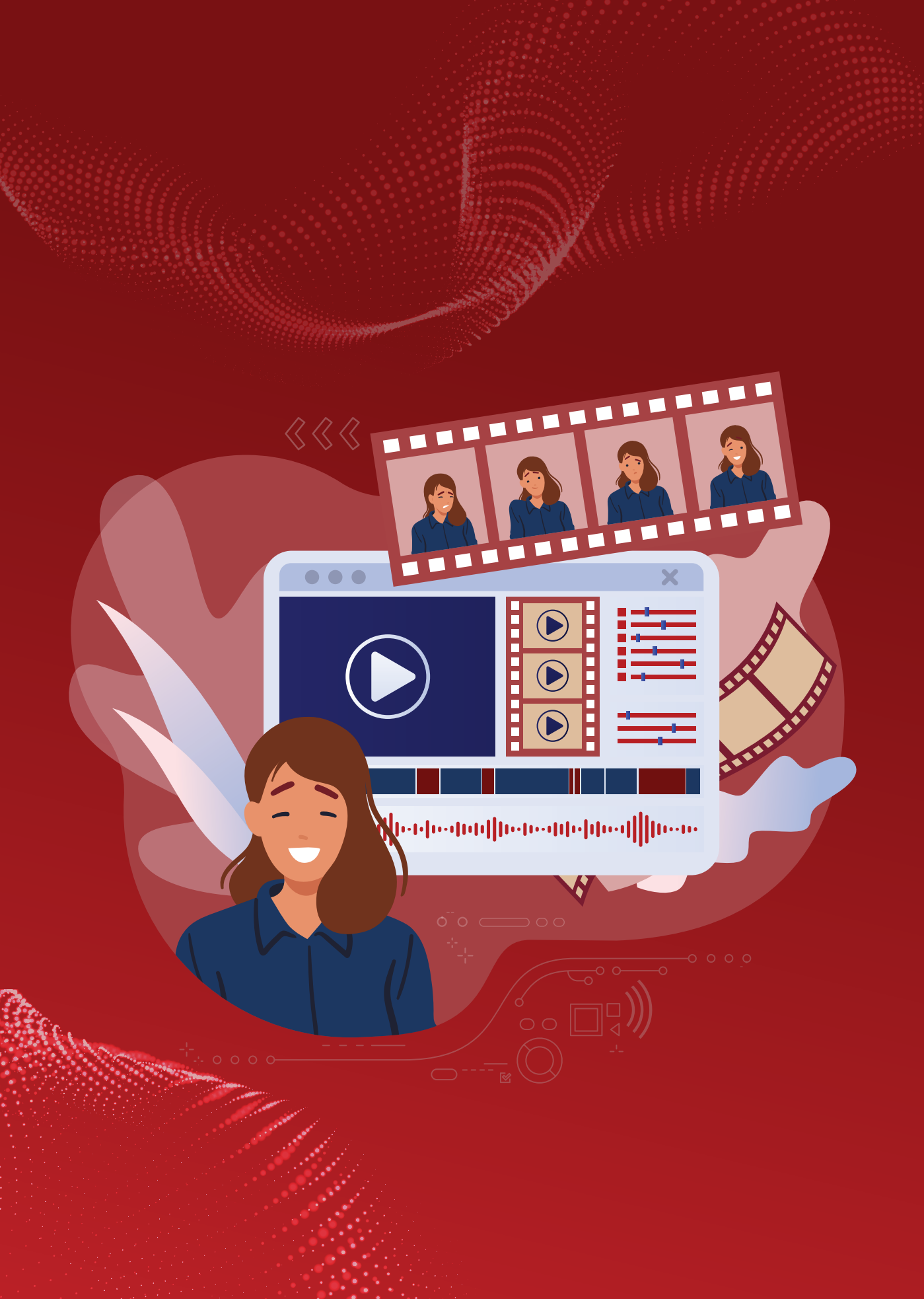




การสร้างฐานข้อมูล
การแสดงออกทางอารมณ์
ผ่านทางอวัจนภาษา
ในบริบทของการบำบัดเชิงจิตวิทยา

NONVERBAL EMOTIONAL
EXPRESSION DATABASE
DEVELOPMENT IN THE CONTEXT OF
PSYCHOLOGICAL THERAPY

กลกรณ์ วงศ์ภาติกะเสรี¹

พนิดา โยมะบุตร²

นกวรรณ มั่นพรรษา³

นริศ หุบทอม⁴

สุเมธ ยืนยง⁵

Konlakorn Wongpatikaseree¹

Panida Yomaboot²

Napawan Munpansa³

Narit Noohom⁴

Sumeth Yuenyong⁵

มหาวิทยาลัยมหิดล นครปฐม 73170^{1 และ 3 ถึง 5}

มหาวิทยาลัยมหิดล กรุงเทพมหานคร 10700²

Mahidol University, Nakhon Pathom 73170 Thailand^{1 and 3 to 5}

Mahidol University, Bangkok 10700 Thailand²

Corresponding E-mail : konlakorn.won@mahidol.ac.th

Received Date August 4, 2023

Revised Date May 30, 2024

Accepted Date July 10, 2024

บทคัดย่อ

งานวิจัยนี้เป็นการพัฒนาการสร้างฐานข้อมูลการแสดงออกทางอารมณ์ผ่านอวัจนภาษาในบริบทการให้คำปรึกษาทางจิตวิทยาใน 2 ด้าน คือ การแสดงออกผ่านกล้ามเนื้อใบหน้า รวมทั้งการเคลื่อนไหวของดวงตา และการแสดงออกผ่านน้ำเสียง ศึกษาด้วยวิธีวิจัยเชิงทดลอง โดยจำลองสถานการณ์การให้คำปรึกษาทางจิตวิทยาในกลุ่มตัวอย่างทั่วไปซึ่งได้รับการกระตุ้นอารมณ์ตามทฤษฎีกลิ่นอายอารมณ์คนละ 1 อารมณ์ และบันทึกการแสดงออกทางอารมณ์ที่ได้จากกระบวนการพูดคุย 15-25 นาที เพื่อนำข้อมูลให้บุคคลทั่วไปและนักจิตวิทยาระบุลักษณะการแสดงออกทางอารมณ์ อ้างอิงระบบถอดรหัสสีหน้าและการแสดงอารมณ์ผ่านเสียง ผลการศึกษาพบว่า ด้านข้อมูลภาพ จากการตัดภาพที่ได้จากคลิปวิดีโอความยาว 15-25 นาที ข้อมูลที่ได้ 1 ชุด จะมีภาพใบหน้าของกลุ่มตัวอย่างซึ่งเป็นภาพต่อเนื่องจำนวน 3,000-8,000 ภาพ ภาพที่ได้ขึ้นอยู่กับเงื่อนไขการกระตุ้นอารมณ์ที่ได้รับ สำหรับข้อมูลเสียงที่วิเคราะห์ได้ ไม่มีความหมายเชิงภาษา เนื่องจากเป็นวลีที่มีใจความไม่ครบถ้วน โดยสัดส่วนอารมณ์ที่วิเคราะห์ได้จะแตกต่างกันไปขึ้นอยู่กับเงื่อนไขการกระตุ้นอารมณ์ที่ได้รับเช่นเดียวกับข้อมูลภาพ ข้อมูลการแสดงออกทั้งภาพและเสียงดังกล่าวสามารถนำไปต่อยอดเพื่อพัฒนาโมเดลปัญญาประดิษฐ์สำหรับวิเคราะห์อารมณ์จากอวัจนภาษาของคนไทยในบริบทการให้บริการทางจิตใจได้

คำสำคัญ: อารมณ์ อวัจนภาษา จิตวิทยา ฐานข้อมูล ปัญญาประดิษฐ์ (เอไอ)

Abstract

This study aims to develop a database of nonverbal emotional expressions in the context of psychological counseling. The study focused on two nonverbal expression parameters: expressions conveyed through facial muscles and eye movements, and expressions conveyed through voice modulation. An experimental research method was employed by simulating a psychological counseling scenario among samples. During the 15-25-minute counselling session, each sample was randomly assigned to one mood induction based on the circumplex model of emotion theory. Their nonverbal emotional expressions were recorded in video clips. The records were then identified by individuals and psychologists according to the Facial Action Coding System and analysis of expressions through voice modulation. From the 15-25 minute video clips, the facial expressions dataset contained 3,000-8,000 sequential frames capturing various emotional states, while vocal expressions contained no linguistic meaning as they were incomplete phrases. The emotions identified from the records varied depending on different emotional inductions given to the samples. The nonverbal expression database could be used to develop an Artificial Intelligence (AI) model for emotion recognition exhibited by the Thai population in the provision of mental health services.

