส่วน คือ 1) การทดสอบ 2) การรายงานผลการทดสอบ และ 3) การจัดการ ซึ่งมีรายละเอียด ดังนี้

1) การทดสอบ เป็นส่วนของการดำเนินการทดสอบด้วยโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์เพื่อวัดความสุขของผู้ทดสอบตามทฤษฎีความอยู่ดีมีสุขเชิงอัตวิสัย ซึ่งประกอบด้วย 3 มาตรวัดย่อย ได้แก่ ความพึงพอใจในชีวิต อารมณ์ความรู้สึกเชิงบวก และอารมณ์ความรู้สึกเชิงลบ โดยมีเงื่อนไขการทดสอบ ดังนี้

เงื่อนไขที่ 1 ผู้ทดสอบต้องตอบคำถามข้อปัจจุบันก่อนจึงจะทำข้อคำถามข้อต่อไปได้

เงื่อนไขที่ 2 ผู้ทดสอบไม่สามารถกลับมาแก้ไขคำตอบในข้อที่ผ่านไปแล้วได้

เงื่อนไขที่ 3 หากผู้ทดสอบยกเลิกการทดสอบก่อนจบการทดสอบ โปรแกรมจะไม่สามารถประมวลผลได้

ผู้ทดสอบสามารถเริ่มการทดสอบได้โดยเข้าที่หน้าจอหลัก และเลือกในส่วนของ “เริ่มการทดสอบ” ซึ่งจะปรากฏหน้าจอใหม่ ดังภาพที่ 4-10

เริ่มทดสอบ !!!

กรุณาเลือกมาตรวัดที่ต้องการทดสอบ
(เลือกมาตรวัดย่อยที่ต้องการทดสอบ หรือเลือกมาตรวัดความสุขโดยรวมเพื่อรับการทดสอบครบทั้งสามมาตรวัดย่อย)

ความพึงพอใจในชีวิต อารมณ์ความรู้สึกเชิงบวก
 อารมณ์ความรู้สึกเชิงลบ มาตรวัดความสุขโดยรวม

คำนำหน้า : นาย นาง นางสาว
 ชื่อ : _____
 สกุล : _____
 เลขประจำตัวประชาชน : _____
 อายุ (ปี) : _____
 เพศ : ชาย หญิง
 กรุณาใส่แคปช่าที่ท่านเห็น :

ภาพที่ 4-10 หน้าจอสำหรับบันทึกรายละเอียดก่อนเริ่มการทดสอบ

จากภาพที่ 4-10 ผู้ทดสอบสามารถเลือกมาตรวัดเพื่อทดสอบได้ตามความต้องการ โดยเลือก “มาตรวัดความสุขโดยรวม” เพื่อดำเนินการทดสอบทั้งสามมาตรวัดย่อย หรือเลือกเจาะจงในแต่ละมาตรวัดย่อยตามความสนใจ ผู้ทดสอบต้องกรอกข้อมูลส่วนบุคคล และใส่รหัสแคปช่า จึงจะสามารถเริ่มการทดสอบได้ โดยสามารถเลือกการทดสอบได้สองรูปแบบ คือ “เริ่มทดสอบ (CAT)” เมื่อต้องการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ และ “เริ่มทดสอบ (ทุกข้อ)” เมื่อต้องการทดสอบด้วยข้อคำถามทั้งหมดในคลัง หลังจากเลือกรูปแบบการทดสอบแล้วจะปรากฏหน้าจอการทดสอบ ซึ่งแสดงข้อคำถามทีละข้อ ดังภาพที่ 4-11

ฉันกังวลว่าจะไม่มีคนดูแลหรือให้ความช่วยเหลือในอนาคต

ไม่เห็นด้วย/เห็นด้วยน้อยที่สุด
 เห็นด้วยน้อย
 เห็นด้วยปานกลาง
 เห็นด้วยมาก
 เห็นด้วยอย่างยิ่ง

ตกลง

ภาพที่ 4-11 หน้าจอการทดสอบของโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย

จากภาพที่ 4-11 ผู้ทดสอบจะต้องตอบข้อคำถามทีละข้อ จึงจะสามารถตอบข้อคำถามต่อไปได้ และไม่สามารถย้อนกลับมาเปลี่ยนแปลงคำตอบในข้อคำถามที่ผ่านไปแล้วได้ การทดสอบจะเริ่มต้นด้วยข้อคำถามที่ให้ค่าสารสนเทศสูงที่สุด (Maximum Information) ณ ตำแหน่งค่าคุณลักษณะแฝงที่กำหนดไว้ โดยกำหนดค่าคุณลักษณะแฝงเริ่มต้นของผู้ทดสอบ ในทั้งสามมาตรวัดย่อย ให้มีค่าคุณลักษณะแฝงเริ่มต้นเท่ากับ 0.00 โปรแกรมจะเลือกข้อคำถามที่ให้ค่าสารสนเทศข้อสอบสูงที่สุด ณ ตำแหน่งค่าคุณลักษณะแฝงดังกล่าว เมื่อผู้ทดสอบตอบข้อคำถาม โปรแกรมจะประมาณค่าคุณลักษณะแฝงของผู้ทดสอบใหม่ โดยใช้วิธีการ Expected a Posteriori (EAP) จากนั้นจึงคัดเลือกข้อคำถามข้อถัดไปซึ่งมีค่าสารสนเทศสูงที่สุด ณ ตำแหน่งค่าคุณลักษณะแฝงที่ประมาณค่าได้ การทดสอบจะดำเนินการไปเรื่อย ๆ จนกระทั่งถึงเกณฑ์การยุติการทดสอบ คือ เมื่อความคลาดเคลื่อนจากการประมาณค่าคุณลักษณะแฝงมีค่าต่ำกว่า 0.30 หรือเมื่อข้อคำถามหมดคลัง (ในกรณีเลือกการทดสอบด้วยข้อคำถามทุกข้อ การทดสอบจะยุติเมื่อข้อคำถามหมดคลังเท่านั้น) จากนั้นโปรแกรมจะรายงานผลการทดสอบ ดังภาพที่ 4-12



ภาพที่ 4-12 หน้าจอรายงานผลการทดสอบ

จากภาพที่ 4-12 หน้าจอรายงานผลการทดสอบจะปรากฏขึ้นทุกครั้งที่การทดสอบสิ้นสุดลง ซึ่งหากผู้ทดสอบต้องการเก็บรายงานผลดังกล่าวในรูปแบบของเอกสาร สามารถสั่งพิมพ์ได้โดยกดปุ่ม “พิมพ์หน้านี้” แต่หากต้องการกลับไปสู่หน้าจอแสดงประวัติการทดสอบโดยรวม สามารถกดปุ่ม “ก่อนหน้า”

ผลการทดสอบจะแสดงประวัติของผู้ทดสอบ และผลการทดสอบในแต่ละมาตรวัดย่อย หรือมาตรวัดความสุขโดยรวม ซึ่งจะระบุคะแนนที่ (คะแนนเต็ม 100 คะแนน) รวมทั้งเกณฑ์การแปลความหมายของคะแนนที่ได้ ดังตารางที่ 4-13

ตารางที่ 4-13 เกณฑ์การแปลความหมายคะแนนที่ ได้จากโปรแกรมการทดสอบแบบ
ปรับเหมาะด้วยคอมพิวเตอร์

คะแนนที่	มาตรวัดย่อย/มาตรวัด			
	ความพึงพอใจ ในชีวิต	อารมณ์ความรู้สึก เชิงบวก	อารมณ์ความรู้สึก เชิงลบ	ความสุขโดยรวม
มากกว่า 68	พึงพอใจอย่างยิ่ง	สูงที่สุด	ต่ำที่สุด	มีความสุขอย่างยิ่ง
63-68	พึงพอใจมาก	สูงมาก	ต่ำมาก	มีความสุขมาก
58-62	พึงพอใจ	สูง	ต่ำ	มีความสุข
53-57	ค่อนข้างพึงพอใจ	สูงกว่าปกติเล็กน้อย	ต่ำกว่าปกติเล็กน้อย	มีความสุขสูงกว่า ปกติเล็กน้อย
48-52	ปานกลาง	ปกติ	ปกติ	มีความสุข ตามปกติ
43-47	พึงพอใจน้อย	ต่ำกว่าปกติเล็กน้อย	สูงกว่าปกติเล็กน้อย	มีความสุขต่ำกว่า ปกติเล็กน้อย
38-42	ไม่ค่อยพึงพอใจ	ต่ำ	สูง	มีความสุขน้อย
32-37	ไม่พึงพอใจ	ต่ำมาก	สูงมาก	มีความสุข น้อยมาก
น้อยกว่า 32	ไม่พึงพอใจ อย่างยิ่ง	ต่ำที่สุด	สูงที่สุด	ไม่มีความสุข

2) การรายงานผลการทดสอบ เป็นส่วนสำหรับค้นหาผลการทดสอบที่ผ่านมา
ทั้งหมดของผู้ทดสอบ ผ่านเมนู “ผลการทดสอบ” ซึ่งจะปรากฏหน้าจอดังภาพที่ 4-13



ภาพที่ 4-13 หน้าจอค้นหาผลการทดสอบ

จากภาพที่ 4-13 ผู้ทดสอบสามารถค้นหาผลการทดสอบที่ผ่านมาแล้วได้ โดยใช้เลขบัตรประจำตัวประชาชนเป็นรหัสในการสืบค้น ซึ่งจะปรากฏผลการทดสอบที่ผ่านมาทั้งหมดของผู้ทดสอบ แสดงได้ดังภาพที่ 4-14



ภาพที่ 4-14 หน้าจอแสดงประวัติการทดสอบโดยรวม ตามรหัสที่สืบค้น

จากภาพที่ 4-14 ประวัติผลการทดสอบที่ผ่านมาทั้งหมดของผู้ทดสอบ จะปรากฏเรียงตามครั้งที่ทดสอบ ระบุวันเวลาที่ทดสอบ ผู้ทดสอบสามารถดูรายละเอียดการทดสอบแต่ละครั้งได้ โดยเลือกปุ่ม “ผลการทดสอบ” จะปรากฏหน้าจอรายงานผลการทดสอบเช่นเดียวกับภาพที่ 4-12 ซึ่งผู้ทดสอบสามารถสั่งพิมพ์ผลการทดสอบได้ ส่วนปุ่ม “รายละเอียด” นั้นเป็นส่วนสำหรับผู้ดูแลระบบเพื่อใช้ดูข้อมูลการทดสอบโดยละเอียด ซึ่งต้องใช้ชื่อและรหัสสำหรับเข้าใช้งานที่ได้รับจากผู้พัฒนาโปรแกรม ในส่วนนี้ผู้ใช้งานไม่สามารถเข้าได้

3) การจัดการ เป็นส่วนสำหรับผู้ดูแลระบบ เพื่อใช้สำหรับการจัดการคลังข้อคำถามและการจัดการประวัติการทดสอบ การเข้าสู่ระบบจะต้องใช้รหัสที่ได้รับจากผู้พัฒนาโปรแกรม โดยลงชื่อเข้าใช้ได้เมนู “เข้าสู่ระบบ” ซึ่งจะปรากฏหน้าจอ ดังภาพที่ 4-15

ภาพที่ 4-15 หน้าจอการเข้าสู่ระบบของผู้ดูแลระบบ

จากภาพที่ 4-15 เมื่อเข้าสู่ระบบแล้วจะปรากฏเมนู “การจัดการ” ขึ้นในแถบเมนู ผู้ดูแลระบบสามารถเข้าไปจัดการคลังข้อคำถามและประวัติการทดสอบได้ผ่านเมนูนี้ ซึ่งจะปรากฏหน้าจอ ดังภาพที่ 4-16



ภาพที่ 4-16 หน้าจอการจัดการ

จากภาพที่ 4-16 หน้าจอการจัดการ แบ่งออกเป็นสองส่วน คือ การจัดการคลังข้อคำถาม และการจัดการประวัติการทดสอบ

1. การจัดการคลังข้อคำถาม แบ่งออกเป็นสองส่วน คือ มาตรฐานวัดย่อยและข้อคำถาม ปรากฏหน้าจอ ดังภาพที่ 4-17



ภาพที่ 4-17 หน้าจอการจัดการคลังข้อคำถาม

จากภาพที่ 4-17 หากเลือกส่วนของมาตรวัดย่อย จะปรากฏหน้าจอ ดังภาพที่ 4-18

THS-CAT

โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์
สำหรับมาตรวัดความสุขของคนไทย

หน้าหลัก ผลการทดสอบ การจัดการ ติดต่อเรา ออกจากระบบ

มาตรวัดย่อย

เพิ่ม

ชื่อมาตรวัดย่อย	ค่า Theta เริ่มต้น	ค่า SEE สำหรับชุดการทดสอบ
ค้นหา		
1 มาตรวัดความพึงพอใจในชีวิต	0	0.3
2 มาตรวัดอารมณ์ความรู้สึกเชิงบวก	0	0.3
3 มาตรวัดอารมณ์ความรู้สึกเชิงลบ	0	0.3

แก้ไข ลบ

แก้ไข ลบ

แก้ไข ลบ

วิทยาลัยวิทยาการวิจัยและวิทยาการปัญญา มหาวิทยาลัยบูรพา

ภาพที่ 4-18 หน้าจอการจัดการมาตรวัดย่อย

จากภาพที่ 4-18 แสดงหน้าจอ “การจัดการมาตรวัดย่อย” ของโปรแกรม ซึ่งผู้ดูแลระบบสามารถเพิ่ม ลบ หรือแก้ไขมาตรวัดย่อยได้ตามความต้องการ ในหน้าจอการจัดการคลังข้อคำถาม หากเลือกที่ข้อคำถาม จะปรากฏหน้าจอ ดังภาพที่ 4-19



ภาพที่ 4-19 หน้าจอการจัดการข้อคำถาม

จากภาพที่ 4-19 หน้าจอ “การจัดการข้อคำถาม” ของโปรแกรม ซึ่งผู้ดูแลระบบสามารถเพิ่ม ลบ หรือแก้ไขข้อคำถามที่ผ่านการวิเคราะห์คุณภาพด้วย Graded Response Model ตามทฤษฎีการตอบสนองข้อสอบเรียบร้อยแล้วเท่านั้น

2. การจัดการประวัติการทดสอบ เป็นส่วนที่เกี่ยวข้องกับรายละเอียดผลการทดสอบของผู้ทดสอบทั้งหมด โดยโปรแกรมจะรวบรวมผลการทดสอบที่ผ่านมา ซึ่งผู้ดูแลระบบสามารถเข้าดูรายละเอียด และลบรายการที่ไม่ต้องการบันทึกได้ในส่วนนี้ เมื่อเลือก “การจัดการประวัติการทดสอบ” ในหน้าจอการจัดการ ดังภาพที่ 4-16 จะปรากฏหน้าจอประวัติการทดสอบขึ้นมา ดังภาพที่ 4-20

THS-CAT

โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์
สำหรับมาตรวัดความสุขของคนไทย

หน้าหลัก ผลการทดสอบ การจัดการ ติดต่อเรา ออกจากระบบ

ประวัติการทดสอบ

ลำดับที่	ชื่อ-นามสกุล	เพศ	อายุ	
1	วัฒน์วี อัมมวรรณ	หญิง	34	รายละเอียด
2	สุมาลย์ ปานคำ	หญิง	37	รายละเอียด
3	thiprapa singthong	หญิง	50	รายละเอียด
4	โสฬส สุขานนท์สวัสดิ์	ชาย	36	รายละเอียด
5	ปัทมา วรรัตน์	หญิง	49	รายละเอียด

ภาพที่ 4-20 หน้าจอการจัดการประวัติการทดสอบ

จากภาพที่ 4-20 หน้าจอ “การจัดการประวัติการทดสอบ” ของโปรแกรม แสดงข้อมูล การทดสอบเรียงตามลำดับครั้งที่ทดสอบ โดยแสดงข้อมูลในภาพรวมเกี่ยวกับประวัติส่วนตัวของ ผู้ทดสอบ ได้แก่ ชื่อ-นามสกุล เพศ อายุ ซึ่งหากต้องการรายละเอียดการทดสอบเพิ่มเติมเกี่ยวกับ มาตรวัดที่ทดสอบ ค่าคุณลักษณะแฝง จำนวนข้อที่ทดสอบ ระยะเวลาที่ใช้ ข้อมูลข้อคำถามรายชื่อ และรายละเอียดอื่น ๆ สามารถคลิกเพื่อดูรายละเอียดเพิ่มเติมได้จากส่วนของรายละเอียด

1.2 ผลการจัดทำคู่มือการใช้งานโปรแกรมการทดสอบแบบปรับเหมาะด้วย คอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย

คู่มือการใช้งานโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัด ความสุขของคนไทย (The Manual of Computerized Adaptive Testing Program for Thai Happiness Scale) แบ่งเนื้อหาออกเป็น 4 ส่วน ดังนี้

1) การดำเนินการทดสอบ เป็นส่วนของผู้ทดสอบ สำหรับวัดความสุขโดยรวม หรือวัดใน แต่ละองค์ประกอบของความสุข ตามแนวคิดของความอยู่ดีมีสุขเชิงอัตวิสัย จัดทำขึ้นในรูปแบบ การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ซึ่งดำเนินการทดสอบไปตามระดับคุณลักษณะแฝงของ แต่ละบุคคล ในส่วนนี้ยังประกอบด้วยข้อมูลทั่วไปของโปรแกรม คำแนะนำในการทดสอบเบื้องต้น รวมถึงแหล่งดาวน์โหลดคู่มือการใช้งาน เพื่อให้ผู้ใช้งานสามารถใช้งานโปรแกรมได้อย่างสะดวก

2) การค้นหาผลการทดสอบ เป็นส่วนของการค้นหาผลการทดสอบที่ผ่านมาแล้วของ ผู้ทดสอบ ซึ่งใช้เลขประจำตัวประชาชนในการสืบค้น

3) การจัดการโปรแกรม เป็นส่วนของผู้ดูแลระบบเพื่อใช้จัดการประวัติการทดสอบโดย รายละเอียด และการจัดการคลังข้อคำถาม โดยสามารถเพิ่ม ลบ หรือแก้ไขข้อคำถามในคลังข้อคำถาม ที่ ผ่านการวิเคราะห์คุณภาพด้วย Graded Response Model ตามทฤษฎีการตอบสนองข้อสอบ

4) ติดต่อเรา เป็นส่วนที่แสดงข้อมูลสถาบันของผู้วิจัย และข้อมูลผู้วิจัย สำหรับผู้สนใจที่ต้องการสอบถามข้อมูลเพิ่มเติมเกี่ยวกับงานวิจัย (รายละเอียดคู่มือการใช้งานโปรแกรม แสดงในภาคผนวก ง)

2. ผลการประเมินความเหมาะสมของโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย

ผลการประเมินความเหมาะสมของโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย นำเสนอเป็นสองส่วน (แบบประเมินความเหมาะสมของโปรแกรม แสดงในภาคผนวก จ) ดังนี้

2.1 ผลการประเมินโดยผู้เชี่ยวชาญ

ผลการประเมินความเหมาะสมของโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย โดยผู้เชี่ยวชาญจำนวน 3 คน มีหัวข้อการประเมิน 4 ด้าน ได้แก่ 1) ด้านความสะดวกในการนำไปใช้ 2) ด้านความถูกต้องในการใช้งาน 3) ด้านลักษณะทั่วไปของโปรแกรม และ 4) ด้านความชัดเจนของคู่มือการใช้งาน ผลการประเมินความเหมาะสมของโปรแกรมทั้ง 4 ด้าน แสดงได้ดังตารางที่ 4-14 (รายละเอียดผลการประเมินความเหมาะสมของโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย โดยผู้เชี่ยวชาญเป็นรายชื่อ แสดงในภาคผนวก ฉ)

ตารางที่ 4-14 ผลการประเมินความเหมาะสมของโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย โดยผู้เชี่ยวชาญ

ด้านที่ประเมิน	M	SD	ระดับความเหมาะสม
1. ความสะดวกในการนำไปใช้	4.33	0.61	มาก
2. ความถูกต้องในการใช้งาน	4.20	0.72	มาก
3. ลักษณะทั่วไปของโปรแกรม	4.47	0.58	มาก
4. ความชัดเจนของคู่มือการใช้โปรแกรม	4.27	0.81	มาก
สรุปผลโดยรวม	4.32	0.65	มาก

จากการประเมินความเหมาะสมของโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทยโดยผู้เชี่ยวชาญ แสดงให้เห็นว่าโปรแกรมโดยรวมมีความเหมาะสมอยู่ในระดับมาก (M=4.32) โดยมีข้อเสนอแนะเพิ่มเติม ดังนี้

1. ผลการทดสอบควรจัดทำในรูปแบบไฟล์ pdf เพราะรูปแบบการรายงานผลมีความสวยงามดี แต่เมื่อสั่งพิมพ์ผลการทดสอบแล้วไม่สวยงามเหมือนที่อยู่ในหน้าจอคอมพิวเตอร์
2. ควรปรับรายละเอียดในคู่มือให้มีความเหมาะสม ถูกต้องชัดเจน และใช้ภาษาให้สอดคล้องกับโปรแกรม
3. สีของปุ่มตัวเลือกมาตรวัดความสุขโดยรวมควรใช้สีที่แตกต่างจากปุ่มอื่น ๆ เพื่อให้เห็นได้ชัดเจนขึ้น

4. ควรมีการป้องกันความผิดพลาดในการใช้งาน ของผู้ใช้งานด้วย เช่น การใส่เลขบัตรประชาชน อายุ ควรใส่ได้เฉพาะตัวเลขเท่านั้น
5. ควรมีการแจ้งเตือนเมื่อผู้ใช้งานใส่ข้อมูลสำคัญไม่ครบ ก่อนที่จะดำเนินการในขั้นต่อไป
6. ผู้ใช้งานไม่สามารถบันทึกผลสอบได้โดยตรง ควรทำให้มีการบันทึกผลสอบของผู้ใช้งานได้ด้วยตนเองหลังการทดสอบเสร็จ
7. หน้าติดต่อเรา ควรใส่ชื่อที่ปรึกษาด้วย และควรระบุสถานะของผู้วิจัยว่าอยู่ในสถานะใด เช่น นิสิตปริญญาเอก สาขาวิชาการวัดและเทคโนโลยีทางวิทยาการปัญญา
8. ในขณะที่ทดสอบ หรือขณะที่โปรแกรมกำลังใช้งานในหน้าจออื่น ควรมีการซ่อนเมนูผลการทดสอบ เมนูติดต่อเรา และเมนูเข้าสู่ระบบ เพื่อไม่ให้เมนูที่ไม่เกี่ยวข้องมีความเด่นชัดตลอดเวลา ยกเว้นเมนูหน้าหลักที่สามารถปรากฏอยู่ตลอดเวลาได้ อย่างไรก็ตาม ควรย้ายตำแหน่งเมนูหน้าหลักไว้มุมใดมุมหนึ่งของหน้าจอ ขณะที่โปรแกรมกำลังทำงาน

2.2 ผลการประเมินโดยผู้ใช้งาน

การประเมินโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย ในด้านความสะดวกในการนำไปใช้ และด้านลักษณะทั่วไปของโปรแกรม โดยผู้ใช้งานจำนวน 30 คนซึ่งเป็นคนไทยอายุตั้งแต่ 18 ปีขึ้นไป สามารถใช้คอมพิวเตอร์เบื้องต้นได้ ระหว่างวันที่ 14-16 มีนาคม 2558 แสดงได้ดังตารางที่ 4-15 (รายละเอียดผลการประเมินความเหมาะสมของโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย โดยผู้ใช้งานเป็นรายข้อ แสดงในภาคผนวก ข)

ตารางที่ 4-15 ผลการประเมินความเหมาะสมของโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย โดยผู้ใช้งาน

ด้านที่ประเมิน	M	SD	ระดับความเหมาะสม
1. ความสะดวกในการนำไปใช้	4.55	0.34	มากที่สุด
2. ลักษณะทั่วไปของโปรแกรม	4.63	0.30	มากที่สุด
สรุปผลโดยรวม	4.59	0.29	มากที่สุด

จากการประเมินความเหมาะสมของโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทยโดยผู้ใช้งาน แสดงให้เห็นว่าโปรแกรมมีความเหมาะสมอยู่ในระดับมากที่สุด (M=4.59) โดยผู้ใช้งานมีข้อคิดเห็น ดังนี้

1. คำศัพท์บางคำในคำถามอาจเข้าใจยาก
2. คำถามที่ว่า ฉันมีชีวิตที่มีความสุขสบายตามสมควร คำว่า “ตามสมควร” ทำให้เกิดความเข้าใจ
3. ขณะทดลองใช้งานโปรแกรม บางครั้งกดปุ่มแล้วค้างอยู่ข้อเดิม ไม่เปลี่ยนเป็นข้อต่อไป
4. บางครั้งเมื่อคลิกพลาดไปโดนส่วนอื่นของหน้าจอ ที่ไม่ใช่ปุ่มตัวเลือก หน้าจอจะหลุดออกจากโปรแกรม และขึ้นหน้าจอผิดพลาด

จากผลการประเมินความเหมาะสมของโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์โดยผู้เชี่ยวชาญและผู้ใช้งาน ผู้วิจัยได้นำข้อคิดเห็นและข้อเสนอแนะต่าง ๆ ไปปรับปรุงโปรแกรมเพื่อให้โปรแกรมมีความเหมาะสมมากยิ่งขึ้น อย่างไรก็ตาม ผลจากการทดลองใช้ แสดงให้เห็นว่าโปรแกรมมีความเหมาะสมเพียงพอสำหรับการนำไปใช้เพื่อวัดความสุขของคนไทย และทำให้การทดสอบมีความรวดเร็ว แม่นยำขึ้น

ตอนที่ 3 ผลการศึกษาความสัมพันธ์ของผลการประมาณค่าความสุขที่ได้จากการทดสอบด้วยโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย กับการทดสอบด้วยข้อคำถามทั้งหมดในคลังข้อคำถาม

กลุ่มตัวอย่างสำหรับศึกษาความสัมพันธ์ของผลการประมาณค่าความสุขที่ได้จากการทดสอบด้วยโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย กับการทดสอบด้วยข้อคำถามทั้งหมดในคลังข้อคำถาม จำนวน 30 คน ปรากฏว่าส่วนใหญ่เป็นเพศหญิง (ร้อยละ 83.30) มีอายุเฉลี่ย 30.83 ปี กลุ่มตัวอย่างแต่ละคนได้รับการทดสอบทั้งสองรูปแบบ ซึ่งปรากฏผลการทดสอบดังตารางที่ 4-16

ตารางที่ 4-16 ระยะเวลาที่ใช้ในการทดสอบ (วินาที) จำนวนข้อคำถามที่ใช้ในการทดสอบ (ข้อ) ค่าความสุขที่ประมาณค่าได้ และคะแนนที่ได้จากการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ กับการทดสอบด้วยข้อคำถามทั้งหมดในคลังข้อคำถามของกลุ่มตัวอย่าง (n=30)

คนที่	ระยะเวลาที่ใช้ในการทดสอบ (วินาที)		จำนวนข้อคำถามที่ใช้ในการทดสอบ (ข้อ)		ค่าความสุขที่ประมาณค่าได้		คะแนนที่	
	CAT	ทุกข้อ	CAT	ทุกข้อ	CAT	ทุกข้อ	CAT	ทุกข้อ
	1	79	460	18	61	1.21	1.24	62
2	78	365	15	61	1.32	1.44	63	64
3	123	254	16	61	1.71	1.12	67	61
4	86	172	16	61	1.89	2.42	69	74
5	107	233	19	61	1.24	1.38	62	64
6	109	112	15	61	1.31	1.23	63	62
7	57	194	15	61	1.84	1.73	68	67
8	67	162	16	61	1.16	1.05	62	61
9	60	239	15	61	1.58	1.48	66	65
10	122	281	18	61	0.73	1.20	57	62

ตารางที่ 4-16 (ต่อ)

คนที่	ระยะเวลาที่ใช้ในการ		จำนวนข้อคำถามที่ใช้ใน		ค่าความสุขที่		คะแนนที่	
	ทดสอบ (วินาที)		การทดสอบ (ข้อ)		ประมาณค่าได้			
	CAT	ทุกข้อ	CAT	ทุกข้อ	CAT	ทุกข้อ	CAT	ทุกข้อ
11	65	198	17	61	1.36	1.58	64	66
12	77	239	18	61	1.90	2.32	69	73
13	57	211	16	61	1.22	1.12	62	61
14	128	341	22	61	0.85	1.21	59	62
15	57	202	17	61	0.79	1.21	58	62
16	85	337	20	61	0.97	1.62	60	66
17	72	238	18	61	1.44	1.64	64	66
18	90	261	19	61	1.55	1.69	65	67
19	55	217	15	61	1.62	1.67	66	67
20	58	167	18	61	1.49	1.74	65	67
21	109	289	17	61	-0.18	-0.14	48	49
22	131	278	17	61	0.87	0.69	59	57
23	74	232	16	61	1.63	1.35	66	64
24	55	321	17	61	0.89	0.97	59	60
25	121	474	15	61	0.78	0.75	58	58
26	143	357	18	61	1.50	1.99	65	70
27	68	425	16	61	1.22	1.39	62	64
28	89	236	15	61	0.37	0.51	54	55
29	67	223	16	61	1.78	1.62	68	66
30	114	181	20	61	1.29	1.55	63	66
Min	55.00	112.00	15.00	61.00	-0.18	-0.14	48.00	49.00
Max	143.00	474.00	22.00	61.00	1.90	2.42	69.00	74.00
M	86.77	263.30	17.00	61.00	1.24	1.36	62.43	63.60
SD	27.00	88.23	1.78	0.00	0.47	0.51	4.64	4.99

จากตารางที่ 4-16 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 มีค่าเท่ากับ 1.24 (คะแนนที่ 62.43 คะแนน) ขณะที่ค่าความสุขเฉลี่ยที่ประมาณค่าได้จากการทดสอบด้วยข้อคำถามทั้งหมดในคลัง มีค่าเท่ากับ 1.36 (คะแนนที่ 63.60 คะแนน) ซึ่งมีค่าแตกต่างกันเพียงเล็กน้อย

ผลการวิเคราะห์ค่าสัมประสิทธิ์สหสัมพันธ์เพียร์สันระหว่างผลการประมาณค่าความสุข (Theta) ที่ได้จากการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ กับการทดสอบด้วยข้อคำถามทั้งหมดในคลังข้อคำถาม แสดงดังตารางที่ 4-17

ตารางที่ 4-17 ค่าสัมประสิทธิ์สหสัมพันธ์เพียร์สันระหว่างผลการประมาณค่าความสุข (Theta) ที่ได้จากการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ กับการทดสอบด้วยข้อคำถามทั้งหมดในคลังข้อคำถาม

	THS-CAT			
	ความพึงพอใจในชีวิต	อารมณ์ความรู้สึกเชิงบวก	อารมณ์ความรู้สึกเชิงลบ	ความสุขโดยรวม
ความพึงพอใจในชีวิต	.82**	-	-	-
THS-ทุกข้อ				
อารมณ์ความรู้สึกเชิงบวก	-	.72**	-	-
อารมณ์ความรู้สึกเชิงลบ	-	-	.44*	-
ความสุขโดยรวม	-	-	-	.85**

* $p < .05$, ** $p < .01$

จากตารางที่ 4-17 ปรากฏว่า ผลการประมาณค่าความสุข (Theta) ที่ได้จากการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ กับการทดสอบด้วยข้อคำถามทั้งหมดในคลังข้อคำถามมีความสัมพันธ์กันทางบวก โดยมีค่าสัมประสิทธิ์สหสัมพันธ์เพียร์สันเท่ากับ 0.85 ที่ระดับนัยสำคัญทางสถิติ .01 ซึ่งเป็นไปตามสมมติฐานการวิจัยที่ตั้งไว้

บทที่ 5

สรุปผลและอภิปรายผล

การวิจัยนี้มีวัตถุประสงค์เพื่อสร้างคลังข้อคำถามที่มีคุณภาพ ด้วยการวิเคราะห์คุณภาพ รายชื่อของข้อคำถามตามทฤษฎีการตอบสนองข้อสอบและตรวจสอบความตรงเชิงโครงสร้างของข้อคำถามในคลังข้อคำถาม นอกจากนี้ ได้พัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทยตามแนวคิดของวงจรการพัฒนาระบบ และศึกษาความสัมพันธ์ของผลการประมาณค่าความสุขที่ได้จากโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย กับการทดสอบโดยใช้ข้อคำถามทั้งหมดในคลังข้อคำถาม การดำเนินการวิจัยแบ่งเป็น 3 ขั้นตอน ได้แก่ 1) การสร้างคลังข้อคำถามสำหรับมาตรวัดความสุขของคนไทย รูปแบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ 2) การพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย และ 3) การศึกษาความสัมพันธ์ของผลการประมาณค่าความสุขที่ได้จากโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย กับการทดสอบโดยใช้ข้อคำถามทั้งหมดในคลังข้อคำถาม กลุ่มตัวอย่างเป็นคนไทย อายุตั้งแต่ 18 ปีขึ้นไป โดยกลุ่มตัวอย่างที่ใช้สำหรับการตรวจสอบคุณภาพของคลังข้อคำถาม มีจำนวน 934 คน และกลุ่มตัวอย่างที่ใช้ในการศึกษาความสัมพันธ์ของผลการประมาณค่าความสุขที่ได้จากโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย กับการทดสอบโดยใช้ข้อคำถามทั้งหมดในคลังข้อคำถาม มีจำนวน 30 คน วิเคราะห์ค่า Item-Total Correlations ด้วยโปรแกรม LERTAP 5.10.4 วิเคราะห์ค่าพารามิเตอร์ของข้อคำถามตามทฤษฎีการตอบสนองข้อสอบด้วยโปรแกรม Xcalibre 4.2.2.1 ตรวจสอบความตรงเชิงโครงสร้างและข้อตกลงเบื้องต้นตามทฤษฎีการตอบสนองข้อสอบด้วยโปรแกรม LISREL 8.8 วิเคราะห์ค่าสถิติพื้นฐานและค่าสัมประสิทธิ์สหสัมพันธ์ด้วยโปรแกรม SPSS

สรุปผลการวิจัย

1. ผลการสร้างคลังข้อคำถามสำหรับมาตรวัดความสุขของคนไทย ตามทฤษฎีความอยู่ดีมีสุขเชิงอัตวิสัย ซึ่งวิเคราะห์คุณภาพรายชื่อของข้อคำถามด้วยการตรวจสอบความตรงเชิงเนื้อหาโดยผู้เชี่ยวชาญ ทดลองใช้ วิเคราะห์ค่า Item-Total Correlations วิเคราะห์คุณภาพรายชื่อตามทฤษฎีการตอบสนองข้อสอบ รวมทั้งตรวจสอบความตรงเชิงโครงสร้างของคลังข้อคำถาม ได้ข้อคำถามในคลังข้อคำถามสุดท้าย จำนวนทั้งสิ้น 61 ข้อ ดังนี้

- 1.1 มาตรวัดย่อยความพึงพอใจในชีวิต จำนวน 13 ข้อ
- 1.2 มาตรวัดย่อยอารมณ์ความรู้สึกเชิงบวก จำนวน 22 ข้อ
- 1.3 มาตรวัดย่อยอารมณ์ความรู้สึกเชิงลบ จำนวน 26 ข้อ

ข้อคำถามทุกข้อผ่านเกณฑ์การตรวจสอบความตรงเชิงเนื้อหาโดยผู้เชี่ยวชาญ โดยมีค่าเฉลี่ยความสอดคล้องของข้อคำถามกับนิยามเชิงปฏิบัติการมากกว่า 4.00 มีค่า Item-Total Correlations มากกว่า 0.30 ผ่านเกณฑ์การตรวจสอบข้อตกลงเบื้องต้นของทฤษฎีการตอบสนอง

ข้อสอบ และเกณฑ์การตรวจสอบความสอดคล้องของข้อมูลเชิงประจักษ์ที่ได้จากการตอบข้อคำถาม กับ Graded Response Model ผลการวิเคราะห์ค่าพารามิเตอร์ของข้อคำถามตามทฤษฎีการตอบสนองข้อสอบ ปรากฏว่า ข้อคำถามมีค่าพารามิเตอร์ความชันอยู่ระหว่าง 0.44 ถึง 1.53 และพารามิเตอร์เทรซโฮลด์อยู่ระหว่าง -4.81 ถึง 4.70

ผลการตรวจสอบความตรงเชิงโครงสร้างของคลังข้อคำถามสำหรับวัดความสุขของคนไทย ด้วยการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสอง ปรากฏว่าโมเดลมีความสอดคล้องกับข้อมูลเชิงประจักษ์อยู่ในเกณฑ์ดี โดยมีค่า Chi-Square เท่ากับ 1067.36; $p = 0.47$ ท้องศาอิสระเท่ากับ 1065 ค่า CFI, GFI และ AGFI มีค่ามากกว่า 0.90 ค่า RMSEA และ SRMR มีค่าน้อยกว่า 0.05 ส่วนค่าน้ำหนักองค์ประกอบแต่ละองค์ประกอบของมาตรวัด มีค่าตั้งแต่ 0.53 – 0.89 จำแนกได้ ดังนี้ ด้านความพึงพอใจในชีวิตมีน้ำหนักองค์ประกอบ 0.89 ด้านอารมณ์ความรู้สึกเชิงบวก มีค่าน้ำหนักองค์ประกอบ 0.53 และด้านอารมณ์ความรู้สึกเชิงลบ มีค่าน้ำหนักองค์ประกอบ 0.64 ค่าน้ำหนักองค์ประกอบมีนัยสำคัญทางสถิติที่ระดับ .01 ทุกค่า แสดงว่าโครงสร้างของคลังข้อคำถามสำหรับวัดความสุขของคนไทยที่พัฒนาขึ้นประกอบด้วย 3 องค์ประกอบ ตามทฤษฎีความอยู่ดีมีสุขเชิงอัตวิสัย

2. ผลการพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย

2.1 โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทยที่พัฒนาขึ้นอยู่ในรูปแบบเว็บแอปพลิเคชัน ซึ่งสามารถเข้าใช้งานได้ที่เว็บไซต์ <http://www.thscat.com/Test/> การประเมินความเหมาะสมของโปรแกรมโดยผู้เชี่ยวชาญปรากฏว่าโปรแกรมมีความเหมาะสมอยู่ในระดับมาก ทั้งในด้านความสะดวกในการใช้งาน ความถูกต้องในการใช้งาน ลักษณะทั่วไปของโปรแกรม และความชัดเจนของคู่มือการใช้โปรแกรม ในขณะเดียวกัน ปรากฏว่าโปรแกรมเป็นที่ยอมรับของผู้ใช้งานที่ได้ทดลองใช้ ทั้งในด้านความสะดวกในการใช้งาน และด้านลักษณะทั่วไปของโปรแกรม

2.2 คลังข้อคำถาม สำหรับมาตรวัดความสุขของคนไทย บรรจุข้อคำถามที่วิเคราะห์ตามทฤษฎีการตอบสนองข้อสอบ แบบ Graded Response Model จำนวน 61 ข้อ จำแนกเป็น 1) มาตรวัดย่อยความพึงพอใจในชีวิต 13 ข้อ 2) มาตรวัดย่อยอารมณ์ความรู้สึกเชิงบวก 22 ข้อ และ 3) มาตรวัดย่อยอารมณ์ความรู้สึกเชิงลบ 26 ข้อ โดยข้อคำถามมีค่าพารามิเตอร์ความชันอยู่ระหว่าง 0.44 ถึง 1.53 และพารามิเตอร์เทรซโฮลด์อยู่ระหว่าง -4.81 ถึง 4.70 คลังข้อคำถามสามารถเพิ่มมาตรวัดย่อยได้อีกไม่เกิน 10 มาตรวัดย่อย และสามารถเพิ่มข้อคำถามแบบมาตรประมาณค่า 5 ระดับได้ไม่จำกัด ขึ้นอยู่กับขนาดของ Server โดยข้อคำถามจะต้องทราบค่าพารามิเตอร์ความชันและค่าพารามิเตอร์เทรซโฮลด์ที่ได้จากการประมาณค่าด้วย Graded Response Model

2.3 กระบวนการทดสอบ ของโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ มีสองรูปแบบ รูปแบบที่หนึ่งเป็นการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ซึ่งประกอบด้วย 4 ขั้นตอน ได้แก่ 1) การคัดเลือกข้อคำถามเริ่มต้นด้วยข้อคำถามที่ให้ค่าสารสนเทศสูงสุด (Maximum Information) ณ ตำแหน่งค่าคุณลักษณะแฝงของผู้ทดสอบที่กำหนดไว้ เท่ากับ 0.00 2) การประมาณค่าคุณลักษณะแฝงของผู้ทดสอบทุกครั้งที่มีการตอบข้อคำถาม โดยใช้วิธีการ Expected a Posteriori (EAP) 3) การคัดเลือกข้อคำถามข้อถัดไป ด้วยข้อคำถามซึ่งมีค่าสารสนเทศ

สูงที่สุด ณ ตำแหน่งค่าคุณลักษณะแฝงที่ประมาณค่าได้ และ 4) การยุติการทดสอบ เมื่อถึงเกณฑ์ที่กำหนดไว้ คือ ความคลาดเคลื่อนจากการประมาณค่าคุณลักษณะแฝงมีค่าต่ำกว่า 0.30 หรือข้อคำถามหมดคลัง รูปแบบที่สองเป็นการทดสอบด้วยข้อคำถามทั้งหมดในคลังข้อคำถาม ซึ่งมีขั้นตอนที่ 1-3 เช่นเดียวกับการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ แต่การยุติการทดสอบมีเกณฑ์เดียว คือ เมื่อข้อคำถามหมดคลัง

3. ผลการศึกษาความสัมพันธ์ของผลการประมาณค่าความสุขที่ได้จากโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย กับการทดสอบโดยใช้ข้อคำถามทั้งหมดในคลังข้อคำถาม

3.1 การทดสอบด้วยโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย ช่วยลดจำนวนข้อคำถามที่ใช้ในการทดสอบคิดเป็นร้อยละ 72.13 เมื่อเทียบกับการทดสอบด้วยข้อคำถามทั้งหมดในคลัง และช่วยลดระยะเวลาในการทดสอบลง คิดเป็นร้อยละ 67.05 เมื่อเทียบกับการทดสอบด้วยข้อคำถามทั้งหมดในคลัง

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

อภิปรายผล

ผู้วิจัยแบ่งการอภิปรายผลการวิจัย ออกเป็น 3 ประเด็น คือ 1) การสร้างคลังข้อคำถาม สำหรับมาตรวัดความสุขของคนไทย รูปแบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ 2) การพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย และ 3) การตรวจสอบประสิทธิภาพของโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย ดังนี้

1. การสร้างคลังข้อคำถามสำหรับมาตรวัดความสุขของคนไทย รูปแบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์

การพัฒนาค้างข้อคำถาม สำหรับมาตรวัดความสุขของคนไทย ตามทฤษฎีความอยู่ดีมีสุขเชิงอัตวิสัย โดยใช้ทฤษฎีการตอบสนองข้อสอบเป็นแนวทางในการตรวจสอบคุณภาพของข้อคำถาม รายข้อ รวมทั้งตรวจสอบความตรงเชิงเนื้อหาและความตรงเชิงโครงสร้างของคลังข้อคำถาม ปรากฏว่าคลังข้อคำถามเริ่มต้นมีข้อคำถามจำนวนทั้งสิ้น 191 ข้อ จากนั้นได้นำข้อคำถามทั้งหมดไปวิเคราะห์และตรวจสอบคุณภาพตามเกณฑ์ต่าง ๆ จนกระทั่งเหลือข้อคำถามในคลังข้อคำถามสุดท้าย จำนวนทั้งสิ้น 61 ข้อ เป็นข้อคำถามในมาตรวัดย่อยความพึงพอใจในชีวิต 13 ข้อ มาตรวัดย่อยอารมณ์ความรู้สึกเชิงบวก 22 ข้อ และมาตรวัดย่อยอารมณ์ความรู้สึกเชิงลบ 26 ข้อ

การตรวจสอบคุณภาพของข้อคำถามในแต่ละชั้นตอนมีการปรับปรุง แก้ไข รวมถึงคัดข้อคำถามที่ไม่เหมาะสมออกจากคลังข้อคำถาม โดยเฉพาะอย่างยิ่ง ในขั้นตอนการพิจารณาความสอดคล้องของข้อมูลเชิงประจักษ์ที่ได้จากการตอบข้อคำถามกับ Graded Response Model ที่ปรากฏว่า มีการคัดข้อคำถามออกจากคลังข้อคำถามมากที่สุด สาเหตุส่วนหนึ่งอาจเนื่องมาจากดัชนีตรวจสอบความสอดคล้องระหว่างข้อมูลเชิงประจักษ์ที่ได้จากการตอบข้อคำถามกับโมเดลการตอบสนองข้อสอบ มักมีความไว (Sensitivity) ต่อกลุ่มตัวอย่างขนาดเล็ก จึงอาจทำให้พบความไม่สอดคล้องระหว่างข้อมูลเชิงประจักษ์ที่ได้จากการตอบข้อคำถามกับโมเดลการตอบสนองข้อสอบมากกว่าที่ควรจะเป็น (Fliege et al., 2005) งานวิจัยหลายเรื่องนิยมสร้างคลังข้อคำถาม ด้วยการรวบรวมข้อคำถามจากมาตรวัดเดิมที่มีใช้อยู่แล้ว โดยมาตรวัดเหล่านั้นอาจเป็นมาตรวัดต่างฉบับแต่มีงวัดคุณลักษณะแฝงเดียวกัน และมีข้อมูลการตอบข้อคำถามอยู่ก่อนแล้ว การนำข้อมูลเดิมไปวิเคราะห์ใหม่ด้วยทฤษฎีการตอบสนองข้อสอบ จึงทำได้ง่ายและไม่ประสบปัญหาเรื่องขนาดของกลุ่มตัวอย่าง (Walter, Becker, & Bjorner, 2007; Rose, Bjorner, Becker, Fries, & Ware, 2006; Kocalevent et al., 2009) ในทางตรงกันข้ามกับมาตรวัดความสุข ถึงแม้จะมีการพัฒนามาตรวัดความสุขขึ้นหลายฉบับ แต่มาตรวัดแต่ละฉบับมีแนวคิดและนิยามความสุขที่แตกต่างกัน จึงไม่สามารถนำข้อมูลเดิมมาใช้ในการวิเคราะห์ได้ ผู้วิจัยจึงต้องสร้างข้อคำถามขึ้นเอง และนำไปเก็บรวบรวมข้อมูลใหม่ รวมทั้งตรวจสอบคุณภาพของข้อคำถามหลายชั้นตอนเพื่อให้ได้ข้อคำถามที่มีความเหมาะสม การสร้างมาตรวัดขึ้นใหม่ไม่ใช่เรื่องที่ทำได้โดยง่าย อีกทั้งด้วยข้อจำกัดด้านเวลาและค่าใช้จ่าย ทำให้ไม่สามารถใช้กลุ่มตัวอย่างขนาดใหญ่ได้เท่ากับงานวิจัยอื่น เมื่อกลุ่มตัวอย่างมีขนาดเล็ก จึงอาจทำให้พบปัญหาความไม่สอดคล้องของข้อมูลเชิงประจักษ์กับโมเดลการตอบสนองข้อสอบได้บ้าง อย่างไรก็ตาม ปัญหาความไม่สอดคล้องของข้อมูลเชิงประจักษ์กับโมเดลการตอบสนองข้อสอบได้บ้าง อย่างไรก็ตาม ขนาดกลุ่มตัวอย่างที่ใช้ในงานวิจัยนี้ก็มีขนาดเพียงพอต่อการประมาณค่าพารามิเตอร์ของข้อคำถามด้วย Graded Response Model (De Ayala, 2009, p. 223)

งานวิจัยหลายเรื่องให้ความสำคัญกับดัชนีความสอดคล้อง (Item Fit Indices) และพยายามแก้ปัญหาความไม่สอดคล้องของข้อคำถามกับโมเดลการตอบสนองข้อสอบด้วยการยุบรวมตัวเลือกรายการคำตอบให้มีจำนวนลดลง หรือคัดข้อคำถามที่ไม่ผ่านเกณฑ์ออกจากคลังข้อคำถาม (Rose, Bjorner, Becker, Fries, & Ware, 2006; Smits, Cuijpers, & Straten, 2010) ดัชนีความสอดคล้องยังสะท้อนให้เห็นถึงปัญหาเกี่ยวกับความไม่เหมาะสมของภาษาที่ใช้ในข้อคำถาม หรือปัญหาที่เกิดจากผลการตอบข้อคำถามของผู้ทดสอบบางคน ซึ่งทำให้ผลการประมาณค่าพารามิเตอร์ของข้อคำถามคลาดเคลื่อนไปจากที่ควรจะเป็น (Embretson & Reise, 2000, pp. 245-246) ผู้วิจัยให้ความสำคัญกับการพิจารณาดัชนีความสอดคล้องระหว่างข้อมูลเชิงประจักษ์กับโมเดลการตอบสนองข้อสอบด้วย เนื่องจากเมื่อพิจารณาข้อคำถามที่ไม่ผ่านเกณฑ์ดัชนีความสอดคล้องอย่างละเอียด พบว่า ข้อคำถามที่ไม่ผ่านเกณฑ์หลายข้อ มีปัญหาเรื่องค่าพารามิเตอร์เทรซโฮลด์ ที่มีค่าต่ำหรือสูงผิดปกติด้วย ดังนั้น ผู้วิจัยจึงตัดสินใจใช้การพิจารณาความสอดคล้องระหว่างข้อมูลเชิงประจักษ์ที่ได้จากการตอบข้อคำถามกับโมเดลการตอบสนองข้อสอบเป็นเกณฑ์หนึ่งในการคัดเลือกข้อคำถาม การวิจัยที่เกี่ยวข้องกับการพัฒนาคลังข้อคำถาม สำหรับการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ในด้านการวัดความสามารถ เป็นแนวคิดที่เกิดขึ้นมาหลายทศวรรษ เช่น ความสามารถทางภาษา ความสามารถทางคณิตศาสตร์ โดยปกติแล้ว ข้อคำถามในคลังข้อคำถามควรมีจำนวนไม่ต่ำกว่า 100

ข้อ อย่างไรก็ดี ในบริบททางด้านการวัดบุคลิกภาพและเจตคตินั้น การพัฒนาข้อคำถามให้มีจำนวน มากเช่นเดียวกับการวัดความสามารถ เป็นเรื่องที่เป็นไปได้ยาก (Embretson & Reise, 2000, pp. 264-266) ปัจจุบันมีการศึกษาวิจัยที่นำการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์มาใช้ในการ ประเมินทางด้านบุคลิกภาพ เจตคติ และตัวแปรทางด้านจิตวิทยาอื่น ๆ เพิ่มมากขึ้น ซึ่งงานวิจัย เหล่านี้มีจำนวนข้อคำถามในคลังข้อคำถามแตกต่างกันไป ถึงแม้จะมีข้อคำถามในคลังข้อคำถามไม่มากนัก แต่ผลการวิจัยก็ปรากฏว่า การพัฒนามาตรวัดรูปแบบการทดสอบแบบปรับเหมาะด้วย คอมพิวเตอร์ช่วยลดจำนวนข้อคำถามและระยะเวลาที่ใช้ในการทดสอบได้เป็นอย่างดี อีกทั้งยังช่วยลด ความเบื่อหน่ายจากการทดสอบได้อีกด้วย (Fliege et al., 2005; Becker et al., 2008; Kocalevent et al., 2009; Smits, Cuijpers, & Straten, 2010; Smits, Zitman, Cuijpers, Hollander-Gijsman, & Carlier, 2012; Zheng, Chang, & Chang, 2013)

การกำหนดขนาดของคลังข้อคำถาม ไม่ได้พิจารณาจากคุณลักษณะแฝง (Latent Trait) หรือความสามารถ (Ability) ที่ต้องการวัดเพียงอย่างเดียว ลักษณะการตอบและการตรวจให้คะแนนก็ มีผลต่อการพิจารณาขนาดของคลังข้อคำถามด้วย แบบสอบที่วัดความสามารถส่วนใหญ่มีการตรวจให้ คะแนนแบบ 2 ค่า (Dichotomous Response) คือ ตอบถูกได้ 1 คะแนน และตอบผิดได้ 0 คะแนน ข้อสอบข้อหนึ่งจึงมีโอกาที่ฟังก์ชันสารสนเทศของข้อสอบจะครอบคลุมช่วงคุณลักษณะแฝงที่ต้องการ วัดน้อยกว่าข้อคำถามที่มีการตรวจให้คะแนนมากกว่า 2 ค่า (Polytomous Response) เนื่องจาก โมเดลการตอบสนองข้อสอบแบบที่มีการตรวจให้คะแนนมากกว่า 2 ค่านั้น (Polytomous IRT Model) จะคำนวณฟังก์ชันสารสนเทศของข้อคำถามข้อหนึ่งจากผลรวมฟังก์ชันสารสนเทศของแต่ละ ตัวเลือกรายการคำตอบในข้อคำถามนั้น ซึ่งแต่ละตัวเลือกรายการคำตอบมักแสดงถึงระดับ คุณลักษณะแฝงที่ต่างกัน ดังนั้น โอกาสที่ข้อคำถามข้อหนึ่งจะมีฟังก์ชันสารสนเทศครอบคลุมช่วง คุณลักษณะแฝงที่ต้องการวัดจึงมีมากกว่า ข้อคำถามที่ใช้ในคลังข้อคำถาม จึงมีจำนวนน้อยกว่าได้ (Embretson & Reise, 2000, p. 264)

ข้อมูลที่ได้จากการวิเคราะห์ข้อคำถามในคลังด้วยทฤษฎีการตอบสนองข้อสอบไม่เพียงแต่ ใช้เป็นข้อมูลสำหรับการดำเนินการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์เท่านั้น แต่สามารถ นำมาใช้เพื่อปรับปรุงคุณภาพของมาตรวัดรูปแบบกระดาศ-ดินสอ ได้หลายกรณี ได้แก่ 1) การเลือก ข้อคำถามที่เหมาะสมกับระดับคุณลักษณะแฝงของบุคคลซึ่งเป็นกลุ่มเป้าหมายที่จะนำมาตรวจวัดไปใช้ โดยพิจารณาจากค่าพารามิเตอร์เทรซโฮลด์ซึ่งสามารถบ่งบอกถึงความเหมาะสมระหว่างข้อคำถามกับ ระดับความสามารถหรือคุณลักษณะแฝงของบุคคลได้ด้วย เช่น ค่าพารามิเตอร์เทรซโฮลด์ของข้อ คำถามทางด้านสุขภาพจิตที่มีค่าอยู่ระหว่าง -3.43 ถึง -0.49 แสดงให้เห็นว่า ข้อคำถามนั้นเหมาะกับการวัดบุคคลที่มีสุขภาพจิตอยู่ในระดับต่ำ (สุขภาพจิตไม่ดี) 2) การพิจารณาจำนวนตัวเลือกรายการ คำตอบที่เหมาะสมกับข้อคำถาม โดยพิจารณาจากโค้งการเลือกรายการคำตอบ และ/หรือ โค้ง คุณลักษณะข้อสอบ เพราะตัวเลือกรายการคำตอบของข้อคำถามบางข้ออาจมีความซ้ำซ้อนกัน ซึ่ง สะท้อนให้เห็นว่า การเลือกรายการคำตอบในตัวเลือกรายการคำตอบที่ซ้ำกันนั้น ให้ข้อมูลเกี่ยวกับ ผู้ทดสอบได้ไม่แตกต่างกัน ดังนั้น จึงสามารถลดจำนวนตัวเลือกรายการคำตอบที่เหลือเพียงตัวเลือกที่ เหมาะสม เช่น ผลการวิเคราะห์ข้อคำถามที่มีตัวเลือกรายการคำตอบแบบมาตรฐานค่า 5 ระดับ อาจพบว่าการเลือกตัวเลือก “เห็นด้วยอย่างยิ่ง” กับ “เห็นด้วยมาก” ให้ข้อมูลสารสนเทศเกี่ยวกับ

ผู้ทดสอบได้ไม่ต่างกัน จึงอาจพิจารณารวมตัวเลือกคำตอบ 2 ระดับนั้นเข้าด้วยกัน ซึ่งจะให้ข้อมูลเกี่ยวกับผู้ทดสอบได้มากกว่า 3) การลดความยาวของมาตรวัด สามารถพิจารณาฟังก์ชันสารสนเทศของข้อคำถามแล้วเลือกตัดข้อคำถามที่ให้ค่าฟังก์ชันสารสนเทศของข้อคำถามไม่แตกต่างกันออกจากมาตรวัด ทำให้มาตรวัดมีจำนวนข้อคำถามลดลง แต่มีคุณภาพไม่แตกต่างจากเดิม 4) การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ เพื่อคัดข้อคำถามที่มีแนวโน้มลำเอียงกับคนกลุ่มใดกลุ่มหนึ่งออกจากมาตรวัด ทำให้มาตรวัดที่พัฒนาขึ้นมีความเหมาะสมกับคนทุกกลุ่ม และ 5) การพัฒนามาตรวัดที่มีความแม่นยำ หรือความเที่ยงในระดับที่ยอมรับได้ แต่มีความยาวของมาตรวัดตามความต้องการ โดยการพิจารณาฟังก์ชันสารสนเทศของมาตรวัดทั้งฉบับ เนื่องจากฟังก์ชันสารสนเทศของมาตรวัดมีความสัมพันธ์โดยตรงกับความคลาดเคลื่อนมาตรฐานในการวัด ถ้ามาตรวัดมีความคลาดเคลื่อนมาตรฐานในการวัดต่ำ มาตรวัดนั้นจะมีความแม่นยำ หรือความเที่ยงสูง การพัฒนามาตรวัดจึงสามารถทำได้โดยการนำข้อคำถามทั้งหมดมาคำนวณหาค่าสารสนเทศของมาตรวัด แล้วจึงประเมินว่ามาตรวัดนั้นมีความเที่ยงถึงระดับที่ต้องการหรือไม่ นอกจากนี้ ยังสามารถคัดข้อคำถามบางข้อออกจากมาตรวัด หรือเพิ่มข้อคำถามบางข้อเข้าสู่มาตรวัด แล้วคำนวณค่าสารสนเทศของมาตรวัดใหม่ จนกระทั่งมาตรวัดนั้นมีคุณสมบัติตรงตามความต้องการ (Reeve & Fayers, 2005) ดังนั้น ข้อมูลที่ได้จากการวิเคราะห์ข้อคำถามในคลังข้อคำถามจึงเป็นข้อมูลที่เป็นประโยชน์และสามารถนำไปใช้ในการพัฒนามาตรวัดได้อีกหลายแนวทาง

การพัฒนาคลังข้อคำถาม สำหรับมาตรวัดความสุขของคนไทยในงานวิจัยนี้ แม้จะมีจำนวนข้อคำถามในคลังข้อคำถามน้อยกว่าคลังข้อคำถามทั่วไป แต่จากกระบวนการตรวจสอบคุณภาพของข้อคำถามอย่างละเอียด ทำให้มั่นใจได้ว่าข้อคำถามทุกข้อมีคุณภาพเพียงพอสำหรับการนำไปใช้เป็นส่วนหนึ่งในการพัฒนามาตรวัดความสุขของคนไทย รูปแบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ต่อไป นอกจากนี้ หลักการของการพัฒนาโปรแกรม ในด้านการบำรุงรักษาระบบ (Thompson & Weiss, 2011) ซึ่งเป็นแนวคิดหลักในการพัฒนามาตรวัดความสุขของคนไทย รูปแบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ยังเอื้อให้สามารถจัดการกับคลังข้อคำถามได้อยู่เสมอ ดังนั้น การเพิ่ม ลบ หรือแก้ไขข้อคำถาม จึงทำได้ไม่ยาก หากในอนาคตมีการพัฒนาหรือตรวจสอบคุณภาพของข้อคำถามเพิ่มเติม โดยประมาณค่าพารามิเตอร์ด้วย Graded Response Model ตามทฤษฎีการตอบสนองข้อสอบ ก็สามารถนำมาบรรจุในคลังข้อคำถามได้

2. การพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย

การพัฒนามาตรวัดความสุขของคนไทย รูปแบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ พัฒนาขึ้นตามแนวคิดของวงจรการพัฒนาระบบ ในรูปแบบ Web Application โดยวิเคราะห์ข้อคำถามและพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ บนพื้นฐานทฤษฎีการตอบสนองข้อสอบแบบเอกมิติ ตามนิยามความอยู่ดีมีสุขเชิงอัตวิสัย ซึ่งระบุว่าคุณภาพประกอบด้วย 3 องค์ประกอบ ได้แก่ ความพึงพอใจในชีวิต อารมณ์ความรู้สึกเชิงบวก และอารมณ์ความรู้สึกเชิงลบ องค์ประกอบทั้งสามเป็นองค์ประกอบที่แยกออกจากกัน การดำเนินการทดสอบจึงแยกออกจากกันตามแต่ละมาตรวัดย่อย แต่หากผู้ทดสอบทดสอบครบทั้งสามมาตรวัดย่อย โปรแกรมจะนำคะแนนที่ได้จากแต่ละมาตรวัดมาหาค่าเฉลี่ยเพื่อรายงานคะแนนความสุขโดยรวม

ตัวแปรทางจิตวิทยาส่วนใหญ่มีลักษณะเป็นพหุมิติ ที่มีหลายองค์ประกอบย่อย แต่การพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับตัวแปรทางจิตวิทยาส่วนใหญ่ ยังอยู่บนพื้นฐานทฤษฎีการตอบสนองข้อสอบแบบเอกมิติ ตัวอย่างเช่น การศึกษาของ Kocalevent et al. (2009) ที่พัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับมาตรวัดความเครียด ซึ่งจำแนกความเครียดออกเป็นสององค์ประกอบหลัก ได้แก่ การเผชิญความเครียด (Exposure to Stress) และการตอบสนองต่อความเครียด (Stress Reaction) งานวิจัยได้วิเคราะห์ข้อคำถามด้วย Generalized Partial Credit Model (GPCM) ซึ่งเป็นโมเดลการตอบสนองข้อสอบแบบเอกมิติที่มีการตรวจให้คะแนนมากกว่าสองค่า การดำเนินการทดสอบแยกออกจากกันตามแต่ละคลังข้อคำถามย่อย ผลการวิจัยปรากฏว่าโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่พัฒนาขึ้น ช่วยลดจำนวนข้อคำถามและระยะเวลาที่ใช้ในการทดสอบ มีความแม่นยำสูงในการประมาณค่าความเครียด แสดงให้เห็นว่าแม้ตัวแปรทางจิตวิทยาจะมีลักษณะเป็นพหุมิติ แต่สามารถพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ บนพื้นฐานทฤษฎีการตอบสนองข้อสอบแบบเอกมิติเพื่อวัดในแต่ละองค์ประกอบย่อยได้ การพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ในลักษณะนี้เป็นแนวทางที่พัฒนาได้ง่ายและซับซ้อนน้อยกว่าการพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ บนพื้นฐานทฤษฎีการตอบสนองข้อสอบแบบพหุมิติ

Gibbons et al. (2008) ได้พัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์สำหรับมาตรวัด The Mood and Anxiety Spectrum Scales (MASS) ที่ประกอบด้วย 4 องค์ประกอบหลัก ได้แก่ อารมณ์ (Mood) ความตื่นตระหนกและความกลัวที่โล่งแจ้ง (Panic-Agoraphobia) อาการย้ำคิดย้ำทำ (Obsessive-Compulsive Disorder) และความกลัวสังคม (Social Phobia) การวิจัยมีการวิเคราะห์เปรียบเทียบความสอดคล้องของข้อมูลที่ได้จากการตอบข้อคำถามกับโมเดลการตอบสนองข้อสอบแบบเอกมิติ และ Bifactor Model ซึ่งเป็นโมเดลการตอบสนองข้อสอบแบบพหุมิติ จากนั้นจึงตรวจสอบความสอดคล้องของโมเดลทั้งสอง ปรากฏว่าข้อมูลเชิงประจักษ์ที่ได้จากการตอบข้อคำถามมีความสอดคล้องกับ Bifactor Model มากกว่าโมเดลการตอบสนองข้อสอบแบบเอกมิติ นับเป็นงานวิจัยเรื่องแรกที่เริ่มนำโมเดลการตอบสนองข้อสอบแบบพหุมิติมาใช้ในการพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ โดยคาดหวังว่า Bifactor Model น่าจะมีความเหมาะสมกับตัวแปรทางด้านจิตวิทยาหรือมาตรวัดทางด้านสุขภาพมากขึ้น อย่างไรก็ตาม Bifactor Model มีข้อจำกัดหลายประการ ได้แก่ 1) การเลือกใช้ Bifactor Model จะต้องมีทฤษฎีหรืองานวิจัยที่แสดงให้เห็นความสัมพันธ์ระหว่างข้อคำถามกับองค์ประกอบต่าง ๆ มาก่อน 2) ถ้าข้อคำถามแต่ละข้อมีความสัมพันธ์กับองค์ประกอบใดองค์ประกอบหนึ่งเพียงองค์ประกอบเดียว ก็ไม่จำเป็นต้องใช้ Bifactor Model ซึ่งสามารถตรวจสอบลักษณะดังกล่าวได้ โดยใช้การวิเคราะห์องค์ประกอบเชิงยืนยัน และ 3) ผลการวิเคราะห์ข้อคำถามจะต้องแสดงให้เห็นว่าข้อคำถามนั้นเกี่ยวข้องกับองค์ประกอบหลัก 1 องค์ประกอบและมีความสัมพันธ์กับองค์ประกอบอื่นอีกเพียงองค์ประกอบเดียว จึงจะสามารถใช้ Bifactor Model ได้ แต่หากข้อคำถามมีความสัมพันธ์กับองค์ประกอบต่าง ๆ ในมาตรวัดมากกว่าสององค์ประกอบ จะต้องเลือกใช้โมเดลการตอบสนองข้อสอบแบบพหุมิติรูปแบบอื่น

การพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทยนี้ ไม่ได้มีการเปรียบเทียบความสอดคล้องของข้อคำถามกับโมเดลการตอบสนองแบบเอกมิตี และโมเดลการตอบสนองข้อสอบแบบพหุมิติโดยตรง แต่ได้วิเคราะห์องค์ประกอบเชิงยืนยันเพื่อตรวจสอบมิติของมาตรวัดเบื้องต้น ผลปรากฏว่า ข้อคำถามแต่ละข้อมีความสัมพันธ์กับองค์ประกอบหลักในมาตรวัดความสุขของคนไทยเพียงองค์ประกอบเดียว การพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย จึงเหมาะสมที่จะพัฒนาขึ้นบนพื้นฐานโมเดลการตอบสนองข้อสอบแบบเอกมิตีมากกว่าโมเดลการตอบสนองข้อสอบแบบพหุมิติ

การพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ โดยปกติแล้วต้องพิจารณาประเด็นที่สำคัญอีก 2 ประการ ได้แก่ 1) การควบคุมอัตราการใช้ข้อคำถาม (Item Exposure Control) ซึ่งเป็นการควบคุมไม่ให้ข้อคำถามข้อใดข้อหนึ่ง ถูกเลือกใช้มากกว่าข้อคำถามอื่น ๆ และ 2) การควบคุมสัดส่วนเนื้อหาของข้อคำถาม (Content Balancing Strategies) ซึ่งเป็นการควบคุมความสมดุลของการวัดตามองค์ประกอบที่กำหนดไว้ และเป็นการระมัดระวังไม่ให้ผู้ทดสอบได้รับการทดสอบในเนื้อหาด้านใดด้านหนึ่งเพียงด้านเดียว (Thompson & Weiss, 2011; Wainer et al., 2000, pp. 108-111) กระบวนการเหล่านี้เป็นกระบวนการที่เกี่ยวข้องกับการเลือกข้อคำถามข้อถัดไป ซึ่งในอดีตมีการพัฒนาวิธีการเลือกข้อคำถามข้อถัดไปไว้เป็นจำนวนมาก การวิจัยนี้ใช้วิธีการเลือกข้อคำถามข้อถัดไป โดยพิจารณาจากฟังก์ชันสารสนเทศของข้อคำถามที่เรียกว่า Maximum Information ซึ่งเป็นวิธีการที่ไม่มีการควบคุมเรื่องอัตราการใช้ข้อคำถามหรือสัดส่วนเนื้อหาของข้อคำถาม เนื่องจากคลังข้อคำถามของมาตรวัดความสุขของคนไทยมีจำนวนไม่มากนัก การควบคุมอัตราการใช้ข้อคำถามอาจส่งผลกระทบต่อกระบวนการทดสอบได้ นอกจากนี้ ข้อคำถามในแต่ละมาตรวัดย่อยเป็นข้อคำถามที่มุ่งวัดคุณลักษณะเดียวกัน ซึ่งไม่ได้ให้ความสำคัญกับการจำแนกข้อคำถามออกเป็นด้านย่อย ๆ การควบคุมสัดส่วนเนื้อหาของข้อคำถามจึงอาจไม่ใช่ประเด็นสำคัญในการวิจัยนี้ อีกทั้งการควบคุมอัตราการใช้ข้อคำถาม และการควบคุมสัดส่วนเนื้อหาของข้อคำถาม อาจส่งผลกระทบต่อประสิทธิภาพในการทดสอบได้ เพราะเป็นการเข้าไปแทรกแซงกระบวนการคัดเลือกข้อคำถาม (Thompson & Weiss, 2011; Wainer et al., 2000, pp. 108-111) ผู้วิจัยจึงไม่ได้นำเทคนิคดังกล่าวมาใช้ในงานวิจัยนี้

ประเด็นสำคัญอีกประการหนึ่งในการพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ คือ การประมาณค่าคุณลักษณะแฝง โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย ใช้วิธีการประมาณค่าคุณลักษณะแฝงด้วยวิธีเบย์สแบบ Expected a Posterior (EAP) เพราะเป็นวิธีการที่เหมาะสมกับการประมาณค่าความสูงมากที่สุด การประมาณค่าคุณลักษณะแฝงด้วยวิธี Maximum Likelihood (ML) ไม่สามารถนำมาใช้ในงานวิจัยนี้ได้ เนื่องจากการประมาณค่าคุณลักษณะแฝงด้วยวิธี ML มีข้อจำกัดที่ไม่สามารถประมาณค่าคุณลักษณะแฝงของผู้ทดสอบที่ตอบข้อคำถามถูกหรือผิดทั้งหมด (ในกรณีแบบทดสอบความสามารถ) หรือการให้คำตอบแบบสุดโต่ง นอกจากนี้ การประมาณค่าคุณลักษณะแฝงด้วยวิธี ML จะประมาณค่าคุณลักษณะแฝงได้ดี เมื่อมีข้อคำถามจำนวนมาก (เช่น 50 ข้อขึ้นไป) การประมาณค่าคุณลักษณะแฝงแบบเบย์ส (Bayesian Estimation Method) จึงเหมาะกับมาตรวัดความสุขมากกว่า วิธีการประมาณ

ค่าคุณลักษณะแฝงแบบเบส์จำแนกเป็นสองประเภท ได้แก่ Maximum a Posteriori (MAP) และ Expected a Posterior (EAP) การประมาณค่าคุณลักษณะแฝงด้วยวิธี MAP มีปัญหาเกี่ยวกับการถดถอยเข้าสู่ค่าเฉลี่ยมากกว่าการประมาณค่าคุณลักษณะแฝงด้วยวิธี EAP และมีค่าเฉลี่ยของความคลาดเคลื่อนกำลังสองในการประมาณค่าพารามิเตอร์มากกว่าการประมาณค่าคุณลักษณะแฝงด้วยวิธี EAP (De Ayala, 2009, pp. 75-79; Bock & Mislevy, 1982; Embretson & Reise, 2000, pp. 158-186; Keller, 2000; Thompson, 2009; Thompson & Weiss, 2011; Wainer et al., 2000, pp. 111-112) ผู้วิจัยจึงเลือกใช้การประมาณค่าคุณลักษณะแฝงด้วยวิธี EAP ซึ่งเป็นวิธีการเดียวกับที่ใช้ในการพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความซึมเศร้า มาตรวัดความวิตกกังวล มาตรวัดความเครียด และมาตรวัดทางด้านจิตใจอื่น ๆ (Fliege et al., 2005; Becker et al., 2008; Kocalevent et al., 2009; Kirisci et al., 2012)

การพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย ดำเนินการตามหลักวงจรการพัฒนาระบบ และมีการกำหนดองค์ประกอบในการดำเนินการทดสอบอย่างเหมาะสม นอกจากนี้ การประเมินความเหมาะสมของโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทยโดยผู้เชี่ยวชาญและผู้ใช้งาน ปรากฏว่า ผู้เชี่ยวชาญและผู้ใช้งานให้การยอมรับโปรแกรมเป็นอย่างดี แสดงให้เห็นว่าโปรแกรมมีคุณภาพเพียงพอสำหรับการนำไปใช้สำรวจความสุขของบุคคล และการศึกษาวิจัยทางด้านความสุข

3. การศึกษาความสัมพันธ์ของผลการประมาณค่าความสุขที่ได้จากโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย กับ การทดสอบโดยใช้ข้อคำถามทั้งหมดในคลังข้อคำถาม

ผลการศึกษาความสัมพันธ์ของผลการประมาณค่าความสุขที่ได้จากโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย กับการทดสอบด้วยข้อคำถามทั้งหมดในคลังข้อคำถาม ซึ่งมีความสัมพันธ์ในทางบวก แสดงให้เห็นว่าการดำเนินการทดสอบทั้งสองรูปแบบสามารถประเมินผลความสุขได้ในทิศทางเดียวกัน อย่างไรก็ตาม โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทยสามารถดำเนินการทดสอบโดยใช้จำนวนข้อคำถามและระยะเวลาในการทดสอบน้อยกว่าการทดสอบด้วยข้อคำถามทั้งหมดในคลังข้อคำถาม อีกทั้งมีความแม่นยำในการประมาณค่าสูง เนื่องจากกำหนดค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าให้มีค่าน้อยกว่า 0.30 เป็นเกณฑ์ในการยุติการทดสอบ ดังนั้น การประมาณค่าคุณลักษณะแฝงของผู้ทดสอบทุกคน จึงมีความแม่นยำมากกว่า 0.90 โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย จึงเป็นโปรแกรมการทดสอบที่มีประสิทธิภาพ สามารถลดทรัพยากรในการทดสอบ (ข้อคำถาม) ประหยัดเวลา และประมาณค่าความสุขได้สอดคล้องกับการประมาณค่าความสุขโดยใช้ข้อคำถามทั้งหมดในคลังข้อคำถาม จึงสามารถใช้เป็นอีกทางเลือกหนึ่งของการวัดความสุขที่มีประสิทธิภาพ ไม่ต่างจากมาตรวัดอื่น ๆ ที่พัฒนาขึ้นในรูปแบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ เช่น มาตรวัดความซึมเศร้า (Fliege et al., 2005) มาตรวัดความวิตกกังวล (Becker et al., 2008) มาตรวัดความเครียด (Kocalevent et al., 2009)

ข้อเสนอแนะในการนำผลการวิจัยไปใช้

การพัฒนามาตรวัดความสุขของคนไทย รูปแบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สามารถใช้ประโยชน์ได้ ดังนี้

1. โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย พัฒนาขึ้นในรูปแบบเว็บแอปพลิเคชัน ผู้สนใจสามารถเชื่อมต่อสัญญาณอินเทอร์เน็ตและเข้าใช้งานได้ที่ <http://www.thscat.com/Test/> เป็นประโยชน์สำหรับบุคคลทั่วไปในการเฝ้าระวังและสำรวจความสุขของตนเอง ขณะเดียวกันผู้ที่สนใจศึกษาวิจัยเกี่ยวกับความอยู่ดีมีสุขเชิงอัตวิสัย สามารถใช้โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย เป็นเครื่องมือในการศึกษาวิจัยได้

2. บุคคลหรือองค์กรสามารถนำผลการทดสอบที่ได้จากมาตรวัดความสุขของคนไทย รูปแบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ มาเป็นแนวทางในการวางแผนการส่งเสริมให้บุคคลมีความสุขเพิ่มขึ้น เนื่องจากผลการทดสอบที่ได้ แสดงข้อมูลทั้งในส่วนของมาตรวัดย่อยและมาตรวัดความสุขโดยรวม ดังนั้นจึงสามารถพิจารณารายละเอียดที่ได้จากผลการทดสอบของแต่ละมาตรวัด ซึ่งช่วยให้เห็นภาพชัดเจนยิ่งขึ้นว่าบุคคลมีจุดเด่นหรือจุดที่ควรส่งเสริมในด้านใด

ข้อเสนอแนะสำหรับการวิจัยต่อไป

1. ข้อคำถามในคลังข้อคำถาม สำหรับการวิจัยนี้ ได้จากการสร้างข้อคำถามขึ้นใหม่ทั้งหมด เนื่องจากตัวแปรที่ผู้วิจัยสนใจ ยังไม่มีมาตรวัดรูปแบบกระดาษ-ดินสอ อย่างไรก็ตาม ผู้ที่สนใจพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับตัวแปรอื่น ๆ ที่มีมาตรวัดรูปแบบกระดาษ-ดินสอ ใช้ก่อนแล้ว เช่น มาตรวัดความเครียด มาตรวัดความซึมเศร้า มาตรวัดความวิตกกังวล สามารถพัฒนาคลังข้อคำถามขึ้นจากข้อคำถามในมาตรวัดเดิม และใช้ผลการตอบข้อคำถามที่มีอยู่เดิม เป็นข้อมูลในการประมาณค่าพารามิเตอร์ด้วยโมเดลการตอบสนองข้อสอบได้ อย่างไรก็ตาม ต้องใช้เทคนิคการ Linking และ Equating เพื่อปรับให้ข้อคำถามจากมาตรวัดต่างฉบับกัน แต่มาตรวัดคุณลักษณะเดียวกันสามารถนำมาใช้ร่วมกันได้ ซึ่งจะช่วยลดระยะเวลาการวิจัยได้มาก

2. ข้อคำถามในคลังข้อคำถาม สำหรับมาตรวัดความสุขที่พัฒนาขึ้นยังมีจำนวนไม่มากนัก เนื่องจากมีการคัดข้อคำถามออกจากคลังเป็นจำนวนมากในขั้นตอนการพิจารณาความสอดคล้องของข้อมูลเชิงประจักษ์ที่ได้จากการตอบข้อคำถามกับ Graded Response Model อย่างไรก็ตาม สถิติที่ใช้ในการตรวจสอบความสอดคล้องดังกล่าวยังมีอยู่ไม่มากนัก และขึ้นอยู่กับโปรแกรมที่ใช้ในการประมาณค่าพารามิเตอร์ของข้อคำถาม จึงควรมีการศึกษาเกี่ยวกับเกณฑ์ในการตรวจสอบความสอดคล้องของข้อมูลเชิงประจักษ์ที่ได้จากการตอบข้อคำถามกับโมเดลการตอบสนองข้อสอบเพิ่มขึ้น

3. โมเดลการตอบสนองข้อสอบที่ใช้ในประมาณค่าพารามิเตอร์ของข้อคำถาม สำหรับมาตรวัดความสุขของคนไทย ยังจำกัดอยู่เพียง Graded Response Model ผู้สนใจสามารถใช้วิธีการประมาณค่าพารามิเตอร์ด้วยโมเดลการตอบสนองข้อสอบรูปแบบอื่น เช่น Generalized Partial Credit Model, Rating Scale Model, Multidimensional IRT Model, Bifactor Model

นอกจากนี้ ควรศึกษาเปรียบเทียบผลการประมาณค่าที่ได้จากแต่ละโมเดล เพื่อตัดสินใจว่าโมเดลใดมีความเหมาะสมกับข้อคำถามที่พัฒนาขึ้นมาที่สุด

4. เกณฑ์การคัดเลือกข้อคำถามที่สำคัญอีกประการหนึ่ง คือ การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ การพัฒนาคลังข้อคำถามในบริบทที่มีความแตกต่างทางวัฒนธรรมมาก ควรศึกษาเกี่ยวกับการทำหน้าที่ต่างกันของข้อสอบ เพื่อใช้เป็นเกณฑ์ในการคัดเลือกข้อคำถามด้วย

5. หลักการของทฤษฎีการตอบสนองข้อสอบ มีแนวคิดที่ว่าพารามิเตอร์ของข้อคำถาม และพารามิเตอร์ของบุคคลจะไม่แปรเปลี่ยนไปตามกลุ่มผู้สอบหรือมาตรวัดที่ใช้ อย่างไรก็ตาม การตรวจสอบคุณสมบัติดังกล่าวอยู่นอกเหนือขอบเขตงานวิจัยนี้ การวิเคราะห์พารามิเตอร์ของข้อคำถามตามทฤษฎีการตอบสนองข้อสอบในครั้งต่อไป จึงควรมีการตรวจสอบคุณสมบัติด้านความไม่แปรเปลี่ยนของพารามิเตอร์ของข้อคำถามและพารามิเตอร์บุคคลด้วย

6. การพัฒนามาตรวัด โดยอาศัยข้อมูลที่ได้จากการวิเคราะห์ด้วยทฤษฎีการตอบสนองข้อสอบ ทั้งในเรื่องพารามิเตอร์ของข้อคำถาม และฟังก์ชันสารสนเทศของข้อคำถามและมาตรวัด เป็นแนวทางหนึ่งที่มีควรมีการศึกษาวิจัยเพิ่มเติม เพื่อให้การพัฒนามาตรวัดมีคุณภาพดียิ่งขึ้น

7. การวิจัยนี้ไม่ได้้นำประเด็นด้านการควบคุมอัตราการใช้ข้อคำถาม และการควบคุมสัดส่วนเนื้อหาของข้อคำถามมาใช้ในการพัฒนาโปรแกรม เนื่องจากข้อจำกัดเรื่องจำนวนข้อคำถามในคลังข้อคำถาม การพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ในบริบทอื่น ๆ ที่มีจำนวนข้อคำถามในคลังข้อคำถามเพียงพอ ควรนำประเด็นนี้ไปพิจารณาในการพัฒนาโปรแกรมด้วย โดยอาจศึกษาด้วยวิธีการจำลองการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์โดยใช้โปรแกรม CATSim ก่อนการดำเนินการทดสอบจริงเพื่อค้นหาวิธีการที่เหมาะสมก่อนก็ได้

8. การพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์แต่ละครั้ง ใช้ระยะเวลา ค่าใช้จ่าย และความทุ่มเทอย่างมาก หากมีการจัดทำโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่มีความยืดหยุ่นต่อการกำหนดเกณฑ์การเริ่มต้นการทดสอบ การคัดเลือกข้อคำถามในการทดสอบ การประมาณค่าคุณลักษณะแฝง และการยุติการทดสอบที่หลากหลาย รวมทั้งมีคลังข้อคำถามที่สามารถปรับเปลี่ยนเนื้อหาข้อคำถามได้ตามความต้องการของผู้วิจัย จะช่วยเอื้อให้นักวิจัยสามารถทุ่มเทความรู้ และระยะเวลากับการพัฒนางานวิจัยในส่วนอื่น ๆ ได้มากขึ้น รวมทั้งทำให้องค์ความรู้ทางด้านนี้มีความเจริญก้าวหน้าอย่างรวดเร็วขึ้นกว่าเดิม

9. แนวทางการพัฒนาโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สามารถนำไปใช้ในการพัฒนามาตรวัดตัวแปรทางจิตวิทยาอื่น ๆ เช่น ความเครียด ความวิตกกังวล ความซึมเศร้า คุณภาพชีวิต พัฒนาการทางร่างกายและจิตใจ พัฒนาการทางสังคม การพัฒนาดังกล่าวจะก่อให้เกิดความเจริญก้าวหน้าทางการวัด และทำให้ได้เครื่องมือที่มีคุณภาพสำหรับการนำไปใช้ประโยชน์ต่อไป

บรรณานุกรม

- ชูฤทธิ์ เต็งไตรสรณ์ พระครุฑชิต คุณวโร และอรุณี วชิราพรทิพย์. (2551). การพัฒนามาตรวัดสุขภาวะ
องค์รวมแนวพุทธ. *วารสารวิชาการสาธารณสุข*, 17(6), 1650-1660.
- ธรรมจักร. (2556). *สู่ความสุข: ความสุขคืออะไร*. วันที่ค้นข้อมูล 6 ธันวาคม 2556, เข้าถึงได้จาก
<http://www.dhammadjak.net/book-dhammaraksa/-3-4.html>
- นรบรรณ พันธุมเมธา. (2549). *คลังคำ* (พิมพ์ครั้งที่ 3). กรุงเทพฯ: อมรินทร์.
- ประเวศ ตันติพิวัฒนสกุล และเอกอนงค์ สีตลาภินันท์. (2554). *คู่มือสร้างสุขระดับจังหวัด*
(พิมพ์ครั้งที่ 2). กรุงเทพฯ: โรงพิมพ์ชุมนุมสหกรณ์การเกษตรแห่งประเทศไทย จำกัด.
- ราชบัณฑิตยสถาน. (2542). *พจนานุกรมฉบับราชบัณฑิตยสถาน พ.ศ. 2542*. กรุงเทพฯ: นานมีบุ๊คส์
พับลิเคชั่น.
- ศิริชัย กาญจนวาสี. (2555). *ทฤษฎีการทดสอบแนวใหม่* (พิมพ์ครั้งที่ 4). กรุงเทพฯ: โรงพิมพ์แห่ง
จุฬาลงกรณ์มหาวิทยาลัย.
- ศูนย์สุขภาวะทางจิต คณะจิตวิทยา จุฬาลงกรณ์มหาวิทยาลัย. (2556). *ความสุขคืออะไร*. วันที่ค้น
ข้อมูล 6 ธันวาคม 2556, เข้าถึงได้จาก [http://www.chulawellness.com/index.php?
lay=show&ac=article&id=539568775&Ntype=11](http://www.chulawellness.com/index.php?lay=show&ac=article&id=539568775&Ntype=11)
- ศรินันท์ กิตติสุขสถิต, กาญจนา ตั้งชลทิพย์, สุภรต์ จรัสสิทธิ์, เฉลิมพล สายประเสริฐ, พอตา
บุญยัตริณะ และ วรรณภา อารีย์. (2555). *คู่มือการวัดความสุขด้วยตนเอง*
HAPPINOMETER (พิมพ์ครั้งที่ 1). นครปฐม: สถาบันวิจัยประชากรและสังคม
มหาวิทยาลัยมหิดล.
- เสรี ชัดแจ่ม. (2544). *เอกสารประกอบการสอนรายวิชา การวัดผลการศึกษา*. ชลบุรี: ภาควิชาวิจัย
และวัดผลการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยบูรพา.
- อรยา ปรีชาพานิช. (2557). *คู่มือเรียน การวิเคราะห์และออกแบบระบบ (System Analysis and
Design) ฉบับสมบูรณ์*. นนทบุรี: อดิษฐ์ พรีเมียร์.
- Allen, M. J., & Yen, W. M. (1979). *Introduction to Measurement Theory*. United States
of America: Waveland Press.
- Becker, J., Fliege, H. Kocalevent, R. D., Bjorner, J. B., Rose, M., Walter, O. B., & Klapp, B.
F. (2008). Functioning and Validity of a Computerized Adaptive Test to
Measure Anxiety (A-Cat). *Depression and Anxiety*, 25, 182-194.
- Bock, R. D., & Mislevy, R. J. (1982). Adaptive EAP Estimation of Ability in a
Microcomputer Environment. *Applied Psychological Measurement*, 6(4), 431-
444.
- Choi, S. W. (2009). Firestar: Computerized adaptive testing (CAT) simulation program
for polytomous IRT Models (Computer software). *Applied Psychological
Measurement*, 33, 644-645.