Keywords: emotion, nonverbal expression, psychology, database, Artificial Intelligence (AI)

1. บทนำ

อารมณ์ คือ สภาวะของความรู้สึกทั้งด้านบวกและลบ เกิดขึ้นเพื่อตอบสนองต่อสถานการณ์บางอย่างที่เข้ามากระทบกับมนุษย์ แต่ละคนเกิดอารมณ์ความรู้สึกต่อเหตุการณ์ต่าง ๆ แตกต่างกัน ขึ้นอยู่กับการตีความและการให้ความหมายต่อสิ่งเรานั้น (นันทวัช สิริธีรักษ์ และคณะ, 2558; Gross, 2015) อารมณ์ที่เกิดขึ้นภายในสามารถนำไปสู่พฤติกรรมการแสดงออกภายนอกบางอย่างที่สอดคล้องกับอารมณ์ในขณะนั้น เรียกการแสดงออกดังกล่าวว่า “การแสดงออกทางอารมณ์ (emotional expression)” ซึ่งสามารถแสดงออกได้ผ่านการใช้ภาษา (verbal) และอวัจนภาษา (nonverbal) การสื่อสารทางอารมณ์ (emotional communication) เป็นวิธีการสื่อสารเชิงลึกที่ไม่เพียงแค่มีวัตถุประสงค์ให้คู่สนทนาทำความเข้าใจเนื้อหาของการสื่อสารเท่านั้น แต่ยังเป็น การสื่อสารถึงสิ่งที่อยู่ภายในจิตใจของผู้พูด เพื่อให้คู่สนทนารับรู้ถึงอารมณ์ความรู้สึกที่เกิดขึ้น ซึ่งจะช่วยให้การสื่อสารเป็นไปอย่างมีประสิทธิภาพ และส่งผลดีต่อสัมพันธภาพระหว่างบุคคล

มนุษย์ใช้วิธีการสื่อสารอารมณ์ผ่านการแสดงออกทางอวัจนภาษา เป็นสัดส่วนที่มากกว่าการใช้คำพูด (Lapakko, 2007) เนื่องจากในหลายสถานการณ์ การสื่อสารความรู้สึกโดยใช้คำพูดที่ตรงไปตรงมา อาจยุ่งยาก และสร้างความลำบากใจได้มากกว่า ทั้งจากบริบททางสังคม ค่านิยม วัฒนธรรม รวมถึงปัจจัยส่วนบุคคล อวัจนภาษา จึงเป็นเครื่องสะท้อนอารมณ์ความรู้สึกของมนุษย์ได้ชัดเจนและแม่นยำกว่าคำพูด หนึ่งในสถานการณ์หรือบริบทที่อวัจนภาษาถือเป็นสิ่งสำคัญ คือ กระบวนการให้ความช่วยเหลือด้านจิตใจ หรือกระบวนการให้คำปรึกษาทางจิตวิทยา (psychological counseling) สำหรับบริบทดังกล่าว ผู้รับบริการมักเข้าพบนักจิตวิทยาด้วยปัญหาอารมณ์ทางจิตใจ โดยลักษณะอารมณ์ที่พบได้บ่อย คือ อารมณ์เศร้า ผิดหวัง ท้อแท้ รู้สึกผิด เครียด กังวล โกรธ หงุดหงิด ไม่พอใจ ซึ่งอารมณ์เหล่านี้ก่อให้เกิดความซับซ้อนด้านการแสดงอารมณ์ความรู้สึกได้มากยิ่งขึ้น โดยทั่วไป นักจิตวิทยาจะถูกฝึกฝนให้สังเกตการเปลี่ยนแปลงอารมณ์ทางอวัจนภาษา (mood shift) โดยละเอียด ทำให้แม้เมื่อมีการเปลี่ยนแปลงบางอย่างเพียงเล็กน้อย เช่น การเปลี่ยนแปลงของน้ำเสียง ความเร็วในการพูด สีหน้า หรือการเคลื่อนไหวของร่างกาย จะสามารถสังเกตและรับรู้อารมณ์ที่เปลี่ยนแปลงของผู้รับบริการได้ และใช้วิธีการตอบสนองทางอารมณ์ที่เหมาะสมต่อสถานการณ์ในขณะนั้น (emphatic response) นอกจากนี้ นักจิตวิทยาจะสังเกตการเปลี่ยนแปลงอารมณ์ในเชิงบวก เช่น มีความสุข สบายใจ มั่นใจ หรือผ่อนคลาย ในระหว่างการสนทนา เพื่อค้นหาปัจจัยที่จะช่วยเหลือผู้เข้ารับบริการได้อีกด้วย

ในปัจจุบันมีการประยุกต์ใช้เทคโนโลยีเพื่อมาช่วยเพิ่มการเข้าถึงและความสะดวกในการเข้ารับบริการช่วยเหลือด้านจิตใจในการบำบัดและการรับคำปรึกษา ซึ่งจากงานวิจัยพบว่า มีประสิทธิภาพใกล้เคียงกับการเข้ารับ การบำบัดหรือการปรึกษาแบบตัวต่อตัว (วรารัณ โสมะนันท์ และคณะ, 2564; Zeren et al., 2020) อย่างไรก็ตามพบว่า มีข้อจำกัดในเรื่องการสื่อสาร โดยเฉพาะอย่างยิ่ง การสื่อสารทางอารมณ์ซึ่งอาจเกิดจากความขัดข้องของสัญญาณอินเทอร์เน็ต ความคมชัดของภาพ และเมื่อประกอบกับข้อจำกัดเรื่องการแสดงอารมณ์ของผู้ที่มีปัญหาด้านสภาพจิตใจ เช่น ผู้ที่มีภาวะอารมณ์เศร้า เครียด หรือกังวล อาจทำให้ความแม่นยำในการรับรู้ อารมณ์ของนักบำบัดลดน้อยลง จึงมีความจำเป็นที่จะต้องพัฒนาเทคโนโลยีหรือเครื่องมือช่วยเหลือในเรื่อง การสื่อสารและการรับรู้ทางอารมณ์เพื่อให้กระบวนการให้ความช่วยเหลือด้านจิตใจผ่านระบบเทคโนโลยีนี้ มีความแม่นยำมากขึ้น

ปัจจุบันพบว่า หลายประเทศได้พัฒนาการใช้เทคโนโลยีปัญญาประดิษฐ์ (เอไอ) (Artificial Intelligence: AI) ในการวิเคราะห์อารมณ์จากอวัจนภาษาของมนุษย์ (emotional recognition model) และนำไปประยุกต์ใช้ตามวัตถุประสงค์ต่าง ๆ เช่น งานบริการ เพื่อช่วยอำนวยความสะดวกให้แก่ผู้ให้บริการ และสร้างประสบการณ์ที่พึงประสงค์ต่อผู้เข้ารับบริการ (Borges et al., 2019) ซึ่งการพัฒนาเทคโนโลยีปัญญาประดิษฐ์จำเป็นต้องอาศัยฐานข้อมูลที่มีความสอดคล้องกับเป้าหมายที่ต้องการใช้งาน เช่น ฐานข้อมูลการแสดงอารมณ์ผ่านใบหน้า หรือ ฐานข้อมูลการแสดงอารมณ์ผ่านน้ำเสียง ซึ่งในปัจจุบันมีฐานข้อมูลการแสดงออกผ่านน้ำเสียงทั้งจากต่างประเทศ และในประเทศไทยที่เป็นที่นิยมอยู่จำนวนหนึ่ง โดยได้ข้อมูลมาจากนักแสดงที่แสดงน้ำเสียงตามอารมณ์ต่าง ๆ ตามบทและแนวทางที่กำหนด เช่น RAVDESS (Luna-Jiménez et al., 2021), Emo DB (Zhang et al., 2016), IEMOCAP (Busso et al., 2008), EMOLA (Kasuriya et al., 2019) และ Thai-SER ส่วนฐานข้อมูลการแสดงออกผ่านใบหน้าที่เป็นที่รู้จักในปัจจุบัน เช่น AffectNet พบว่า ข้อมูลภาพใบหน้าเป็นลักษณะภาพที่มีการแสดงออก

ทางอารมณ์ที่หลากหลาย แต่ภาพดังกล่าวเป็นภาพที่มีการแสดงออกทางอารมณ์ที่ชัดเจน ซึ่งฐานข้อมูลที่มีอยู่เหล่านี้ไม่สอดคล้องกับการนำไปใช้ในการพัฒนาโมเดลวิเคราะห์อารมณ์ที่เกิดขึ้นในกระบวนการบำบัด เนื่องจากลักษณะการแสดงอารมณ์ในกระบวนการบำบัดมีความซับซ้อนและคลุมเครือมากกว่าอารมณ์ที่ได้จากการแสดงที่แสดงออกชัดเจนทั้งใบหน้าและน้ำเสียง การจะพัฒนาโมเดลวิเคราะห์อารมณ์ที่เกิดขึ้นในกระบวนการบำบัด จึงควรมีฐานข้อมูลที่สร้างขึ้นจากสถานการณ์ที่มีความเฉพาะเจาะจง เพื่อให้ได้ข้อมูลที่เหมาะสมแก่การนำไปใช้ เพื่อให้โมเดลสามารถตรวจจับอารมณ์ที่ซับซ้อนของมนุษย์ในวัฒนธรรมและสังคมไทยได้

การศึกษาในครั้งนี้มีวัตถุประสงค์เพื่อสร้างฐานข้อมูลการแสดงออกทางอวัจนภาษาของคนไทย ในบริบทของการให้คำปรึกษาทางจิตวิทยา ผ่านกระบวนการวิจัยเชิงทดลอง (experimental research) ซึ่งเป็นสถานการณ์จำลองกระบวนการบำบัด โดยมีการกระตุ้นอารมณ์แก่กลุ่มตัวอย่างซึ่งเป็นบุคคลทั่วไป จากนั้นเก็บข้อมูลการแสดงออกทางอารมณ์ จากลักษณะทางอารมณ์ที่พบได้บ่อยในการให้คำปรึกษาดังกล่าวข้างต้น ผู้วิจัยได้จัดกลุ่มอารมณ์ตามแนวคิดวงล้ออารมณ์ (circumplex model of emotion) (Russell, 1980) ซึ่งประกอบด้วยกลุ่มอารมณ์ทางลบที่มีพลังงาน (negative-activation) อารมณ์ทางลบที่ไม่มีพลังงาน (negative-deactivation) อารมณ์ทางบวกที่มีพลังงาน (positive-activation) อารมณ์ทางบวกที่ไม่มีพลังงาน (positive-deactivation) และอารมณ์เป็นกลาง (neutral) โดยการกระตุ้นอารมณ์กลุ่มตัวอย่าง แต่ละคนจะได้รับการสุ่มเข้าเงื่อนไขการกระตุ้นอารมณ์ใดอารมณ์หนึ่ง และบันทึกภาพและเสียงที่ได้จากการสนทนา เพื่อนำไปสกัดเป็นข้อมูลภาพการแสดงออกทางอารมณ์ผ่านทางใบหน้าและน้ำเสียง และนำไปประมวลผลตามกลุ่มต่าง ๆ โดยภาพใบหน้าจะอ้างอิงจากระบบถอดรหัสสีหน้า (Facial Action Coding System หรือ FACS) (Ekman & Friesen, 1978) และน้ำเสียงจะอ้างอิงตามการศึกษาเรื่องการแสดงอารมณ์ผ่านน้ำเสียง (emotion in speech) (Sobin & Alpert, 1999) เพื่อนำไปฝึกสอนให้โมเดลวิเคราะห์อารมณ์ได้เรียนรู้อารมณ์ต่าง ๆ ของคนไทยที่เกิดขึ้นจริงตามกระบวนการให้คำปรึกษาทางจิตวิทยาต่อไป

2. วัตถุประสงค์

เพื่อสร้างฐานข้อมูลการแสดงอารมณ์ผ่านอวัจนภาษาของคนไทยในบริบทการให้คำปรึกษาทางจิตวิทยา ใน 2 ด้าน คือ การแสดงออกผ่านกล้ามเนื้อบนใบหน้า รวมทั้งการเคลื่อนไหวของดวงตา และการแสดงออกผ่านน้ำเสียง

3. วิธีการศึกษา

การศึกษานี้ใช้วิธีวิจัยเชิงทดลอง จำลองสถานการณ์การให้คำปรึกษาทางจิตวิทยาในรูปแบบทางไกล (tele-counseling) ผ่านโปรแกรม Zoom ในกลุ่มตัวอย่างบุคคลทั่วไป จำนวน 125 คน โดยสุ่มกลุ่มตัวอย่างเข้ารับเงื่อนไขการกระตุ้นอารมณ์ตามทฤษฎีกล้ออารมณ์ ซึ่งครอบคลุมอารมณ์ที่พบบ่อยในบริบทการบำบัด ได้แก่ อารมณ์ทางลบที่มีพลังงาน อารมณ์ทางลบที่ไม่มีพลังงาน อารมณ์ทางบวกที่มีพลังงาน และอารมณ์ทางบวกที่ไม่มีพลังงาน คนละ 1 อารมณ์ และบันทึกการแสดงออกทางอารมณ์ผ่านอวัจนภาษาที่ได้จากกระบวนการพูดคุย 15-25 นาที เพื่อนำข้อมูลที่ได้ให้บุคคลทั่วไปและนักจิตวิทยาทำการระบุลักษณะการแสดงออกทางอารมณ์อ้างอิงระบบ FACS และการแสดงอารมณ์ผ่านน้ำเสียง

โดยการศึกษานี้ได้รับการรับรองด้านจริยธรรมการวิจัยในมนุษย์ จากคณะกรรมการจริยธรรมการวิจัยในคนชุดกลาง มหาวิทยาลัยมหิดล รหัสโครงการเลขที่ MU-IRB 2021/131/1903 หนังสือรับรองโครงการวิจัยหมายเลข 2021/195.0690

3.1 กลุ่มตัวอย่าง

การศึกษานี้คัดเลือกจากกลุ่มผู้ใช้งานอินเทอร์เน็ตทั่วไปที่สนใจและสมัครเข้าร่วมโครงการผ่านการประชาสัมพันธ์ออนไลน์ และใช้การสุ่มแบบเจาะจง โดยมีเกณฑ์การคัดเลือกผู้เข้าร่วมงานวิจัย ดังต่อไปนี้

3.1.1 เกณฑ์การคัดเลือก

3.1.1.1 เพศหญิงหรือชาย อายุตั้งแต่ 18 ปีบริบูรณ์ขึ้นไป

3.1.1.2 เป็นผู้ที่มีสุขภาพจิตดี โดยได้คะแนนประเมินสุขภาพจิตของคนไทย (The Thai Mental Health Questionnaire: TMHQ) อยู่ในเกณฑ์ปกติ

3.1.1.3 ไม่เคยมีประวัติความผิดปกติทางด้านการรู้คิดหรือความผิดปกติด้านอารมณ์

3.1.1.4 ไม่เคยมีประวัติความผิดปกติบริเวณกล้ามเนื้อใบหน้า

3.1.1.5 สามารถใช้งานอุปกรณ์คอมพิวเตอร์หรือแท็บเล็ต (tablet)