- Cohen, R. J., Swerdlik, M. E., & Sturman, E. D. (2013). *Psychological Testing and Assessment: An Introduction to Tests and Measurement* (8th ed.). United States of America: McGraw-Hill.
- Compton, W. C., & Hoffman, E. (2013). *Positive Psychology: The science of happiness and flourishing* (2nd ed.). United States of America: Wadsworth Cengage Learning.
- Crocker, L., & Algina, J. (2008). *Introduction to Classical and Modern Test Theory*. United States of America: Cengage Learning.
- De Ayala, R. J. (2009). *The Theory and Practice of Item Response Theory*. United States of America: The Guildford Press.
- Deci, E. L., & Ryan R. M. (2008). Self-Determination Theory: A Macrotheory of Human Motivation, Development, and Health. *Canadian Psychology*, 49(3), 182–185.
- Diener, E. (1984). Subjective well-being. *Psychological Bulletin*, 95(3), 542-575.
- Diener, E. (2000). Subjective well-being: The science of happiness and a proposal for a national index. *American Psychologist*, 55(1), 34-43.
- Diener, E., & Emmons, R. A. (1985). The independence of positive and negative affect. *Journal of Personality and Social Psychology*, 47(5), 1105-1117.
- Diener, E., Emmons, R. A., Larsen, R.J., & Griffin, S. (1985). Satisfaction with Life Scale (SWLS). *Journal of Personality Assessment*, 49, 71-75.
- Diener, E., Sandvik, E., & Pavot, W. (1991). Happiness is the frequency, not the intensity, of positive versus negative affect. In F.Strack, M. Argyle, & N. Schwarz (Eds.), *Subjective well-being: An interdisciplinary perspective* (pp. 119-139). New York: Pergamon.
- Diener, E., Suh, E. M., Lucas, R. E., & Smith, H. E. (1999). Subjective well-being: Three Decades of Progress. *Psychological Bulletin*, 125, 276-302.
- Diener, E., Wirtz, D., Tov, W., Kim-Prieto, C., Choi, D., Oishi, S., & Biswas-Diener, R. (2009). New measures of well-being: Flourishing and positive and negative feelings. *Social Indicators Research*, 39, 247-266.
- Edelen, M. O., & Reeve, B. R. (2007). Applying item response theory (IRT) modeling to questionnaire development, evaluation, and refinement. *Quality of Life Research*, 16, 5-18.
- Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists*. Mahwah, NJ: Lawrence Erlbaum.
- Fliege, H., Becker, J., Walter, O. B., Bjorner, J. B., Klapp, B. F., & Rose, M. (2005). Development of a computer-adaptive test for depression (D-CAT). *Quality of Life Research*, 14, 2277–2291.

- Frisch, M. B., Cornell, J., Villanueva, M., & Retzlaff, P. J. (1992). 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.
- Gravetter, F. J., & Forzano, L. B. (2012). *Research Methods for the Behavioral Sciences* (4th ed.). Canada: Wadsworth Cengage Learning.
- Gregory, R. J. (2007). *Psychological testing: history, principles, and applications*. Boston: Pearson/Allyn and Bacon.
- 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. (2008). Using Computerized Adaptive Testing to Reduce the Burden of Mental Health Assessment. *Psychiatric Services, 59*(4), 361–368.
- Han, K. T. (2007). WinGen: Windows software that generates IRT parameters and item responses. *Applied Psychological Measurement, 31*(5), 457-459.
- Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of item response theory*. United States of America: Sage Publications.
- Hooper, D., Coughlan, J., & Mullen, M. R. (2008). Structural Equation Modelling: Guidelines for Determining Model Fit. *Electronic Journal of Business Research Methods, 6*(1), 53-60.
- Keyes, C. L. M. (1998). Social Well-Being. *Social Psychology Quarterly, 61*(2), 121-140.
- Keller, L. A. (2000). Ability Estimation Procedures in Computerized Adaptive Testing. Retrieved August 12, 2014, from American Institute of CPAs: http://www.aicpa.org/BecomeACPA/CPAExam/PsychometricsandScoring/TechnicalReports/DownloadableDocuments/Keller_AbilityEstimation.pdf
- Kirisci, L., Tarter, R., Reynolds, M., Ridenour, T., Stone, C., & Vanyukov, M. (2012). Computer adaptive testing of liability to addiction: Identifying individuals at risk. *Drug and Alcohol Dependence, 123S*, S79-S86.
- Kocalevent, R. D., Rosec, M., Becker, J., Walterd, O. B., Fliegea, H., Bjorner, J. B., Kleiberb, D., & Klappa, 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.
- Lai, J. S., Butt, Z., Zelko, F., Cella, D., Krull, K. R., Kieran, M. W., & Goldman, S. (2011). Development of a Parent-Report Cognitive Function Item Bank Using Item Response Theory and Exploration of its Clinical Utility in Computerized Adaptive Testing. *Journal of Pediatric Psychology, 36*(7), 766–779.

- Lyubomirsky, S., & Lepper, H. S. (1999). A measure of subjective happiness: Preliminary reliability and construct validation. *Social Indicators Research*, *46*, 137-155.
- McCullough, M. E., Emmons, R. A., & Tsang, J. (2002). The Grateful Disposition: A conceptual and Empirical Topography. *Journal of Personality and Social Psychology*, *82*, 112-127.
- Miller, L. A., Lovler, R. L., & McIntire, S. A. (2013). *Foundations of Psychological Testing: A Practical Approach* (4th ed.). Los Angeles: SAGE.
- Myers, D. G., & Diener, E. (1995). Who is happy? *Psychological Science*, *6*(1), 10-19.
- Naragon, K., & Watson, D. (2013). Positive Affectivity. In S. J. Lopez (Eds.), *the Encyclopedia of Positive Psychology*. (pp. 707-711). Singapore: Wiley-Blackwell.
- Olino, T. M., Yu, L., Klein, D. N., Rohde, P., Seeley, J. R., Pilkonis, P. A., & Lewinsohn, P. M. (2012). Measuring depression using item response theory: an examination of three measures of depressive symptomatology. *International Journal of Methods in Psychiatric Research*, *21*(1), 76-85.
- Pavot, W., & Diener, E. (2013). Happiness Experienced: The Science of Subjective Well-being. In S. David, I. Boniwell, and A.C. Ayers (Eds.), *Oxford Handbook of Happiness*. (pp. 134-145). Oxford University Press.
- Pilkonis, P. A., Choi, S. W., Reise, S. P., Stover, A. M., Riley, W. T., & Cella, D. (2011). Item Banks for Measuring Emotional Distress from the Patient-Reported Outcomes Measurement Information System (PROMIS®): Depression, Anxiety, and Anger. *Assessment*, *18*(3), 263-283.
- Patient-Reported Outcome Measurement Information System (PROMIS). (2013, November). *PROMIS Overview*. Retrieved November 25, 2013, from the PROMIS Dynamic Tools to Measure Health Outcomes for the Patient Perspective Web site: <http://www.nihpromis.org/about/overview>
- Reeve, B. B., & Fayers, P. (2005). Applying item response theory modeling for evaluating questionnaire item and scale properties. In P. Fayers and R. D. Hays (Eds.), *Assessing Quality of Life in Clinical Trials: Methods of Practice* (2nd ed.) (pp. 55-73). Oxford University Press.
- Reynolds, C. R., & Livingston, R. B. (2012). *Mastering Modern Psychological Testing: Theory and Methods*. United States of America: Pearson.

- Rose, M., Bjorner, J. B., Becker, J., Fries, J. F., & Ware, J. E. (2006). 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.
- Ruch, W., Kohler, G., & Van Thriel (1996). Assessing the "humorous temperament": Construction of the facet and standard trait forms of the State-Trait-Cheerfulness-Inventory - STCI. In W. Ruch (Ed.), *Measurement of the sense of humor* [special issue]. *Humor: International Journal of Humor Research*, *9*, 303-339.
- Ryff, C. D. (1989). Happiness Is Everything, or Is It? Explorations on the Meaning of Psychological Well-Being. *Journal of Personality and Social Psychology*, *57*(6), 1069-1081.
- Ryff, C. D., & Singer, B. (1998). The contours of positive human health. *Psychological Inquiry*, *9*, 1-28.
- Seligman, M. E. P. (1998). *Building human strength: Psychology's forgotten mission*. *APA Monitor*, *29*(1).
- Shiota, M. N., & Kalat, J. W. (2012). *Emotion* (2nd ed.). Australia: Thomson Wadsworth.
- Schumacker, R. E., & Lomax, R. G. (2004). *A Beginner's Guide to Structural Equation Modeling* (2nd ed.). Mahwah, N.J.: Lawrence Erlbaum Associates.
- Simms L. J., Goldberg, L. R., Roberts, J. E., Watson, D., Welte, J., & Rotterman, J. H. (2011). Computerized Adaptive Assessment of Personality Disorder: Introducing the CAT-PD Project. *Personality Assessment*, *93*(4), 380-389.
- Smits, N., Cuijpers, P., & Straten, A. V. (2011). Applying computerized adaptive testing to the CES-D scale: A simulation study. *Psychiatry Research*, *188*, 147-155.
- Smits, N., Zitman, F. G., Cuijpers, P., Hollander-Gijsman, M. E., & Carlier, I. V. (2012). A proof of principle for using adaptive testing in routine Outcome Monitoring: the efficiency of the Mood and Anxiety Symptoms Questionnaire -Anhedonic Depression CAT. *BMC Medical Research Methodology*, *12*(4).
- Thomas, M. L. (2011). The Value of Item Response Theory in Clinical Assessment: A Review. *Assessment*, *18*(3), 291-307.
- Thompson, N. A. (2009). Ability Estimation with Item Response Theory [White paper]. Retrieved August 9, 2014, from Assessment Systems Corporation: [http://www.assessment.com/docs/Thompson_\(2009\)_-_Ability_estimation_with_IRT.pdf](http://www.assessment.com/docs/Thompson_(2009)_-_Ability_estimation_with_IRT.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>.

- Thrash, T. M., & Elliot, A. J. (2003). Inspiration as a psychological construct. *Journal of Personality and Social Psychology, 84*, 871-889.
- United Nations. (2012, 12 July). *Resolution adopted by the General Assembly. A/66/L.48/Rev.1.*
- Vogels, A. G., Jacobusse, G. W., & Reijneveld, S. A. (2011). An accurate and efficient identification of children with psychosocial problems by means of computerized adaptive testing. *BMC Medical Research Methodology, 11*, 1-9.
- Wainer, H., Dorans, N. J., Daniel, E., Flaugher, R., Green, B. F., Mislevy, R. J., Steinberg, L., & Thissen, D. (2000). *Computerized Adaptive Testing: A Primer* (2nd ed.). United States of America: Lawrence Erlbaum Associates.
- Walker, J., Böhnke, J. R., Cerny, T., & Strasser, F. (2010). Development of symptom assessments utilising item response theory and computer adaptive testing—a practical method based on a systematic review. *Critical Reviews in Oncology/Hematology, 73*, 47–67.
- Walter, O. B., Becker, J., Bjorner, J. B., Fliege, H., Klapp, B. F., & Rose M. (2007). Development and evaluation of a computer adaptive test for anxiety (Anxiety-CAT). *Quality of Life Research, 16*, 143-155.
- Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS Scales. *Journal of Personality and Social Psychology, 47*, 1063–1070.
- Weiss, D. J. (2011). Better Data From Better Measurements Using Computerized Adaptive Testing. *Journal of Methods and Measurement in the Social Sciences, 2*(1), 1-27.
- Weiss, D. J., & Guyer, R. (2012). *Manual for CATSim: Comprehensive simulation of computerized adaptive testing*. St. Paul MN: Assessment Systems Corporation.
- Wiersma, W., & Jurs, S.G. (1990). *Educational Measurement and Testing* (2nd ed.). Boston: Allyn and Bacon.
- Wikipedia, the free encyclopedia. (2013). *Happiness*. Retrieved December 6, 2013, from <https://en.wikipedia.org/wiki/Happiness>
- Yoes, M. (1997). *PARDSIM* (Computer software). Saint Paul, MN: Assessment Systems Corporation.
- Zheng, Y., Chang, C. H., & Chang, H. H. (2013). Content-balancing strategy in bifactor computerized adaptive patient-reported outcome measurement. *Quality of life research, 22*(3), 491-499.

ภาคผนวก

ภาคผนวก ก
รายชื่อผู้เชี่ยวชาญ

รายชื่อผู้เชี่ยวชาญในการตรวจสอบความสอดคล้องของข้อความ
กับนิยามเชิงปฏิบัติการ ในส่วนของการสร้างคลังข้อความ

- 1) รศ.กนกรัตน์ สุขะตุงคะ ที่ปรึกษาสาขาวิชาจิตวิทยาคลินิก ภาควิชาจิตเวชศาสตร์ คณะแพทยศาสตร์ศิริราชพยาบาล มหาวิทยาลัยมหิดล
- 2) รศ.ดร.สุชีรา ภัทรายุตวรรัตน์ อาจารย์ประจำสาขาวิชาจิตวิทยาคลินิก ภาควิชาจิตเวชศาสตร์ คณะแพทยศาสตร์ศิริราชพยาบาล มหาวิทยาลัยมหิดล
- 3) รศ.ดร.สุพิมพ์ ศรีพันธ์วรสกุล อาจารย์ประจำวิทยาลัยวิทยาการวิจัยและวิทยาการปัญญา มหาวิทยาลัยบูรพา
- 4) อาจารย์นัฐธิยา บุญอาพัทธ์เจริญ อาจารย์ประจำสาขาวิชามนุษยศาสตร์และสังคมศาสตร์ คณะศิลปศาสตร์ มหาวิทยาลัยหัวเฉียวเฉลิมพระเกียรติ
- 5) อาจารย์สร้อยสุดา อิมอรุณรักษ์ อาจารย์ประจำสาขาวิชาจิตวิทยาคลินิก ภาควิชาจิตเวชศาสตร์ คณะแพทยศาสตร์ศิริราชพยาบาล มหาวิทยาลัยมหิดล

ภาคผนวก ข
การพัฒนามาตรวัดความสุขของคนไทย รูปแบบกระดาษ-ดินสอ

1. การพัฒนามาตรวัดความสุขของคนไทย รูปแบบกระดาษ-ดินสอ
2. มาตรวัดความสุขของคนไทย รูปแบบกระดาษ-ดินสอ
3. คู่มือการใช้มาตรวัดความสุขของคนไทย รูปแบบกระดาษ-ดินสอ

การพัฒนามาตรวัดความสุขของคนไทย รูปแบบกระดาษ-ดินสอ

มาตรวัดความสุขของคนไทย รูปแบบกระดาษ-ดินสอพัฒนาขึ้นจากการรวบรวมข้อคำถามจากคลังข้อคำถามสำหรับวัดความสุขที่ผ่านการตรวจสอบคุณภาพแล้วมาจัดพิมพ์ในรูปแบบแบบรายงานตนเอง ซึ่งมีลักษณะการตอบแบบมาตราประมาณค่า 5 ระดับ โดยมีเกณฑ์การตรวจให้คะแนน ดังนี้

- 1) ความพึงพอใจในชีวิต มีเกณฑ์การตรวจให้คะแนน ดังนี้

ไม่เห็นด้วย/เห็นด้วยน้อยที่สุด	ให้คะแนน 1
เห็นด้วยน้อย	ให้คะแนน 2
เห็นด้วยปานกลาง	ให้คะแนน 3
เห็นด้วยมาก	ให้คะแนน 4
เห็นด้วยอย่างยิ่ง	ให้คะแนน 5
- 2) อารมณ์ความรู้สึกเชิงบวก มีเกณฑ์การตรวจให้คะแนน ดังนี้

ไม่เคยหรือแทบจะไม่เคยเกิดอารมณ์ความรู้สึกดังกล่าว	ให้คะแนน 1
เกิดอารมณ์ความรู้สึกดังกล่าวน้อยครั้ง	ให้คะแนน 2
เกิดอารมณ์ความรู้สึกดังกล่าวบางครั้ง	ให้คะแนน 3
เกิดอารมณ์ความรู้สึกดังกล่าวบ่อยครั้ง	ให้คะแนน 4
เกิดอารมณ์ความรู้สึกดังกล่าวเป็นประจำ	ให้คะแนน 5
- 3) อารมณ์ความรู้สึกเชิงลบ มีเกณฑ์การตรวจให้คะแนน ดังนี้

ไม่เคยหรือแทบจะไม่เคยเกิดอารมณ์ความรู้สึกดังกล่าว	ให้คะแนน 5
เกิดอารมณ์ความรู้สึกดังกล่าวน้อยครั้ง	ให้คะแนน 4
เกิดอารมณ์ความรู้สึกดังกล่าวบางครั้ง	ให้คะแนน 3
เกิดอารมณ์ความรู้สึกดังกล่าวบ่อยครั้ง	ให้คะแนน 2
เกิดอารมณ์ความรู้สึกดังกล่าวเป็นประจำ	ให้คะแนน 1

การคำนวณคะแนนผลการทดสอบ ดำเนินการดังนี้

1) คะแนนมาตรวัดย่อย ได้จากผลรวมคะแนนรายข้อในแต่ละมาตรวัดย่อย ได้แก่ ความพึงพอใจในชีวิต อารมณ์ความรู้สึกเชิงบวก และอารมณ์ความรู้สึกเชิงลบ ซึ่งจะได้เป็นผลรวมคะแนนดิบในแต่ละมาตรวัดย่อย จากนั้นสามารถนำไปเทียบเป็นคะแนนมาตรฐานจากตารางปกติวิสัย ในรูปแบบตำแหน่งเปอร์เซ็นต์ไทล์ คะแนนซี คะแนนที และคะแนนสเตโนน

2) คะแนนความสุขโดยรวม ได้จากการผลรวมคะแนนจาก 3 มาตรวัดย่อย ได้แก่ มาตรวัดย่อยความพึงพอใจในชีวิต มาตรวัดย่อยอารมณ์ความรู้สึกเชิงบวก และมาตรวัดย่อยอารมณ์ความรู้สึกเชิงลบ ซึ่งสามารถนำคะแนนดิบรวมที่ได้ ไปเทียบเป็นคะแนนมาตรฐานจากตารางปกติวิสัย ในรูปแบบตำแหน่งเปอร์เซ็นต์ไทล์ คะแนนซี คะแนนที และคะแนนสเตโนน

การแปลความหมายคะแนน แบ่งออกเป็นสองส่วน คือ การแปลความหมายคะแนนในแต่ละมาตรวัดย่อย และการแปลความหมายคะแนนความสุขโดยรวม โดยมีการแปลความหมายคะแนนแบ่งออกเป็น 9 ช่วงตามคะแนนสเตโนน ทั้งในส่วนของมาตรวัดย่อย และมาตรวัดความสุขโดยรวม

จากนั้นจึงนำมาตรวัดความสุขของคนไทย รูปแบบกระดาษ-ดินสอที่จัดพิมพ์ขึ้นมาวิเคราะห์ความเที่ยงทั้งฉบับ และจัดทำปกติวิสัยเพื่อใช้สำหรับเปรียบเทียบคะแนนด้วย ดังนี้

1) การวิเคราะห์ค่าความเที่ยงของมาตรวัดความสุขของคนไทยทั้งฉบับ ดำเนินการโดยตรวจสอบความสอดคล้องภายใน (Internal Consistency) ด้วยวิธีครอนบาค แอลฟา (Cronbach's Alpha Reliability) ทั้งในส่วนของมาตรวัดย่อย และมาตรวัดความสุขโดยรวม โดยใช้โปรแกรม LERTAP Version 5.10.4

2) การสร้างปกติวิสัย เป็นการหาปกติวิสัยในรูปของตำแหน่งเปอร์เซ็นต์ไทล์ (Percentile Rank) และคะแนนมาตรฐาน (คะแนนซี คะแนนที และคะแนน สเตโนน) โดยคำนวณตำแหน่งเปอร์เซ็นต์ไทล์ และคะแนนซีจาก

การแจกแจงของผลการทดสอบที่ได้จากกลุ่มตัวอย่างจำนวน 934 คน โดยใช้โปรแกรม SPSS ในการวิเคราะห์ จากนั้นจึงแปลงคะแนนซี (Z-scores) เป็นคะแนนที (T-scores) ด้วยสูตร $T = 10z + 50$ โดย T หมายถึง คะแนนที และ z หมายถึง คะแนนซี (Gregory, 2007, p. 87) สำหรับคะแนนสเตโนไนน์ใช้วิธีการเปรียบเทียบจาก ตำแหน่งเปอร์เซ็นต์ไทล์ ดังตารางที่ ข-1

ตารางที่ ข-1 การเปรียบเทียบคะแนนสเตโนไนน์กับตำแหน่งเปอร์เซ็นต์ไทล์
(Miller, Lovler & McIntire, 2013, p. 140)

คะแนนสเตโนไนน์	ตำแหน่งเปอร์เซ็นต์ไทล์
9	มากกว่า 96.00
8	มากกว่า 89.00 และน้อยกว่าหรือเท่ากับ 96.00
7	มากกว่า 77.00 และน้อยกว่าหรือเท่ากับ 89.00
6	มากกว่า 60.00 และน้อยกว่าหรือเท่ากับ 77.00
5	มากกว่า 40.00 และน้อยกว่าหรือเท่ากับ 60.00
4	มากกว่า 23.00 และน้อยกว่าหรือเท่ากับ 40.00
3	มากกว่า 11.00 และน้อยกว่าหรือเท่ากับ 23.00
2	มากกว่า 4.00 และน้อยกว่าหรือเท่ากับ 11.00
1	น้อยกว่าหรือเท่ากับ 4.00

นอกจากนี้ ได้จัดทำคู่มือการใช้งาน ซึ่งระบุถึงจุดมุ่งหมายของมาตรวัด ลักษณะของมาตรวัด การดำเนินการทดสอบ วิธีการตรวจให้คะแนน วิธีแปลความหมาย และตารางปกติวิสัยเพื่อให้ง่ายต่อการนำไปใช้ด้วย ผลการพัฒนามาตรวัดความสุขของคนไทย รูปแบบกระดาษ-ดินสอ มีดังนี้

1. ผลการวิเคราะห์ค่าสถิติพื้นฐาน

ผลการวิเคราะห์ค่าสถิติพื้นฐานของมาตรวัดความสุขของคนไทย แสดงดังตารางที่ ข-2

ตารางที่ ข-2 ผลการวิเคราะห์ค่าสถิติพื้นฐานของมาตรวัดความสุขของคนไทย (61 ข้อ)

ค่าสถิติ	มาตรวัดย่อย			มาตรวัดความสุข โดยรวม (61 ข้อ)
	ความพึงพอใจ ในชีวิต (13 ข้อ)	อารมณ์ความรู้สึก เชิงบวก (22 ข้อ)	อารมณ์ความรู้สึก เชิงลบ (26 ข้อ)	
MaxPos	65.00	110.00	130.00	305.00
MinPos	13.00	22.00	26.00	61.00
Max	65.00	110.00	130.00	300.00
Min	20.00	40.00	41.00	128.00
M	47.25	78.28	100.38	224.24
SD	7.32	12.73	16.20	26.63
Variance	53.55	162.04	262.39	709.26
Skewness	-0.26	0.07	-0.59	-0.35
Kurtosis	0.16	-0.15	0.20	0.35

จากตารางที่ ข-2 ปรากฏว่า คะแนนจากมาตรวัดความสุขโดยรวม (61 ข้อ) มีคะแนนที่เป็นไปได้ต่ำสุดเท่ากับ 61 คะแนน คะแนนที่เป็นไปได้สูงสุดเท่ากับ 305 คะแนน ในขณะที่คะแนนที่ได้จากกลุ่มตัวอย่าง มีคะแนนต่ำสุด เท่ากับ 128 คะแนน และคะแนนสูงสุดเท่ากับ 300 คะแนน ค่าเฉลี่ยเท่ากับ 224.24 ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 26.63 ความแปรปรวนเท่ากับ 709.26 มีการแจกแจงแบบเบ้ซ้ายเล็กน้อย (-0.35) แสดงว่าส่วนใหญ่ของกลุ่มตัวอย่างได้คะแนนสูงกว่าค่าเฉลี่ย การแจกแจงของคะแนนมีความโด่งสูงกว่าปกติเล็กน้อย (0.35)

เมื่อพิจารณาในแต่ละมาตรวัดย่อย ปรากฏว่า

มาตรวัดย่อยความพึงพอใจในชีวิต (13 ข้อ) มีคะแนนที่เป็นไปได้ต่ำสุดเท่ากับ 13 คะแนน คะแนนที่เป็นไปได้สูงสุดเท่ากับ 65 คะแนน ในขณะที่คะแนนที่ได้จากกลุ่มตัวอย่าง มีคะแนนต่ำสุด เท่ากับ 20 คะแนน และคะแนนสูงสุดเท่ากับ 65 คะแนน ค่าเฉลี่ยเท่ากับ 47.25 ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 7.32 ความแปรปรวนเท่ากับ 53.55 มีการแจกแจงแบบเบ้ซ้ายเล็กน้อย (-0.26) แสดงว่าส่วนใหญ่ของกลุ่มตัวอย่างได้คะแนนสูงกว่าค่าเฉลี่ย การแจกแจงของคะแนนมีความโด่งสูงกว่าปกติเล็กน้อย (0.16)

มาตรวัดย่อยอารมณ์ความรู้สึกเชิงบวก (22 ข้อ) มีคะแนนที่เป็นไปได้ต่ำสุดเท่ากับ 22 คะแนน คะแนนที่เป็นไปได้สูงสุดเท่ากับ 110 คะแนน ในขณะที่คะแนนที่ได้จากกลุ่มตัวอย่าง มีคะแนนต่ำสุด เท่ากับ 40 คะแนน และคะแนนสูงสุดเท่ากับ 110 คะแนน ค่าเฉลี่ยเท่ากับ 78.28 ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 12.73 ความแปรปรวนเท่ากับ 162.04 มีการแจกแจงแบบเบ้ขวาเล็กน้อย (0.07) แสดงว่าส่วนใหญ่ของกลุ่มตัวอย่างได้คะแนนต่ำกว่าค่าเฉลี่ย การแจกแจงของคะแนนมีความโด่งต่ำกว่าปกติเล็กน้อย (-0.15)

มาตรวัดย่อยอารมณ์ความรู้สึกเชิงลบ (26 ข้อ) มีคะแนนที่เป็นไปได้ต่ำสุดเท่ากับ 26 คะแนน คะแนนที่เป็นไปได้สูงสุดเท่ากับ 130 คะแนน ในขณะที่คะแนนที่ได้จากกลุ่มตัวอย่าง มีคะแนนต่ำสุด เท่ากับ 41 คะแนน และคะแนนสูงสุดเท่ากับ 130 คะแนน ค่าเฉลี่ยเท่ากับ 100.38 ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 16.20 ความแปรปรวนเท่ากับ 262.39 มีการแจกแจงแบบเบ้ซ้ายเล็กน้อย (-0.59) แสดงว่าส่วนใหญ่ของกลุ่มตัวอย่างได้คะแนนสูงกว่าค่าเฉลี่ย การแจกแจงของคะแนนมีความโด่งสูงกว่าปกติเล็กน้อย (0.20)

2. ผลการวิเคราะห์ค่าความเที่ยงทั้งฉบับ

ผลการวิเคราะห์ค่าความเที่ยงโดยตรวจสอบความสอดคล้องภายใน ด้วยวิธีครอนบาค แอลฟา

ปรากฏผลดังตารางที่ ข-3

ตารางที่ ข-3 ความเที่ยงของมาตรวัดย่อยและมาตรวัดความสุขโดยรวม

มาตรวัดย่อย	ความเที่ยง
ความพึงพอใจในชีวิต	0.84
อารมณ์ความรู้สึกเชิงบวก	0.92
อารมณ์ความรู้สึกเชิงลบ	0.94
มาตรวัดความสุขโดยรวม	0.94

จากตารางที่ ข-3 ปรากฏว่า ค่าความเที่ยงโดยรวมของมาตรวัดความสุข มีค่าเท่ากับ 0.94 และเมื่อพิจารณาในแต่ละมาตรวัดย่อย ปรากฏว่า มาตรวัดย่อยความพึงพอใจในชีวิต (13 ข้อ) มีค่าความเที่ยงเท่ากับ 0.84 มาตรวัดย่อยอารมณ์ความรู้สึกเชิงบวก (22 ข้อ) มีค่าความเที่ยงเท่ากับ 0.92 และมาตรวัดย่อยอารมณ์ความรู้สึกเชิงลบ (26 ข้อ) มีค่าความเที่ยงเท่ากับ 0.94

3. ผลการสร้างปกติวิสัย

ผู้วิจัยสร้างปกติวิสัยของมาตรวัดความสุขของคนไทย ตามทฤษฎีความอยู่ดีมีสุขเชิงอัตวิสัย (Pavot & Diener, 2013) จากกลุ่มตัวอย่างที่เป็นคนไทย อายุ 18 ปีขึ้นไป จำนวน 934 คน โดยคำนวณหาตำแหน่งเปอร์เซน

ไทล์ และคะแนนมาตรฐาน 3 ประเภท ได้แก่ คะแนนซี (Z-score) คะแนนที (T-score) และคะแนนสเตโนน (Stanine) ปกติวิสัยของมาตรวัดย่อยความพึงพอใจในชีวิต แสดงได้ดังตารางที่ ข-4

ตารางที่ ข-4 เกณฑ์สำหรับเปรียบเทียบคะแนนดิบกับตำแหน่งเปอร์เซ็นต์ไทล์ คะแนนซี คะแนนที และคะแนนสเตโนนของมาตรวัดย่อยความพึงพอใจในชีวิต (n=934)

คะแนนดิบ	PR	Z-score	T-score	Stanine
65	100.00	2.43	74	9
64	99.46	2.29	73	9
63	99.04	2.15	72	9
62	98.29	2.02	70	9
61	97.97	1.88	69	9
60	97.22	1.74	67	9
59	96.04	1.61	66	9
58	94.75	1.47	65	8
57	92.83	1.33	63	8
56	90.47	1.20	62	8
55	87.37	1.06	61	7
54	83.94	0.92	59	7
53	80.41	0.79	58	7
52	75.37	0.65	56	6
51	70.45	0.51	55	6
50	66.17	0.38	54	6
49	60.81	0.24	52	6
48	55.25	0.10	51	5
47	49.25	-0.03	50	5
46	44.75	-0.17	48	5
45	39.19	-0.31	47	4
44	33.83	-0.44	46	4
43	30.84	-0.58	44	4
42	25.37	-0.72	43	4
41	21.63	-0.85	41	3
40	18.42	-0.99	40	3
39	14.35	-1.13	39	3
38	11.35	-1.26	37	3
37	8.89	-1.40	36	2
36	6.96	-1.54	35	2
35	5.46	-1.67	33	2
34	4.28	-1.81	32	2
33	2.89	-1.95	31	1
32	2.57	-2.08	29	1

ตารางที่ ข-4 (ต่อ)

คะแนนดิบ	PR	Z-score	T-score	Stanine
31	2.25	-2.22	28	1
30	1.71	-2.36	26	1
29	1.39	-2.49	25	1
28	0.96	-2.63	24	1
27	0.75	-2.77	22	1
26	0.64	-2.90	21	1
25	0.64	-3.04	20	1
24	0.32	-3.18	18	1
23	0.21	-3.31	17	1
22	0.21	-3.45	15	1
21	0.21	-3.59	14	1
13-20	0.11	-3.72	13	1

จากตารางที่ ข-4 ปกติวิสัยของมาตรวัดย่อยความพึงพอใจในชีวิต ผู้ทดสอบที่มีความพึงพอใจในชีวิตอย่างยิ่ง ได้คะแนนดิบตั้งแต่ 59 – 65 คะแนน มีตำแหน่งเปอร์เซ็นต์ไทล์ตั้งแต่ 96.04 – 100.00 ผู้ทดสอบที่มีความพึงพอใจในชีวิตมาก ได้คะแนนดิบตั้งแต่ 56 – 58 คะแนน มีตำแหน่งเปอร์เซ็นต์ไทล์ตั้งแต่ 90.47 – 94.75 ผู้ทดสอบที่มีความพึงพอใจในชีวิต ได้คะแนนดิบตั้งแต่ 53 – 55 คะแนน มีตำแหน่งเปอร์เซ็นต์ไทล์ตั้งแต่ 80.41 – 87.37 ผู้ทดสอบที่ค่อนข้างพึงพอใจในชีวิต ได้คะแนนดิบตั้งแต่ 49 – 52 คะแนน มีตำแหน่งเปอร์เซ็นต์ไทล์ตั้งแต่ 60.81 – 75.37 ผู้ทดสอบที่มีความพึงพอใจในชีวิตอยู่ในระดับปานกลาง ได้คะแนนดิบตั้งแต่ 46 – 48 คะแนน มีตำแหน่งเปอร์เซ็นต์ไทล์ตั้งแต่ 44.75 – 55.25 ผู้ทดสอบที่มีความพึงพอใจในชีวิตน้อย ได้คะแนนดิบตั้งแต่ 42 – 45 คะแนน มีตำแหน่งเปอร์เซ็นต์ไทล์ตั้งแต่ 25.37 – 39.19 ผู้ทดสอบที่ไม่ค่อยพึงพอใจในชีวิต ได้คะแนนดิบตั้งแต่ 38 – 41 คะแนน มีตำแหน่งเปอร์เซ็นต์ไทล์ตั้งแต่ 11.35 – 21.63 ผู้ทดสอบที่ไม่พึงพอใจในชีวิต ได้คะแนนดิบตั้งแต่ 34 – 37 คะแนน มีตำแหน่งเปอร์เซ็นต์ไทล์ตั้งแต่ 4.28 – 8.89 ผู้ทดสอบที่ไม่พึงพอใจในชีวิตอย่างยิ่ง ได้คะแนนดิบตั้งแต่ 13 – 33 คะแนน มีตำแหน่งเปอร์เซ็นต์ไทล์ตั้งแต่ 0.11 - 2.89

เกณฑ์การแปลความหมายช่วงคะแนนดิบ ที่ได้จากมาตรวัดย่อยความพึงพอใจในชีวิต แบ่งออกเป็น 9 ระดับตามคะแนนสแตนิเน่ ดังตารางที่ ข-5

ตารางที่ ข-5 เกณฑ์การเทียบช่วงคะแนนดิบ กับระดับความพึงพอใจในชีวิต

ช่วงคะแนนดิบ	ระดับความพึงพอใจในชีวิต
59 – 65	พึงพอใจอย่างยิ่ง
56 – 58	พึงพอใจมาก
53 – 55	พึงพอใจ
49 – 52	ค่อนข้างพึงพอใจ
46 – 48	ปานกลาง
42 – 45	พึงพอใจน้อย
38 – 41	ไม่ค่อยพึงพอใจ
34 – 37	ไม่พึงพอใจ

ตารางที่ ข-5 (ต่อ)

ช่วงคะแนนดิบ	ระดับความพึงพอใจในชีวิต
13 – 33	ไม่พึงพอใจอย่างยิ่ง

จากตารางที่ ข-5 แสดงให้เห็นว่าผู้ทดสอบที่ได้คะแนนดิบ 59 – 65 มีความพึงพอใจในชีวิตอย่างยิ่ง ผู้ทดสอบที่ได้คะแนนดิบ 56 – 58 มีความพึงพอใจในชีวิตมาก ผู้ทดสอบที่ได้คะแนนดิบ 53 – 55 มีความพึงพอใจในชีวิต ผู้ทดสอบที่ได้คะแนนดิบ 49 – 52 ค่อนข้างพึงพอใจในชีวิต ผู้ทดสอบที่ได้คะแนนดิบ 46 – 48 มีความพึงพอใจในชีวิตอยู่ในระดับปานกลาง ผู้ทดสอบที่ได้คะแนนดิบ 42 – 45 มีความพึงพอใจในชีวิตน้อย ผู้ทดสอบที่ได้คะแนนดิบ 38 – 41 ไม่ค่อยพึงพอใจในชีวิต ผู้ทดสอบที่ได้คะแนนดิบ 34 – 37 ไม่พึงพอใจในชีวิต ผู้ทดสอบที่ได้คะแนนดิบ 13 – 33 ไม่พึงพอใจในชีวิตอย่างยิ่ง