3.1.1.6 สามารถใช้โปรแกรม Zoom

3.1.1.7 ยินยอมเข้าร่วมงานวิจัย

3.1.2 เกณฑ์การคัดออก

3.1.2.1 คะแนนจากแบบประเมินสุขภาพจิตสูงเกินกว่าที่กำหนด

3.1.2.2 มีความผิดปกติด้านการสื่อสารด้วยภาษาพูด

3.1.3 เกณฑ์การยุติการเข้าร่วม (withdrawal criteria)

3.1.3.1 ผู้เข้าร่วมการวิจัยสามารถถอนตัวจากการวิจัยได้ทุกเมื่อ โดยไม่จำเป็นต้องแจ้งให้ทราบล่วงหน้า

3.1.3.2 ผู้วิจัยสามารถยุติการดำเนินการเก็บข้อมูลกับผู้เข้าร่วมวิจัยได้ หากเกิดกรณีใดกรณีหนึ่งตามเงื่อนไขดังต่อไปนี้

- 1) ผู้เข้าร่วมวิจัยไม่สามารถอยู่ในพื้นที่ส่วนตัวได้
- 2) ผู้เข้าร่วมวิจัยอยู่ในพื้นที่ที่มีเสียงรบกวนและไม่สามารถกำจัดเสียงดังกล่าวดังได้
- 3) สัญญาณอินเทอร์เน็ตของผู้ร่วมวิจัยขัดข้องและไม่สามารถแก้ไขได้
- 4) ผู้เข้าร่วมวิจัยเกิดอารมณ์ที่รุนแรงในระหว่างการดำเนินการเก็บข้อมูล

3.2 เครื่องมือที่ใช้ในการศึกษา

3.2.1 เครื่องมือกระตุ้นอารมณ์

เครื่องมือสำหรับการกระตุ้นอารมณ์ที่ใช้ในการศึกษานี้ ดัดแปลงมาจากเทคนิคของ Siedlecka & Denson (2019) โดยอารมณ์เป้าหมายที่ต้องการกระตุ้น ผู้วิจัยจัดกลุ่มตามทฤษฎีกล้ออารมณ์ ได้แก่ อารมณ์โกรธและเครียดเป็นตัวแทนกลุ่มอารมณ์ทางลบที่มีพลังงาน อารมณ์เศร้าเป็นตัวแทนกลุ่มอารมณ์ทางลบที่ไม่มีพลังงาน อารมณ์มีความสุขเป็นตัวแทนกลุ่มอารมณ์ทางบวกที่มีพลังงาน และอารมณ์ผ่อนคลายเป็นตัวแทนกลุ่มอารมณ์ทางบวกที่ไม่มีพลังงาน เนื่องจากอารมณ์ทั้ง 5 นี้เป็นอารมณ์ที่พบได้บ่อยในบริบทการให้คำปรึกษาทางจิตวิทยา การกระตุ้นอารมณ์แต่ละอารมณ์นั้นจะใช้เครื่องมือกระตุ้นอารมณ์ 2 ชนิด ควบคู่ไปกับกระบวนการจินตนาการ เครื่องมือดังกล่าว ได้แก่

3.2.1.1 คำบรรยายสถานการณ์ที่สอดคล้องกับแต่ละอารมณ์ พัฒนาขึ้นโดยผู้วิจัย โดยการศึกษาและทบทวนเอกสารที่เกี่ยวข้องกับเหตุการณ์ที่ส่งผลกระทบต่ออารมณ์ต่าง ๆ ของกลุ่มคนตั้งแต่วัยผู้ใหญ่ตอนต้นจนถึงวัยผู้ใหญ่ตอนกลาง เพื่อให้ครอบคลุมกับกลุ่มตัวอย่างให้ได้มากที่สุด โดยพบว่า ในอารมณ์บุคคล

แต่ละช่วงวัยสามารถมีประสบการณ์ที่คล้ายคลึงกันได้ ตัวอย่างเหตุการณ์ที่กระตุ้นอารมณ์ที่ไม่มีพลังงานทั้งบวกและลบ ทุกช่วงวัยสามารถถูกกระตุ้นอารมณ์เศร้าด้วยเหตุการณ์การสูญเสีย หรือเหตุการณ์ที่เกี่ยวข้องกับการพักผ่อน สามารถกระตุ้นให้เกิดอารมณ์ผ่อนคลายในทุกช่วงวัยได้เช่นกัน ในขณะที่การกระตุ้นอารมณ์ที่มีพลังงาน สถานการณ์ที่ใช้จะต้องมีความเฉพาะเจาะจงมากขึ้น และสอดคล้องกับประสบการณ์ตามช่วงวัย จึงจะสามารถกระตุ้นให้เกิดความตื่นตัวของอารมณ์ได้ โดยคำบรรยายสถานการณ์ทั้งหมดได้รับการตรวจสอบเนื้อหาและความเหมาะสมต่อการนำไปใช้งานโดยนักจิตวิทยา 3 ท่าน

3.2.1.2 คนตรีที่เหมาะสมกับการกระตุ้นอารมณ์ในแต่ละกลุ่ม คัดเลือกโดยนักดนตรีบำบัด ซึ่งศึกษาและอ้างอิงเอกสารที่เกี่ยวข้องเกี่ยวกับเสียงเพลงที่ส่งผลต่ออารมณ์ต่าง ๆ ของมนุษย์ โดยเสียงดนตรีทั้งหมดได้รับการตรวจสอบความเหมาะสมต่อการนำไปใช้งานโดยนักดนตรีบำบัดอีก 3 ท่าน

3.2.2 นักจิตวิทยา

ในกระบวนการทดลอง นักจิตวิทยาจะต้องดำเนินการกระตุ้นอารมณ์เป้าหมาย สังเกตอารมณ์ การแสดงออก และพฤติกรรมที่เกิดขึ้น รวมทั้งดำเนินการสนทนากึ่งโครงสร้างเพื่อให้ผู้เข้าร่วมวิจัยได้แสดงออกทางอารมณ์ทั้งทางวจนภาษาและอวจนภาษา เพื่อให้ระบบบันทึกข้อมูลสามารถบันทึกข้อมูลได้ นอกจากนี้ หากพบว่า ผู้เข้าร่วมวิจัยมีความไม่สบายใจอันเกิดมาจากการบวนการกระตุ้นอารมณ์และกระบวนการระลึกถึงเหตุการณ์ของตน นักจิตวิทยาจะต้องสามารถให้การช่วยเหลือด้านจิตใจและอารมณ์เบื้องต้น และส่งต่อข้อมูลให้คณะผู้วิจัยให้การดูแลและช่วยเหลือต่อไป

3.2.3 แบบสัมภาษณ์กึ่งโครงสร้าง

แบบสัมภาษณ์กึ่งโครงสร้างสำหรับใช้ในการบวนการให้คำปรึกษานี้ พัฒนาขึ้นโดยจำลองโครงสร้างการพูดคุยของกระบวนการให้คำปรึกษา มีวัตถุประสงค์เพื่อให้ผู้เข้าร่วมวิจัยเกิดการแสดงออกทางอารมณ์ที่สอดคล้องกับกระบวนการกระตุ้นอารมณ์ที่ได้รับ (อารมณ์ตอนนี้เป็นอย่างไร) โดยแบบสัมภาษณ์กึ่งโครงสร้างจะประกอบไปด้วยข้อคำถามที่ครอบคลุมหัวข้อต่าง ๆ ดังนี้

3.2.3.1 เหตุการณ์ที่ได้อ่านไปเมื่อสักครู่ ใกล้เคียงกับเรื่องราวของคุณอย่างไร

3.2.3.2 หลังจากที่ได้ฟังเพลงและอ่านคำบรรยายสถานการณ์แล้ว คุณนึกถึงเรื่องราวหรือภาพเหตุการณ์อะไร

3.2.3.3 เหตุการณ์ที่นึกถึงนั้น เรื่องราวเป็นอย่างไร เกิดขึ้นได้อย่างไร เกิดขึ้นที่ไหน มีใครเกี่ยวข้องบ้าง และอารมณ์ ณ ตอนนั้นเป็นอย่างไร