ปกติวิสัยของมาตรวัดย่อยอารมณ์ความรู้สึกเชิงบวก แสดงได้ดังตารางที่ ข-6

ตารางที่ ข-6 เกณฑ์สำหรับเปรียบเทียบคะแนนดิบกับตำแหน่งเปอร์เซ็นต์ไทล์ คะแนนซี คะแนนที และคะแนนสเตโนนซ์ของมาตรวัดย่อยอารมณ์ความรู้สึกเชิงบวก (n=934)

คะแนนดิบ	PR	Z-score	T-score	Stanine
110	100.00	2.49	75	9
109	99.25	2.41	74	9
108	99.04	2.33	73	9
107	98.61	2.26	73	9
106	98.29	2.18	72	9
105	97.86	2.10	71	9
104	97.32	2.02	70	9
103	96.79	1.94	69	9
102	96.47	1.86	69	9
101	95.40	1.79	68	8
100	94.97	1.71	67	8
99	93.90	1.63	66	8
98	93.25	1.55	65	8
97	92.51	1.47	65	8
96	92.08	1.39	64	8
95	90.79	1.31	63	8
94	90.15	1.24	62	8
93	88.76	1.16	62	7
92	87.58	1.08	61	7
91	85.22	1.00	60	7
90	82.55	0.92	59	7
89	81.26	0.84	58	7
88	78.91	0.76	58	7
87	77.41	0.69	57	7

ตารางที่ ข-6 (ต่อ)

คะแนนดิบ	PR	Z-score	T-score	Stanine
86	75.59	0.61	56	6
85	72.59	0.53	55	6
84	69.27	0.45	54	6
83	65.95	0.37	54	6
82	62.74	0.29	53	6
81	59.42	0.21	52	5
80	56.32	0.14	51	5
79	53.85	0.06	51	5
78	50.54	-0.02	50	5
77	48.29	-0.10	49	5
76	45.72	-0.18	48	5
75	42.61	-0.26	47	5
74	40.04	-0.34	47	5
73	36.62	-0.41	46	4
72	33.19	-0.49	45	4
71	30.30	-0.57	44	4
70	27.84	-0.65	43	4
69	25.27	-0.73	43	4
68	22.70	-0.81	42	3
67	20.13	-0.89	41	3
66	17.99	-0.96	40	3
65	15.10	-1.04	40	3
64	13.06	-1.12	39	3
63	11.35	-1.20	38	3
62	10.49	-1.28	37	2
61	9.21	-1.36	36	2
60	7.82	-1.44	36	2
59	7.28	-1.51	35	2
58	6.42	-1.59	34	2
57	5.14	-1.67	33	2
56	4.18	-1.75	32	2
55	3.10	-1.83	32	1
54	2.78	-1.91	31	1
53	2.36	-1.99	30	1
52	1.82	-2.06	29	1
51	1.39	-2.14	29	1
50	1.28	-2.22	28	1
49	0.86	-2.30	27	1

ตารางที่ ข-6 (ต่อ)

คะแนนดิบ	PR	Z-score	T-score	Stanine
48	0.64	-2.38	26	1
47	0.54	-2.46	25	1
46	0.54	-2.54	25	1
45	0.32	-2.61	24	1
44	0.32	-2.69	23	1
43	0.32	-2.77	22	1
42	0.32	-2.85	22	1
41	0.21	-2.93	21	1
22-40	0.11	-3.01	20	1

จากตารางที่ ข-6 ปกติวิสัยของมาตรวัดย่อยอารมณ์ความรู้สึกเชิงบวก ผู้ทดสอบที่มีอารมณ์ความรู้สึกเชิงบวก อยู่ในระดับสูงที่สุด ได้คะแนนดิบตั้งแต่ 102 - 110 คะแนน มีตำแหน่งเปอร์เซ็นต์ไทล์ตั้งแต่ 96.47 - 100.00 ผู้ทดสอบที่มีอารมณ์ความรู้สึกเชิงบวกอยู่ในระดับสูงมาก ได้คะแนนดิบตั้งแต่ 94-101 คะแนน มีตำแหน่งเปอร์เซ็นต์ไทล์ตั้งแต่ 90.15 - 95.40 ผู้ทดสอบที่มีอารมณ์ความรู้สึกเชิงบวกอยู่ในระดับสูง ได้คะแนนดิบตั้งแต่ 87 - 93 คะแนน มีตำแหน่งเปอร์เซ็นต์ไทล์ตั้งแต่ 77.41 - 88.76 ผู้ทดสอบที่มีอารมณ์ความรู้สึกเชิงบวกอยู่ในระดับสูงกว่าปกติเล็กน้อย ได้คะแนนดิบตั้งแต่ 82 - 86 คะแนน มีตำแหน่งเปอร์เซ็นต์ไทล์ตั้งแต่ 62.74 - 75.59 ผู้ทดสอบที่มีอารมณ์ความรู้สึกเชิงบวกอยู่ในระดับปกติ ได้คะแนนดิบตั้งแต่ 74 - 81 คะแนน มีตำแหน่งเปอร์เซ็นต์ไทล์ตั้งแต่ 40.04 - 59.42 ผู้ทดสอบที่มีอารมณ์ความรู้สึกเชิงบวกอยู่ในระดับต่ำกว่าปกติเล็กน้อย ได้คะแนนดิบตั้งแต่ 69 - 73 คะแนน มีตำแหน่งเปอร์เซ็นต์ไทล์ตั้งแต่ 25.27 - 36.62 ผู้ทดสอบที่มีอารมณ์ความรู้สึกเชิงบวกอยู่ในระดับต่ำ ได้คะแนนดิบตั้งแต่ 63 - 68 คะแนน มีตำแหน่งเปอร์เซ็นต์ไทล์ตั้งแต่ 11.35 - 22.70 ผู้ทดสอบที่มีอารมณ์ความรู้สึกเชิงบวกอยู่ในระดับต่ำมาก ได้คะแนนดิบตั้งแต่ 56 - 62 คะแนน มีตำแหน่งเปอร์เซ็นต์ไทล์ตั้งแต่ 4.18 - 10.49 ผู้ทดสอบที่มีอารมณ์ความรู้สึกเชิงบวกอยู่ในระดับต่ำที่สุด ได้คะแนนดิบตั้งแต่ 22 - 55 คะแนน มีตำแหน่งเปอร์เซ็นต์ไทล์ตั้งแต่ 0.11 - 3.10

เกณฑ์การแปลความหมายช่วงคะแนนดิบ ที่ได้จากมาตรวัดย่อยอารมณ์ความรู้สึกเชิงบวก แบ่งออกเป็น 9 ระดับตามคะแนนสเตโนई ดังตารางที่ ข-7

ตารางที่ ข-7 เกณฑ์การเทียบช่วงคะแนนดิบ กับระดับอารมณ์ความรู้สึกเชิงบวก

ช่วงคะแนนดิบ	ระดับอารมณ์ความรู้สึกเชิงบวก
102 - 110	สูงที่สุด
94 - 101	สูงมาก
87 - 93	สูง
82 - 86	สูงกว่าปกติเล็กน้อย
74 - 81	ปกติ
69 - 73	ต่ำกว่าปกติเล็กน้อย
63 - 68	ต่ำ
56 - 62	ต่ำมาก
22 - 55	ต่ำที่สุด

จากตารางที่ ข-7 แสดงให้เห็นว่าผู้ทดสอบที่ได้คะแนนดิบ 102 – 110 มีอารมณ์ความรู้สึกเชิงบวกอยู่ในระดับสูงสุด ผู้ทดสอบที่ได้คะแนนดิบ 94 – 101 มีอารมณ์ความรู้สึกเชิงบวกอยู่ในระดับสูงมาก ผู้ทดสอบที่ได้คะแนนดิบ 87 – 93 มีอารมณ์ความรู้สึกเชิงบวกอยู่ในระดับสูง ผู้ทดสอบที่ได้คะแนนดิบ 82 – 86 มีอารมณ์ความรู้สึกเชิงบวกอยู่ในระดับสูงกว่าปกติเล็กน้อย ผู้ทดสอบที่ได้คะแนนดิบ 74 – 81 มีอารมณ์ความรู้สึกเชิงบวกอยู่ในระดับปกติ ผู้ทดสอบที่ได้คะแนนดิบ 69 – 73 มีอารมณ์ความรู้สึกเชิงบวกอยู่ในระดับต่ำกว่าปกติเล็กน้อย ผู้ทดสอบที่ได้คะแนนดิบ 63 – 68 มีอารมณ์ความรู้สึกเชิงบวกอยู่ในระดับต่ำ ผู้ทดสอบที่ได้คะแนนดิบ 56 – 62 มีอารมณ์ความรู้สึกเชิงบวกอยู่ในระดับต่ำมาก และผู้ทดสอบที่ได้คะแนนดิบ 22 – 55 มีอารมณ์ความรู้สึกเชิงบวกอยู่ในระดับต่ำที่สุด

ปกติวิสัยของมาตรฐานวัดย่อยอารมณ์ความรู้สึกเชิงลบ แสดงได้ดังตารางที่ ข-8

ตารางที่ ข-8 เกณฑ์สำหรับเปรียบเทียบคะแนนดิบกับตำแหน่งเปอร์เซ็นต์ไทล์ คะแนนซี คะแนนที และคะแนนสเตโนนซ์ของมาตรฐานวัดย่อยอารมณ์ความรู้สึกเชิงลบ (n=934)

คะแนนดิบ	PR	Z-score	T-score	Stanine
130	100.00	1.83	68	9
129	99.68	1.77	68	9
128	98.93	1.71	67	9
127	98.50	1.64	66	9
126	97.75	1.58	66	9
125	96.57	1.52	65	9
124	95.50	1.46	65	8
123	94.22	1.40	64	8
122	93.36	1.33	63	8
121	91.76	1.27	63	8
120	89.94	1.21	62	8
119	88.87	1.15	61	7
118	87.58	1.09	61	7
117	84.69	1.03	60	7
116	83.40	0.96	60	7
115	81.69	0.90	59	7
114	78.59	0.84	58	7
113	76.34	0.78	58	6
112	74.20	0.72	57	6
111	71.84	0.66	57	6
110	70.99	0.59	56	6
109	68.95	0.53	55	6
108	66.92	0.47	55	6
107	64.45	0.41	54	6
106	62.74	0.35	53	6
105	59.96	0.29	53	5
104	56.42	0.22	52	5

ตารางที่ ข-8 (ต่อ)

คะแนนดิบ	PR	Z-score	T-score	Stanine
103	54.60	0.16	52	5
102	50.86	0.10	51	5
101	48.39	0.04	50	5
100	46.36	-0.02	50	5
99	44.00	-0.08	49	5
98	41.01	-0.15	49	5
97	39.51	-0.21	48	4
96	37.26	-0.27	47	4
95	34.48	-0.33	47	4
94	32.23	-0.39	46	4
93	30.41	-0.46	45	4
92	28.16	-0.52	45	4
91	26.02	-0.58	44	4
90	24.63	-0.64	44	4
89	23.13	-0.70	43	4
88	21.84	-0.76	42	3
87	20.66	-0.83	42	3
86	19.38	-0.89	41	3
85	17.77	-0.95	41	3
84	16.60	-1.01	40	3
83	15.52	-1.07	39	3
82	14.24	-1.13	39	3
81	13.60	-1.20	38	3
80	12.74	-1.26	37	3
79	11.46	-1.32	37	3
78	10.28	-1.38	36	2
77	8.89	-1.44	36	2
76	8.24	-1.50	35	2
75	7.82	-1.57	34	2
74	7.07	-1.63	34	2
73	6.21	-1.69	33	2
72	5.46	-1.75	32	2
71	5.25	-1.81	32	2
70	4.39	-1.88	31	2
69	3.96	-1.94	31	1
68	3.64	-2.00	30	1
67	3.32	-2.06	29	1
66	2.89	-2.12	29	1

ตารางที่ ข-8 (ต่อ)

คะแนนดิบ	PR	Z-score	T-score	Stanine
65	2.78	-2.18	28	1
64	2.57	-2.25	28	1
63	2.36	-2.31	27	1
62	2.03	-2.37	26	1
61	2.03	-2.43	26	1
60	1.71	-2.49	25	1
59	1.50	-2.55	24	1
58	1.39	-2.62	24	1
57	1.39	-2.68	23	1
56	1.28	-2.74	23	1
55	1.07	-2.80	22	1
54	1.07	-2.86	21	1
53	1.07	-2.92	21	1
52	0.86	-2.99	20	1
51	0.75	-3.05	20	1
50	0.64	-3.11	19	1
49	0.54	-3.17	18	1
48	0.32	-3.23	18	1
47	0.21	-3.30	17	1
46	0.21	-3.36	16	1
45	0.11	-3.42	16	1
44	0.11	-3.48	15	1
43	0.11	-3.54	15	1
42	0.11	-3.60	14	1
26 - 41	0.11	-3.67	13	1

จากตารางที่ ข-8 ปกติวิสัยของมาตรวัดย่อยอารมณ์ความรู้สึกเชิงลบ ผู้ทดสอบที่มีอารมณ์ความรู้สึกเชิงลบ อยู่ในระดับต่ำที่สุด ได้คะแนนดิบตั้งแต่ 125 - 130 คะแนน มีตำแหน่งเปอร์เซ็นต์ตั้งแต่ 96.57 - 100.00 ผู้ทดสอบที่มีอารมณ์ความรู้สึกเชิงลบอยู่ในระดับต่ำมาก ได้คะแนนดิบตั้งแต่ 120 - 124 คะแนน มีตำแหน่งเปอร์เซ็นต์ตั้งแต่ 89.94 - 95.50 ผู้ทดสอบที่มีอารมณ์ความรู้สึกเชิงลบอยู่ในระดับต่ำ ได้คะแนนดิบตั้งแต่ 114 - 119 คะแนน มีตำแหน่งเปอร์เซ็นต์ตั้งแต่ 78.59 - 88.87 ผู้ทดสอบที่มีอารมณ์ความรู้สึกเชิงลบอยู่ในระดับต่ำกว่าปกติเล็กน้อย ได้คะแนนดิบตั้งแต่ 106 - 113 คะแนน มีตำแหน่งเปอร์เซ็นต์ตั้งแต่ 62.74 - 76.34 ผู้ทดสอบที่มีอารมณ์ความรู้สึกเชิงลบอยู่ในระดับปกติ ได้คะแนนดิบตั้งแต่ 98 - 105 คะแนน มีตำแหน่งเปอร์เซ็นต์ตั้งแต่ 41.01 - 59.96 ผู้ทดสอบที่มีอารมณ์ความรู้สึกเชิงลบอยู่ในระดับสูงกว่าปกติเล็กน้อย ได้คะแนนดิบตั้งแต่ 89 - 97 คะแนน มีตำแหน่งเปอร์เซ็นต์ตั้งแต่ 23.13 - 39.51 ผู้ทดสอบที่มีอารมณ์ความรู้สึกเชิงลบอยู่ในระดับสูง ได้คะแนนดิบตั้งแต่ 79 - 88 คะแนน มีตำแหน่งเปอร์เซ็นต์ตั้งแต่ 11.46 - 21.84 ผู้ทดสอบที่มีอารมณ์ความรู้สึกเชิงลบอยู่ในระดับสูงมาก ได้คะแนนดิบตั้งแต่ 70 - 78 คะแนน มีตำแหน่งเปอร์เซ็นต์ตั้งแต่ 4.39 - 10.28 ผู้ทดสอบที่มีอารมณ์

ความรู้สึกเชิงลบอยู่ในระดับสูงที่สุด ได้คะแนนดิบตั้งแต่ 26 – 69 คะแนน มีตำแหน่งเปอร์เซ็นต์ตั้งแต่ 0.11 - 3.96

เกณฑ์การแปลความหมายช่วงคะแนนดิบ ที่ได้จากมาตรวัดย่อยอารมณ์ความรู้สึกเชิงลบ แบ่งออกเป็น 9 ระดับตามคะแนนสแตนดีน ดังตารางที่ ข-9

ตารางที่ ข-9 เกณฑ์การเทียบช่วงคะแนนดิบ กับระดับอารมณ์ความรู้สึกเชิงลบ

ช่วงคะแนนดิบ	ระดับอารมณ์ความรู้สึกเชิงลบ
125 – 130	ต่ำที่สุด
120 – 124	ต่ำมาก
114 – 119	ต่ำ
106 – 113	ต่ำกว่าปกติเล็กน้อย
98 – 105	ปกติ
89 – 97	สูงกว่าปกติเล็กน้อย
79 – 88	สูง
70 – 78	สูงมาก
26 – 69	สูงที่สุด

จากตารางที่ ข-9 แสดงให้เห็นว่า ผู้ทดสอบที่ได้คะแนนดิบ 125 – 130 มีอารมณ์ความรู้สึกเชิงลบ อยู่ในระดับต่ำที่สุด ผู้ทดสอบที่ได้คะแนนดิบ 120 -124 มีอารมณ์ความรู้สึกเชิงลบอยู่ในระดับต่ำมาก ผู้ทดสอบที่ได้คะแนนดิบ 114 – 119 มีอารมณ์ความรู้สึกเชิงลบอยู่ในระดับต่ำ ผู้ทดสอบที่ได้คะแนนดิบ 106 – 113 มีอารมณ์ความรู้สึกเชิงลบอยู่ในระดับต่ำกว่าปกติเล็กน้อย ผู้ทดสอบที่ได้คะแนนดิบ 98 – 105 มีอารมณ์ความรู้สึกเชิงลบอยู่ในระดับปกติ ผู้ทดสอบที่ได้คะแนนดิบ 89 – 97 มีอารมณ์ความรู้สึกเชิงลบอยู่ในระดับสูงกว่าปกติเล็กน้อย ผู้ทดสอบที่ได้คะแนนดิบ 79 – 88 มีอารมณ์ความรู้สึกเชิงลบอยู่ในระดับสูง ผู้ทดสอบที่ได้คะแนนดิบ 70 – 78 มีอารมณ์ความรู้สึกเชิงลบอยู่ในระดับสูงมาก และผู้ทดสอบที่ได้คะแนนดิบ 26 – 69 มีอารมณ์ความรู้สึกเชิงลบอยู่ในระดับสูงที่สุด

ปกติวิสัยของมาตรวัดความสุขโดยรวม แสดงได้ดังตารางที่ ข-10

ตารางที่ ข-10 เกณฑ์สำหรับเปรียบเทียบคะแนนดิบกับตำแหน่งเปอร์เซ็นต์ คะแนนซี คะแนนที และคะแนนสแตนดีนของมาตรวัดความสุขโดยรวม (n=934)

คะแนนดิบ	PR	Z-score	T-score	Stanine
300 -305	100.00	2.84	78	9
299	99.89	2.81	78	9
298	99.89	2.77	78	9
297	99.89	2.73	77	9
296	99.89	2.69	77	9
295	99.89	2.66	77	9
294	99.89	2.62	76	9
293	99.79	2.58	76	9
292	99.79	2.54	75	9

ตารางที่ ข-10 (ต่อ)

คะแนนดิบ	PR	Z-score	T-score	Stanine
291	99.79	2.51	75	9
290	99.79	2.47	75	9
289	99.79	2.43	74	9
288	99.79	2.39	74	9
287	99.57	2.36	74	9
286	99.36	2.32	73	9
285	99.14	2.28	73	9
284	98.82	2.24	72	9
283	98.61	2.21	72	9
282	98.50	2.17	72	9
281	98.29	2.13	71	9
280	98.07	2.09	71	9
279	97.97	2.06	71	9
278	97.86	2.02	70	9
277	97.54	1.98	70	9
276	97.54	1.94	69	9
275	97.43	1.91	69	9
274	97.22	1.87	69	9
273	97.22	1.83	68	9
272	96.57	1.79	68	9
271	96.47	1.76	68	9
270	96.47	1.72	67	9
269	96.36	1.68	67	9
268	96.36	1.64	66	9
267	95.93	1.61	66	8
266	95.82	1.57	66	8
265	95.82	1.53	65	8
264	95.40	1.49	65	8
263	95.18	1.46	65	8
262	94.86	1.42	64	8
261	94.00	1.38	64	8
260	93.47	1.34	63	8
259	93.25	1.31	63	8
258	92.51	1.27	63	8
257	91.97	1.23	62	8
256	90.79	1.19	62	8
255	89.94	1.16	62	8
254	89.61	1.12	61	8

ตารางที่ ข-10 (ต่อ)

คะแนนดิบ	PR	Z-score	T-score	Stanine
253	88.87	1.08	61	7
252	88.12	1.04	60	7
251	86.83	1.00	60	7
250	85.44	0.97	60	7
249	84.90	0.93	59	7
248	83.83	0.89	59	7
247	82.12	0.85	59	7
246	81.26	0.82	58	7
245	79.87	0.78	58	7
244	77.94	0.74	57	7
243	76.34	0.70	57	6
242	74.95	0.67	57	6
241	73.45	0.63	56	6
240	71.73	0.59	56	6
239	70.45	0.55	56	6
238	68.95	0.52	55	6
237	67.77	0.48	55	6
236	66.38	0.44	54	6
235	64.99	0.40	54	6
234	64.03	0.37	54	6
233	62.63	0.33	53	6
232	60.81	0.29	53	6
231	58.89	0.25	53	5
230	57.07	0.22	52	5
229	54.71	0.18	52	5
228	53.10	0.14	51	5
227	51.61	0.10	51	5
226	50.00	0.07	51	5
225	48.93	0.03	50	5
224	47.54	-0.01	50	5
223	45.29	-0.05	50	5
222	44.22	-0.08	49	5
221	42.61	-0.12	49	5
220	40.79	-0.16	48	5
219	39.61	-0.20	48	4
218	38.12	-0.23	48	4
217	37.04	-0.27	47	4
216	35.44	-0.31	47	4

ตารางที่ ข-10 (ต่อ)

คะแนนดิบ	PR	Z-score	T-score	Stanine
215	33.40	-0.35	47	4
214	32.12	-0.38	46	4
213	31.05	-0.42	46	4
212	29.66	-0.46	45	4
211	28.69	-0.50	45	4
210	27.62	-0.53	45	4
209	26.55	-0.57	44	4
208	25.05	-0.61	44	4
207	24.20	-0.65	44	4
206	23.23	-0.68	43	4
205	21.84	-0.72	43	3
204	20.88	-0.76	42	3
203	20.02	-0.80	42	3
202	19.06	-0.84	42	3
201	18.31	-0.87	41	3
200	17.34	-0.91	41	3
199	16.38	-0.95	41	3
198	15.74	-0.99	40	3
197	15.20	-1.02	40	3
196	14.56	-1.06	39	3
195	13.49	-1.10	39	3
194	13.17	-1.14	39	3
193	12.74	-1.17	38	3
192	11.88	-1.21	38	3
191	11.24	-1.25	38	3
190	11.13	-1.29	37	3
189	10.28	-1.32	37	2
188	9.85	-1.36	36	2
187	9.64	-1.40	36	2
186	9.31	-1.44	36	2
185	9.10	-1.47	35	2
184	8.67	-1.51	35	2
183	8.24	-1.55	35	2
182	7.71	-1.59	34	2
181	7.49	-1.62	34	2
180	7.07	-1.66	33	2
179	6.64	-1.70	33	2
178	5.89	-1.74	33	2

ตารางที่ ข-10 (ต่อ)

คะแนนดิบ	PR	Z-score	T-score	Stanine
177	5.35	-1.77	32	2
176	5.03	-1.81	32	2
175	4.82	-1.85	32	2
174	4.50	-1.89	31	2
173	4.07	-1.92	31	2
172	3.75	-1.96	30	1
171	3.10	-2.00	30	1
170	2.89	-2.04	30	1
169	2.57	-2.07	29	1
168	2.57	-2.11	29	1
167	2.46	-2.15	29	1
166	2.25	-2.19	28	1
165	2.03	-2.22	28	1
164	1.82	-2.26	27	1
163	1.71	-2.30	27	1
162	1.71	-2.34	27	1
161	1.61	-2.37	26	1
160	1.61	-2.41	26	1
159	1.61	-2.45	26	1
158	1.39	-2.49	25	1
157	1.39	-2.52	25	1
156	1.39	-2.56	24	1
155	1.39	-2.60	24	1
154	1.28	-2.64	24	1
153	1.28	-2.67	23	1
152	1.18	-2.71	23	1
151	1.18	-2.75	22	1
150	0.96	-2.79	22	1
149	0.75	-2.83	22	1
148	0.64	-2.86	21	1
147	0.43	-2.90	21	1
146	0.43	-2.94	21	1
145	0.43	-2.98	20	1
144	0.32	-3.01	20	1
143	0.32	-3.05	19	1
142	0.32	-3.09	19	1
141	0.32	-3.13	19	1
140	0.32	-3.16	18	1

ตารางที่ ข-10 (ต่อ)

คะแนนดิบ	PR	Z-score	T-score	Stanine
139	0.32	-3.20	18	1
138	0.21	-3.24	18	1
137	0.21	-3.28	17	1
136	0.21	-3.31	17	1
135	0.11	-3.35	16	1
134	0.11	-3.39	16	1
133	0.11	-3.43	16	1
132	0.11	-3.46	15	1
131	0.11	-3.50	15	1
130	0.11	-3.54	15	1
129	0.11	-3.58	14	1
61 - 128	0.11	-3.61	14	1

จากตารางที่ ข-10 ปกติวิสัยของมาตรวัดความสุขโดยรวม ผู้ทดสอบที่มีความสุขอย่างยิ่ง ได้คะแนนดิบ ตั้งแต่ 268 – 305 คะแนน มีตำแหน่งเปอร์เซ็นต์ไทล์ตั้งแต่ 96.36 - 100.00 ผู้ทดสอบที่มีความสุขมาก ได้คะแนนดิบ ตั้งแต่ 254 – 267 คะแนน มีตำแหน่งเปอร์เซ็นต์ไทล์ตั้งแต่ 89.61 - 95.93 ผู้ทดสอบที่มีความสุข ได้คะแนนดิบตั้งแต่ 244 – 253 คะแนน มีตำแหน่งเปอร์เซ็นต์ไทล์ตั้งแต่ 77.94 - 88.87 ผู้ทดสอบที่มีความสุขสูงกว่าปกติเล็กน้อย ได้คะแนนดิบตั้งแต่ 232 – 243 คะแนน มีตำแหน่งเปอร์เซ็นต์ไทล์ตั้งแต่ 60.81 - 76.34 ผู้ทดสอบที่มีความสุขตามปกติ ได้คะแนนดิบตั้งแต่ 220 – 231 คะแนน มีตำแหน่งเปอร์เซ็นต์ไทล์ตั้งแต่ 40.79 - 58.89 ผู้ทดสอบที่มีความสุขต่ำกว่าปกติเล็กน้อย ได้คะแนนดิบตั้งแต่ 206 – 219 คะแนน มีตำแหน่งเปอร์เซ็นต์ไทล์ตั้งแต่ 23.23 - 39.61 ผู้ทดสอบที่มีความสุขน้อย ได้คะแนนดิบตั้งแต่ 190 – 205 คะแนน มีตำแหน่งเปอร์เซ็นต์ไทล์ตั้งแต่ 11.13 - 21.84 ผู้ทดสอบที่มีความสุขน้อยมาก ได้คะแนนดิบตั้งแต่ 173 – 189 คะแนน มีตำแหน่งเปอร์เซ็นต์ไทล์ตั้งแต่ 4.07 - 10.28 ผู้ทดสอบที่ไม่มีมีความสุข ได้คะแนนดิบตั้งแต่ 61 – 172 คะแนน มีตำแหน่งเปอร์เซ็นต์ไทล์ตั้งแต่ 0.11 - 3.75

เกณฑ์การแปลความหมายช่วงคะแนนดิบ ที่ได้จากมาตรวัดความสุขโดยรวม แบ่งออกเป็น 9 ระดับตามคะแนนสแตโนร์ ดังตารางที่ ข-11

ตารางที่ ข-11 เกณฑ์การเทียบช่วงคะแนนดิบ กับระดับความสุขโดยรวม

ช่วงคะแนนดิบ	ระดับความสุข
268 – 305	มีความสุขอย่างยิ่ง
254 – 267	มีความสุขมาก
244 – 253	มีความสุข
232 – 243	มีความสุขสูงกว่าปกติเล็กน้อย
220 – 231	มีความสุขตามปกติ
206 – 219	มีความสุขต่ำกว่าปกติเล็กน้อย
190 – 205	มีความสุขน้อย
173 – 189	มีความสุขน้อยมาก
61 – 172	ไม่มีความสุข

จากตารางที่ ข-11 แสดงให้เห็นว่า ผู้ทดสอบที่ได้คะแนนดิบ 268 – 305 มีความสุขอย่างยิ่ง ผู้ทดสอบที่ได้คะแนนดิบ 254 – 267 มีความสุขมาก ผู้ทดสอบที่ได้คะแนนดิบ 244 – 253 มีความสุข ผู้ทดสอบที่ได้คะแนนดิบ 232 – 243 มีความสุขสูงกว่าปกติเล็กน้อย ผู้ทดสอบที่ได้คะแนนดิบ 220 – 231 มีความสุขตามปกติ ผู้ทดสอบที่ได้คะแนนดิบ 206 – 219 มีความสุขตามปกติ ผู้ทดสอบที่ได้คะแนนดิบ 190 – 205 มีความสุขต่ำกว่าปกติเล็กน้อย ผู้ทดสอบที่ได้คะแนนดิบ 193 – 208 มีความสุขน้อย ผู้ทดสอบที่ได้คะแนนดิบ 173 – 189 มีความสุขน้อยมาก และผู้ทดสอบที่ได้คะแนนดิบ 61 – 172 ไม่มีความสุข

มาตรวัดความสุขของคนไทย

ชื่อ-สกุล (นาย/นาง/นางสาว)..... อายุ..... ปี เพศ ชาย หญิง

มาตรวัดย่อยความพึงพอใจในชีวิต

คำชี้แจง กรุณาอ่านข้อคำถามต่อไปนี้ แล้วทำเครื่องหมายถูก (✓) ในช่องที่ตรงกับความคิดเห็นของท่านมากที่สุด โดย

- | | | |
|---|---------|--------------------------------|
| 1 | หมายถึง | ไม่เห็นด้วย/เห็นด้วยน้อยที่สุด |
| 2 | หมายถึง | เห็นด้วยน้อย |
| 3 | หมายถึง | เห็นด้วยปานกลาง |
| 4 | หมายถึง | เห็นด้วยมาก |
| 5 | หมายถึง | เห็นด้วยอย่างยิ่ง |

ส่วนนี้สำหรับเจ้าหน้าที่

Raw score	
Z-score	
T-score	
Stanine	
PR	
Level	

คำถาม	ระดับความคิดเห็น				
	1	2	3	4	5
1. ถ้าฉันย้อนกลับไประบบเปลี่ยนตัวเองได้ ชีวิตฉันจะดีกว่านี้					
2. ฉันพึงพอใจในสภาพชีวิตที่เป็นอยู่					
3. ฉันดำเนินชีวิตได้อย่างที่ต้องการ					
4. ฉันมีชีวิตที่มีความสุขสบายตามสมควร					
5. ฉันรู้สึกว่ารันทเวลาที่ผ่านไปในแต่ละวันช่างยาวนานเสียเหลือเกิน					
6. ที่ผ่านมา พ่อแม่หรือคนรอบข้างคอยช่วยเหลือให้กำลังใจฉัน					
7. ฉันรู้สึกภาคภูมิใจในชีวิตที่ผ่านมาของฉัน					
8. ชีวิตฉันที่ผ่านมาดีแต่เหตุการณ์ที่เลวร้าย					
9. ฉันมีความพร้อมสำหรับการใช้ชีวิตที่ดีในอนาคตข้างหน้า					
10. ฉันกังวลว่าจะไม่มีคนดูแลหรือให้ความช่วยเหลือในอนาคต					
11. ฉันไม่แน่ใจว่าจะใช้ชีวิตต่อไปได้อย่างมีความสุข					
12. ฉันประสบความสำเร็จในชีวิตอย่างที่ครอบครัวคาดหวัง					
13. ผู้ที่ฉันนับถือเห็นว่าชีวิตของฉันล้มเหลว					

มาตรวัดย่อยอารมณ์ความรู้สึกเชิงบวก

คำชี้แจง กรุณาอ่านคำเหล่านี้ แล้วพิจารณาว่า“ในช่วงระยะเวลา 1 เดือนที่ผ่านมา” ท่านเกิดอารมณ์ความรู้สึกดังกล่าวมากน้อยเพียงใด โดยทำเครื่องหมายถูก (✓) ในช่องที่ตรงกับความรู้สึกในการเกิดอารมณ์ความรู้สึกของท่านมากที่สุด โดย

- | | | |
|---|---------|--------------------|
| 1 | หมายถึง | ไม่เคย/แทบจะไม่เคย |
| 2 | หมายถึง | น้อยครั้ง |
| 3 | หมายถึง | บางครั้ง |
| 4 | หมายถึง | บ่อยครั้ง |
| 5 | หมายถึง | เป็นประจำ |

ส่วนนี้สำหรับเจ้าหน้าที่

Raw score	
PR	
Z-score	
T-score	
Stanine	
Level	

คำถาม	ระดับความถี่				
	1	2	3	4	5
1. กรุณา					
2. เข้าอกเข้าใจ					
3. ซื่นซม					
4. ซาบซึ้งใจ					
5. ดีใจ					
6. เด็ดเดี่ยว					
7. ตื่นตาตื่นใจ					
8. เบิกบานใจ					
9. ผาสุก					
10. พอใจ					
11. มั่นคง					

คำถาม	ระดับความถี่				
	1	2	3	4	5
12. มีเสน่ห์					
13. มุ่งมั่น					
14. ยินดี					
15. โส่งใจ					
16. ไร้วางใจ					
17. สนุกสนม					
18. สนุกสนาน					
19. สบายใจ					
20. สุขใจ					
21. ให้เกียรติ					
22. อ่อนน้อม					

มาตรวัดย่อยอารมณ์ความรู้สึกเชิงลบ

คำชี้แจง กรุณาอ่านคำเหล่านี้ แล้วพิจารณาว่า“ในช่วงระยะเวลา 1 เดือนที่ผ่านมา” ท่านเกิดอารมณ์ความรู้สึกดังกล่าวมากน้อยเพียงใด โดยทำเครื่องหมายถูก (✓) ในช่องที่ตรงกับความรู้สึกในการเกิดอารมณ์

ความรู้สึกของท่านมากที่สุด โดย

- 1 หมายถึง เป็นประจำ
- 2 หมายถึง บ่อยครั้ง
- 3 หมายถึง บางครั้ง
- 4 หมายถึง น้อยครั้ง
- 5 หมายถึง ไม่เคย/แทบจะไม่เคย

ส่วนนี้สำหรับเจ้าหน้าที่

Raw score	
PR	
Z-score	
T-score	
Stanine	
Level	

คำถาม	ระดับความถี่				
	1	2	3	4	5
1. กังวล					
2. แค้นใจ					
3. ผูกผูกเหยียดหยาม					
4. โดดเดี่ยว					
5. ต่อด้าน					
6. ตึงเครียด					
7. น้อยใจ					
8. บ้าคลั่ง					
9. เบื่อหน่าย					
10. ประหม่า					
11. ผิดหวัง					
12. พลุ่ณส่วน					
13. ไม่มีความสุข					

คำถาม	ระดับความถี่				
	1	2	3	4	5
14. รำคาญ					
15. รู้สึกผิด					
16. เศร้าใจ					
17. สยดสยอง					
18. สลดใจ					
19. สะเทือนใจ					
20. สิ้นหวัง					
21. เสียใจ					
22. เหน็ดเหนื่อย					
23. อึดอั้นตันใจ					
24. อับอาย					
25. อังว้าง					
26. อิจฉา					

ความสุขโดยรวม (ส่วนนี้สำหรับเจ้าหน้าที่)

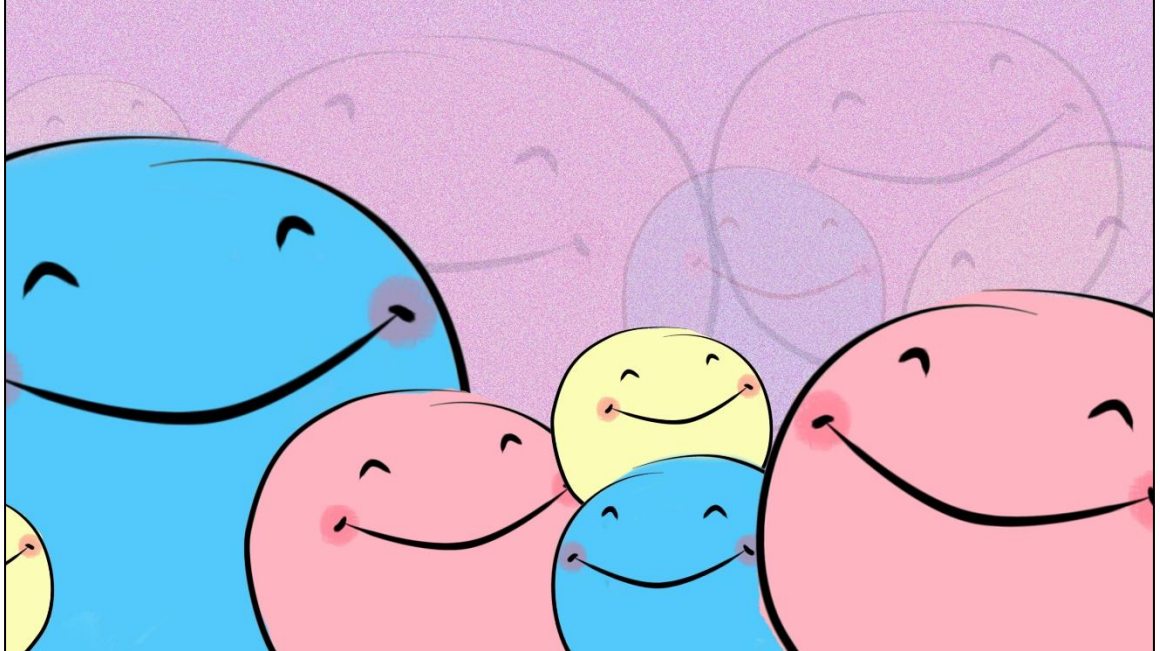
Raw score	PR	Z-score	T-score	Stanine	Level



คู่มือการใช้
มาตรวัด

ความสุข

ของคนไทย



คำนำ

คู่มือการใช้มาตรวัดความสุขของคนไทยฉบับนี้ ได้รวบรวมข้อมูลที่เกี่ยวข้องกับมาตรวัดทั้งในส่วนของความเป็นมา ขั้นตอนการพัฒนาและตรวจสอบคุณภาพของมาตรวัด รวมถึงแนวทางการดำเนินการทดสอบ การตรวจให้คะแนน การแปลความหมายคะแนนที่ได้ สำหรับผู้สนใจที่ต้องการนำมาตรวัดนี้ไปใช้เพื่อการศึกษา การวิจัย หรือแม้แต่การสำรวจความสุขของตนเอง ผู้เขียนหวังเป็นอย่างยิ่งว่า คู่มือการใช้งานเล่มนี้ จะสามารถตอบสนองความต้องการของผู้ที่สนใจได้เป็นอย่างดี หากมีข้อผิดพลาดประการใด ผู้เขียนขอน้อมรับไว้เพียงผู้เดียวและต้องขออภัยมา ณ ที่นี้ด้วย

สุชาดา สกลกิจรุ่งโรจน์



สารบัญ

จุดมุ่งหมาย	• หน้า 1
ลักษณะของมาตรฐานวัด	• หน้า 1
นิยามเชิงปฏิบัติการ	• หน้า 2
การพัฒนามาตรฐานวัด	• หน้า 3
คุณภาพของมาตรฐานวัด	• หน้า 6
การสร้างปกติวิสัย	• หน้า 7
การดำเนินการทดสอบ	• หน้า 7
การตรวจให้คะแนน	• หน้า 8
ปกติวิสัย	• หน้า 10
เกณฑ์การแปลความหมาย	• หน้า 19



1

จุดมุ่งหมาย



มาตรวัดความสุขของคนไทย รูปแบบกระดาษ-ดินสอ พัฒนาขึ้นเพื่อวัดความอยู่ดีมีสุขเชิงอัตวิสัย ซึ่งประกอบด้วย 3 องค์ประกอบย่อย ได้แก่ ความพึงพอใจในชีวิต อารมณ์ความรู้สึกเชิงบวก และอารมณ์ความรู้สึกเชิงลบ เพื่อให้ทราบถึงระดับความสุขของบุคคล ซึ่งนำไปสู่การส่งเสริมให้บุคคลมีคุณภาพชีวิตที่ดียิ่งขึ้น รวมทั้งเป็นเครื่องมือสำคัญสำหรับการสร้างองค์ความรู้และงานวิจัยที่เกี่ยวข้องกับความสุขต่อไป

ลักษณะของมาตรวัด

มาตรวัดความสุขของคนไทย รูปแบบกระดาษ-ดินสอ มีลักษณะเป็นมาตรประมาณค่า 5 ระดับ (Rating Scale) จำนวน 61 ข้อ ประกอบด้วยองค์ประกอบย่อย 3 ด้าน ได้แก่

ด้านที่ 1 ความพึงพอใจในชีวิต	จำนวน 13 ข้อ
ด้านที่ 2 อารมณ์ความรู้สึกเชิงบวก	จำนวน 22 ข้อ
ด้านที่ 3 อารมณ์ความรู้สึกเชิงลบ	จำนวน 26 ข้อ



นิยามเชิงปฏิบัติการ

ความสุข (Happiness) หรือ ความอยู่ดีมีสุขเชิงอัตวิสัย (Subjective Well-being) เป็นกระบวนการในการตัดสินและประเมินคุณภาพชีวิตของบุคคลตามมุมมองส่วนตัว ใน 3 องค์ประกอบย่อย ได้แก่ ความพึงพอใจในชีวิต อารมณ์ความรู้สึกเชิงบวก และอารมณ์ความรู้สึกเชิงลบ แต่ละบุคคลอาจมีเกณฑ์ในการประเมินที่แตกต่างกัน การระบุว่าคุณภาพชีวิตของตนดีหรือไม่ จึงขึ้นอยู่กับ การตัดสินใจของบุคคลนั้นๆ เอง

มาตรวัดความสุขของคนไทย (Thai Happiness Scale) หมายถึง มาตรวัดความสุขที่ผู้วิจัยสร้างขึ้น สำหรับคนไทย ที่มีอายุตั้งแต่ 18 ปีขึ้นไป มีลักษณะเป็นมาตรประมาณค่า 5 ระดับ ซึ่งประกอบด้วย 3 องค์ประกอบย่อย ตามทฤษฎีความอยู่ดีมีสุขเชิงอัตวิสัย

ความพึงพอใจในชีวิต (Life Satisfaction) หมายถึง การประเมินความพึงพอใจในชีวิตตนเองในภาพรวม ไม่เจาะจงที่ด้านใดด้านหนึ่ง โดยเฉพาะ ซึ่งแบ่งมุมมองความพึงพอใจในชีวิต ออกเป็น 5 มุมมอง ได้แก่ 1) ความต้องการเปลี่ยนแปลงชีวิตของตน 2) ความพึงพอใจในชีวิตปัจจุบัน 3) ความพึงพอใจในชีวิตอดีต 4) ความพึงพอใจในชีวิตอนาคต และ 5) ความคิดเห็นจากคนสำคัญที่มีต่อชีวิตของคุณ

อารมณ์ความรู้สึกเชิงบวก (Positive Affect) หมายถึง ประสบการณ์ที่แตกต่างกับบุคคล ในการมีปฏิสัมพันธ์กับบุคคล หรือสิ่งรอบข้างด้วยอารมณ์ความรู้สึกที่ดี ซึ่งทำให้บุคคลเกิดความพึงพอใจ

อารมณ์ความรู้สึกเชิงลบ (Negative Affect) หมายถึง ประสบการณ์ที่แตกต่างกับบุคคลในการมีปฏิสัมพันธ์กับบุคคล หรือสิ่งรอบข้างด้วยอารมณ์ความรู้สึกที่ไม่ดี ซึ่งทำให้บุคคลเกิดความรู้สึกไม่พึงพอใจ และมีภาพพจน์ต่อตนเองในเชิงลบ



การพัฒนามาตรวัด

มาตรวัดความสุขของคนไทย รูปแบบกระดาษ-ดินสอ เป็นมาตรวัดที่พัฒนาคู่ขนานกับมาตรวัดความสุขของคนไทย รูปแบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ซึ่งมีการดำเนินการ ดังนี้

1. การพัฒนามาตรวัดความสุขของคนไทย เริ่มต้นจากการศึกษา และวิเคราะห์ทฤษฎี เอกสารงานวิจัยที่เกี่ยวข้องกับทฤษฎีความอยู่ดีมีสุขเชิงอัตวิสัย (Pavot & Diener, 2013) เมื่อใช้เป็นแนวทางในการพัฒนามาตรวัด กำหนดนิยามเชิงปฏิบัติการของมาตรวัดตามองค์ประกอบทั้ง 3 ด้าน ได้แก่ ความพึงพอใจในชีวิต อารมณ์ความรู้สึกเชิงบวก และอารมณ์ความรู้สึกเชิงลบ

2. เมื่อได้นิยามเชิงปฏิบัติการแล้วจึงพัฒนามาตรวัดความสุขของคนไทย ซึ่งประกอบด้วยข้อคำถามที่มุ่งวัดองค์ประกอบทั้งสามด้าน ในลักษณะมาตรประมาณค่า 5 ระดับ จำนวนทั้งสิ้น 191 ข้อ โดยเป็นข้อคำถามในองค์ประกอบด้านความพึงพอใจในชีวิต จำนวน 30 ข้อ อารมณ์ความรู้สึกเชิงบวก จำนวน 67 ข้อ และอารมณ์ความรู้สึกเชิงลบ จำนวน 94 ข้อ จากนั้นจึงให้ผู้เชี่ยวชาญ 5 คนตรวจสอบความตรงเชิงเนื้อหา กำหนดระดับการแสดงความคิดเห็นในลักษณะมาตรประมาณค่า 5 ระดับ

รายชื่อผู้เชี่ยวชาญ มีดังนี้

- 1) รศ.กนกรัตน์ สุระตุลคะ ที่ปรึกษาสาขาวิชาจิตวิทยาคลินิก ภาควิชาจิตเวชศาสตร์ คณะแพทยศาสตร์ศิริราชพยาบาล มหาวิทยาลัยมหิดล
- 2) รศ.ดร.สุชีรา ภัทรายุตวรรตน์ อาจารย์ประจำสาขาวิชาจิตวิทยาคลินิก ภาควิชาจิตเวชศาสตร์ คณะแพทยศาสตร์ศิริราชพยาบาล มหาวิทยาลัยมหิดล
- 3) รศ.ดร.สุพินธ์ ศรีนันทน์อรสกุล อาจารย์ประจำวิทยาลัยวิทยาการวิจัยและวิทยาการปัญญา มหาวิทยาลัยบูรพา
- 4) อาจารย์นัฐธิดา บุญอานักธีเจริญ อาจารย์ประจำสาขาวิชามนุษยศาสตร์และสังคมศาสตร์ คณะศิลปศาสตร์ มหาวิทยาลัยหัวเวียเฉลิมนระเกียรติ
- 5) อาจารย์สร้อยสุดา อิ่มอรุณรักษ์ อาจารย์ประจำสาขาวิชาจิตวิทยาคลินิก ภาควิชาจิตเวชศาสตร์ คณะแพทยศาสตร์ศิริราชพยาบาล มหาวิทยาลัยมหิดล



การพัฒนามาตรวัด

3. คัดเลือกข้อคำถามที่มีค่าเฉลี่ยความสอดคล้องของข้อคำถามกับนิยามเชิงปฏิบัติการ ตั้งแต่ 4.00 ขึ้นไป และปรับแก้ข้อคำถามตามคำแนะนำของผู้เชี่ยวชาญ ได้ข้อคำถาม 188 ข้อ ประกอบด้วย ข้อคำถามในองค์ประกอบด้านความพึงพอใจในชีวิต จำนวน 30 ข้อ อารมณ์ความรู้สึกเชิงบวก จำนวน 64 ข้อ และอารมณ์ความรู้สึกเชิงลบ จำนวน 94 ข้อ

4. นำข้อคำถามที่ผ่านเกณฑ์การตรวจสอบความตรงตามเนื้อหาโดยผู้เชี่ยวชาญ ไปทดลองใช้กับกลุ่มตัวอย่างที่มีลักษณะใกล้เคียงกับประชากรที่ต้องการศึกษาวิจัย จำนวน 30 คน เมื่อตรวจสอบการใช้ภาษา คำชี้แจงในการทำมาตรวัด ระยะเวลาที่ใช้ รูปแบบการจัดพิมพ์ และข้อคิดเห็นอื่นๆ จากนั้น จึงวิเคราะห์ค่าอำนาจจำแนกเป็นรายข้อ ปรับปรุงข้อคำถามและมาตรวัดตามข้อคิดเห็นที่ได้จากกลุ่มทดลองใช้ แล้วจึงคัดเลือกข้อคำถามที่มีคุณภาพ จำนวน 182 ข้อ ประกอบด้วย ข้อคำถามในองค์ประกอบด้านความพึงพอใจในชีวิต จำนวน 29 ข้อ อารมณ์ความรู้สึกเชิงบวก จำนวน 63 ข้อ และอารมณ์ความรู้สึกเชิงลบ จำนวน 90 ข้อ

5. นำมาตรวัดความสุขของคนไทย จำนวน 182 ข้อ ไปเก็บข้อมูลกับกลุ่มตัวอย่างซึ่งเป็นคนไทย อายุ 18 ปีขึ้นไป จำนวน 1,000 คน จากนั้นจึงนำผลที่ได้มาวิเคราะห์คุณภาพตามทฤษฎีการทดสอบแบบดั้งเดิม และทฤษฎีการตอบสนองข้อสอบ โดยคัดมีเกณฑ์การคัดเลือกข้อคำถาม ดังนี้

5.1 คัดเลือกข้อคำถามที่มีค่าอำนาจจำแนกมากกว่า 0.30 ตามทฤษฎีการทดสอบแบบดั้งเดิม

5.2 คัดเลือกข้อคำถามที่มีความถี่ในการเลือกรายการคำตอบทุกรายการ



การพัฒนามาตรวัด

5.3 คัดเลือกข้อคำถามที่มีความเป็นเอกมิติ และมีความเป็นอิสระของข้อคำถาม ตามทฤษฎีการตอบสนองข้อสอบ

5.4 คัดเลือกข้อคำถามที่ข้อมูลเชิงประจักษ์มีความสอดคล้องกับ Graded Response Model ตามทฤษฎีการตอบสนองข้อสอบ

จากการวิเคราะห์คุณภาพของข้อคำถามตามทฤษฎีการทดสอบแบบดั้งเดิม และทฤษฎีการตอบสนองข้อสอบ ได้ข้อคำถามทั้งสิ้น 61 ข้อ ประกอบด้วย ข้อคำถามในองค์ประกอบด้านความพึงพอใจในชีวิต จำนวน 13 ข้อ อารมณ์ความรู้สึกเชิงบวก จำนวน 22 ข้อ และอารมณ์ความรู้สึกเชิงลบ จำนวน 26 ข้อ

6. ตรวจสอบความตรงเชิงโครงสร้างของมาตรวัดความสุขของคนไทย 3 องค์ประกอบ จำนวน 61 ข้อ ด้วยการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสอง (Second Order Confirmatory Factor Analysis) เพื่อตรวจสอบว่ามาตรวัดความสุขของคนไทย ประกอบด้วย 3 องค์ประกอบตามทฤษฎีความอยู่ดีมีสุขเชิงอัตวิสัยหรือไม่

7. วิเคราะห์ค่าความเที่ยงทั้งฉบับ ของมาตรวัดความสุขของคนไทย โดยตรวจสอบความสอดคล้องภายใน (Internal Consistency) ด้วยวิธีครอนบาค แอลฟา

8. สร้างปกติวิสัยของมาตรวัดความสุขของคนไทยในรูปของตารางตำแหน่งเปอร์เซ็นต์ไทล์ (Percentile Rank) และคะแนนมาตรฐาน (คะแนนซี คะแนนที และคะแนนสเตโนว์)



คุณภาพของมาตรวัดความสุขของคนไทย

1. มาตรวัดความสุขของคนไทยทั้ง 61 ข้อ มีค่าอำนาจจำแนกตามทฤษฎีการทดสอบแบบดั้งเดิมมากกว่า 0.30 โดยองค์ประกอบด้านความพึงพอใจในชีวิตมีค่าอำนาจจำแนกอยู่ระหว่าง 0.35 - 0.64 อารมณ์ความรู้สึกเชิงบวกมีค่าอำนาจจำแนกอยู่ระหว่าง 0.32 - 0.58 และอารมณ์ความรู้สึกเชิงลบมีค่าอำนาจจำแนกอยู่ระหว่าง 0.31 - 0.59

2. มาตรวัดความสุขของคนไทยทั้ง 61 ข้อ มีความตรงเชิงเนื้อหา ซึ่งพิจารณาจากค่าเฉลี่ยความสอดคล้องของข้อคำถามกับนิยามเชิงปฏิบัติการที่มีค่าตั้งแต่ 4.00 ขึ้นไป

3. ความเที่ยงของมาตรวัดความสุขของคนไทย มีค่าเท่ากับ 0.94 พิจารณาในแตละองค์ประกอบ ดังนี้ ด้านความพึงพอใจในชีวิต มีค่าความเที่ยงเท่ากับ 0.84 ด้านอารมณ์ความรู้สึกเชิงบวก มีค่าความเที่ยงเท่ากับ 0.92 และด้านอารมณ์ความรู้สึกเชิงลบ มีค่าความเที่ยงเท่ากับ 0.94

4. ความตรงเชิงโครงสร้างของมาตรวัดความสุขของคนไทยจากการวิเคราะห์องค์ประกอบเชิงยืนยันอันดับสอง 3 องค์ประกอบ ข้อคำถาม 61 ข้อ ปรากฏว่าโมเดลมีความสอดคล้องกับข้อมูลเชิงประจักษ์ ซึ่งพิจารณาได้จากค่าไค-สแควร์ มีค่าเท่ากับ 1067.36 ($p=0.47$) ที่องศาอิสระเท่ากับ 1065 แสดงให้เห็นจากค่า และดัชนีวัดระดับความกลมกลืน (GFI) เท่ากับ 0.96 ดัชนีวัดระดับความกลมกลืนที่ปรับแก้แล้ว (AGFI) เท่ากับ 0.94 ดัชนีวัดระดับความสอดคล้องเปรียบเทียบ (CFI) เท่ากับ 1.00 ดัชนีรากของค่าเฉลี่ยกำลังสองของความคลาดเคลื่อนมาตรฐาน (Standardized RMR) เท่ากับ 0.04 ค่าความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ (RMSEA) เท่ากับ 0.00 แสดงว่าโมเดลมีความสอดคล้องกับข้อมูลเชิงประจักษ์ และแสดงให้เห็นอีกด้วย มาตรวัดความสุขของคนไทยมีความตรงเชิงโครงสร้าง



การสร้างปกติวิสัย

ปกติวิสัยเป็นค่าสถิติที่ได้จากผลการตอบมาตรวัดของกลุ่มตัวอย่าง จำนวน 934 คน ซึ่งเป็นตัวแทนของประชากรเป้าหมาย

ปกติวิสัยของมาตรวัดความสุขของคนไทยที่ผู้วิจัยพัฒนาขึ้น เป็นการหาปกติวิสัยในรูปแบบของตำแหน่งเปอร์เซ็นต์ไทล์ (Percentile Rank) และคะแนนมาตรฐาน (คะแนนซิกมา คะแนนที และคะแนนสเตไดอันท์) โดยนำเสนอในรูปแบบตาราง จากนั้นจึงใช้คะแนนสเตไดอันท์ในการกำหนดเกณฑ์ระดับความสุข

การดำเนินการทดสอบ

การดำเนินการทดสอบควรจัดในสถานที่ที่มีอากาศถ่ายเทสะดวก เงียบสงบ ปราศจากสิ่งรบกวน มีที่นั่งสำหรับการเขียนได้อย่างสะดวก โดยมีการจัดเตรียมอุปกรณ์สำหรับการทดสอบ ได้แก่ มาตรวัดความสุขของคนไทย รูปแบบกระดาษ-ดินสอ และดินสอ 2B (หรือปากกา) ผู้ทดสอบสามารถอ่านคำชี้แจงในการทดสอบได้โดยตรง จากมาตรวัดความสุขของคนไทย ซึ่งจะใช้เวลาในการทดสอบประมาณ 15-20 นาที สำหรับการทดสอบครบทั้งสามมาตรวัดย่อย (ความพึงพอใจในชีวิต อารมณ์ความรู้สึกเชิงบวก และอารมณ์ความรู้สึกเชิงลบ) หรือใช้ระยะเวลาในการทดสอบประมาณ 5-10 นาที สำหรับการทดสอบเฉพาะแต่ละมาตรวัดย่อย



การตรวจให้คะแนน

มาตรวัดความสุขของคนไทย มีข้อคำถาม 61 ข้อ คะแนนเต็ม 305 คะแนน วัดความสุขในสามองค์ประกอบย่อย แต่ละมาตรวัดย่อยมีจำนวนข้อคำถาม คะแนนเต็ม และวิธีการให้คะแนน ดังนี้

ความพึงพอใจในชีวิต

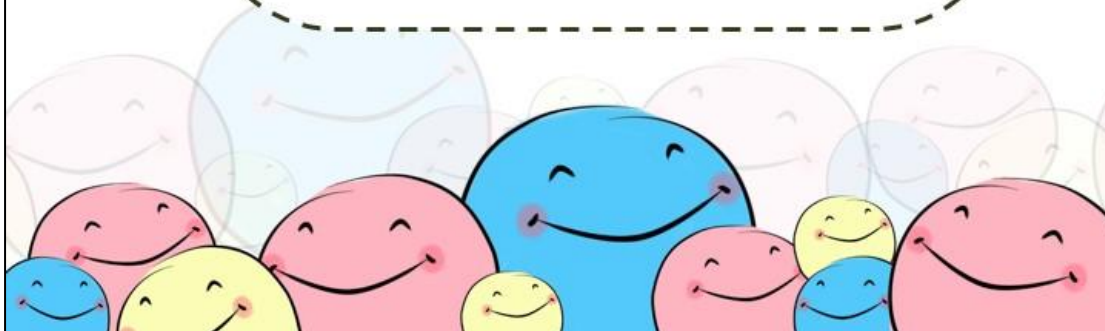
มีข้อคำถาม 13 ข้อ คะแนนเต็ม 65 คะแนน มีเกณฑ์การตรวจให้คะแนน ดังนี้

1) ข้อคำถามทางบวก

ไม่เห็นด้วย/เห็นด้วยน้อยที่สุด	1	คะแนน
เห็นด้วยน้อย	2	คะแนน
เห็นด้วยปานกลาง	3	คะแนน
เห็นด้วยมาก	4	คะแนน
เห็นด้วยอย่างยิ่ง	5	คะแนน

2) ข้อคำถามทางบวก

ไม่เห็นด้วย/เห็นด้วยน้อยที่สุด	5	คะแนน
เห็นด้วยน้อย	4	คะแนน
เห็นด้วยปานกลาง	3	คะแนน
เห็นด้วยมาก	2	คะแนน
เห็นด้วยอย่างยิ่ง	1	คะแนน



อารมณ์ความรู้สึกเชิงบวก

มีข้อคำถาม 22 ข้อ คะแนนเต็ม 110 คะแนน มีเกณฑ์การตรวจให้คะแนน ดังนี้

ไม่เคยหรือแทบจะไม่เคยเกิดอารมณ์ความรู้สึกดังกล่าว	1	คะแนน
เกิดอารมณ์ความรู้สึกดังกล่าวบ่อยครั้ง	2	คะแนน
เกิดอารมณ์ความรู้สึกดังกล่าวบางครั้ง	3	คะแนน
เกิดอารมณ์ความรู้สึกดังกล่าวบ่อยครั้ง	4	คะแนน
เกิดอารมณ์ความรู้สึกดังกล่าวเป็นประจำ	5	คะแนน

อารมณ์ความรู้สึกเชิงลบ

มีข้อคำถาม 26 ข้อ คะแนนเต็ม 130 คะแนน มีเกณฑ์การตรวจให้คะแนน ดังนี้

ไม่เคยหรือแทบจะไม่เคยเกิดอารมณ์ความรู้สึกดังกล่าว	5	คะแนน
เกิดอารมณ์ความรู้สึกดังกล่าวบ่อยครั้ง	4	คะแนน
เกิดอารมณ์ความรู้สึกดังกล่าวบางครั้ง	3	คะแนน
เกิดอารมณ์ความรู้สึกดังกล่าวบ่อยครั้ง	2	คะแนน
เกิดอารมณ์ความรู้สึกดังกล่าวเป็นประจำ	1	คะแนน

การคำนวณคะแนนสามารถทำได้สองรูปแบบ ดังนี้

1) คะแนนมาตรฐานย่อย ได้จากผลรวมคะแนนในแต่ละมาตรฐานย่อย ได้แก่ ความพึงพอใจในชีวิต อารมณ์ความรู้สึกเชิงบวก และอารมณ์ความรู้สึกเชิงลบ

2) คะแนนความสุขโดยรวม ได้จากการผลรวมคะแนนจาก 3 มาตรฐานย่อย ได้แก่ มาตรฐานย่อยความพึงพอใจในชีวิต มาตรฐานย่อยอารมณ์ความรู้สึกเชิงบวก และมาตรฐานย่อยอารมณ์ความรู้สึกเชิงลบ



การแปลความหมายจากปกติวิสัย

ผลรวมคะแนนในแต่ละมาตรวัดย่อย หรือผลรวมคะแนนความสุขโดยรวมซึ่งเป็นคะแนนดิบสามารถนำไปเทียบเป็นคะแนนมาตรฐานจากตารางปกติวิสัย ในรูปแบบตำแหน่งเปอร์เซ็นต์ไทล์ คะแนนซิก และคะแนนที และคะแนนสเตนโน รวมทั้งสามารถแปลความหมายของช่วงคะแนนได้ตามเกณฑ์การแปลความหมายที่ผู้วิจัยพัฒนาขึ้น

ปกติวิสัยของมาตรวัดย่อยความพึงพอใจในชีวิต

คะแนนดิบ	PR	Z-score	T-score	Stanline	คะแนนดิบ	PR	Z-score	T-score	Stanline
65	100.00	2.43	74	9	42	25.37	-0.72	43	4
64	99.46	2.29	73	9	41	21.63	-0.85	41	3
63	99.04	2.15	72	9	40	18.42	-0.99	40	3
62	98.29	2.02	70	9	39	14.35	-1.13	39	3
61	97.97	1.88	69	9	38	11.35	-1.26	37	3
60	97.22	1.74	67	9	37	8.89	-1.40	36	2
59	96.04	1.61	66	9	36	6.96	-1.54	35	2
58	94.75	1.47	65	8	35	5.46	-1.67	33	2
57	92.83	1.33	63	8	34	4.28	-1.81	32	2
56	90.47	1.20	62	8	33	2.89	-1.95	31	1
55	87.37	1.06	61	7	32	2.57	-2.08	29	1
54	83.94	0.92	59	7	31	2.25	-2.22	28	1
53	80.41	0.79	58	7	30	1.71	-2.36	26	1
52	75.37	0.65	56	6	29	1.39	-2.49	25	1
51	70.45	0.51	55	6	28	0.96	-2.63	24	1
50	66.17	0.38	54	6	27	0.75	-2.77	22	1
49	60.81	0.24	52	6	26	0.64	-2.90	21	1
48	55.25	0.10	51	5	25	0.64	-3.04	20	1
47	49.25	-0.03	50	5	24	0.32	-3.18	18	1
46	44.75	-0.17	48	5	23	0.21	-3.31	17	1
45	39.19	-0.31	47	4	22	0.21	-3.45	15	1
44	33.83	-0.44	46	4	21	0.21	-3.59	14	1
43	30.84	-0.58	44	4	13-20	0.11	-3.72	13	1

ปกติวิสัยของมาตรฐานด้วยอารมณ์ความรู้สึกเชิงบวก

คะแนนดิบ	PR	Z-score	T-score	StanIne	คะแนนดิบ	PR	Z-score	T-score	StanIne
110	100.00	2.49	75	9	89	81.26	0.84	58	7
109	99.25	2.41	74	9	88	78.91	0.76	58	7
108	99.04	2.33	73	9	87	77.41	0.69	57	7
107	98.61	2.26	73	9	86	75.59	0.61	56	6
106	98.29	2.18	72	9	85	72.59	0.53	55	6
105	97.86	2.10	71	9	84	69.27	0.45	54	6
104	97.32	2.02	70	9	83	65.95	0.37	54	6
103	96.79	1.94	69	9	82	62.74	0.29	53	6
102	96.47	1.86	69	9	81	59.42	0.21	52	5
101	95.40	1.79	68	8	80	56.32	0.14	51	5
100	94.97	1.71	67	8	79	53.85	0.06	51	5
99	93.90	1.63	66	8	78	50.54	-0.02	50	5
98	93.25	1.55	65	8	77	48.29	-0.10	49	5
97	92.51	1.47	65	8	76	45.72	-0.18	48	5
96	92.08	1.39	64	8	75	42.61	-0.26	47	5
95	90.79	1.31	63	8	74	40.04	-0.34	47	5
94	90.15	1.24	62	8	73	36.62	-0.41	46	4
93	88.76	1.16	62	7	72	33.19	-0.49	45	4
92	87.58	1.08	61	7	71	30.30	-0.57	44	4
91	85.22	1.00	60	7	70	27.84	-0.65	43	4
90	82.55	0.92	59	7	69	25.27	-0.73	43	4



ปกติวิสัยของมาตรวัดย่อยอารมณ์ความรู้สึกเชิงบวก

คะแนนดิบ	PR	Z-score	T-score	StanIne	คะแนนดิบ	PR	Z-score	T-score	StanIne
68	22.70	-0.81	42	3	53	2.36	-1.99	30	1
67	20.13	-0.89	41	3	52	1.82	-2.06	29	1
66	17.99	-0.96	40	3	51	1.39	-2.14	29	1
65	15.10	-1.04	40	3	50	1.28	-2.22	28	1
64	13.06	-1.12	39	3	49	0.86	-2.30	27	1
63	11.35	-1.20	38	3	48	0.64	-2.38	26	1
62	10.49	-1.28	37	2	47	0.54	-2.46	25	1
61	9.21	-1.36	36	2	46	0.54	-2.54	25	1
60	7.82	-1.44	36	2	45	0.32	-2.61	24	1
59	7.28	-1.51	35	2	44	0.32	-2.69	23	1
58	6.42	-1.59	34	2	43	0.32	-2.77	22	1
57	5.14	-1.67	33	2	42	0.32	-2.85	22	1
56	4.18	-1.75	32	2	41	0.21	-2.93	21	1
55	3.10	-1.83	32	1	22-40	0.11	-3.01	20	1
54	2.78	-1.91	31	1					



ปกติวิสัยของมาตรวัดย่อยอารมณ์ความรู้สึกเชิงลบ

คะแนนดิบ	PR	Z-score	T-score	Stanine	คะแนนดิบ	PR	Z-score	T-score	Stanine
130	100.00	1.83	68	9	109	68.95	0.53	55	6
129	99.68	1.77	68	9	108	66.92	0.47	55	6
128	98.93	1.71	67	9	107	64.45	0.41	54	6
127	98.50	1.64	66	9	106	62.74	0.35	53	6
126	97.75	1.58	66	9	105	59.96	0.29	53	5
125	96.57	1.52	65	9	104	56.42	0.22	52	5
124	95.50	1.46	65	8	103	54.60	0.16	52	5
123	94.22	1.40	64	8	102	50.86	0.10	51	5
122	93.36	1.33	63	8	101	48.39	0.04	50	5
121	91.76	1.27	63	8	100	46.36	-0.02	50	5
120	89.94	1.21	62	8	99	44.00	-0.08	49	5
119	88.87	1.15	61	7	98	41.01	-0.15	49	5
118	87.58	1.09	61	7	97	39.51	-0.21	48	4
117	84.69	1.03	60	7	96	37.26	-0.27	47	4
116	83.40	0.96	60	7	95	34.48	-0.33	47	4
115	81.69	0.90	59	7	94	32.23	-0.39	46	4
114	78.59	0.84	58	7	93	30.41	-0.46	45	4
113	76.34	0.78	58	6	92	28.16	-0.52	45	4
112	74.20	0.72	57	6	91	26.02	-0.58	44	4
111	71.84	0.66	57	6	90	24.63	-0.64	44	4
110	70.99	0.59	56	6	89	23.13	-0.70	43	4



ปกติวิสัยของมาตรวัดย่อยอารมณ์ความรู้สึกเชิงลบ

คะแนนดิบ	PR	Z-score	T-score	Stanine	คะแนนดิบ	PR	Z-score	T-score	Stanine
88	21.84	-0.76	42	3	64	2.57	-2.25	28	1
87	20.66	-0.83	42	3	63	2.36	-2.31	27	1
86	19.38	-0.89	41	3	62	2.03	-2.37	26	1
85	17.77	-0.95	41	3	61	2.03	-2.43	26	1
84	16.60	-1.01	40	3	60	1.71	-2.49	25	1
83	15.52	-1.07	39	3	59	1.50	-2.55	24	1
82	14.24	-1.13	39	3	58	1.39	-2.62	24	1
81	13.60	-1.20	38	3	57	1.39	-2.68	23	1
80	12.74	-1.26	37	3	56	1.28	-2.74	23	1
79	11.46	-1.32	37	3	55	1.07	-2.80	22	1
78	10.28	-1.38	36	2	54	1.07	-2.86	21	1
77	8.89	-1.44	36	2	53	1.07	-2.92	21	1
76	8.24	-1.50	35	2	52	0.86	-2.99	20	1
75	7.82	-1.57	34	2	51	0.75	-3.05	20	1
74	7.07	-1.63	34	2	50	0.64	-3.11	19	1
73	6.21	-1.69	33	2	49	0.54	-3.17	18	1
72	5.46	-1.75	32	2	48	0.32	-3.23	18	1
71	5.25	-1.81	32	2	47	0.21	-3.30	17	1
70	4.39	-1.88	31	2	46	0.21	-3.36	16	1
69	3.96	-1.94	31	1	45	0.11	-3.42	16	1
68	3.64	-2.00	30	1	44	0.11	-3.48	15	1
67	3.32	-2.06	29	1	43	0.11	-3.54	15	1
66	2.89	-2.12	29	1	42	0.11	-3.60	14	1
65	2.78	-2.18	28	1	26 - 41	0.11	-3.67	13	1



ปกติวิสัยของมาตรวัดความสุขโดยรวม

คะแนนดิบ	PR	Z-score	T-score	Stanine	คะแนนดิบ	PR	Z-score	T-score	Stanine
300-305	100.00	2.84	78	9	279	97.97	2.06	71	9
299	99.89	2.81	78	9	278	97.86	2.02	70	9
298	99.89	2.77	78	9	277	97.54	1.98	70	9
297	99.89	2.73	77	9	276	97.54	1.94	69	9
296	99.89	2.69	77	9	275	97.43	1.91	69	9
295	99.89	2.66	77	9	274	97.22	1.87	69	9
294	99.89	2.62	76	9	273	97.22	1.83	68	9
293	99.79	2.58	76	9	272	96.57	1.79	68	9
292	99.79	2.54	75	9	271	96.47	1.76	68	9
291	99.79	2.51	75	9	270	96.47	1.72	67	9
290	99.79	2.47	75	9	269	96.36	1.68	67	9
289	99.79	2.43	74	9	268	96.36	1.64	66	9
288	99.79	2.39	74	9	267	95.93	1.61	66	8
287	99.57	2.36	74	9	266	95.82	1.57	66	8
286	99.36	2.32	73	9	265	95.82	1.53	65	8
285	99.14	2.28	73	9	264	95.40	1.49	65	8
284	98.82	2.24	72	9	263	95.18	1.46	65	8
283	98.61	2.21	72	9	262	94.86	1.42	64	8
282	98.50	2.17	72	9	261	94.00	1.38	64	8
281	98.29	2.13	71	9	260	93.47	1.34	63	8
280	98.07	2.09	71	9	259	93.25	1.31	63	8



ปกติวิสัยของมาตรวัดความสุขโดยรวม

คะแนนดิบ	PR	Z-score	T-score	Stanine	คะแนนดิบ	PR	Z-score	T-score	Stanine
258	92.51	1.27	63	8	237	67.77	0.48	55	6
257	91.97	1.23	62	8	236	66.38	0.44	54	6
256	90.79	1.19	62	8	235	64.99	0.40	54	6
255	89.94	1.16	62	8	234	64.03	0.37	54	6
254	89.61	1.12	61	8	233	62.63	0.33	53	6
253	88.87	1.08	61	7	232	60.81	0.29	53	6
252	88.12	1.04	60	7	231	58.89	0.25	53	5
251	86.83	1.00	60	7	230	57.07	0.22	52	5
250	85.44	0.97	60	7	229	54.71	0.18	52	5
249	84.90	0.93	59	7	228	53.10	0.14	51	5
248	83.83	0.89	59	7	227	51.61	0.10	51	5
247	82.12	0.85	59	7	226	50.00	0.07	51	5
246	81.26	0.82	58	7	225	48.93	0.03	50	5
245	79.87	0.78	58	7	224	47.54	-0.01	50	5
244	77.94	0.74	57	7	223	45.29	-0.05	50	5
243	76.34	0.70	57	6	222	44.22	-0.08	49	5
242	74.95	0.67	57	6	221	42.61	-0.12	49	5
241	73.45	0.63	56	6	220	40.79	-0.16	48	5
240	71.73	0.59	56	6	219	39.61	-0.20	48	4
239	70.45	0.55	56	6	218	38.12	-0.23	48	4
238	68.95	0.52	55	6	217	37.04	-0.27	47	4



ปกติวิสัยของมาตรวัดความสุขโดยรวม

คะแนนดิบ	PR	Z-score	T-score	Stanine	คะแนนดิบ	PR	Z-score	T-score	Stanine
216	35.44	-0.31	47	4	195	13.49	-1.10	39	3
215	33.40	-0.35	47	4	194	13.17	-1.14	39	3
214	32.12	-0.38	46	4	193	12.74	-1.17	38	3
213	31.05	-0.42	46	4	192	11.88	-1.21	38	3
212	29.66	-0.46	45	4	191	11.24	-1.25	38	3
211	28.69	-0.50	45	4	190	11.13	-1.29	37	3
210	27.62	-0.53	45	4	189	10.28	-1.32	37	2
209	26.55	-0.57	44	4	188	9.85	-1.36	36	2
208	25.05	-0.61	44	4	187	9.64	-1.40	36	2
207	24.20	-0.65	44	4	186	9.31	-1.44	36	2
206	23.23	-0.68	43	4	185	9.10	-1.47	35	2
205	21.84	-0.72	43	3	184	8.67	-1.51	35	2
204	20.88	-0.76	42	3	183	8.24	-1.55	35	2
203	20.02	-0.80	42	3	182	7.71	-1.59	34	2
202	19.06	-0.84	42	3	181	7.49	-1.62	34	2
201	18.31	-0.87	41	3	180	7.07	-1.66	33	2
200	17.34	-0.91	41	3	179	6.64	-1.70	33	2
199	16.38	-0.95	41	3	178	5.89	-1.74	33	2
198	15.74	-0.99	40	3	177	5.35	-1.77	32	2
197	15.20	-1.02	40	3	176	5.03	-1.81	32	2
196	14.56	-1.06	39	3	175	4.82	-1.85	32	2



ปกติวิสัยของมาตรวัดความสุขโดยรวม

คะแนนดิบ	PR	Z-score	T-score	Stanine	คะแนนดิบ	PR	Z-score	T-score	Stanine
174	4.50	-1.89	31	2	150	0.96	-2.79	22	1
173	4.07	-1.92	31	2	149	0.75	-2.83	22	1
172	3.75	-1.96	30	1	148	0.64	-2.86	21	1
171	3.10	-2.00	30	1	147	0.43	-2.90	21	1
170	2.89	-2.04	30	1	146	0.43	-2.94	21	1
169	2.57	-2.07	29	1	145	0.43	-2.98	20	1
168	2.57	-2.11	29	1	144	0.32	-3.01	20	1
167	2.46	-2.15	29	1	143	0.32	-3.05	19	1
166	2.25	-2.19	28	1	142	0.32	-3.09	19	1
165	2.03	-2.22	28	1	141	0.32	-3.13	19	1
164	1.82	-2.26	27	1	140	0.32	-3.16	18	1
163	1.71	-2.30	27	1	139	0.32	-3.20	18	1
162	1.71	-2.34	27	1	138	0.21	-3.24	18	1
161	1.61	-2.37	26	1	137	0.21	-3.28	17	1
160	1.61	-2.41	26	1	136	0.21	-3.31	17	1
159	1.61	-2.45	26	1	135	0.11	-3.35	16	1
158	1.39	-2.49	25	1	134	0.11	-3.39	16	1
157	1.39	-2.52	25	1	133	0.11	-3.43	16	1
156	1.39	-2.56	24	1	132	0.11	-3.46	15	1
155	1.39	-2.60	24	1	131	0.11	-3.50	15	1
154	1.28	-2.64	24	1	130	0.11	-3.54	15	1
153	1.28	-2.67	23	1	129	0.11	-3.58	14	1
152	1.18	-2.71	23	1	61 - 128	0.11	-3.61	14	1
151	1.18	-2.75	22	1					



เกณฑ์การเทียบค่าคะแนน

เกณฑ์การเทียบช่วงคะแนนดิบของมาตรวัดย่อยความพึงพอใจในชีวิต

ช่วงคะแนนดิบ	ระดับความพึงพอใจในชีวิต
59 - 65	พึงพอใจอย่างยิ่ง
56 - 58	พึงพอใจมาก
53 - 55	พึงพอใจ
49 - 52	ค่อนข้างพึงพอใจ
46 - 48	ปานกลาง
42 - 45	พึงพอใจน้อย
38 - 41	ไม่ค่อยพึงพอใจ
34 - 37	ไม่พึงพอใจ
13 - 33	ไม่พึงพอใจอย่างยิ่ง

เกณฑ์การเทียบช่วงคะแนนดิบของมาตรวัดย่อยอารมณ์ความรู้สึกเชิงบวก

ช่วงคะแนนดิบ	ระดับอารมณ์ความรู้สึกเชิงบวก
102 - 110	สูงที่สุด
94 - 101	สูงมาก
87 - 93	สูง
82 - 86	สูงกว่าปกติเล็กน้อย
74 - 81	ปกติ
69 - 73	ต่ำกว่าปกติเล็กน้อย
63 - 68	ต่ำ
56 - 62	ต่ำมาก
22 - 55	ต่ำที่สุด



เกณฑ์การเทียบค่าคะแนน

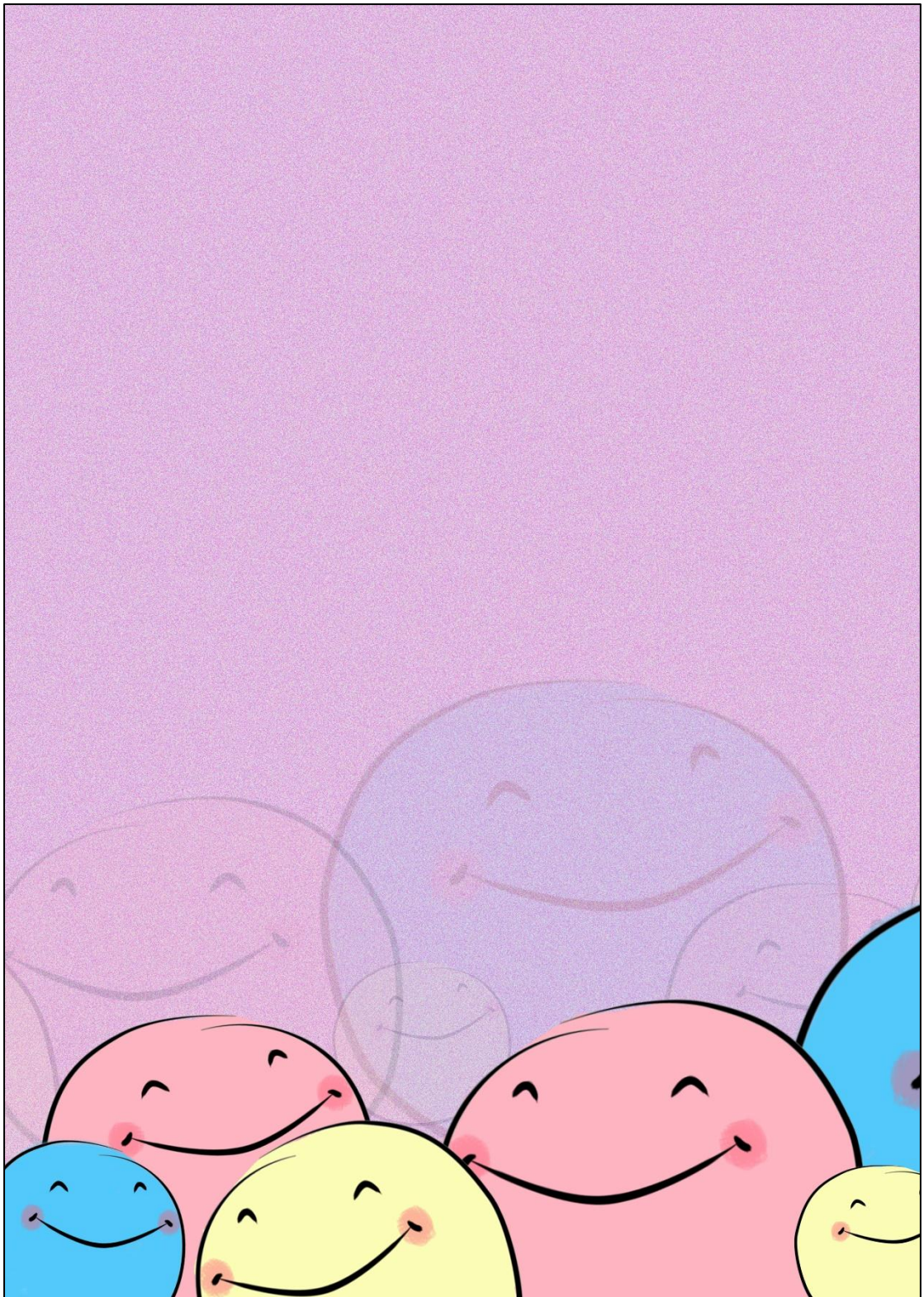
เกณฑ์การเทียบช่วงคะแนนดิบของมาตรวัดย่อยอารมณ์ความรู้สึกเชิงลบ