3.2.3.4 เครื่องมือบันทึกข้อมูลภาพและเสียง

- 1) คอมพิวเตอร์หรือแท็บเล็ตที่มีกล้องหน้าความละเอียด 720p ขึ้นไป
- 2) ไมโครโฟนที่สามารถตัดเสียงรบกวนได้และให้เสียงพูดที่ชัดเจนต่อเนื่อง
- 3) หูฟังที่สามารถได้ยินเสียงจากคู่สนทนาได้ชัดเจน ต่อเนื่อง โดยสามารถปรับความดังของเสียงได้ตามความเหมาะสมของผู้ใช้งาน
- 4) ห้องสนทนาที่มีความเป็นส่วนตัว เจียบ ไม่มีเสียงรบกวนจากภายนอก และมีแสงสว่างเพียงพอที่ทำให้คู่สนทนามองเห็นภาพจากกล้องได้ชัดเจน

3.2.4 สัญญาณอินเทอร์เน็ตความเร็วสูง เพื่อการควบคุมคุณภาพของข้อมูลที่ได้ ในงานวิจัยนี้ได้กำหนดความเร็วขั้นต่ำของอินเทอร์เน็ตไว้ที่ความเร็วในการดาวน์โหลดและอัปโหลดข้อมูลต้องไม่ต่ำกว่า 50 เมกะบิตต่อวินาที และความเร็วในการเชื่อมต่อต้องไม่เกิน 10 มิลลิวินาที และคุณภาพของวิดีโอที่ได้ต้องสั่นไหว ไม่กระตุก หรือล่าช้า เสียงและภาพของคู่สนทนาต้องมีความต่อเนื่อง

3.3 ขั้นตอนการวิจัย

กระบวนการวิจัยในแต่ละครั้งใช้เวลาดำเนินการประมาณ 90-120 นาที ผ่านโปรแกรม Zoom โดยกระบวนการดังกล่าวแบ่งออกเป็น 3 ช่วง แต่ละช่วงใช้เวลาตามความเหมาะสม ซึ่งกลุ่มตัวอย่างที่ผ่านคุณสมบัติการเข้าร่วมแต่ละคนจะได้รับการสุ่มเข้าเงื่อนไขการกระตุ้นอารมณ์ คนละ 1 อารมณ์

3.3.1 ก่อนการทดลอง

3.3.1.1 ก่อนเข้าสู่กระบวนการทดลอง นักจิตวิทยาจะเริ่มต้นทักทายและพูดคุยเบื้องต้นเพื่อให้กลุ่มตัวอย่างลดความตื่นตัวต่อสถานการณ์ใหม่ และเข้าสู่ภาวะอารมณ์ที่เป็นกลาง จากนั้นจึงถูกขอให้แสดงสีหน้าท่าทาง และพูดด้วยน้ำเสียงเป็นกลาง โดยข้อความที่จะต้องพูดคือ “อ่านหนังสือพิมพ์วันนี้หรือยัง พรุ่งนี้เป็นวันหยุดราชการนะ” ซึ่งเป็นข้อความที่เป็นกลาง ไม่มีการสื่อสารทางอารมณ์ใดอารมณ์หนึ่ง เพื่อให้ผู้วิจัยได้เก็บบันทึกข้อมูลการแสดงออกทางอวัจนภาษาผ่านสีหน้าและน้ำเสียงที่เป็นกลางหรือเป็นอารมณ์ปกติของแต่ละคนไว้

3.3.1.2 กลุ่มตัวอย่างจะได้รับคำแนะนำเกี่ยวกับจินตนาการ

3.3.2 ระยะทดลอง

ระยะนี้ แบ่งออกเป็น 2 ขั้นตอนหลัก ได้แก่

3.3.2.1 ขั้นกระตุ้นอารมณ์ กลุ่มตัวอย่างจะได้รับคำบรรยายสถานการณ์ 4-6 สถานการณ์ที่สอดคล้องกับช่วงวัยและอารมณ์ตามเงื่อนไขที่ได้รับ จากนั้นเลือกสถานการณ์ที่ใกล้เคียงกับประสบการณ์ทางอารมณ์ของตนเองมากที่สุด และเข้าสู่งานจินตนาการถึงสถานการณ์ดังกล่าว โดยให้กลุ่มตัวอย่างได้จินตนาการย้อนกลับไปถึงเหตุการณ์ ให้รับรู้ถึงประสบการณ์ผ่านประสาทสัมผัสทั้งหมดของร่างกาย ทั้งภาพที่เห็น เสียงที่ได้ยิน กลิ่นที่รับรู้ สัมผัสทางร่างกาย และความคิด ความรู้สึกที่เกิดขึ้นกับตนเองในสถานการณ์นั้น ระหว่างการจินตนาการจะใช้เพลงเพื่อช่วยกระตุ้นอารมณ์ ใช้เวลาจินตนาการถึงเหตุการณ์ 10 นาที

3.3.2.2 ขั้นพูดคุย หลังจากผ่านช่วงกระตุ้นอารมณ์โดยใช้จินตนาการแล้วจะเข้าสู่กระบวนการพูดคุย ใช้เวลา 20-25 นาที โดยนักจิตวิทยาจะใช้คำถามตามแบบสัมภาษณ์กึ่งโครงสร้างเพื่อเน้นให้กลุ่มตัวอย่างได้สำรวจอารมณ์ความรู้สึกที่เกิดขึ้น พูดคุย และมีการแสดงออกทางอารมณ์ทั้งทางวัจนภาษาและอวัจนภาษา เพื่อให้ระบบบันทึกข้อมูลสามารถบันทึกข้อมูลได้

3.3.3 หลังการทดลอง

สำหรับกลุ่มตัวอย่างที่ได้รับการกระตุ้นอารมณ์เชิงบวก (มีความสุข ผ่อนคลาย) หลังจากเก็บข้อมูล ผู้วิจัยได้อธิบายกระบวนการวิจัยทั้งหมดให้กลุ่มตัวอย่างรับทราบก่อนสิ้นสุดการเก็บข้อมูล

สำหรับกลุ่มตัวอย่างที่ได้รับการกระตุ้นอารมณ์เชิงลบ (โกรธ เครียด หรือเศร้า) หลังจากเก็บข้อมูล ผู้วิจัยจะปรับอารมณ์โดยกระตุ้นอารมณ์เชิงบวกเพื่อให้กลุ่มตัวอย่างกลับเข้าสู่ภาวะอารมณ์ที่เป็นกลาง และฝึกเทคนิคผ่อนคลายความเครียดเพิ่มเติมในกรณีที่เป็น เช่น เมื่อการกระตุ้นอารมณ์เชิงบวกไม่สามารถปรับอารมณ์ได้ มีการติดตามผลกระทบทางอารมณ์ที่เกิดจากการเข้าร่วมการวิจัยภายใน 1-2 สัปดาห์ ในกรณีที่กลุ่มตัวอย่างยังมีอารมณ์เชิงลบที่เป็นผลมาจากการเข้าร่วมโครงการหรือได้รับผลกระทบทางจิตใจอย่างต่อเนื่อง ผู้วิจัยจะพิจารณาประสานงานด้านการดูแลสุขภาพจิตจากผู้เชี่ยวชาญของโรงพยาบาลศิริราชต่อไป โดยผู้วิจัยจะเป็นผู้รับผิดชอบค่าใช้จ่ายทั้งหมด

3.4 การจัดทำข้อมูล

การจัดทำข้อมูลหลังจากการเก็บข้อมูลเสร็จสิ้น สามารถแบ่งออกเป็น 2 ขั้นตอนหลัก โดยแต่ละขั้นตอนมีรายละเอียด ดังนี้

3.4.1 การเตรียมข้อมูล

ภายหลังจากการเก็บข้อมูล จะได้ข้อมูลดิบเป็นคลิปวิดีโอ การพูดคุยของกลุ่มตัวอย่างในแต่ละอารมณ์ ความยาว 15-25 นาที ผู้วิจัยนำข้อมูลดังกล่าว สกัดออกมาเป็นชุดข้อมูล 2 รูปแบบ เพื่อนำไปใช้ในการระบุอารมณ์ ต่อไป โดยข้อมูลดังกล่าว ได้แก่