ช่วงคะแนนดิบ	ระดับอารมณ์ความรู้สึกเชิงลบ
125 - 130	ต่ำที่สุด
120 - 124	ต่ำมาก
114 - 119	ต่ำ
106 - 113	ต่ำกว่าปกติเล็กน้อย
98 - 105	ปกติ
89 - 97	สูงกว่าปกติเล็กน้อย
79 - 88	สูง
70 - 78	สูงมาก
26 - 69	สูงที่สุด

เกณฑ์การเทียบช่วงคะแนนดิบของมาตรวัดความสุขโดยรวม

ช่วงคะแนนดิบ	ระดับความสุข
268 - 305	มีความสุขอย่างยิ่ง
254 - 267	มีความสุขมาก
244 - 253	มีความสุข
232 - 243	มีความสุขสูงกว่าปกติเล็กน้อย
220 - 231	มีความสุขตามปกติ
206 - 219	มีความสุขต่ำกว่าปกติเล็กน้อย
190 - 205	มีความสุขน้อย
173 - 189	มีความสุขน้อยมาก
61 - 172	ไม่มีความสุข





ภาคผนวก ค
ผลการสร้างคลังข้อคำถาม สำหรับมาตรวัดความสุขของคนไทย

1. ค่าเฉลี่ยความสอดคล้องของข้อคำถามกับนิยามเชิงปฏิบัติการ
2. ค่า Item-Total Correlations รายข้อ และข้อคิดเห็นที่ได้จากการทดลองใช้
3. ค่า Item-Total Correlations รายข้อ และความถี่ของการเลือกรายการคำตอบ
4. ค่าความสอดคล้องของข้อมูลเชิงประจักษ์ที่ได้จากการตอบข้อคำถามกับ Graded Response Model ค่าพารามิเตอร์ความชัน และค่าพารามิเตอร์เทรชโฮลด์

ตารางที่ ค-1 ค่าเฉลี่ยความสอดคล้องของข้อความกับนิยามเชิงปฏิบัติการ

	มาตรวัดย่อย/ข้อความ	ค่าเฉลี่ย
ความพึงพอใจในชีวิต		
1	ทุกอย่างที่เกิดขึ้นในชีวิตของฉันในเวลาที่ผ่านมาเป็นสิ่งที่เหมาะสมแล้ว	4.40
2	ฉันไม่อยากจะเปลี่ยนแปลงสิ่งใดในชีวิต	4.40
3	หากย้อนเวลาได้ ฉันอยากจะแก้ไขความผิดพลาดที่เกิดขึ้นในชีวิต (-)	4.80
4	ฉันอยากกลับไปเริ่มต้นชีวิตใหม่อีกครั้ง (-)	4.80
5	ถ้าฉันย้อนกลับไปปรับเปลี่ยนตัวเองได้ ชีวิตฉันจะดีกว่านี้ (-)	4.80
6	ฉันพึงพอใจในสภาพชีวิตที่เป็นอยู่	4.80
7	ฉันมีสิ่งอำนวยความสะดวกในชีวิตตามที่หวัง	4.60
8	ฉันดำเนินชีวิตได้อย่างที่ใจต้องการ	4.60
9	ฉันได้รับความรักจากคนรอบข้างอยู่เสมอ	4.20
10	ฉันมีชีวิตที่มีความสุขสบาย	4.40
11	ฉันรู้สึกวุ่นวายเวลาที่ผ่านไปในแต่ละวันช่างยาวนานเสียเหลือเกิน (-)	4.60
12	ฉันเบื่อหน่ายกับชีวิตที่เป็นอยู่ (-)	4.60
13	ฉันทำงานสำเร็จทุกช่วงชีวิตที่ผ่านมาอย่างน่าชื่นชม	4.40
14	ฉันมีบทเรียนของชีวิตในอดีตที่มีคุณค่า	4.00
15	ที่ผ่านมา พ่อแม่หรือคนรอบข้างคอยช่วยเหลือฉันเสมอ	4.60
16	ฉันรู้สึกภาคภูมิใจในชีวิตที่ผ่านมาของฉัน	4.80
17	ฉันพยายามลืมความผิดพลาดในอดีต (-)	4.20
18	ชีวิตฉันที่ผ่านมาไม่มีเหตุการณ์ที่เลวร้าย (-)	4.60
19	ฉันมั่นใจว่าในอนาคตแม้จะมีอุปสรรคในชีวิตฉันก็สามารถแก้ปัญหาเหล่านั้นได้	4.80
20	ฉันเชื่อมั่นว่าจะมีชีวิตที่ดีต่อไป	4.80
21	ฉันมีความพร้อมสำหรับการใช้ชีวิตในอนาคตให้ดีกว่านี้	4.00
22	ฉันมั่นใจว่าจะมีคนคอยสนับสนุนฉันอยู่เสมอ	4.80
23	ฉันกังวลว่าจะไม่มีคนดูแลหรือให้ความช่วยเหลือในอนาคต (-)	4.60
24	ฉันไม่แน่ใจว่าจะใช้ชีวิตต่อไปได้อย่างมีความสุข (-)	4.60
25	คนรอบข้างมักจะชื่นชมการใช้เวลาในการทำงานอย่างคุ้มค่าของฉัน	4.40
26	ชีวิตของฉันเป็นที่น่าพอใจของคนรอบข้าง	4.20
27	ฉันประสบความสำเร็จในชีวิตอย่างที่ครอบครัวคาดหวัง	4.60
28	เพื่อน ๆ มักจะบอกว่าฉันเป็นคนที่มีความสุข	4.60
29	ผู้ที่ฉันนับถือมักเป็นห่วงว่าชีวิตของฉันจะล้มเหลว (-)	4.20
30	ครอบครัวฉันอยากให้ฉันมีชีวิตที่ดีกว่าเดิม (-)	4.60

ตารางที่ ค-1 (ต่อ)

	มาตรวัดย่อย/ข้อความ	ค่าเฉลี่ย
อารมณ์ความรู้สึกเชิงบวก		
1	กระฉับกระเฉง	4.60
2	กระตือรือร้น	4.60
3	กระปรี้กระเปร่า	4.60
4	กล้าหาญ	4.60
5	เข้มแข็ง	4.60
6	เข้าอกเข้าใจ	4.60
7	ครี้มอกครี้มใจ	4.40
8	คล่องแคล่ว	4.60
9	เคารพนับถือ	4.00
10	จงรักภักดี	3.80
11	ใจกล้า	4.40
12	ใจดีใจจ่อ	4.00
13	ชื่นชม	4.60
14	ซาบซึ้งใจ	4.60
15	ดีใจ	4.60
16	เด็ดเดี่ยว	4.60
17	ตั้งอกตั้งใจ	4.60
18	ตื่นตัว	4.20
19	ตื่นตาดำใจ	4.40
20	ตื่นเต้น	4.60
21	ถูกใจ	4.60
22	ทะนงตัว	4.20
23	ท้าทาย	4.40
24	เบิกบานใจ	4.60
25	ประทับใจ	4.60
26	ประหลาดใจ	4.40
27	ปลาบปลื้ม	4.60
28	เป็นมิตร	4.60
29	ผ่อนคลาย	4.60
30	ผาสุก	4.60
31	ใฝ่ฝัน/ปรารถนา	4.60

ตารางที่ ค-1 (ต่อ)

	มาตรวัดย่อย/ข้อความ	ค่าเฉลี่ย
32	พอใจ	4.60
33	เพติดเพลิน	4.60
34	ภาคภูมิใจ	4.80
35	มองโลกในแง่ดี	4.80
36	มั่นคง	4.80
37	มั่นใจ	4.80
38	มีความหวัง	4.80
39	มีแรงบันดาลใจ	4.80
40	มีเสน่ห์	4.40
41	มีอารมณ์ขัน	4.60
42	ยินดี	4.80
43	รัก	4.80
44	ร่าเริง	4.80
45	รื่นเริงใจ	4.80
46	เลื่อมใสศรัทธา	4.80
47	โล่งใจ	4.40
48	ไว้วางใจ	4.80
49	สงบ	4.80
50	สงสาร	4.40
51	สดชื่น	4.80
52	สนใจ	4.40
53	สนิทสนม	4.60
54	สนุกสนาน	4.80
55	สบายใจ	4.80
56	สุขใจ	4.80
57	หลงใหล	3.80
58	ห่วงใย	4.40
59	เห็นอกเห็นใจ	4.80
60	อบอุ่นใจ	4.80
61	อยากรู้ อยากเห็น	4.20
62	อ่อนน้อม	4.80
63	อ่อนโยน	4.80

ตารางที่ ค-1 (ต่อ)

	มาตรวัดย่อย/ข้อคำถาม	ค่าเฉลี่ย
64	อัศจรรย์ใจ	4.40
65	อารมณ์ดี	4.80
66	อิมเมจใจ	4.60
67	เอื้ออาทร	4.60
อารมณ์ความรู้สึกเชิงลบ		
1	กระวนกระวายใจ	4.60
2	กระสับกระส่าย	4.60
3	กลัดกลุ้ม	4.60
4	กลัว	4.60
5	กังวล	4.80
6	เกลียดชัง	4.60
7	โกรธ	4.80
8	ขมขื่น	4.60
9	ขยะแขยง	4.60
10	ขวยเขิน	4.00
11	ขวัญเสีย	4.40
12	ขัดใจ	4.60
13	ขายหน้า	4.60
14	ชี้เกี้ยว	4.60
15	ชี้ขลาด	4.60
16	ขุ่นเคืองใจ	4.60
17	กลางแคลงใจ	4.60
18	คับข้องใจ	4.80
19	แค้นใจ	4.60
20	ใจคอห่อเหี่ยว	4.60
21	ฉุนเฉียว	4.60
22	เซื่องซึม	4.60
23	เดือดดาล	4.60
24	โศกเศร้า	4.80
25	ตกใจ	4.60
26	ต่อต้าน	4.60
27	ตะขิดตะขวงใจ	4.40

ตารางที่ ค-1 (ต่อ)

	มาตรวัดย่อย/ข้อคำถาม	ค่าเฉลี่ย
28	ตึงเครียด	4.60
29	ตื่นตระหนก	4.60
30	ถูกดูถูกเหยียดหยาม	4.60
31	ถูกตำหนิ	4.60
32	ถูกปฏิเสธ	4.60
33	ท้อแท้	4.80
34	ทุกข์ทรมาน	4.80
35	น้อยใจ	4.80
36	น่าขนลุก	4.40
37	น่ารังเกียจ	4.60
38	บ้าคลั่ง	4.40
39	เบื่อหน่าย	4.80
40	ประหม่า	4.60
41	ผิดหวัง	4.80
42	พลุ่งพล่าน	4.40
43	พ่ายแพ้	4.60
44	โมโห	4.80
45	ไม่ชอบ	4.60
46	ไม่พอใจ	4.80
47	ไม่มั่นคง	4.60
48	ไม่มีความสุข	4.80
49	ไม่ยอมรับ	4.80
50	ไม่สบายอารมณ์	4.40
51	ไม่สบายใจ	4.80
52	รบกวนใจ	4.60
53	ร้อนรน	4.60
54	ร้อนกร้อนใจ	4.60
55	ระทึกใจ	4.40
56	ระอาใจ	4.40
57	รังเกียจ	4.60
58	รำคาญ	4.60
59	รู้สึกผิด	4.60

ตารางที่ ค-1 (ต่อ)

	มาตรวัดย่อย/ข้อคำถาม	ค่าเฉลี่ย
60	รู้สึกแย่	4.60
61	ไร้ชีวิตชีวา	4.60
62	ไร้เรี่ยวแรง	4.40
63	ไร้อารมณ์	4.60
64	ละอายใจ	4.60
65	ลำบากใจ	4.60
66	เศร้าใจ	4.80
67	สยดสยอง	4.40
68	สลดใจ	4.60
69	สะเทือนใจ	4.80
70	สะอิดสะเอียน	4.40
71	สิ้นหวัง	4.80
72	เสียใจ	4.80
73	เสียตายน	4.60
74	แสบใจ	4.40
75	หงุดหงิด	4.80
76	หดหู่	4.80
77	หนักใจ	4.60
78	หม่นหมอง	4.80
79	หมั่นไส้	4.40
80	หัวน้ใจ	4.40
81	หวาดผวา	4.40
82	หวาดเสียว	4.40
83	หึ่งหวง	4.40
84	เหน็ดเหนื่อย	4.80
85	เห็นห่าง	4.80
86	อัดอั้นตันใจ	4.40
87	อับอาย	4.60
88	อับยศอดสู	4.40
89	อาฆาตพยาบาท	4.40
90	อ้างว้าง	4.80
91	อารมณ์เปลี่ยนแปลงง่าย	4.40

ตารางที่ ค-1 (ต่อ)

มาตรวัดย่อย/ข้อคำถาม		ค่าเฉลี่ย
92	อารมณ์เสีย	4.80
93	อิจฉา	4.60
94	อีต้อตใจ	4.60

หมายเหตุ (-) หมายถึง ข้อคำถามทางลบ

ตารางที่ ค-2 ค่า Item-Total Correlations รายข้อ จากการทดลองใช้ (30 คน)

มาตรวัดย่อย/ข้อคำถาม		ค่า Item-Total Correlations
ความพึงพอใจในชีวิต		
1	ส่วนใหญ่สิ่งที่เกิดขึ้นในชีวิตของฉันในเวลาที่ผ่านมาเป็นสิ่งที่เหมาะสมแล้ว	.203
2	ฉันไม่ยากเปลี่ยนแปลงสิ่งใดในชีวิต	.387
3	หากย้อนเวลาได้ ฉันอยากจะแก้ไขความผิดพลาดที่เกิดขึ้นในชีวิต (-)	.095
4	ฉันอยากกลับไปเริ่มต้นชีวิตใหม่อีกครั้ง (-)	.195
5	ถ้าฉันย้อนกลับไปปรับเปลี่ยนตัวเองได้ ชีวิตฉันจะดีกว่านี้ (-)	.217
6	ฉันพึงพอใจในสภาพชีวิตที่เป็นอยู่	.255
7	ฉันมีสิ่งอำนวยความสะดวกในชีวิตตามที่คาดหวัง	.103
8	ฉันดำเนินชีวิตได้อย่างที่ต้องการ	.475
9	ฉันได้รับความรักความปรารถนาดีจากคนรอบข้างอยู่เสมอ	.392
10	ฉันมีชีวิตที่มีความสุขสบายตามสมควร	.371
11	ฉันรู้สึกว่ามันผ่านไปในแต่ละวันช่างยาวนานเสียเหลือเกิน (-)	.306
12	ฉันเบื่อหน่ายกับชีวิตที่เป็นอยู่ (-)	.297
13	ฉันทำงานสำเร็จทุกช่วงชีวิตที่ผ่านมาอย่างน่าพอใจ	.301
14	ฉันมีบทเรียนของชีวิตในอดีตที่มีคุณค่า	.240
15	ที่ผ่านมา พ่อแม่หรือคนรอบข้างคอยช่วยเหลือให้กำลังใจฉัน	.414
16	ฉันรู้สึกภาคภูมิใจในชีวิตที่ผ่านมาของฉัน	.476
17	ฉันพยายามลืมความผิดหวังในอดีต (-)	.123
18	ชีวิตฉันที่ผ่านมาแม้แต่เหตุการณ์ที่เลวร้าย (-)	.319
19	ฉันมั่นใจว่าในอนาคตแม้จะมีอุปสรรคในชีวิต ฉันก็สามารถแก้ปัญหาเหล่านั้นได้	.127
20	ฉันเชื่อมั่นว่าจะมีชีวิตที่ดีต่อไป	.222
21	ฉันมีความพร้อมสำหรับการใช้ชีวิตที่ดีในอนาคตข้างหน้า	.429
22	ฉันมั่นใจว่าจะมีคนคอยสนับสนุนฉันอยู่เสมอ	.373
23	ฉันกังวลว่าจะไม่มีคนดูแลหรือให้ความช่วยเหลือในอนาคต (-)	.137
24	ฉันไม่แน่ใจว่าจะใช้ชีวิตต่อไปได้อย่างมีความสุข (-)	.260
25	คนรอบข้างมักจะชื่นชมการใช้เวลาในการทำงานของฉัน	.412
26	ใคร ๆ มักจะบอกว่าอิจฉาฉัน	.425
27	ฉันประสบความสำเร็จในชีวิตอย่างที่ครอบครัวคาดหวัง	.245
28	เพื่อน ๆ มักจะบอกว่าฉันเป็นคนที่มีความสุข	.358
29	ผู้ที่ฉันนับถือเห็นว่าชีวิตของฉันล้มเหลว (-)	.201
30	ครอบครัวฉันอยากให้ฉันมีชีวิตที่ดีกว่าเดิม	-.014
อารมณ์ความรู้สึกเชิงบวก		
1	กระฉับกระเฉง	.638
2	กระตือรือร้น	.718
3	กระปรี้กระเปร่า	.810
4	กล้าหาญ	.572
5	เข้มแข็ง	.387

ตารางที่ ค-2 (ต่อ)

	มาตรวัดย่อย/ข้อความ	ค่า Item-Total Correlations
6	เข้าอกเข้าใจ	.375
7	ครีมนอกครีมนใจ	.579
8	คล่องแคล่ว	.580
9	ให้เกียรติ	.405
10	ใจกล้า	.613
11	มุ่งมั่น	.644
12	ชื่นชม	.669
13	ซาบซึ้งใจ	.751
14	ดีใจ	.861
15	เด็ดเดี่ยว	.559
16	ตั้งอกตั้งใจ	.654
17	มีสติ	.558
18	ตื่นตาตื่นใจ	.810
19	ตื่นเต้น	.688
20	ถูกใจ	.810
21	ทะนงตัว	.417
22	ทำท่าย	.450
23	เบิกบานใจ	.874
24	ประทับใจ	.830
25	ประหลาดใจ	.556
26	ปลาบปลื้ม	.832
27	เป็นมิตร	.817
28	ผ่อนคลาย	.744
29	ผาสุก	.774
30	ไฟฝัน/ปรารถนา	.510
31	พอใจ	.656
32	เพลิดเพลิน	.619
33	ภาคภูมิใจ	.806
34	มองโลกในแง่ดี	.668
35	มั่นคง	.818
36	มั่นใจ	.769
37	มีความหวัง	.830
38	มีแรงบันดาลใจ	.775
39	มีเสน่ห์	.757
40	มีอารมณ์ขัน	.846
41	ยินดี	.620
42	รัก	.700

ตารางที่ ค-2 (ต่อ)

มาตรวัดย่อย/ข้อความ		ค่า Item-Total Correlations
43	รำเริง	.881
44	รึนเริงใจ	.855
45	เลื้อมใส่ศรัทธา	.662
46	โล่งใจ	.710
47	ไว้วางใจ	.735
48	สงบ	.534
49	เมตตา	.727
50	สดชื่น	.771
51	สนใจ	.720
52	สนิทสนม	.808
53	สนุกสนาน	.871
54	สบายใจ	.882
55	สุขใจ	.820
56	กรุณา	.806
57	เห็นอกเห็นใจ	.641
58	อบอุ่นใจ	.784
59	อ่อนน้อม	.747
60	อ่อนโยน	.765
61	อัศจรรย์ใจ	.400
62	อารมณ์ดี	.830
63	อิมเอมใจ	.890
64	เอื้ออาทร	.601
อารมณ์ความรู้สึกเชิงลบ		
1	กระวนกระวายใจ	.621
2	กระสับกระส่าย	.664
3	กลัดกลุ้ม	.645
4	กลัว	.595
5	กังวล	.643
6	เกลียดชัง	.687
7	โกรธ	.678
8	ขมขื่น	.725
9	ขยะแฉียง	.697
10	ขวยเขิน	.531
11	ขวัญเสีย	.762
12	ขัดใจ	.691
13	ขายหน้า	.661
14	ขี้เกียจ	.640

ตารางที่ ค-2 (ต่อ)

	มาตรวัดย่อย/ข้อความ	ค่า Item-Total Correlations
15	ซีซลาด	.626
16	ชุ่นเคื่องใจ	.801
17	กลางแกลงใจ	.509
18	คับข้องใจ	.591
19	แค้นใจ	.501
20	ใจคอท้อเหี่ยว	.710
21	ฉุนเฉียว	.701
22	เซื่องซึม	.682
23	เดือดดาล	.618
24	โอดเคี้ยว	.622
25	ตกใจ	.590
26	ต่อต้าน	.603
27	ตะขิดตะขวงใจ	.733
28	ตึงเครียด	.551
29	ตื่นตระหนก	.731
30	ถูกดูถูกเหยียดหยาม	.553
31	ถูกตำหนิ	.741
32	ถูกปฏิเสธ	.813
33	ท้อแท้	.661
34	ทุกข์ทรมาน	.806
35	น้อยใจ	.647
36	น่าขนลุก	.794
37	น่ารังเกียจ	.748
38	บ้าคลั่ง	.671
39	เปื้อนอาย	.747
40	ประหม่า	.655
41	ผิดหวัง	.616
42	พลุ่งพล่าน	.409
43	พ่ายแพ้	.615
44	โมโห	.701
45	ไม่ชอบ	.736
46	ไม่พอใจ	.711
47	ไม่มั่นคง	.763
48	ไม่มีความสุข	.610
49	ไม่ยอมรับ	.814
50	ไม่สบายใจ	.751
51	ไม่สบายใจ	.746

ตารางที่ ค-2 (ต่อ)

	มาตรวัดย่อย/ข้อความ	ค่า Item-Total Correlations
52	รบกวนใจ	.580
53	ร้อนรน	.746
54	ร้อนออกร้อนใจ	.745
55	ระทึกใจ	.426
56	ระอาใจ	.716
57	รังเกียจ	.724
58	รำคาญ	.820
59	รู้สึกผิด	.804
60	รู้สึกแย้	.780
61	ไร้ชีวิตชีวา	.697
62	ไร้เรี่ยวแรง	.665
63	ไร้อารมณ์	.813
64	ละอายใจ	.755
65	ลำบากใจ	.631
66	เศร้าใจ	.652
67	สยดสยอง	.671
68	สลดใจ	.786
69	สะเทือนใจ	.692
70	สะอิดสะเอียน	.802
71	สิ้นหวัง	.856
72	เสียใจ	.811
73	เสียตาย	.809
74	แสบใจ	.815
75	หงุดหงิด	.735
76	หดหู่	.783
77	หนักใจ	.684
78	หม่นหมอง	.805
79	หม่นไส้	.730
80	หวั่นใจ	.795
81	หวาดผวา	.792
82	หวาดเสียว	.830
83	หึ่งหวง	.628
84	เหน็ดเหนื่อย	.460
85	เหินห่าง	.786
86	อัดอั้นตันใจ	.530
87	อับอาย	.844
88	อับยศอดสู	.822

ตารางที่ ค-2 (ต่อ)

มาตรวัดย่อย/ข้อความ		ค่า Item-Total Correlations
89	อาฆาตพยาบาท	.507
90	อ้างว้าง	.752
91	อารมณ์ไม่มั่นคง	.579
92	อารมณ์เสีย	.673
93	อิจฉา	.701
94	อึดอัดใจ	.781

หมายเหตุ (-) หมายถึง ข้อคำถามทางลบ

สรุปข้อคิดเห็นจากการทดลองใช้

มาตรวัดย่อยความพึงพอใจในชีวิต

1. การให้ความคิดเห็นในรูปแบบของร้อยละ ถ้าเป็นชาวบ้านอาจไม่เข้าใจ น่าจะใช้ตัวเลข 1 2 3 4 5 และบอกระดับความคิดเห็น
2. รายการตัวเลือกในช่องระดับความคิดเห็น ส่วนของความพึงพอใจในชีวิตมีขนาดเล็กเกินไป
3. ข้อที่ 1 ควรเปลี่ยนเป็น “ส่วนใหญ่สิ่งที่เกิดขึ้นกับชีวิตของฉันที่ผ่านมาเป็นสิ่งที่เหมาะสมแล้ว”
4. ข้อที่ 2 ควรเปลี่ยนเป็น “หากย้อนเวลาได้ ฉันอยากแก้ไขความผิดพลาดที่เกิดขึ้นในชีวิต” (ตัดคำว่า “จะ” ออก)
5. ข้อที่ 26 และ 28 ควรตัด “ ๆ” ออก
6. ข้อ 30 คำว่า “ชีวิตที่ดี” หมายถึงอะไรได้หลายอย่าง เช่น มีงาน มีเงิน มีครอบครัว มีบ้าน มีรถ

มาตรวัดย่อยอารมณ์ความรู้สึกเชิงบวกและมาตรวัดย่อยอารมณ์ความรู้สึกเชิงลบ

1. จำนวนคำแสดงอารมณ์เชิงลบและอารมณ์เชิงบวก ควรมีจำนวนที่สมดุลกัน และควรมีการ rating คำก่อนนำมาใส่ น่าจะให้ผลดีกว่า
2. ข้อคำถามที่เกี่ยวข้องกับอารมณ์ความรู้สึก บางข้อเข้าใจยาก
3. คำว่า “กลัดกลุ้ม” ควรใช้คำว่า “กลุ้มใจ” เพื่อให้เข้าใจง่าย
4. คำว่า “ขวยเขิน” ควรใช้คำว่า “เขิน” เพื่อให้เข้าใจง่าย
5. ไม่เข้าใจคำว่า “เดือดดาล”
6. คำว่า “ถูกดูถูกเหยียดหยาม” หมายถึง ผู้ตอบถูกกระทำ หรือ ผู้ตอบเป็นผู้กระทำ คำอื่น ๆ ต้องการสื่อการถูกกระทำหรือเป็นผู้กระทำ
7. คำว่า “น่ารังเกียจ” (ข้อ 59) และ คำว่า “รังเกียจ” (ข้อ 99) ความหมายเดียวกันหรือไม่
8. คำว่า “ระอาใจ” ควรเปลี่ยนเป็นคำว่า “เอือมระอา” เพื่อให้เข้าใจง่าย
9. ไม่เข้าใจความหมายของคำว่า “ครี้มอกครี้มใจ”
10. ไม่เข้าใจความหมายของคำว่า “น่าขมลูก”

อื่น ๆ

1. คำที่ต้องการเน้น ควรใช้ตัวหนา น่าจะทำให้เห็นชัดเจนกว่าการขีดเส้นใต้
2. ควรมีการคั่นด้วยสีระหว่างแถวเพื่อให้เห็นชัดเจน จะได้ไม่พลาดการตอบคำถาม เช่น แถวคี่เป็นสีทึบ แถวคู่เป็นโปร่งใส
3. จำนวนข้อคำถามมากเกินไป บางที่ทำให้สับสน เพราะบางข้อมีความคล้ายคลึงกัน
4. คำถามอ่านเข้าใจง่าย ตรงไปตรงมาดี

ตารางที่ ค-3 ค่า Item-Total Correlations รายข้อ และความถี่ของการเลือกรายการคำตอบ
ในแต่ละข้อคำถามจากกลุ่มตัวอย่าง (934 คน)

มาตรวัดย่อย/ข้อคำถาม	ความถี่ของ การเลือกรายการคำตอบ (คน)					ค่า Item- Total Correlations
	1	2	3	4	5	
ความพึงพอใจในชีวิต						
1. ส่วนใหญ่สิ่งที่เกิดขึ้นในชีวิตของฉันในเวลาที่ผ่านมาเป็นสิ่งที่เหมาะสมแล้ว	36	133	450	273	42	0.51
2. ฉันไม่อยากจะเปลี่ยนแปลงสิ่งใดในชีวิต	210	263	300	133	28	0.25
3. หากย้อนเวลาได้ ฉันอยากแก้ไขความผิดพลาดที่เกิดขึ้นในชีวิต (-)	66	154	221	258	235	0.36
4. ฉันอยากกลับไปเริ่มต้นชีวิตใหม่อีกครั้ง (-)	174	234	261	142	123	0.51
5. ถ้าฉันย้อนกลับไปปรับเปลี่ยนตัวเองได้ชีวิตฉันจะดีกว่านี้ (-)	89	200	253	216	176	0.43
6. ฉันพึงพอใจในสภาพชีวิตที่เป็นอยู่	35	82	364	334	119	0.59
7. ฉันมีสิ่งอำนวยความสะดวกในชีวิตตามที่คาดหวัง	29	105	387	334	79	0.58
8. ฉันดำเนินชีวิตได้อย่างที่ต้องการ	31	125	398	309	71	0.58
9. ฉันได้รับความรักความปรารถนาดีจากคนรอบข้างอยู่เสมอ	5	51	300	425	153	0.50
10. ฉันมีชีวิตที่มีความสุขสบายตามสมควร	10	70	308	439	107	0.64
11. ฉันรู้สึกว่าวันเวลาที่ผ่านไปในแต่ละวันช่างยาวนานเสียเหลือเกิน (-)	141	376	270	114	33	0.38
12. ฉันเบื่อหน่ายกับชีวิตที่เป็นอยู่ (-)	273	353	193	87	28	0.50
13. ฉันทำงานสำเร็จทุกช่วงชีวิตที่ผ่านมาอย่างน่าพอใจ	29	129	512	225	39	0.48
14. ฉันมีบทเรียนของชีวิตในอดีตที่มีคุณค่า	7	74	295	370	188	0.20
15. ที่ผ่านมา พ่อแม่หรือคนรอบข้างคอยช่วยเหลือให้กำลังใจฉัน	13	40	173	339	369	0.38
16. ฉันรู้สึกภาคภูมิใจในชีวิตที่ผ่านมาของฉัน	12	51	310	386	175	0.61
17. ฉันพยายามลืมความผิดหวังในอดีต (-)	88	213	300	217	116	0.22
18. ชีวิตฉันที่ผ่านมามีแต่เหตุการณ์ที่เลวร้าย (-)	293	330	207	73	31	0.48
19. ฉันมั่นใจว่าในอนาคตแม้จะมีอุปสรรคในชีวิต ฉันก็สามารถแก้ปัญหาเหล่านั้นได้	6	43	234	411	240	0.39
20. ฉันเชื่อมั่นว่าจะมีชีวิตที่ดีต่อไป	3	30	137	422	342	0.46
21. ฉันมีความพร้อมสำหรับการใช้ชีวิตที่ดีในอนาคตข้างหน้า	8	42	211	405	268	0.51

ตารางที่ ค-3 (ต่อ)

มาตรวัดย่อย/ข้อความถาม	ความถี่ของ การเลือกรายการคำตอบ (คน)					ค่า Item- Total Correlations
	1	2	3	4	5	
	22. ฉันมั่นใจว่าจะมีคนคอยสนับสนุนฉันอยู่เสมอ	8	69	310	351	
23. ฉันกังวลว่าจะไม่มีคนดูแลหรือให้ความช่วยเหลือในอนาคต (-)	177	355	309	67	26	0.43
24. ฉันไม่แน่ใจว่าจะใช้ชีวิตต่อไปได้อย่างมีความสุข (-)	275	359	223	66	11	0.51
25. คนรอบข้างมักจะชื่นชมการใช้เวลาในการทำงานของฉัน	23	167	498	205	41	0.35
26. ใคร ๆ มักจะบอกว่าอัจฉาฉัน	158	354	302	105	15	0.14
27. ฉันประสบความสำเร็จในชีวิตอย่างที่ครอบครัวคาดหวัง	29	124	397	291	93	0.61
28. เพื่อนมักจะบอกว่าฉันเป็นคนที่มีความสุข	27	117	385	305	100	0.43
29. ผู้ที่ฉันนับถือเห็นว่าชีวิตของฉันล้มเหลว (-)	463	303	121	40	7	0.46
อารมณ์ความรู้สึกเชิงบวก						
1 กระฉับกระเฉง	16	70	385	342	121	0.08
2 กระตือรือร้น	9	69	352	362	142	0.13
3 กระปรี้กระเปร่า	11	84	393	342	104	0.09
4 กรุณา	132	329	297	124	52	0.47
5 กล้าหาญ	236	343	254	77	24	0.44
6 เข้มแข็ง	17	81	324	351	161	0.16
7 เข้าอกเข้าใจ	193	379	260	76	26	0.42
8 คล่องแคล่ว	10	73	400	332	119	0.10
9 ใจกล้า	138	315	294	130	57	0.47
10 ชื่นชม	126	303	298	147	60	0.48
11 ซาบซึ้งใจ	327	353	195	41	18	0.45
12 ดีใจ	136	363	320	87	28	0.49
13 เด็ดเดี่ยว	479	280	122	38	15	0.44
14 ตั้งอกตั้งใจ	470	294	125	27	18	0.44
15 ตื่นตาดิ้นใจ	397	337	167	22	11	0.50
16 ตื่นเต้น	136	350	323	99	26	0.56
17 ถูกใจ	374	382	143	27	8	0.51
18 ทะนงตัว	114	265	346	127	82	0.29

ตารางที่ ค-3 (ต่อ)

มาตรวัดย่อย/ข้อความ	ความถี่ของ การเลือกรายการคำตอบ (คน)					ค่า Item- Total
						Correlations
	1	2	3	4	5	
19 ทำทนาย	363	361	169	28	13	0.40
20 เบิกบานใจ	209	418	237	49	21	0.55
21 ประทับใจ	12	57	300	376	189	0.13
22 ประหลาดใจ	7	48	274	418	187	0.18
23 ปลาบปลื้ม	7	56	337	388	146	0.15
24 เป็นมิตร	171	361	300	81	21	0.49
25 ผ่อนคลาย	211	391	254	61	17	0.58
26 ผาสุก	410	315	149	39	21	0.54
27 ใฝ่ฝัน/ปรารถนา	23	80	417	305	109	0.22
28 พอใจ	303	378	184	48	21	0.50
29 เพื่อดูเพลิน	213	368	252	83	18	0.52
30 ภาคภูมิใจ	33	84	304	368	145	0.24
31 มองโลกในแง่ดี	19	80	343	356	136	0.24
32 มั่นคง	348	347	165	53	21	0.49
33 มั่นใจ	7	44	245	394	244	0.14
34 มีความหวัง	538	251	115	23	7	0.44
35 มีแรงบันดาลใจ	65	139	385	233	112	0.29
36 มีสติ	319	280	197	82	56	0.45
37 มีเสน่ห์	198	389	277	53	17	0.50
38 มีอารมณ์ขัน	230	338	271	73	22	0.50
39 มุ่งมั่น	259	371	232	56	16	0.58
40 เมตตา	18	64	273	410	169	0.21
41 ยินดี	155	327	288	120	44	0.52
42 รัก	258	388	221	54	13	0.52
43 ร่าเริง	74	226	370	195	69	0.34
44 รื่นเริงใจ	54	182	419	202	77	0.39
45 เลื่อมใสศรัทธา	14	74	344	356	146	0.18
46 โโล่งใจ	207	396	252	51	28	0.51
47 ใ่ว่างใจ	219	416	228	51	20	0.52
48 สงบ	268	346	209	80	31	0.53
49 สดชื่น	332	290	247	48	17	0.32

ตารางที่ ค-3 (ต่อ)

มาตรวัดย่อย/ข้อคำถาม	ความถี่ของ การเลือกรายการคำตอบ (คน)					ค่า Item- Total Correlations
	1	2	3	4	5	
	50 สนใจ	126	228	367	159	54
51 สนุกสนม	486	276	114	39	19	0.50
52 สนุกสนาน	169	365	253	103	44	0.52
53 สบายใจ	566	274	73	15	6	0.39
54 สุขใจ	582	200	103	33	16	0.46
55 เห็นอกเห็นใจ	12	92	325	346	159	0.14
56 ให้เกียรติ	176	348	264	90	56	0.51
57 อบอุนใจ	19	93	404	298	120	0.25
58 อ่อนน้อม	211	419	236	65	3	0.50
59 อ่อนโยน	151	362	339	69	13	0.46
60 อัจฉริยะใจ	26	143	403	276	86	0.24
61 อารมณ์ดี	8	52	231	362	281	0.16
62 อิ่มเอมใจ	9	81	346	340	158	0.09
63 เลือ้อาหาร	27	121	337	313	136	0.06
อารมณ์ความรู้สึกเชิงลบ						
1 กระวนกระวายใจ	24	59	195	423	233	0.52
2 กระสับกระส่าย	151	270	364	102	47	0.29
3 ก้าว	21	63	224	353	273	0.40
4 กลุ้มใจ	123	352	376	76	7	0.20
5 กังวล	12	31	147	372	372	0.48
6 เกลียดชัง	168	358	331	66	11	0.26
7 โกรธ	157	358	355	55	9	0.21
8 ขมขื่น	267	376	244	40	7	0.15
9 ขยะแขยง	174	332	352	67	9	0.12
10 ขวัญเสีย	169	354	332	73	6	0.14
11 ชัดใจ	199	379	295	52	9	0.21
12 ขายหน้า	202	366	293	63	10	0.21
13 ขี้เกียจ	220	413	259	37	5	0.11
14 ขี้ลาด	122	218	397	148	49	0.22
15 ขุ่นเคืองใจ	228	324	291	86	5	0.28
16 คลางแคลงใจ	196	367	309	58	4	0.25

ตารางที่ ค-3 (ต่อ)

มาตรวัดย่อย/ข้อความ	ความถี่ของ การเลือกรายการคำตอบ (คน)					ค่า Item- Total Correlations
	1	2	3	4	5	
	17	251	384	240	53	6
18	36	166	255	376	151	0.48
19	27	97	293	384	133	0.50
20	25	100	293	401	115	0.50
21	13	58	247	383	233	0.46
22	19	54	207	377	277	0.47
23	12	38	205	427	252	0.47
24	23	58	253	405	195	0.51
25	26	84	274	405	145	0.52
26	170	390	309	55	10	0.20
27	16	84	262	405	167	0.52
28	12	59	195	388	280	0.52
29	12	55	180	380	307	0.52
30	12	57	244	359	262	0.47
31	299	332	237	49	17	0.24
32	277	355	258	34	10	0.31
33	37	109	256	382	150	0.49
34	181	329	346	64	14	0.29
35	29	110	284	360	151	0.52
36	34	68	224	376	22	0.59
37	25	446	186	332	345	0.50
38	26	57	187	334	330	0.53
39	24	47	194	337	332	0.52
40	9	60	193	377	295	0.48
41	26	78	243	379	208	0.57
42	118	243	325	160	8	0.28
43	81	248	418	149	38	0.29
44	94	304	402	117	17	0.23
45	28	67	213	394	232	0.57
46	97	250	415	136	36	0.11
47	157	379	320	65	13	0.16

ตารางที่ ค-3 (ต่อ)

มาตรวัดย่อย/ข้อคำถาม	ความถี่ของ การเลือกรายการคำตอบ (คน)					ค่า Item- Total Correlations
	1	2	3	4	5	
	48 ไม่สบายใจ	137	337	363	80	17
49 รบกวนใจ	134	326	370	84	20	0.28
50 ร้อนรน	223	373	278	54	6	0.26
51 ร้อนออกร้อนใจ	192	378	295	62	7	0.21
52 ระทึกใจ	11	16	93	285	529	0.39
53 รำคาญ	9	34	113	354	424	0.47
54 รู้สึกผิด	10	36	156	375	357	0.49
55 รู้สึกแย้	7	12	99	291	525	0.47
56 ไร้ชีวิตชีวา	13	28	119	275	499	0.45
57 ไร้เรี่ยวแรง	219	375	254	68	18	0.08
58 ไร้อารมณ์	24	48	256	378	228	0.56
59 ละอายใจ	27	90	306	344	167	0.52
60 ลำบากใจ	5	21	132	351	425	0.50
61 เศร้าใจ	41	101	287	350	155	0.54
62 สยดสยอง	19	44	150	393	328	0.56
63 สลดใจ	34	71	274	386	169	0.55
64 สะเทือนใจ	20	41	157	404	312	0.54
65 สะอิดสะเอียน	22	62	202	360	288	0.51
66 สิ้นหวัง	16	47	212	399	260	0.58
67 เสียใจ	10	21	110	338	455	0.49
68 เสียตาย	6	24	144	367	393	0.48
69 แสลงใจ	45	70	174	296	349	0.46
70 หงุดหงิด	100	163	352	234	85	0.45
71 หดหู่	115	323	364	98	34	0.38
72 หนักใจ	25	73	234	370	232	0.48
73 หม่นหมอง	235	378	248	59	14	0.27
74 หมั่นไส้	210	365	289	61	9	0.24
75 หัวน้ใจ	234	348	275	67	10	0.31
76 หวาดผวา	210	360	301	53	10	0.26
77 หวาดเสียว	45	96	232	342	219	0.50
78 หึ่งหวง	11	31	131	326	435	0.43

ตารางที่ ค-3 (ต่อ)

มาตรวัดย่อย/ข้อความ	ความถี่ของ การเลือกรายการคำตอบ (คน)					ค่า Item- Total Correlations
	1	2	3	4	5	
	79	10	10	55	161	
80	40	87	342	285	180	0.29
81	12	25	55	194	648	0.37
82	30	35	156	307	406	0.49
83	286	357	241	43	7	0.23
84	46	108	254	329	197	0.47
85	27	77	294	366	170	0.51
86	10	19	141	348	416	0.41
87	126	266	361	132	49	0.20
88	30	53	233	378	240	0.57
89	27	50	188	352	317	0.50
90	187	341	297	74	35	0.23

หมายเหตุ

- ก. ตัวเลือกรายการคำตอบของมาตรวัดย่อยความพึงพอใจในชีวิต สำหรับข้อความทางบวก ระดับ 1= ไม่เห็นด้วย/เห็นด้วยน้อยที่สุด 2= เห็นด้วยน้อย 3=เห็นด้วยปานกลาง 4= เห็นด้วยมาก และ 5= เห็นด้วยอย่างยิ่ง สำหรับข้อความทางลบ ระดับ 5= ไม่เห็นด้วย/เห็นด้วยน้อยที่สุด 4= เห็นด้วยน้อย 3=เห็นด้วยปานกลาง 2= เห็นด้วยมาก และ 1= เห็นด้วยอย่างยิ่ง
- ข. ตัวเลือกรายการคำตอบของมาตรวัดย่อยอารมณ์ความรู้สึกเชิงบวก ระดับ 1= ไม่เคย/แทบจะไม่เคย 2 = น้อยครั้ง 3 = บางครั้ง 4 = ป่อยครั้ง และ 5= เป็นประจำ
- ค. ตัวเลือกรายการคำตอบของมาตรวัดย่อยอารมณ์ความรู้สึกเชิงลบ ระดับ 5= ไม่เคย/แทบจะไม่เคย 4 = น้อยครั้ง 3 = บางครั้ง 2 = ป่อยครั้ง และ 1= เป็นประจำ
- ง. ข้อคำถามจะต้องมีค่า Item-Total Correlations มากกว่า 0.30 จึงจะคัดเลือกเข้าสู่การวิเคราะห์ขั้นต่อไป

ตารางที่ ค-4 ค่าพารามิเตอร์ของข้อคำถามและค่าความสอดคล้องของข้อมูลเชิงประจักษ์ที่ได้จากการ
ตอบข้อคำถามกับ Graded Response Model

มาตรวัดย่อย/ข้อคำถาม	ค่าพารามิเตอร์ของข้อคำถาม					P - value
	α (SE)	β_1 (SE)	β_2 (SE)	β_3 (SE)	β_4 (SE)	
ความพึงพอใจในชีวิต						
1. ส่วนใหญ่สิ่งที่เกิดขึ้นใน ชีวิตของฉันในเวลาที่ผ่านมา ผ่านมาเป็นสิ่งที่ เหมาะสมแล้ว	0.7592 (0.0419)	-3.0537 (0.1297)	-1.5533 (0.0665)	0.6324 (0.0574)	2.9589 (0.1299)	0.0000
2. หากย้อนเวลาได้ ฉัน อยากแก้ไขความ ผิดพลาดที่เกิดขึ้นในชีวิต (-)	0.3678 (0.0213)	4.3831 (0.1828)	1.9955 (0.096)	0.1486 (0.0805)	-1.9137 (0.1064)	0.0044
3. ฉันอยากกลับไปเริ่มต้น ชีวิตใหม่อีกครั้ง (-)	0.5784 (0.0332)	1.7706 (0.0806)	0.2931 (0.0575)	-1.1540 (0.0593)	-2.2714 (0.0855)	0.0064
4. ถ้าฉันย้อนกลับไป ปรับเปลี่ยนตัวเองได้ ชีวิตฉันจะดีกว่านี้ (-)	0.4608 (0.0260)	3.1790 (0.1318)	1.1236 (0.0745)	-0.4985 (0.0672)	-2.0964 (0.0943)	0.4389
5. ฉันพึงพอใจในสภาพ ชีวิตที่เป็นอยู่	0.9930 (0.0564)	-2.6308 (0.1017)	-1.6769 (0.0594)	0.0272 (0.0442)	1.6477 (0.0668)	0.8550
6. ฉันมีสิ่งอำนวยความสะดวก ในชีวิตตามที่ คาดหวัง	1.0021 (0.0568)	-2.7519 (0.1153)	-1.5357 (0.0582)	0.1818 (0.0446)	2.0016 (0.0795)	0.0004
7. ฉันดำเนินชีวิตได้อย่างที่ ต้องการ	0.9696 (0.0544)	-2.7491 (0.1154)	-1.4157 (0.0569)	0.3147 (0.0460)	2.1288 (0.0851)	0.1410
8. ฉันได้รับความรักความ ปรารถนาดีจากคนรอบ ข้างอยู่เสมอ	0.8060 (0.0467)	-4.4390 (0.3238)	-2.5493 (0.1013)	-0.4889 (0.0533)	1.5795 (0.0713)	0.0007
9. ฉันมีชีวิตที่มีความสุข สบายตามสมควร	1.2509 (0.0756)	-3.0465 (0.1612)	-1.8004 (0.0627)	-0.2792 (0.0389)	1.5843 (0.0610)	0.1671
10. ฉันรู้สึกว่าคุณเวลาที่ผ่านไป ในแต่ละวันช่าง ยาวนานเสียเหลือเกิน (-)	0.4423 (0.0241)	2.5466 (0.1192)	-0.3232 (0.0787)	-2.4575 (0.1017)	-4.7265 (0.2161)	0.4940
11. ฉันเบื่อหน่ายกับชีวิตที่ เป็นอยู่ (-)	0.6189 (0.0241)	1.0299 (0.0696)	-0.8088 (0.0596)	-2.2111 (0.0808)	-3.7823 (0.1701)	0.0148
12. ฉันทำงานสำเร็จทุกช่วง ชีวิตที่ผ่านมาอย่างน่า พอใจ	0.7414 (0.0362)	-3.2761 (0.147)	-1.6063 (0.0699)	0.9331 (0.0615)	3.0711 (0.1354)	0.0001

ตารางที่ ค-4 (ต่อ)

มาตรวัดย่อย/ข้อคำถาม	ค่าพารามิเตอร์ของข้อคำถาม					p - value
	α (SE)	β_1 (SE)	β_2 (SE)	β_3 (SE)	β_4 (SE)	
13. ที่ผ่านมา พ่อแม่หรือคนรอบข้างคอยช่วยเหลือให้กำลังใจฉัน	0.6144 (0.0386)	-4.5804 (0.2438)	-3.1354 (0.1174)	-1.3170 (0.068)	0.5049 (0.0651)	0.7870
14. ฉันรู้สึกภาคภูมิใจในชีวิตที่ผ่านมาของฉัน	1.0960 (0.0651)	-3.1246 (0.1554)	-2.0540 (0.0724)	-0.3592 (0.0423)	1.2150 (0.0538)	0.1217
15. ชีวิตฉันที่ผ่านมามีแต่เหตุการณ์ที่เลวร้าย (-)	0.6134 (0.0365)	0.8920 (0.0678)	-0.8484 (0.0606)	-2.3757 (0.0841)	-3.6989 (0.157)	0.4781
16. ฉันมั่นใจว่าในอนาคตแม้จะมีอุปสรรคในชีวิต ฉันก็สามารถแก้ปัญหา นั้นได้	0.6471 (0.0381)	-5.0910 (0.3574)	-3.0813 (0.1255)	-0.9353 (0.0645)	1.1975 (0.0716)	0.0471
17. ฉันเชื่อมั่นว่าจะมีชีวิตที่ดีต่อไป	0.7766 (0.0490)	-4.9753 (0.4359)	-3.0672 (0.1282)	-1.4607 (0.0644)	0.5637 (0.0570)	0.0098
18. ฉันมีความพร้อมสำหรับการใช้ชีวิตที่ดีในอนาคตข้างหน้า	0.8669 (0.0521)	-3.9104 (0.2377)	-2.5097 (0.0949)	-0.8572 (0.0522)	0.8722 (0.0553)	0.0888
19. ฉันมั่นใจว่าจะมีคนคอยสนับสนุนฉันอยู่เสมอ	0.8332 (0.0484)	-4.0260 (0.2547)	-2.2194 (0.0854)	-0.3350 (0.0498)	1.2613 (0.0618)	0.0005
20. ฉันกังวลว่าจะไม่มีใครดูแลหรือให้ความช่วยเหลือในอนาคต (-)	0.5332 (0.0298)	1.8571 (0.0916)	-0.3810 (0.0695)	-2.7770 (0.1047)	-4.3432 (0.1965)	0.5593
21. ฉันไม่แน่ใจว่าจะใช้ชีวิตต่อไปได้อย่างมีความสุข (-)	0.6923 (0.0410)	0.9315 (0.0638)	-0.8414 (0.0582)	-2.5130 (0.0952)	-4.3132 (0.2498)	0.2550
22. คนรอบข้างมักจะชื่นชมการใช้เวลาในการทำงานของฉัน	0.5309 (0.0284)	-4.4867 (0.2284)	-1.7559 (0.0885)	1.3059 (0.0810)	3.8229 (0.1726)	0.0028
23. ฉันประสบความสำเร็จในชีวิตอย่างที่ครอบครัวคาดหวัง	1.0261 (0.0578)	-2.7082 (0.1140)	-1.3901 (0.0552)	0.3027 (0.0438)	1.8422 (0.0721)	0.1585
24. เพื่อนมักจะบอกว่าฉันเป็นคนที่มีความสุข	0.6450 (0.0351)	-3.7103 (0.1729)	-1.8855 (0.0792)	0.3186 (0.0606)	2.3479 (0.099)	0.0000
25. ผู้ที่ฉันนับถือเห็นว่าชีวิตของฉันล้มเหลว (-)	0.6721 (0.0447)	0.0107 (0.0589)	-1.6428 (0.0676)	-3.0400 (0.1155)	-4.8119 (0.3173)	0.9259

ตารางที่ ค-4 (ต่อ)

มาตรวัดย่อย/ข้อคำถาม	ค่าพารามิเตอร์ของข้อคำถาม					P - value
	α (SE)	β_1 (SE)	β_2 (SE)	β_3 (SE)	β_4 (SE)	
อารมณ์ความรู้สึกเชิงบวก						
1 กรุณา	0.7372 (0.0407)	-1.8332 (0.0740)	0.0159 (0.0522)	1.5004 (0.0612)	2.7544 (0.1072)	0.4401
2 กล้าหาญ	0.7858 (0.0455)	-1.0462 (0.0604)	0.5235 (0.0506)	2.0194 (0.073)	3.2939 (0.1461)	0.0028
3 เข้าอกเข้าใจ	0.7780 (0.0442)	-1.3283 (0.0664)	0.4642 (0.0517)	2.0174 (0.0731)	3.2679 (0.1427)	0.9767
4 ใจกล้า	0.9048 (0.0506)	-1.5740 (0.0667)	-0.0199 (0.0449)	1.2610 (0.0514)	2.3435 (0.0872)	0.0078
5 ชื่นชม	0.9314 (0.0519)	-1.6519 (0.0680)	-0.1095 (0.0441)	1.1302 (0.0491)	2.2536 (0.0838)	0.0707
6 ซาบซึ้งใจ	0.8085 (0.0494)	-0.5866 (0.0544)	0.9719 (0.0535)	2.4623 (0.0849)	3.4357 (0.1527)	0.4904
7 ดีใจ	0.8110 (0.0454)	-1.6893 (0.0733)	0.1544 (0.0503)	1.8313 (0.0682)	3.0799 (0.1310)	0.4186
8 ดีดเดี่ยว	0.9522 (0.0633)	0.0778 (0.0448)	1.2794 (0.0503)	2.3226 (0.0776)	3.2633 (0.1525)	0.2108
9 ต้งอกต้งใจ	0.8736 (0.0581)	0.0311 (0.0481)	1.3719 (0.055)	2.5611 (0.0823)	3.2575 (0.1325)	0.0270
10 ตื่นตาคือใจ	1.1122 (0.0722)	-0.2318 (0.0419)	1.0488 (0.0464)	2.4484 (0.0858)	3.1629 (0.1489)	0.4517
11 ตื่นเต้น	0.9631 (0.0545)	-1.5394 (0.0649)	0.1175 (0.0445)	1.6119 (0.0597)	2.8175 (0.1192)	0.0080
12 ถูกใจ	0.9990 (0.0638)	-0.3378 (0.0462)	1.2210 (0.0517)	2.5303 (0.0924)	3.5111 (0.1915)	0.0032
13 ทำทาย	0.8802 (0.0554)	-0.4161 (0.0504)	1.1202 (0.0535)	2.6321 (0.0927)	3.4876 (0.1634)	0.0441
14 เบิกบานใจ	1.1859 (0.0712)	-0.9805 (0.0487)	0.5987 (0.0406)	1.8998 (0.0625)	2.7024 (0.1107)	0.8420
15 เป็นมิตร	0.9084 (0.0519)	-1.3300 (0.0616)	0.3081 (0.0464)	1.8471 (0.0672)	3.0911 (0.1392)	0.0012
16 ผ่อนคลาย	1.2000 (0.0717)	-0.9541 (0.0478)	0.5194 (0.0396)	1.8049 (0.0611)	2.7892 (0.1233)	0.0006
17 ผาสุก	1.2637 (0.0818)	-0.1622 (0.0381)	0.9852 (0.0409)	1.9434 (0.0596)	2.5769 (0.0979)	0.2328
18 พอใจ	1.2213 (0.0752)	-0.5504 (0.0421)	0.8075 (0.0405)	1.8642 (0.0595)	2.6317 (0.1052)	0.3508

ตารางที่ ค-4 (ต่อ)

มาตรวัดย่อย/ข้อความ	ค่าพารามิเตอร์ของข้อความ					P - value
	α (SE)	β_1 (SE)	β_2 (SE)	β_3 (SE)	β_4 (SE)	
19 เพลิดเพลิน	0.9428 (0.0547)	-1.0517 (0.0553)	0.4882 (0.0450)	1.8111 (0.0652)	3.1282 (0.1468)	0.0365
20 มั่นคง	1.1558 (0.0719)	-0.3911 (0.0418)	0.8782 (0.0415)	1.8598 (0.0596)	2.6942 (0.1098)	0.9319
21 มีความหวัง	0.9781 (0.0677)	0.2843 (0.0440)	1.4294 (0.0538)	2.6905 (0.1009)	3.6684 (0.2113)	0.0005
22 มีสติ	0.8365 (0.0509)	-0.5874 (0.0511)	0.6107 (0.0452)	1.6656 (0.0552)	2.4849 (0.0871)	0.0000
23 มีเสน่ห์	0.8881 (0.0516)	-1.1880 (0.0598)	0.5096 (0.0484)	2.1702 (0.0771)	3.2611 (0.1497)	0.0797
24 มีอารมณ์ขัน	0.8275 (0.0480)	-1.0454 (0.0585)	0.4447 (0.0489)	2.0108 (0.0727)	3.2522 (0.1462)	0.0072
25 มุ่งมั่น	1.1487 (0.0693)	-0.7480 (0.0457)	0.6259 (0.0409)	1.9056 (0.0649)	2.8973 (0.1317)	0.0755
26 ยินดี	0.9670 (0.0546)	-1.3823 (0.0605)	0.1101 (0.0429)	1.3513 (0.0520)	2.4580 (0.0942)	0.1622
27 รัก	0.9145 (0.0544)	-0.8416 (0.0535)	0.7715 (0.0484)	2.1822 (0.0799)	3.4058 (0.1717)	0.0178
28 ร่าเริง	0.1802 (0.0095)	-8.0731 (0.3574)	-2.4130 (0.1920)	3.1504 (0.1977)	8.3825 (0.3661)	0.0002
29 รื่นเริงใจ	0.2737 (0.0143)	-6.1660 (0.2741)	-2.3807 (0.1407)	1.9481 (0.1322)	5.3591 (0.2296)	0.0050
30 โล่งใจ	0.9013 (0.0524)	-1.1348 (0.0584)	0.5611 (0.0477)	2.0593 (0.0687)	2.9031 (0.1138)	0.9897
31 ใจกว้าง	1.0291 (0.0609)	-0.9847 (0.0524)	0.6735 (0.0446)	2.0106 (0.0679)	2.9221 (0.1251)	0.7102
32 สงบ	1.1896 (0.0711)	-0.7006 (0.0439)	0.5536 (0.0381)	1.5384 (0.0508)	2.4279 (0.0925)	0.0248
33 สดชื่น	0.4043 (0.0252)	-0.8849 (0.0904)	1.1788 (0.0872)	4.0635 (0.1566)	6.1083 (0.3106)	0.0000
34 สนใจ	0.2435 (0.0130)	-4.5655 (0.2014)	-1.1551 (0.1338)	3.0966 (0.1584)	6.9228 (0.3036)	0.0000
35 สนุกสนม	1.2858 (0.0864)	0.0874 (0.0366)	1.1119 (0.0411)	1.9289 (0.0594)	2.6477 (0.1071)	0.7941
36 สนุกสนาน	0.9833 (0.0560)	-1.2952 (0.0587)	0.2836 (0.0425)	1.4363 (0.0521)	2.4382 (0.0916)	0.1041

ตารางที่ ค-4 (ต่อ)

มาตรวัดย่อย/ข้อคำถาม	ค่าพารามิเตอร์ของข้อคำถาม					P - value
	α (SE)	β_1 (SE)	β_2 (SE)	β_3 (SE)	β_4 (SE)	
37 สบายใจ	0.9963 (0.0715)	0.3934 (0.0450)	1.8148 (0.0637)	2.8994 (0.1097)	3.6813 (0.2072)	0.5551
38 สุขใจ	0.9243 (0.0650)	0.4782 (0.0449)	1.4642 (0.0518)	2.4475 (0.0793)	3.2523 (0.1429)	0.0883
39 ให้เกียรติ	1.0203 (0.0586)	-1.2241 (0.056)	0.2412 (0.0416)	1.4351 (0.0499)	2.2371 (0.0781)	0.1900
40 อ่อนน้อม	0.8247 (0.0485)	-1.1739 (0.0624)	0.7284 (0.0522)	2.3189 (0.0900)	4.6966 (0.4020)	0.3106
41 อ่อนโยน	0.5305 (0.0297)	-2.0612 (0.0975)	0.3386 (0.0710)	2.9615 (0.1188)	5.1196 (0.2879)	0.0000
อารมณ์ความรู้สึกเชิงลบ						
1 กระวนกระวายใจ	1.2020 (0.0725)	2.5799 (0.1014)	1.7469 (0.0563)	0.6888 (0.0401)	-0.8616 (0.0463)	0.0090
2 กลัว	0.6551 (0.0387)	3.8760 (0.1795)	2.4912 (0.0899)	0.8391 (0.0596)	-0.9386 (0.0663)	0.0006
3 กังวล	1.0522 (0.0669)	3.2024 (0.1527)	2.3472 (0.0806)	1.1251 (0.0485)	-0.3375 (0.0447)	0.1782
4 แค้นใจ	0.8822 (0.0497)	2.7522 (0.1114)	1.5020 (0.0571)	0.2691 (0.0461)	-1.4965 (0.0660)	0.1138
5 ใจคอห่อเหี่ยว	0.9983 (0.0569)	2.7593 (0.1150)	1.5856 (0.0578)	0.2223 (0.0436)	-1.5282 (0.0636)	0.0097
6 ฉุนเฉียว	1.0066 (0.0574)	2.7884 (0.1185)	1.5525 (0.0572)	0.2063 (0.0435)	-1.6625 (0.0673)	0.0001
7 เชื่องซึม	1.0686 (0.0638)	3.1203 (0.1530)	1.9647 (0.0689)	0.5706 (0.0429)	-0.9009 (0.0495)	0.0393
8 ตูถูกเหยียดหยาม	1.2665 (0.0777)	2.6771 (0.1116)	1.8263 (0.0594)	0.6823 (0.0391)	-0.6555 (0.0423)	0.5946
9 โดดเดี่ยว	1.1564 (0.0710)	3.0356 (0.1439)	2.1413 (0.0734)	0.7988 (0.0432)	-0.7964 (0.0466)	0.3497
10 ตกใจ	1.2654 (0.0762)	2.5474 (0.1000)	1.7343 (0.0564)	0.4738 (0.0384)	-1.0325 (0.0475)	0.0062
11 ต่อด้าน	1.2638 (0.0749)	2.4796 (0.0976)	1.5235 (0.0516)	0.3098 (0.0378)	-1.3028 (0.0528)	0.3055
12 ตึงเครียด	1.1236 (0.0660)	2.9222 (0.1363)	1.6834 (0.0591)	0.4234 (0.0409)	-1.2192 (0.0538)	0.0804
13 ตื่นตระหนก	1.2304 (0.0755)	2.9573 (0.1442)	1.8537 (0.0625)	0.7372 (0.0401)	-0.6555 (0.0431)	0.0246

ตารางที่ ค-4 (ต่อ)

มาตรวัดย่อย/ข้อคำถาม	ค่าพารามิเตอร์ของข้อคำถาม					P - value
	α (SE)	β_1 (SE)	β_2 (SE)	β_3 (SE)	β_4 (SE)	
14 ถูกตำหนิ	1.2524 (0.0776)	2.9321 (0.1414)	1.8880 (0.0632)	0.8023 (0.0402)	-0.5539 (0.0416)	0.0034
15 ถูกปฏิเสธ	0.7257 (0.0430)	4.0487 (0.2187)	2.5050 (0.0938)	0.7633 (0.0562)	-0.9393 (0.0622)	0.0000
16 ทุกข์ทรมาน	0.0480 (0.0029)	10.5582 (0.7843)	-9.0462 (0.7331)	-36.8242 (1.5982)	-55.4316 (3.4329)	0.0003
17 น้อยใจ	0.9749 (0.0555)	2.5605 (0.0995)	1.4701 (0.0538)	0.2967 (0.0434)	-1.4263 (0.0618)	0.0529
18 บ้าคลั่ง	0.9147 (0.0516)	2.8233 (0.1180)	1.5406 (0.0583)	0.2076 (0.0454)	-1.4622 (0.0639)	0.2092
19 เบื่อหน่าย	1.5338 (0.0944)	2.1567 (0.0745)	1.4474 (0.0448)	0.4828 (0.0335)	-0.7853 (0.0396)	0.1373
20 ประหม่า	1.4390 (0.0912)	2.3923 (0.0876)	1.7384 (0.0530)	0.7331 (0.0362)	-0.3685 (0.0366)	0.9194
21 ผิดหวัง	1.3786 (0.0860)	2.4055 (0.0901)	1.6572 (0.0515)	0.6972 (0.0364)	-0.4258 (0.0379)	0.1057
22 หลงพลาถน	1.4113 (0.0889)	2.4193 (0.0894)	1.7544 (0.0539)	0.7158 (0.0365)	-0.4217 (0.0374)	0.0777
23 พ่ายแพ้	0.9580 (0.0581)	3.5064 (0.1942)	2.0935 (0.0744)	0.8744 (0.0473)	-0.6541 (0.0498)	0.0021
24 โหมโห	1.2112 (0.0718)	2.5086 (0.0987)	1.5604 (0.0523)	0.4569 (0.0385)	-0.9726 (0.0474)	0.0199
25 ไม่มีความสุข	1.3934 (0.0851)	2.3359 (0.086)	1.5641 (0.0492)	0.5786 (0.0360)	-0.8109 (0.0421)	0.3037
26 ระทกใจ	0.8545 (0.0592)	3.6277 (0.1618)	2.9676 (0.1013)	1.7251 (0.0638)	0.2633 (0.0498)	0.0000
27 รำคาญ	1.1082 (0.0724)	3.2397 (0.1675)	2.2646 (0.0770)	1.2795 (0.0486)	-0.1362 (0.0423)	0.0537
28 รู้สึกผิด	1.0264 (0.0646)	3.3295 (0.1699)	2.3427 (0.0822)	1.0951 (0.0487)	-0.3996 (0.0458)	0.2589
29 รู้สึกแย	1.1142 (0.0780)	3.3432 (0.1639)	2.7488 (0.0996)	1.5157 (0.0548)	0.2244 (0.0415)	0.0039
30 ไร้ชีวิตชีวา	1.3678 (0.0935)	2.7651 (0.1179)	2.1060 (0.0671)	1.1559 (0.0420)	0.1323 (0.0357)	0.0459
31 ไร้อารมณ์	1.2523 (0.0761)	2.5192 (0.0944)	1.8310 (0.0580)	0.5267 (0.0388)	-0.8680 (0.0451)	0.0416

ตารางที่ ค-4 (ต่อ)

มาตรวัดย่อย/ข้อคำถาม	ค่าพารามิเตอร์ของข้อคำถาม					p - value
	α (SE)	β_1 (SE)	β_2 (SE)	β_3 (SE)	β_4 (SE)	
32 ละอายใจ	0.8889 (0.0505)	2.9390 (0.1240)	1.7351 (0.0637)	0.2196 (0.0467)	-1.3737 (0.0625)	0.0264
33 ลำบากใจ	1.1525 (0.0764)	3.5238 (0.2198)	2.5559 (0.0979)	1.2315 (0.0488)	-0.1349 (0.0411)	0.0011
34 เศร้าใจ	1.1004 (0.0632)	2.3486 (0.0867)	1.4075 (0.0502)	0.1828 (0.0403)	-1.3107 (0.0557)	0.1950
35 สยดสยอง	1.5270 (0.0983)	2.4652 (0.0951)	1.7775 (0.0540)	0.8960 (0.0369)	-0.4251 (0.0363)	0.6424
36 สลดใจ	1.1433 (0.0669)	2.4226 (0.0894)	1.5988 (0.0532)	0.3447 (0.0400)	-1.1949 (0.0527)	0.0643
37 สะเทือนใจ	1.4940 (0.0956)	2.4579 (0.0933)	1.8022 (0.0548)	0.8721 (0.0373)	-0.4877 (0.0373)	0.7683
38 สะอิดสะเอียน	0.9127 (0.0547)	3.0236 (0.1280)	2.0111 (0.0680)	0.7722 (0.0473)	-0.7100 (0.0515)	0.0069
39 สิ้นหวัง	1.3026 (0.0805)	2.7328 (0.1176)	1.9003 (0.0622)	0.6999 (0.0390)	-0.7103 (0.0424)	0.5732
40 เสียใจ	1.1763 (0.0790)	3.0919 (0.1444)	2.4004 (0.0813)	1.3302 (0.0492)	-0.0167 (0.0403)	0.5770
41 เสียตาย	0.9374 (0.0604)	3.8359 (0.2338)	2.7141 (0.1040)	1.2835 (0.0546)	-0.2688 (0.0479)	0.0283
42 แสลงใจ	0.7204 (0.0444)	2.8927 (0.1060)	1.9542 (0.0643)	0.8269 (0.0514)	-0.5356 (0.0566)	0.0454
43 หงุดหงิด	0.6513 (0.0356)	2.2948 (0.0873)	1.0574 (0.0586)	-0.6911 (0.0599)	-2.4736 (0.1009)	0.0251
44 หดหู่	0.1387 (0.0074)	8.4000 (0.3882)	0.5732 (0.2406)	-7.6568 (0.3285)	-13.945 (0.6478)	0.0051
45 หนักใจ	0.8562 (0.0499)	3.0629 (0.1294)	1.9540 (0.0681)	0.6122 (0.0485)	-1.0039 (0.0574)	0.0001
46 หัวนใจ	0.0480 (0.0028)	13.4139 (0.8262)	-6.1375 (0.6924)	-29.478 (1.2696)	-55.4327 (3.6437)	0.0012
47 หวาดเสียว	0.9257 (0.0537)	2.5098 (0.0933)	1.5437 (0.0542)	0.4399 (0.0439)	-1.0198 (0.0548)	0.0000
48 หึ่งหวง	0.9856 (0.0645)	3.3734 (0.1676)	2.4562 (0.0853)	1.2856 (0.0517)	-0.0859 (0.0451)	0.0079
49 เหน็ดเหนื่อย	1.0479 (0.0801)	3.2898 (0.1371)	2.8355 (0.0922)	1.9095 (0.0620)	0.9151 (0.0458)	0.8554

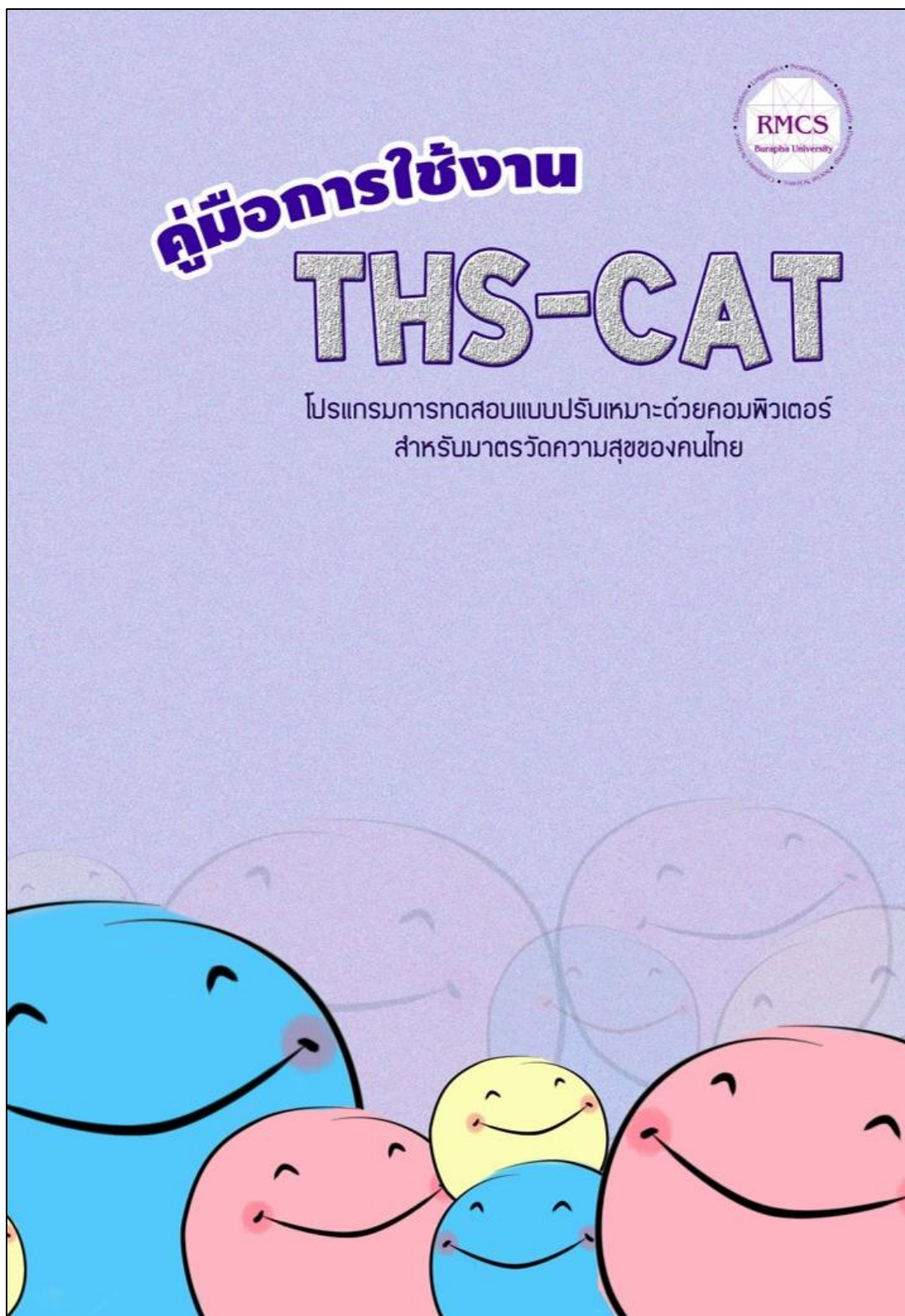
ตารางที่ ค-4 (ต่อ)

มาตรวัดย่อย/ข้อคำถาม	ค่าพารามิเตอร์ของข้อคำถาม					<i>p</i> - value
	α (SE)	β_1 (SE)	β_2 (SE)	β_3 (SE)	β_4 (SE)	
50 อัดอั้นตันใจ	0.9175 (0.0671)	3.4190 (0.1585)	2.6112 (0.0821)	1.8723 (0.0591)	0.7383 (0.0483)	0.0789
51 อับอาย	1.0301 (0.0660)	2.6314 (0.0906)	2.0672 (0.0610)	1.0023 (0.0452)	-0.1957 (0.0435)	0.3703
52 อาฆาตพยาบาท	0.8120 (0.0463)	2.6766 (0.1025)	1.5612 (0.0584)	0.2971 (0.0477)	-1.2412 (0.0621)	0.0106
53 อ่างว้าง	1.0012 (0.0576)	2.7425 (0.1113)	1.7186 (0.0603)	0.2781 (0.0435)	-1.2823 (0.0576)	0.8072
54 อารมณ์ไม่มั่นคง	0.8657 (0.0563)	3.6841 (0.1803)	2.8714 (0.1039)	1.3458 (0.0577)	-0.1966 (0.0500)	0.0362
55 อิจฉา	1.2356 (0.0748)	2.4221 (0.0872)	1.7464 (0.0546)	0.5597 (0.0389)	-0.8145 (0.0447)	0.3508
56 อี๊ดอืดใจ	1.0616 (0.0651)	2.6506 (0.0995)	1.9177 (0.0604)	0.8162 (0.0434)	-0.5298 (0.0453)	0.0003

หมายเหตุ

- α คือ ค่าพารามิเตอร์ความชัน
- $\beta_1, \beta_2, \beta_3, \beta_4$ คือ ค่าพารามิเตอร์เทรซไฮลด์ที่ 1 2 3 และ 4
- SE* คือ ค่าความคลาดเคลื่อนมาตรฐานของค่าเฉลี่ย
- p* - value คือ คำนัยสำคัญทางสถิติ ของสถิติทดสอบ χ^2 ที่ใช้ตรวจสอบความสอดคล้องของข้อมูลเชิงประจักษ์ที่ได้จากการตอบข้อคำถามกับ Graded Response Model เป็นรายข้อ

ภาคผนวก ง
คู่มือการใช้โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์
สำหรับมาตรวัดความสุขของคนไทย



สารบัญ

THS-CAT คืออะไร

• หน้า 1 - 3

การดำเนินการทดสอบ

• หน้า 4 - 10

การค้นหาผลการทดสอบ

• หน้า 11-12

การจัดการโปรแกรม

• หน้า 13 - 19

ส่วนที่เกี่ยวข้องอื่นๆ

• หน้า 20



สารบัญภาพ

ภาพที่ 1	หน้าจอลึกของ THS-CAT.....	หน้า 4
ภาพที่ 2	หน้าจอ THS-CAT คืออะไร.....	หน้า 5
ภาพที่ 3	หน้าจอคำแนะนำการใช้โปรแกรม.....	หน้า 6
ภาพที่ 4	หน้าจอเริ่มทดสอบ.....	หน้า 7
ภาพที่ 5	หน้าจอการทดสอบ.....	หน้า 8
ภาพที่ 6	หน้าจอรายงานผลการทดสอบ.....	หน้า 9
ภาพที่ 7	หน้าจอเมนูผลการทดสอบ.....	หน้า 11
ภาพที่ 8	หน้าจอแสดงข้อมูลผลการทดสอบตามรหัสที่สืบค้น.....	หน้า 12
ภาพที่ 9	หน้าจอแสดงรายละเอียดการทดสอบในแต่ละมาตรวัดย่อย.....	หน้า 12
ภาพที่ 10	หน้าจอการเข้าสู่ระบบของผู้ดูแลระบบ.....	หน้า 13



สารบัญภาพ (ต่อ)

ภาพที่ 11	หน้าจอบริการจัดการ.....	หน้า 14
ภาพที่ 12	หน้าจอบริการจัดการคลังข้อสอบ.....	หน้า 15
ภาพที่ 13	หน้าจอบริการจัดการมาตรฐานวัดย่อย.....	หน้า 15
ภาพที่ 14	หน้าจอการเพิ่มมาตรฐานวัดย่อย.....	หน้า 16
ภาพที่ 15	หน้าจอการแก้ไขมาตรฐานวัดย่อย.....	หน้า 16
ภาพที่ 16	หน้าจอบริการจัดการข้อคำถาม.....	หน้า 17
ภาพที่ 17	หน้าจอการเพิ่มข้อคำถาม.....	หน้า 18
ภาพที่ 18	หน้าจอการแก้ไขข้อคำถาม.....	หน้า 18
ภาพที่ 19	หน้าจอบริการจัดการประวัติการทดสอบ.....	หน้า 19
ภาพที่ 20	หน้าจอติดต่อเรา.....	หน้า 20

สารบัญตาราง

ตารางที่ 1	เกณฑ์การแปลความหมายคะแนนที่ได้จาก THS-CAT.....	หน้า 10
------------	------------------------------------------------	---------



THS-CAT คืออะไร?



THS-CAT เป็นชื่อย่อของโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย (Thai Happiness Scale: Computerized Adaptive Testing Version) ซึ่งสร้างขึ้นจากนิยามความอยู่ดีมีสุขเชิงอัตวิสัย (Subjective Well-being) ที่เชื่อว่าความสุขเป็นกระบวนการตัดสินใจและประเมินคุณภาพชีวิตตามมุมมองส่วนตัว แต่ละคนมีเกณฑ์ในการประเมินที่แตกต่างกัน ความสุขในมุมมองนี้ ประกอบด้วย สามองค์ประกอบย่อย ได้แก่

ความพึงพอใจในชีวิต

เป็นการประเมินชีวิตในภาพรวมว่าคุณคนรู้สึกพึงพอใจในชีวิตมากน้อยเพียงใด

อารมณ์ความรู้สึกเชิงบวก

เป็นการประเมินว่าคุณคนเกิดอารมณ์ความรู้สึกที่ดีซึ่งทำให้เกิดความรู้สึกพอใจบ่อยครั้งเพียงใด

อารมณ์ความรู้สึกเชิงลบ

เป็นการประเมินว่าคุณคนเกิดอารมณ์ความรู้สึกที่ไม่ดีซึ่งทำให้เกิดความรู้สึกไม่พอใจบ่อยครั้งเพียงใด

บุคคลที่มีความสุข คือ บุคคลที่มีความรู้สึกพึงพอใจในชีวิต มีประสบการณ์ทางอารมณ์ความรู้สึกเชิงบวกบ่อยครั้ง ขณะที่ประสบการณ์ทางอารมณ์ความรู้สึกเชิงลบน้อยครั้งหรือแทบไม่มีเลย ซึ่งหากประเมินด้วยมาตรวัดควรได้คะแนนของมาตรวัดย่อยด้านความพึงพอใจในชีวิตและอารมณ์ความรู้สึกเชิงบวกอยู่ในระดับสูง ในขณะที่ได้คะแนนของมาตรวัดย่อยด้านอารมณ์ความรู้สึกเชิงลบอยู่ในระดับต่ำ



THS-CAT คืออะไร?



THS-CAT จัดทำขึ้นบนพื้นฐานทฤษฎีการตอบสนองข้อสอบ (Item Response Theory) กับการบริหารการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ (Computerized Adaptive Testing) ซึ่งทำให้กระบวนการทดสอบมีความเฉพาะเจาะจงตามระดับคุณลักษณะแฝงของบุคคล ผู้รับการทดสอบไม่จำเป็นต้องได้รับการทดสอบด้วยข้อคำถามชุดเดียวกันที่มีจำนวนข้อคำถามเท่ากัน แต่จะได้รับชุดข้อคำถามที่แตกต่างกันไปตามระดับคุณลักษณะแฝงของผู้รับการทดสอบ ซึ่งทำให้ใช้ข้อคำถามในจำนวนที่แตกต่างกันด้วย อย่างไรก็ตาม คะแนนที่ได้จากการทดสอบซึ่งใช้ข้อคำถามที่แตกต่างกัน สามารถนำมาเปรียบเทียบกันได้โดยใช้ทฤษฎีการตอบสนองข้อสอบ นอกจากนี้ กระบวนการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์มีเกณฑ์การยุติการทดสอบที่พิจารณาจากความคลาดเคลื่อนมาตรฐานของการประมาณค่าคุณลักษณะแฝง ซึ่งความคลาดเคลื่อนมาตรฐานดังกล่าวนี้ มีความสัมพันธ์โดยตรงกับความแม่นยำ (Precision) ของการทดสอบ ส่งผลให้ THS-CAT ใช้ระยะเวลาในการทดสอบน้อย เพราะในการทดสอบจะเลือกเฉพาะข้อคำถามที่เหมาะสมกับแต่ละบุคคล ขณะที่มีความแม่นยำในการทดสอบสูง เพราะพิจารณาจากความคลาดเคลื่อนมาตรฐานที่ได้จากการประมาณค่าคุณลักษณะแฝง ก่อนยุติการทดสอบ



คู่มือการใช้งาน THS-CAT



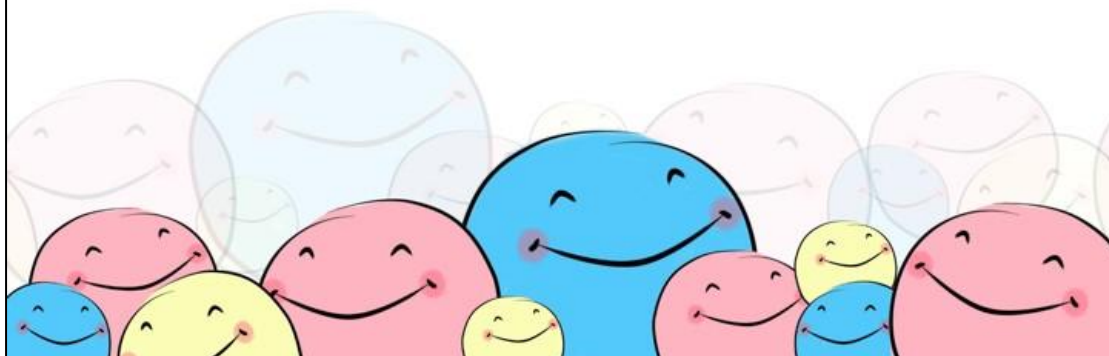
มีอะไรบ้าง?