3.4.1.1 ภาพการแสดงออกทางอารมณ์ผ่านทางใบหน้า ลักษณะเป็นภาพนิ่งความละเอียด อยู่ที่ 5 วินาทีต่อเฟรม

3.4.1.2 เสียง ลักษณะเป็นคลิปเสียง ที่แบ่งช่วงตามระยะเวลาที่หยุดพูด 0.5 วินาทีขึ้นไป ทำให้ คลิปเสียงที่ได้ไม่ได้แบ่งตามเนื้อหาคำพูด จะได้คลิปเสียงที่มีความยาวตั้งแต่ 2-20 วินาที ขึ้นอยู่กับลักษณะการพูด

3.4.2 การระบุอารมณ์

เมื่อได้ชุดข้อมูลของกลุ่มตัวอย่างแต่ละรายแล้ว ผู้วิจัยนำข้อมูลทั้ง 2 ประเภทไประบุอารมณ์ ข้อมูล 1 ชุด โดยใช้ผู้ระบุอารมณ์ 2 กลุ่ม คือ 1) กลุ่มนักจิตวิทยา ซึ่งเป็นผู้มีพื้นฐานความรู้ด้านการวิเคราะห์ อารมณ์ และ 2) กลุ่มบุคคลทั่วไป การระบุอารมณ์ข้อมูลการแสดงออกทางอวัจนภาษามีการแบ่งกลุ่มอารมณ์ อ้างอิงจากทฤษฎีของอารมณ์ โดยแบ่งอารมณ์ที่แสดงออกที่สังเกตเห็นได้เป็น 5 กลุ่มหลัก ได้แก่ กลุ่มอารมณ์ ทางลบที่มีพลังงาน อารมณ์ทางลบที่ไม่มีพลังงาน อารมณ์ทางบวกที่มีพลังงาน อารมณ์ทางบวกที่ไม่มีพลังงาน อารมณ์เป็นกลาง และมีการเพิ่มกลุ่มอารมณ์อื่น ๆ (other) ที่มีความหมายสำหรับงานวิจัยนี้โดยเฉพาะ ซึ่งความหมายของแต่ละกลุ่มและตัวอย่างการแสดงออกทางอารมณ์ในแต่ละกลุ่ม ดังตารางที่ 1

ตารางที่ 1 การแบ่งกลุ่มอารมณ์ของการระบุอารมณ์ของข้อมูล

การจัดกลุ่มอารมณ์	ความหมาย	ตัวอย่างอารมณ์	ประเภทข้อมูล
1. อารมณ์ทางลบที่มีพลังงาน	ใช้สำหรับระบุข้อมูลที่มี การแสดงอารมณ์ทางลบ ที่มีความตื่นตัว-มีพลังงานสูง	โกรธ หงุดหงิด เครียด รังเกียจ กลัว กังวล กระวนกระวาย ตกใจกลัว	ภาพและเสียง
2. อารมณ์ทางลบที่ไม่มี พลังงาน	ใช้สำหรับระบุข้อมูลที่มี การแสดงอารมณ์ทางลบ ที่ไม่ตื่นตัว-มีพลังงานต่ำ	เศร้า เสียใจ หดหู่ เบื่อหน่าย เหนื่อย รู้สึกผิด	ภาพและเสียง
3. อารมณ์ทางบวกที่มีพลังงาน	ใช้สำหรับระบุข้อมูลที่มี การแสดงอารมณ์ทางบวก ที่มีความตื่นตัว-มีพลังงานสูง	มีความสุข ตื่นเต้น กระตือรือร้น สนใจ มีความหวัง	ภาพและเสียง
4. อารมณ์ทางบวกที่ไม่มี พลังงาน	ใช้สำหรับระบุข้อมูลที่มี การแสดงอารมณ์ทางบวก ที่ไม่ตื่นตัว-มีพลังงานต่ำ	ผ่อนคลาย สงบ สงบ รู้สึกว่า ทุกอย่างอยู่ในการควบคุม บั่นใจ	ภาพและเสียง

การจัดกลุ่มอารมณ์	ความหมาย	ตัวอย่างอารมณ์	ประเภทข้อมูล
5. อารมณ์เป็นกลาง	ใช้สำหรับระบุข้อมูลที่มี การแสดงออกถึงอารมณ์เฉย ๆ หรือไม่มีการแสดงออก ทางอารมณ์	การแสดงออกมีลักษณะ ใกล้เคียงกับข้อมูล neutral	ภาพและเสียง
6. อารมณ์อื่น ๆ	ใช้สำหรับระบุข้อมูลที่มี การแสดงอารมณ์อื่น ๆ นอกเหนือจากกลุ่มอารมณ์ ทั้ง 5	ตื่นตัว (มีพลัง แต่ไม่มีหัว) ง่วง (ไม่มีพลัง และไม่มีหัว) การระลึกถึง หรือนึกคิด (อาจมีความตื่นตัว แต่ไม่มีหัว) อารมณ์ผสมระหว่างบวก และลบ จนไม่สามารถบอกได้ว่า เป็นหัวใดมากกว่ากัน อารมณ์ที่มีความซับซ้อน จนไม่สามารถจัดหมวดหมู่ ใน 5 กลุ่มได้	ภาพและเสียง
7. ไม่สามารถระบุได้ (none)	ใช้สำหรับเมื่อข้อมูลนั้น ไม่สามารถระบุอารมณ์ได้ เนื่องจากลักษณะข้อมูล เป็นตามเกณฑ์ฟัลส์เสีย	-	ภาพและเสียง
8. อารมณ์หลากหลาย (multi)	ใช้สำหรับระบุเมื่อข้อมูลนั้น มีลักษณะการแสดงออกทาง อารมณ์ที่มากกว่า 1 อารมณ์	-	เสียง

หมายเหตุ

A การระบุอารมณ์จากภาพ “none” มีเกณฑ์ ดังนี้

1. อวัยวะสำคัญบนใบหน้า (ได้แก่ คิ้ว 2 ข้าง ลูกตา 2 ข้าง โหนกแก้ม 2 ข้าง จมูก ปาก) ไม่ครบ จากการหันข้างจนมองไม่เห็นใบหน้าอีกฝั่ง มีวัตถุบังใบหน้า
บางส่วน การหลับตา
2. ภาพเบลอจนไม่สามารถระบุตำแหน่งอวัยวะที่ชัดเจนได้

B การระบุอารมณ์หลากหลายจากเสียง มีเกณฑ์ ดังนี้

1. มีเสียงคนพูดมากกว่า 1 คน พูดพร้อมกัน
2. ในคลิปนั้นมีเสียงคนพูดมากกว่า 1 คน ในคลิปเดียวกัน
3. มีการออกเสียงที่ไม่เกี่ยวข้องกับการแสดงอารมณ์ เช่น เสียงไอ เสียงจาม
4. มีเสียงรบกวนอื่น ๆ จากสิ่งแวดล้อม เช่น เสียงนก เสียงรถ เสียงลมพัด เสียงช้างของอุปกรณ์อิเล็กทรอนิกส์ เสียงทุบโต๊ะ

3.4.2.1 ข้อมูลภาพการแสดงออกทางอารมณ์ผ่านทางใบหน้า จะใช้วิธีระบุโดยสังเกตจาก
ลักษณะการเปลี่ยนแปลงของกล้ามเนื้อต่าง ๆ บนใบหน้า (Action Unit หรือ AU) อ้างอิงลักษณะการแสดงอารมณ์
ผ่านทางสีหน้าตาม FACS อารมณ์ที่ระบุ แบ่งเป็น 7 กลุ่ม โดยภาพที่นำมาใช้ระบุอารมณ์ได้ต้องมีเกณฑ์ ดังนี้