การดำเนินการทดสอบ

การค้นหาผลการทดสอบ

การจัดการโปรแกรม (ผู้ดูแลระบบ)

ส่วนที่เกี่ยวข้องอื่นๆ





การดำเนินการทดสอบ

ผู้ใช้งานสามารถเข้าใช้งานโปรแกรมฯ ได้ที่เว็บไซต์

<http://www.thscat.com/Test/>

ซึ่งจะพบหน้าจอแรก ดังภาพที่ 1



ภาพที่ 1 หน้าจอหลักของ THS-CAT



จากภาพที่ 1 แสดงหน้าจอหลักของ THS-CAT สำหรับผู้ใช้งานทั่วไป ซึ่งประกอบด้วย 4 เมนู ดังนี้

- 1) หน้าหลัก เป็นเมนูสำหรับแสดงเนื้อหา 3 ส่วนย่อย ได้แก่ THS-CAT คืออะไร คำแนะนำในการใช้โปรแกรม และการเริ่มการทดสอบ
 - 2) ผลการทดสอบ เป็นเมนูสำหรับการค้นหาผลการทดสอบที่ผ่านมาของผู้ใช้งาน
 - 3) ติดต่อเรา เป็นเมนูสำหรับแสดงข้อมูลที่เกี่ยวข้องกับผู้วิจัย
 - 4) เข้าสู่ระบบ เป็นเมนูสำหรับการเข้าสู่ระบบของผู้ดูแลระบบ
- ในหน้าจอหลัก หากผู้ใช้งานเลือกปุ่ม "THS-CAT คืออะไร" จะปรากฏหน้าจอ ดังภาพที่ 2

THS-CAT คืออะไร?

THS-CAT เป็นชื่อภาษาอังกฤษของโปรแกรมการทดสอบแบบเป็นภาษา
คอมพิวเตอร์ สำหรับวัดความพึงพอใจในชีวิต

Thai Happiness Scale Computerized Adaptive Testing Version

โปรแกรมนี้สร้างขึ้นจากแนวคิดของทฤษฎีการตอบสนองต่อข้อสอบ และการ
ทดสอบแบบเป็นภาษาคอมพิวเตอร์ (CAT) ใช้ตรวจสอบระดับของความสุข
ของผู้ใช้ในชีวิตจริง (Subjective Well-being) ซึ่งสามารถวัดเป็นกระบวนการ
คิดและประเมินคุณค่าชีวิตของผู้ทดสอบแต่ละส่วน และสามารถนำมา
ใช้การวิจัยเชิงทดลองกับ จำนวนผู้ทดลองที่มี 3 องค์ประกอบ ได้แก่

- ความพึงพอใจในชีวิต** เป็นการประเมินถึงคุณค่าในชีวิต
[มีระดับค่าได้ 100-150 คะแนน]
- อารมณ์ความรู้สึกบวก** เป็นการประเมินถึงอารมณ์ความรู้สึก
ดีดีที่เกิดขึ้นในชีวิตประจำวัน
- อารมณ์ความรู้สึกลบ** เป็นการประเมินถึงอารมณ์ความรู้สึก
ไม่ดีที่เกิดขึ้นในชีวิตประจำวัน

โปรแกรมนี้เหมาะสำหรับวัย 18 ปีขึ้นไป ซึ่งสามารถใช้ทดสอบ
ได้ถึง 10 ครั้งในโปรแกรม 1 ครั้ง

- ✓ ผู้ทดสอบสามารถประเมินค่าชีวิตประจำวันของตนเองได้
- ✓ ผู้ทดสอบสามารถประเมินค่าชีวิตประจำวันของตนเองได้
- ✓ สามารถประเมินค่าชีวิตประจำวันของตนเองได้

เมื่อทำการทดสอบ จะมีค่าออกมาบ้างซึ่งถ้ามีค่าสูง เป็นค่าที่ดีกว่า
เมื่อ ผู้ทดสอบสามารถประเมินค่าชีวิตประจำวันของตนเองได้ โดยพบว่ามีค่า
ได้มากถึง 100

ระดับความพึงพอใจในชีวิต ความพึงพอใจในชีวิตของตนเอง	ระดับอารมณ์ความรู้สึกบวก ระดับอารมณ์ความรู้สึกลบ
ไม่มีความพึงพอใจเลย	ไม่เลย/แทบไม่เลย
เล็กน้อย	น้อยไป
ปานกลาง	ปานกลาง
ดีเล็กน้อย	ดีไป
ดีมาก	ดีมาก

ภาพที่ 2
หน้าจอ THS-CAT คืออะไร



ในหน้าจอหลัก หากผู้ใช้งานเลือกปุ่ม “คำแนะนำการใช้โปรแกรม” จะปรากฏหน้าจอ ดังภาพที่ 3



ภาพที่ 3 หน้าจอคำแนะนำการใช้โปรแกรม

จากภาพที่ 3 โปรแกรมแสดงวิธีการใช้งานเบื้องต้น เพื่อให้ผู้ใช้งานสามารถเข้าใช้โปรแกรมได้อย่างสะดวกรวดเร็ว แต่หากผู้ใช้งานมีข้อสงสัยหรือต้องการทราบรายละเอียดส่วนใดเพิ่มเติม สามารถเลือกดาวน์โหลดคู่มือการใช้โปรแกรมในรูปแบบเอกสารได้ทันที



ในหน้าจอหลัก หากผู้ใช้งานเลือกปุ่ม "เริ่มทดสอบ" จะปรากฏหน้าจอ
ดังภาพที่ 4

เริ่มทดสอบ !!!

กรุณาเลือกมาตรวัดที่ต้องการทดสอบ
(เลือกมาตรวัดย่อยที่ต้องการทดสอบ หรือเลือกมาตรวัดความสุขโดยรวมเพื่อรับการทดสอบครบทั้งสามมาตรวัดย่อย)

ความพึงพอใจในชีวิต อารมณ์ความรู้สึกเชิงบวก
 อารมณ์ความรู้สึกเชิงลบ มาตรวัดความสุขโดยรวม

คำนำหน้า : นาย นาง นางสาว
ชื่อ :
สกุล :
เลขประจำตัวประชาชน :
อายุ (ปี) :
เพศ : ชาย หญิง
กรุณาใส่รหัสผ่านที่ท่านเห็น :

ภาพที่ 4 หน้าจอเริ่มทดสอบ



จากภาพที่ 4 ผู้ทดสอบสามารถเลือกมาตรวัดเพื่อทดสอบได้ตามความต้องการ โดยเลือก "มาตรวัดความสุขโดยรวม" เพื่อดำเนินการทดสอบทั้งสามมาตรวัดย่อย หรือเลือกเจาะจงในแต่ละมาตรวัดย่อยตามความสนใจ จากนั้นจึงกรอกข้อมูลส่วนบุคคล และใส่รหัสแคปทชา นอกจากนี้ผู้ทดสอบสามารถเลือกทดสอบได้สองรูปแบบ ได้แก่

- 1) ทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ เลือกปุ่ม "เริ่มทดสอบ (CAT)"
- 2) ทดสอบด้วยข้อคำถามทั้งหมดในคลังข้อคำถาม (แบบปกติ) เลือกปุ่ม "เริ่มทดสอบ (ทุกข้อ)"

เมื่อเลือกรูปแบบการทดสอบแล้วจะปรากฏหน้าจอการทดสอบ ดังภาพที่ 5

ฉันกังวลว่าจะไม่มีคนดูแลหรือให้ความช่วยเหลือในอนาคต

ไม่เห็นด้วย/เห็นด้วยน้อยที่สุด

เห็นด้วยน้อย

เห็นด้วยปานกลาง

เห็นด้วยมาก

เห็นด้วยอย่างยิ่ง

ตกลง

ภาพที่ 5
หน้าจอการทดสอบ

จากภาพที่ 5 ผู้ทดสอบจะต้องตอบข้อคำถามทีละข้อ จึงจะสามารถตอบข้อคำถามต่อไปได้ และไม่สามารถย้อนกลับมาเปลี่ยนแปลงคำตอบในข้อคำถามที่ผ่านไปแล้วได้ การทดสอบจะดำเนินการเช่นนี้ไปเรื่อยๆ จนกระทั่งถึงเกณฑ์ยุติการทดสอบ โปรแกรมจะแสดงหน้าจอรายงานผลการทดสอบ ดังภาพที่ 6



ภาพที่ 6
หน้าจอรายงานผลการทดสอบ



จากภาพที่ 6 หน้าจอรายงานผลการทดสอบจะปรากฏขึ้นทุกครั้งที่การทดสอบสิ้นสุดลง ซึ่งหากผู้ใช้งานต้องการเก็บรายงานผลดังกล่าวในรูปแบบของเอกสาร สามารถสั่งพิมพ์ได้โดยกดปุ่ม "พิมพ์หน้านี้" แต่หากต้องการกลับหน้าก่อนหน้าก็สามารถกดปุ่ม "ก่อนหน้า"



ผลการทดสอบจะระบุประวัติของผู้ทดสอบ และผลการทดสอบในแต่ละมาตรวัดย่อย ซึ่งหากทดสอบครบสามมาตรวัดย่อยจะมีรายงานผลของมาตรวัดความสุขโดยรวมด้วย คะแนนในรายงานผลการทดสอบแสดงในรูปของคะแนนที่ (คะแนนเต็ม 100 คะแนน) ซึ่งสามารถแปลความหมายช่วงคะแนนเป็น 9 ระดับตามค่าคะแนนสเตรนจ์ รายละเอียดดังตารางที่ 1

ตารางที่ 1 เกณฑ์การแปลความหมายคะแนนที่ได้จาก THS-CAT

คะแนนที่	มาตรวัดย่อย/มาตรวัด			
	ความพึงพอใจในชีวิต	อารมณ์ความรู้สึกเชิงบวก	อารมณ์ความรู้สึกเชิงลบ	ความสุขโดยรวม
มากกว่า 68	พึงพอใจอย่างยิ่ง	สูงที่สุด	ต่ำที่สุด	มีความสุขอย่างยิ่ง
63-68	พึงพอใจมาก	สูงมาก	ต่ำมาก	มีความสุขมาก
58-62	พึงพอใจ	สูง	ต่ำ	มีความสุข
53-57	ค่อนข้างพึงพอใจ	สูงกว่าปกติเล็กน้อย	ต่ำกว่าปกติเล็กน้อย	มีความสุขสูงกว่าปกติเล็กน้อย
48-52	ปานกลาง	ปกติ	ปกติ	มีความสุขตามปกติ
43-47	พึงพอใจน้อย	ต่ำกว่าปกติเล็กน้อย	สูงกว่าปกติเล็กน้อย	มีความสุขต่ำกว่าปกติเล็กน้อย
38-42	ไม่ค่อยพึงพอใจ	ต่ำ	สูง	มีความสุขน้อย
32-37	ไม่พึงพอใจ	ต่ำมาก	สูงมาก	มีความสุขน้อยมาก
น้อยกว่า 32	ไม่พึงพอใจอย่างยิ่ง	ต่ำที่สุด	สูงที่สุด	ไม่มีความสุข





การค้นหาผลการทดสอบ

ผู้ใช้งานสามารถค้นหาผลการทดสอบที่ผ่านมาทั้งหมดได้ โดยเลือกเมนู "ผลการทดสอบ" ซึ่งจะปรากฏหน้าจอ ดังภาพที่ 7

The screenshot shows the THS-CAT system interface. At the top left is the THS-CAT logo. To its right, the text reads: "โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรฐานวัดความสุขของคนไทย". Below this are four buttons: "หน้าหลัก", "ผลการทดสอบ", "ติดต่อเรา", and "เข้าสู่ระบบ". The main area contains a form with the text: "กรุณาใส่เลขประจำตัวประชาชน เพื่อค้นหาผลการทดสอบ". There is a white input field for the ID number and a green button labeled "ตกลง". At the bottom, it says "วิทยาลัยการวิจัยและวิทยาการปัญญา มหาวิทยาลัยบูรพา".

ภาพที่ 7 หน้าจอเมนูผลการทดสอบ

จากภาพที่ 7 ผู้ใช้งานสามารถค้นหาผลการทดสอบที่ผ่านมา โดยใช้เลขประจำตัวประชาชนเป็นรหัสในการสืบค้น ซึ่งจะปรากฏผลการทดสอบที่ผ่านมาทั้งหมดของผู้ใช้งาน ดังภาพที่ 8



การจัดการ > การทดสอบ

ครั้งที่	วันที่ทดสอบ	
1	03/11/2015 18:12:15	รายละเอียด ผลการทดสอบ
2	03/13/2015 18:49:22	รายละเอียด ผลการทดสอบ

ภาพที่ 8 หน้าจอแสดงข้อมูลผลการทดสอบตามรหัสที่สืบค้น

จากภาพที่ 8 แสดงครั้งที่และวันเวลาที่ทดสอบทั้งหมดของผู้ใช้งาน ผู้ใช้งานสามารถดูรายละเอียดการทดสอบแต่ละครั้งได้ โดยเลือกปุ่ม “ผลการทดสอบ” จะปรากฏหน้าจอรายงานผลการทดสอบเช่นเดียวกับภาพที่ 6 ซึ่งผู้ใช้งานสามารถสั่งพิมพ์ผลการทดสอบในรูปแบบเอกสารได้ แต่หากผู้ทดสอบเลือกปุ่ม “รายละเอียด” จะปรากฏหน้าจอดังภาพที่ 9

การจัดการ > การทดสอบ

ครั้งที่	วันที่ทดสอบ	มาตรวัดที่ทดสอบ	ระยะเวลาที่ใช้	จำนวนข้อที่ใช้	
1	03/11/2015 18:12:22	มาตรวัดความพึงพอใจในชีวิต	00:00:17	7	รายละเอียด ลบ
2	03/11/2015 18:12:42	มาตรวัดอารมณ์ความรู้สึกเชิงบวก	00:00:19	5	รายละเอียด ลบ
3	03/11/2015 18:13:04	มาตรวัดอารมณ์ความรู้สึกเชิงลบ	00:00:08	4	รายละเอียด ลบ

ภาพที่ 9 หน้าจอแสดงรายละเอียดการทดสอบในแต่ละมาตรวัดย่อย

จากภาพที่ 9 แสดงวันเวลาที่ทดสอบ ระยะเวลาและจำนวนข้อคำถามที่ใช้ในการทดสอบ สำหรับแต่ละมาตรวัดย่อย ซึ่งเป็นรายละเอียดในภาพรวม ส่วนของปุ่ม “รายละเอียด” และปุ่ม “ลบ” ในหน้าจอนี้ ผู้ใช้งานไม่สามารถเข้าไปใช้งานได้ เนื่องจากเป็นส่วนของผู้ดูแลระบบเท่านั้น





การจัดการโปรแกรม (ผู้ดูแลระบบ)

การจัดการโปรแกรม เป็นส่วนสำหรับผู้ดูแลระบบ เพื่อใช้สำหรับการจัดการคลังข้อความ และการจัดการประวัติการทดสอบ การเข้าสู่ระบบจะต้องใช้รหัสที่ได้รับจากทีมพัฒนาโปรแกรม โดยลงชื่อเข้าใช้ได้ในเมนู "เข้าสู่ระบบ" ซึ่งจะปรากฏหน้าจอ ดังภาพที่ 10

ภาพที่ 10 หน้าจอการเข้าสู่ระบบของผู้ดูแลระบบ



จากภาพที่ 10 เมื่อเข้าสู่ระบบแล้ว จะปรากฏเมนู "การจัดการ" ขึ้นในแถบเมนู ผู้ดูแลระบบสามารถเข้าไปจัดการคลังข้อคำถามและประวัติการทดสอบได้ผ่านเมนูนี้ ซึ่งจะปรากฏหน้าจอ ดังภาพที่ 11



ภาพที่ 11 หน้าจอการจัดการ

จากภาพที่ 11 ในส่วนของหน้าจอการจัดการ เมื่อเลือกในส่วนของคลังข้อคำถาม จะปรากฏหน้าจอดังภาพที่ 12





ภาพที่ 12
หน้าจอการจัดการ
คลังข้อคำถาม

จากภาพที่ 12 หากเลือกส่วนของมาตรฐานวัดย่อย จะปรากฏหน้าจอ ดังภาพที่ 13

ภาพที่ 13
หน้าจอการจัดการ
มาตรฐานวัดย่อย



จากภาพที่ 13 หน้าจอการจัดการมาตรฐานวัดย่อย เป็นส่วนที่ผู้ดูแลระบบสามารถ
เพิ่ม ลบ หรือแก้ไขมาตรฐานวัดย่อยได้ตามความต้องการ



เมื่อผู้ดูแลระบบเลือกในส่วนของการเพิ่ม หรือแก้ไขมาตรฐานวัดย่อย จะปรากฏหน้าจอ สำหรับระบุชื่อมาตรฐานวัดย่อย ค่า Theta (θ) ที่ใช้เริ่มต้นการทดสอบ และค่า SEE สำหรับยุติการทดสอบ ซึ่งเมื่อเพิ่มหรือแก้ไขเรียบร้อยแล้วข้อมูลจะถูกจัดเก็บในฐานข้อมูล ในขณะที่เมื่อเลือกกลับมาตราวัดย่อย ข้อมูลที่เกี่ยวข้องดังกล่าวจะถูกลบออกไปจากฐานข้อมูล หน้าจอในส่วนของการเพิ่มและแก้ไขมาตรฐานวัดย่อย แสดงดังภาพที่ 14 และภาพที่ 15

ชื่อมาตรฐานวัดย่อย :

ค่า Theta เริ่มต้น :

ค่า SEE สำหรับยุติการทดสอบ :

เพิ่ม

ภาพที่ 14 หน้าจอการเพิ่มมาตรฐานวัดย่อย

ชื่อมาตรฐานวัดย่อย :

ค่า Theta เริ่มต้น :

ค่า SEE สำหรับยุติการทดสอบ :

แก้ไข

ภาพที่ 15 หน้าจอการแก้ไขมาตรฐานวัดย่อย



จากภาพที่ 12 หากเลือกส่วนของข้อคำถาม จะปรากฏหน้าจอ ดังภาพที่ 16

ภาพที่ 16
หน้าจอการจัดการ
ข้อคำถาม



จากภาพที่ 16 หน้าจอการจัดการข้อคำถาม เป็นส่วนที่ผู้ดูแลระบบสามารถเพิ่ม ลบ หรือแก้ไขข้อคำถามที่ผ่านการวิเคราะห์คุณภาพเรียบร้อยแล้ว เมื่อผู้ดูแลระบบเลือก ในส่วนของกรเพิ่ม หรือแก้ไขข้อคำถาม จะปรากฏหน้าจอสำหรับระบุค่าและข้อมูลต่างๆ ที่เกี่ยวข้อง ได้แก่ ชื่อมาตราวัดย่อย คำถาม ตัวเลือกคำตอบ คำพารามิเตอร์ความชันและ คำพารามิเตอร์เทรซไฮลด์ที่ได้จากการประมาณค่าข้อคำถามด้วย Grade Response Model (GRM) ในขณะที่เมื่อเลือกลบข้อคำถาม ข้อมูลที่เกี่ยวข้องกับข้อคำถามดังกล่าวจะ ถูกลบออกไปจากฐานข้อมูล หน้าจอในส่วนของการเพิ่มและแก้ไขข้อคำถาม แสดงดังภาพ ที่ 17 และภาพที่ 18



ชื่อมาควรีต

คำถาม

คำตอบที่ 1

คำตอบที่ 2

คำตอบที่ 3

คำตอบที่ 4

คำตอบที่ 5

Alpha

Beta 1

Beta 2

Beta 3

Beta 4

เพิ่ม

ภาพที่ 17
หน้าจอการเพิ่ม
ข้อคำถาม

ชื่อมาควรีต

คำถาม

คำตอบที่ 1

คำตอบที่ 2

คำตอบที่ 3

คำตอบที่ 4

คำตอบที่ 5

Alpha

Beta 1

Beta 2

Beta 3

Beta 4

แก้ไข

ภาพที่ 18
หน้าจอการแก้ไข
ข้อคำถาม



จากภาพที่ 11 ในส่วนของหน้าจอการจัดการ เมื่อเลือกในส่วนของประวัติการทดสอบ จะปรากฏหน้าจอดังภาพที่ 19



The screenshot shows the THS-CAT program interface. At the top left is the THS-CAT logo, a blue smiley face with a red and yellow smiley face below it. To the right of the logo is the text: "โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย". Below this text are five buttons: "หน้าหลัก", "ผลการทดสอบ", "การจัดการ", "ติดต่อเรา", and "ออกจากระบบ". Below the buttons is the heading "ประวัติการทดสอบ" and a table with the following data:

ลำดับที่	ชื่อ-นามสกุล	เพศ	อายุ	
1	วิมลภาวี อัมภวรรณ	หญิง	34	รายละเอียด
2	สุนามภรณ์ ปานคำ	หญิง	37	รายละเอียด
3	thiprapa singthong	หญิง	50	รายละเอียด
4	โสมส สุซานท์สวัสดิ์	ชาย	36	รายละเอียด
5	ปัทมา วรรัตน์	หญิง	49	รายละเอียด

ภาพที่ 19 หน้าจอการจัดการประวัติการทดสอบ

จากภาพที่ 19 หน้าจอ "การจัดการประวัติการทดสอบ" ของโปรแกรม แสดงข้อมูลการทดสอบเรียงตามลำดับครั้งที่ทดสอบ โดยแสดงข้อมูลในภาพรวมเกี่ยวกับประวัติส่วนตัวของผู้ทดสอบ ได้แก่ ชื่อ-นามสกุล เพศ อายุ ซึ่งหากต้องการรายละเอียดการทดสอบเพิ่มเติมเกี่ยวกับมาตรวัดที่ทดสอบ ค่าคุณลักษณะแฝง จำนวนข้อที่ทดสอบ ระยะเวลาที่ใช้ ข้อมูลข้อคำถามรายข้อ และอื่นๆ สามารถคลิกดูรายละเอียดเพิ่มเติมได้จากส่วนของรายละเอียด





ส่วนที่เกี่ยวข้องอื่นๆ

ผู้ใช้งานที่ต้องการข้อมูลเพิ่มเติมเกี่ยวกับงานวิจัย สถาบันของผู้วิจัย และข้อมูลผู้วิจัย สามารถเข้าดูรายละเอียดได้ในเมนู "ติดต่อเรา" ดังภาพที่ 20

THS-CAT

โปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์
สำหรับมาตรฐานวัดความสุขของคนไทย

หน้าหลัก | ผลการทดสอบ | การจัดการ | ติดต่อเรา | ออกจากระบบ

RMCS วิทยาลัยวิทยาการวิจัยและวิทยาการปัญญา มหาวิทยาลัยบูรพา

วิทยาลัยวิทยาการวิจัยและวิทยาการปัญญา เป็นหน่วยงานต้นแบบวิชา มีวิสัยทัศน์การพัฒนาระบบปัญญา
วิจัยและวิชาการด้านวิทยาการวิจัยและวิทยาการปัญญาเพื่อคนไทย โดยสนับสนุนและประสานความร่วมมือกับภาคี
วิชาการผู้เกี่ยวข้อง และส่งเสริมโดยนโยบายรัฐบาล เป็นวิทยาลัยต้นแบบของประเทศไทยที่เน้นศาสตร์ปัญญา
วิทยาการปัญญา (Cognitive Science) มุ่งพัฒนาบัณฑิต และคุณวุฒิบัณฑิตด้านวิทยาการปัญญา เพื่อการส่งเสริมสังคม
ด้านการศึกษา สู่สังคมที่ชาญฉลาดและมีความก้าวหน้าอย่างยั่งยืน

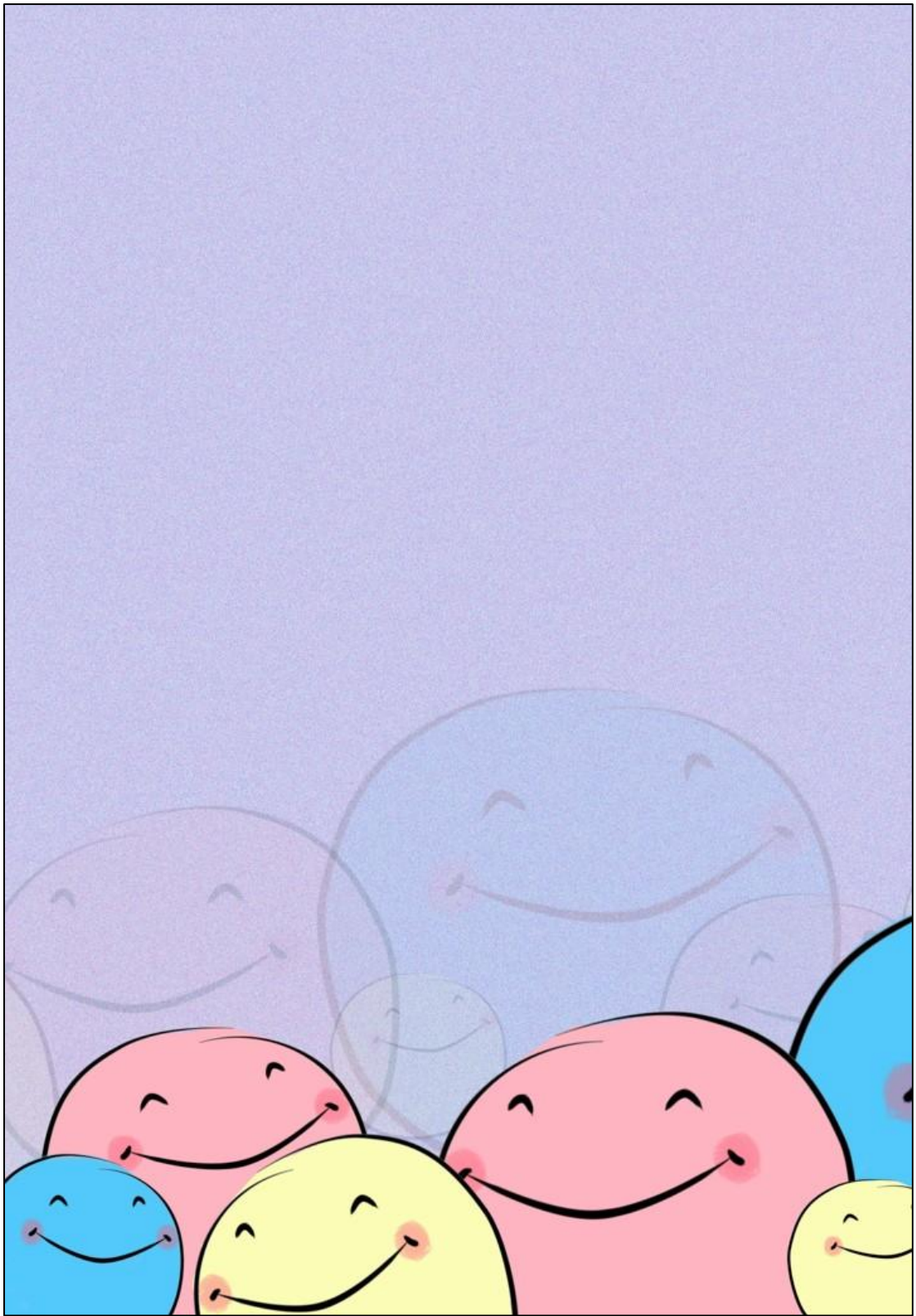
ที่อยู่ 169 ถนนพหลโยธิน แขวงสามยุค อ.เมือง จ.ชลบุรี 20131
โทร. 0-3810-2077 ต่อ 114-116, 08-17827490 โทรสาร 0-3839-3484
อีเมล: rmcs@buu.ac.th

ผู้วิจัย
สุชาดา สกอกิจรุ่งโรจน์
อีเมล: suchada.rmcs@gmail.com

วิทยาลัยวิทยาการวิจัยและวิทยาการปัญญา มหาวิทยาลัยบูรพา

ภาพที่ 20
หน้าจอติดต่อเรา





ภาคผนวก จ
แบบประเมินความเหมาะสมของโปรแกรมการทดสอบแบบปรับเหมาะ
ด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย

THS-CAT



สำหรับผู้เชี่ยวชาญ

แบบประเมินความเหมาะสมของโปรแกรมการทดสอบแบบปรับเหมาะ
ด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย (THS-CAT)

ตอนที่ 1

โปรดทำเครื่องหมาย (✓)
ในช่องที่ตรงกับระดับ
ความคิดเห็นของท่านมากที่สุด
โดยมีเกณฑ์การประเมิน ดังนี้

- 5 โปรแกรมมีความเหมาะสมมากที่สุด
- 4 โปรแกรมมีความเหมาะสมมาก
- 3 โปรแกรมมีความเหมาะสมปานกลาง
- 2 โปรแกรมมีความเหมาะสมน้อย
- 1 โปรแกรมมีความเหมาะสมน้อยที่สุด

รายการประเมิน		ระดับความคิดเห็น				
		5	4	3	2	1
ความสะดวกในการใช้โปรแกรม						
1.	เมื่อมีข้อสงสัยในการใช้โปรแกรม ผู้ใช้สามารถดูวิธีการแก้ไขจากคู่มือการใช้โปรแกรมได้					
2.	โปรแกรมมีเมนูที่เข้าถึงส่วนต่างๆ ของโปรแกรมได้ง่ายและสะดวก					
3.	ผู้ใช้สามารถเลือกมาตรวัดย่อยในการทดสอบได้ตามความต้องการส่วนบุคคล					
4.	โปรแกรมมีการแสดงผลการทดสอบบนจอภาพ เมื่อทดสอบเสร็จสิ้น					
5.	ผู้ใช้สามารถบันทึกผลการทดสอบ แล้วสั่งพิมพ์ผลการทดสอบได้สะดวก					
ความถูกต้องในการใช้งาน						
1.	โปรแกรมสามารถเพิ่ม ลบ หรือแก้ไข ข้อคำถามได้					
2.	โปรแกรมสามารถเพิ่ม ลบ หรือแก้ไข มาตรวัดย่อยได้					
3.	โปรแกรมสามารถจัดการทดสอบได้ตรงตามวัตถุประสงค์ หรือเงื่อนไขของการทดสอบ					
4.	โปรแกรมสามารถคำนวณค่าคะแนนของผู้ใช้ได้					
5.	โปรแกรมสามารถบันทึกผลการทดสอบ แล้วสั่งพิมพ์ผลการทดสอบได้					

รายการประเมิน		ระดับความคิดเห็น				
		5	4	3	2	1
ลักษณะทั่วไปของโปรแกรม						
1.	การออกแบบโปรแกรมมีความน่าสนใจ					
2.	การจัดรูปแบบหน้าจอต่อการใช้งาน					
3.	การแสดงผลข้อมูลเป็นไปอย่างรวดเร็ว					
4.	การเรียกใช้งานโปรแกรมทำได้ง่าย					
5.	โปรแกรมมีระบบป้องกันการทำงานผิดพลาดของผู้ใช้ทุกขั้นตอน					
ความชัดเจนของคู่มือการใช้โปรแกรม						
1.	มีการกล่าวถึงความเป็นมาของการพัฒนาโปรแกรม ในคู่มือการใช้โปรแกรมอย่างชัดเจน					
2.	คู่มือการใช้โปรแกรมแสดงวิธีการใช้งานอย่างมีลำดับขั้นตอน					
3.	ภาษาที่ใช้ในคู่มือการใช้โปรแกรมเข้าใจง่าย					
4.	คู่มือมีการใช้ภาพประกอบการอธิบายกระบวนการต่างๆอย่างชัดเจน					
5.	หลังจากอ่านคู่มือแล้ว ผู้ใช้มีความมั่นใจว่าสามารถใช้โปรแกรมได้					

ตอนที่ 2 ข้อเสนอแนะอื่นๆ

.....

.....

.....

.....

.....

.....

ลงชื่อ

.....
(.....)

ผู้เชี่ยวชาญ

THS-CAT



สำหรับผู้ใช้งาน

แบบประเมินความเหมาะสมของโปรแกรมการทดสอบแบบปรับเหมาะ
ด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย (THS-CAT)

ตอนที่ 1

โปรดทำเครื่องหมาย (✓)
ในช่องที่ตรงกับระดับ
ความคิดเห็นของท่านมากที่สุด
โดยมีเกณฑ์การประเมิน ดังนี้

- 5 โปรแกรมมีความเหมาะสมมากที่สุด
- 4 โปรแกรมมีความเหมาะสมมาก
- 3 โปรแกรมมีความเหมาะสมปานกลาง
- 2 โปรแกรมมีความเหมาะสมน้อย
- 1 โปรแกรมมีความเหมาะสมน้อยที่สุด

รายการประเมิน		ระดับความคิดเห็น				
		5	4	3	2	1
ความสะดวกในการใช้โปรแกรม						
1.	เมื่อมีข้อสงสัยในการใช้โปรแกรม ผู้ใช้สามารถดูวิธีการแก้ไขจากคู่มือการใช้โปรแกรมได้					
2.	โปรแกรมมีเมนูที่เข้าถึงส่วนต่างๆ ของโปรแกรมได้ง่ายและสะดวก					
3.	ผู้ใช้สามารถเลือกมาตรวัดย่อยในการทดสอบได้ตามความต้องการส่วนบุคคล					
4.	โปรแกรมมีการแสดงผลการทดสอบบนจอภาพ เมื่อทดสอบเสร็จสิ้น					
5.	ผู้ใช้สามารถบันทึกผลการทดสอบ แล้วสั่งพิมพ์ผลการทดสอบได้สะดวก					
ลักษณะทั่วไปของโปรแกรม						
1.	การออกแบบโปรแกรมมีความน่าสนใจ					
2.	การจัดรูปแบบหน้าจอต่อการใช้งาน					
3.	การแสดงผลข้อมูลเป็นไปอย่างรวดเร็ว					
4.	การเรียกใช้งานโปรแกรมทำได้ง่าย					
5.	โปรแกรมมีระบบป้องกันการดำเนินงานผิดพลาดของผู้ใช้ทุกขั้นตอน					

ตอนที่ 2 ข้อเสนอแนะอื่นๆ

Handwriting practice area with 10 horizontal lines.

ขอขอบพระคุณสำหรับ
ความร่วมมือ



ภาคผนวก ฉ

ผลการประเมินความเหมาะสมของโปรแกรมการทดสอบแบบปรับเหมาะด้วย
คอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย โดยผู้เชี่ยวชาญเป็นรายข้อ

ตารางที่ ฉ-1 ผลการประเมินความเหมาะสมของโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย โดยผู้เชี่ยวชาญ เป็นรายชื่อ

ข้อคำถาม	M	SD	ระดับความเหมาะสม
ด้านความสะดวกในการนำไปใช้			
1. เมื่อมีข้อสงสัยในการใช้โปรแกรม ผู้ใช้สามารถดูวิธีการแก้ไขจากคู่มือการใช้โปรแกรมได้	4.33	0.58	มาก
2. โปรแกรมมีเมนูที่เข้าถึงส่วนต่าง ๆ ของโปรแกรมได้ง่ายและสะดวก	4.33	0.58	มาก
3. ผู้ใช้สามารถเลือกมาตรวัดย่อยในการทดสอบได้ตามความต้องการส่วนบุคคล	4.67	0.58	มากที่สุด
4. โปรแกรมมีการแสดงผลการทดสอบบนจอภาพ เมื่อทดสอบเสร็จสิ้น	4.67	0.58	มากที่สุด
5. ผู้ใช้สามารถบันทึกผลการทดสอบ แล้วสั่งพิมพ์ผลการทดสอบได้สะดวก	3.67	1.16	มาก
ด้านความถูกต้องในการใช้งาน			
1. โปรแกรมสามารถเพิ่ม ลบ หรือแก้ไข ข้อคำถามได้	4.33	0.58	มาก
2. โปรแกรมสามารถเพิ่ม ลบ หรือแก้ไข มาตรวัดย่อยได้	4.00	1.00	มาก
3. โปรแกรมสามารถจัดการทดสอบได้ตรงตามวัตถุประสงค์ หรือเงื่อนไขของการทดสอบ	4.00	1.00	มาก
4. โปรแกรมสามารถคำนวณค่าคะแนนของผู้ใช้ได้	4.67	0.58	มากที่สุด
5. โปรแกรมสามารถบันทึกผลการทดสอบ แล้วสั่งพิมพ์ผลการทดสอบได้	4.00	1.00	มาก
ด้านลักษณะทั่วไปของโปรแกรม			
1. การออกแบบโปรแกรมมีความน่าสนใจ	4.67	0.58	มากที่สุด
2. การจัดรูปแบบหน้าจอง่ายต่อการใช้งาน	5.00	0.00	มากที่สุด
3. การแสดงผลข้อมูลเป็นไปอย่างรวดเร็ว	4.67	0.58	มากที่สุด
4. การเรียกใช้งานโปรแกรมทำได้ง่าย	4.33	0.58	มาก
5. โปรแกรมมีระบบป้องกันการดำเนินงานผิดพลาดของผู้ใช้ทุกขั้นตอน	3.67	1.53	มาก
ด้านความชัดเจนของคู่มือการใช้โปรแกรม			
1. มีการกล่าวถึงความเป็นมาของการพัฒนาโปรแกรมอย่างชัดเจนในคู่มือการใช้โปรแกรม	4.33	0.58	มาก
2. คู่มือการใช้โปรแกรมแสดงวิธีการใช้งานอย่างมีลำดับขั้นตอน	4.33	1.16	มาก
3. ภาษาที่ใช้ในคู่มือการใช้โปรแกรมเข้าใจง่าย	4.33	1.16	มาก
4. คู่มือมีการใช้ภาพประกอบอธิบายกระบวนการต่าง ๆ อย่างชัดเจน	4.00	1.00	มาก
5. หลังจากอ่านคู่มือแล้ว ผู้ใช้มีความมั่นใจว่าสามารถใช้โปรแกรมได้	4.33	0.58	มาก

ภาคผนวก ช

ผลการประเมินความเหมาะสมของโปรแกรมการทดสอบแบบปรับเหมาะด้วย
คอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย โดยผู้ใช้งานเป็นรายข้อ

ตารางที่ ข-1 ผลการประเมินความเหมาะสมของโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับมาตรวัดความสุขของคนไทย โดยผู้ใช้งาน เป็นรายข้อ

ข้อคำถาม	M	SD	ระดับความเหมาะสม
ด้านความสะดวกในการนำไปใช้			
1. เมื่อมีข้อสงสัยในการใช้โปรแกรม ผู้ใช้สามารถดูวิธีการแก้ไขจากคู่มือการใช้โปรแกรมได้	4.33	0.58	มาก
2. โปรแกรมมีเมนูที่เข้าถึงส่วนต่าง ๆ ของโปรแกรมได้ง่ายและสะดวก	4.33	0.58	มาก
3. ผู้ใช้สามารถเลือกมาตรวัดย่อยในการทดสอบได้ตามความต้องการส่วนบุคคล	4.67	0.58	มากที่สุด
4. โปรแกรมมีการแสดงผลการทดสอบบนจอภาพ เมื่อทดสอบเสร็จสิ้น	4.67	0.58	มากที่สุด
5. ผู้ใช้สามารถบันทึกผลการทดสอบ แล้วสั่งพิมพ์ผลการทดสอบได้สะดวก	3.67	1.16	มาก
ด้านลักษณะทั่วไปของโปรแกรม			
1. การออกแบบโปรแกรมมีความน่าสนใจ	4.67	0.55	มากที่สุด
2. การจัดรูปแบบหน้าจอง่ายต่อการใช้งาน	4.77	0.43	มากที่สุด
3. การแสดงผลข้อมูลเป็นไปอย่างรวดเร็ว	4.77	0.43	มากที่สุด
4. การเรียกใช้งานโปรแกรมทำได้ง่าย	4.83	0.38	มากที่สุด
5. โปรแกรมมีระบบป้องกันการทำงานผิดพลาดของผู้ใช้ทุกขั้นตอน	4.10	0.66	มาก