- 1) มืองค์ประกอบของใบหน้าครบถ้วน ได้แก่ ตาทั้ง 2 ข้าง คิ้วทั้ง 2 ข้าง โหนกแก้ม
ทั้ง 2 ข้าง จมูก ปาก และคาง โดยไม่มีสิ่งใดปิดบัง
- 2) อวัยวะสำคัญต้องมีความคมชัดเพียงพอ

3.4.2.2 ข้อมูลการแสดงออกทางอารมณ์ผ่านทางน้ำเสียง จะใช้วิธีระบุโดยสังเกตจากลักษณะของความสูงต่ำของเสียง ความดังของเสียง จำนวนครั้งที่หยุดพูด ความเร็วในการพูด ความเร็วในการออกเสียง การเน้นเสียง และเสียงอื่น ๆ อ้างลักษณะการแสดงอารมณ์ผ่านทางน้ำเสียงจากแนวคิด emotion in speech อารมณ์ที่ระบุจะแบ่งเป็น 8 กลุ่ม โดยคลิปเสียงที่นำมาใช้ระบุอารมณ์ได้จะต้องมีเกณฑ์ ดังนี้

- 1) ต้องเป็นเสียงที่เปล่งออกมาจากกลุ่มตัวอย่างเพียงคนเดียว โดยไม่มีเสียงของผู้อื่น หรือสิ่งแวดล้อมรอบตัวแทรกเข้ามา
- 2) เสียงดังกล่าวไม่จำเป็นต้องเป็นคำพูดหรือคำที่มีความหมาย

3.5 สถิติและการวิเคราะห์ข้อมูล

ใช้สถิติเชิงพรรณนาเพื่อรายงานค่าความถี่และร้อยละของจำนวนที่ข้อมูลที่ได้ในแต่ละกลุ่มอารมณ์ รวมถึงค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานอายุของกลุ่มตัวอย่างในชุดข้อมูลนี้

4. การทบทวนวรรณกรรม

การศึกษาครั้งนี้ มีเป้าหมายเพื่อสร้างฐานข้อมูลการแสดงออกทางอารมณ์ผ่านทางอวัจนภาษาในบริบทการบำบัดทางจิตวิทยาของคนไทย โดยผู้วิจัยได้ทบทวนและรวบรวมแนวคิด ทฤษฎี หรือองค์ความรู้ที่เกี่ยวข้อง 3 หัวข้อ ได้แก่ ทฤษฎีอารมณ์ที่ใช้กำหนดหมวดหมู่กลุ่มอารมณ์ ลักษณะการแสดงออกทางอารมณ์ผ่านทางอวัจนภาษา และฐานข้อมูลการแสดงออกทางอวัจนภาษาที่มีอยู่ในปัจจุบัน โดยมีรายละเอียด ดังนี้

4.1 ทฤษฎีอารมณ์

การศึกษานี้อ้างอิงทฤษฎีกล่ออารมณ์ ซึ่งพัฒนาโดย Russell (1980) เป็นการจำลองอารมณ์ออกเป็น 2 มิติ ดังภาพที่ 1 ประกอบด้วย 2 แกน ได้แก่ แกน X บอกความพึงพอใจ (valence) และแกน Y บอกความตื่นตัว (arousal) ทำให้อารมณ์ต่าง ๆ ถูกจัดกลุ่มออกเป็น 4 กลุ่ม โดยแต่ละกลุ่มมีตัวอย่างอารมณ์ที่สามารถเข้าใจได้ง่าย ดังนี้

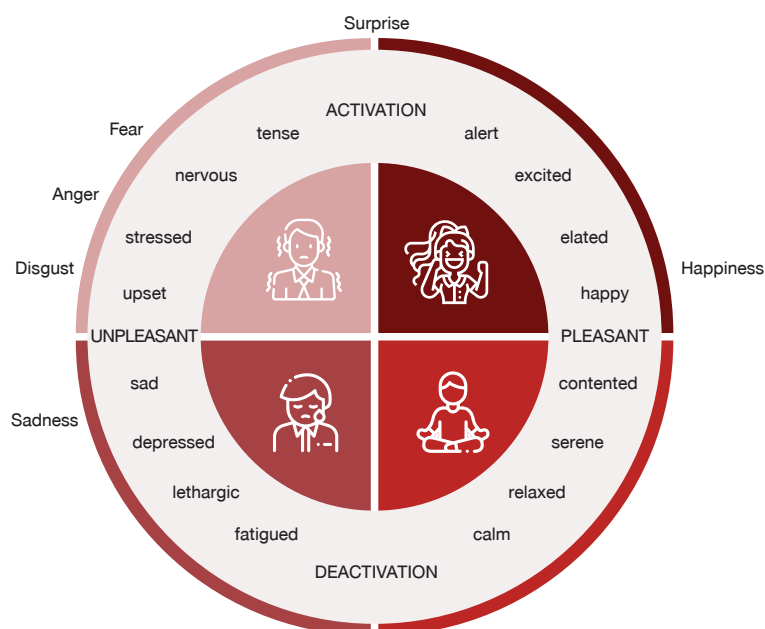
4.1.1 อารมณ์ทางลบที่มีพลังงาน เช่น โกรธ โมโห หงุดหงิด กังวล ตึงเครียด หวาดกลัว

4.1.2 อารมณ์ทางลบที่ไม่มีพลังงาน เช่น เศร้า เสียใจ เบื่อหน่าย ท้อแท้

4.1.3 อารมณ์ทางบวกที่มีพลังงาน เช่น มีความสุข ร่าเริง ดีใจ สนุกสนาน

4.1.4 อารมณ์ทางบวกที่ไม่มีพลังงาน เช่น ผ่อนคลาย มั่นใจ สงบ

การจัดกลุ่มอารมณ์ตามแนวคิดนี้ ถือเป็นการระบุอารมณ์ที่พบอย่างกว้าง ๆ ทำให้ครอบคลุมลักษณะการแสดงอารมณ์ต่าง ๆ ได้ดีกว่าการระบุชื่ออารมณ์อย่างเฉพาะเจาะจง อีกทั้งอารมณ์ต่าง ๆ ที่อยู่ในกลุ่มเดียวกัน เช่น โกรธและเคียด หรือเศร้าและท้อแท้ เมื่ออยู่ในบริบทของการบำบัดหรือการให้คำปรึกษาทางจิตวิทยา ผู้บำบัดแสดงการตอบสนองต่ออารมณ์ไม่แตกต่างกัน



ภาพที่ 1 วงล้ออารมณ์

ที่มา: ปรับปรุงจาก Russell (1980)

4.2 ลักษณะการแสดงอารมณ์ผ่านอวัจนภาษา

ในการศึกษานี้ ผู้วิจัยได้ทบทวนและรวบรวมแนวคิดหรือทฤษฎีที่เกี่ยวข้องกับลักษณะการแสดงออกทางน้ำเสียง และสีหน้าในแต่ละอารมณ์ จากการงานวิจัยของ Cordaro et al. (2018); Keltner et al. (2019); Matsumoto & Willingham (2009) และ Sobin & Alpert (1999) และคัดเลือกเฉพาะลักษณะการแสดงออกที่มีความเป็นสากล หรือกล่าวได้ว่า มีการแสดงออกที่สอดคล้องกันทั้งในวัฒนธรรมตะวันออกและตะวันตก เพื่อให้ได้เกณฑ์การแสดงออกทางอารมณ์ผ่านอวัจนภาษาเบื้องต้นสำหรับแต่ละอารมณ์ ดังตารางที่ 2

ตารางที่ 2 เกณฑ์การแสดงออกผ่านใบหน้าและน้ำเสียงตามอารมณ์ต่าง ๆ ที่กำหนดในการศึกษา

อารมณ์	การแสดงออกทางใบหน้าอ้างอิง FACS	การแสดงออกทางน้ำเสียง
โกรธ	AU04 คิ้วกดลง-หมวดเข้าหากัน AU07 เปลือกตาบน-ล่างตึงขึ้น (ขยับเข้าหากัน) AU24 ปากกดเข้าหากัน	- เสียงสูงขึ้น - ดังขึ้น - พูดเร็วมากขึ้น
เครียด	AU01 หัวคิ้วด้านในยกขึ้น AU17 คางยกสูงขึ้น AU25 ริมฝีปากแยกออกจากกัน AU26 สันกรมตกล	- เสียงสูงขึ้น - มีการลงน้ำหนักเสียงหนักขึ้น (กระแทกเสียง) - พูดเร็วมากขึ้น - ไม่ค่อยมีจังหวะเว้นช่วงในการพูดแต่ละคำ
เศร้า	AU01 หัวคิ้วด้านในยกขึ้น AU04 คิ้วกดลง หมวดเข้าหากัน AU15 มุมปากกดลง	- เสียงเบาลง - เสียงต่ำลง - พูดช้าลง - มีการเว้นช่วงในการพูดแต่ละคำนาน - เสียงมีความสั่น
มีความสุข	AU06 โหนกแก้มยกขึ้นไปดันตา AU07 เปลือกตาบน-ล่างตึงขึ้น (ขยับเข้าหากัน) AU12 มุมปากยกขึ้น AU25 ริมฝีปากแยกออกจากกัน	- เสียงดังขึ้น - เสียงมีความเหวี่ยง เคี้ยวสูง เคี้ยวต่ำ - ความเร็วในการพูดปกติ ไม่ได้เปลี่ยนจากช่วงไม่แสดงอารมณ์
ผ่อนคลาย	AU12 มุมปากยกขึ้น AU25 ริมฝีปากแยกออกจากกัน AU26 สันกรมตกล AU43 เปลือกตาปิดลง (มีการกะพริบตาช้าลงกว่าเดิม)	- เสียงต่ำลง - เสียงเบาลง - มีการเว้นช่วงในแต่ละคำนาน

4.3 ฐานข้อมูลการแสดงออกทางอวัจนภาษาในปัจจุบัน

การวิเคราะห์อารมณ์จากอวัจนภาษาสามารถนำไปประยุกต์ใช้ได้ค่อนข้างหลากหลาย ซึ่งการสร้างโมเดลสำหรับวิเคราะห์อารมณ์ จำเป็นต้องมีฐานข้อมูลที่สามารถจำแนกเป็นอารมณ์ต่าง ๆ ได้ ในปัจจุบันฐานข้อมูลการแสดงออกทางอวัจนภาษา ทั้งข้อมูลน้ำเสียงและภาพการแสดงออกทางใบหน้าที่สื่อถึงอารมณ์ต่าง ๆ ผู้วิจัยได้รวบรวมมา ดังนี้

4.3.1 ฐานข้อมูลเสียง

ปัจจุบันมีการสร้างฐานข้อมูลเสียงที่ใช้เฉพาะในแต่ละประเทศหรือภูมิภาค โดยแต่ละฐานข้อมูลมีการเก็บรวบรวมข้อมูลด้วยวิธีการต่าง ๆ โดยฐานข้อมูลที่ได้รับความสนใจ ดังนี้

4.3.1.1 RAVDESS เป็นฐานข้อมูลเสียงพูดภาษาอังกฤษสำเนียงอเมริกัน อัตราความเร็วในการพูดปานกลาง ออกเสียงชัดเจน ไม่ค่อยพบลักษณะการรวบคำ โดยแต่ละอารมณ์มีการพูดในประโยคเดียวกัน แต่เปลี่ยนอารมณ์เพื่อให้เห็นได้ชัดเจน ชุดข้อมูลแบ่งออกเป็นเสียงพูดปกติและเสียงร้องเพลง เก็บข้อมูลโดยใช้นักแสดง อัดเสียงจากการจำลองภาพแวดล้อมการพูดแบบปิด (close talk) จำแนกอารมณ์ตามทฤษฎีอารมณ์พื้นฐาน (Luna-Jiménez et al., 2021)

4.3.1.2 Emo DB เป็นฐานข้อมูลเสียงพูดภาษาเยอรมันสำเนียงเบอร์ลิน อัตราความเร็วในการพูดสูง พบลักษณะการรวบคำค่อนข้างบ่อย เก็บข้อมูลโดยใช้นักแสดง อัดเสียงจากการจำลองภาพแวดล้อมการพูดแบบปิด (Zhang et al., 2016)

4.3.1.3 EMOCAP เป็นฐานข้อมูลเสียงพูดภาษาอังกฤษ สำเนียงหลากหลาย อัตราความเร็วในการพูดปานกลางจนสูง การออกเสียงหลากหลาย เป็นข้อมูลที่นำมาจากยูทูบ (YouTube) และสื่อออนไลน์เป็นหลัก (Busso et al., 2008)

4.3.1.4 Thai-SER เป็นฐานข้อมูลเสียงพูดภาษาไทย เก็บข้อมูลโดยใช้นักแสดง อัดเสียงจากสภาพแวดล้อมแบบปิดและแบบเปิด

4.3.1.5 EMOLA เป็นฐานข้อมูลเสียงพูดภาษาไทย เป็นข้อมูลที่นำมาจากสื่อละครไทย (Kasuriya et al., 2019)

4.3.2 ฐานข้อมูลภาพ

ฐานข้อมูลการแสดงออกทางสีหน้าที่มีขนาดใหญ่ และมีข้อมูลจำนวนมากในปัจจุบัน ได้แก่ AffectNet (Mollahosseini et al., 2019) ซึ่งมีการเก็บรวบรวมภาพใบหน้าที่เป็นภาพนิ่ง (still image) จากอินเทอร์เน็ต และระบุอารมณ์ตามทฤษฎีอารมณ์พื้นฐานร่วมกับการบอกระดับของอารมณ์

จากฐานข้อมูลการแสดงออกทางอารมณ์ผ่านสีหน้าและน้ำเสียงที่มีอยู่ในปัจจุบัน นับว่าเป็นข้อมูลที่มีความชัดเจนและนำไปใช้ประโยชน์ได้หลากหลาย แต่บริบทของการแสดงอารมณ์อาจมีความคลาดเคลื่อนไปจากการพูดคุยตามรูปแบบทั่วไปที่มักไม่ได้แสดงอารมณ์อย่างชัดเจนนัก ผู้วิจัยจึงเปรียบเทียบลักษณะข้อมูลจากฐานข้อมูลที่มีอยู่กับฐานข้อมูลที่ได้จากการจำลองกระบวนการให้คำปรึกษาในการศึกษานี้ ดังตารางที่ 3

ตารางที่ 3 แสดงความแตกต่างระหว่างฐานข้อมูลการแสดงอารมณ์ที่มีในปัจจุบัน เปรียบเทียบกับฐานข้อมูลจากงานวิจัยครั้งนี้

	ฐานข้อมูลเสียง		ฐานข้อมูลภาพ	
	Thai-SER	งานวิจัยครั้งนี้	AffectNet	งานวิจัยครั้งนี้
ผู้ให้ข้อมูล	นักแสดง (ไทย)	บุคคลทั่วไป (ไทย)	บุคคลที่สามารถสืบค้นภาพใบหน้าได้จากเว็บไซต์	บุคคลทั่วไป (ไทย)
บริบทของข้อมูล	การแสดงแบบมีประโยคให้และการแสดงแบบบทกบถสมมติ	กระบวนการให้คำปรึกษาตามภาวะอารมณ์	ไม่จำกัด	กระบวนการให้คำปรึกษาตามภาวะอารมณ์

	ฐานข้อมูลเสียง		ฐานข้อมูลภาพ	
	Thai-SER	งานวิจัยครั้งนี้	AffectNet	งานวิจัยครั้งนี้
การบันทึกข้อมูล	สตูดิโอ อินเทอร์เน็ตผ่าน Zoom	อินเทอร์เน็ตผ่าน Zoom	-	อินเทอร์เน็ตผ่าน Zoom
ลักษณะของข้อมูล	เป็นประโยคย่อย ๆ	เป็นการแสดงน้ำเสียงหรือประโยค	ภาพนิ่งที่สีหน้าชัดเจน	ภาพนิ่งที่แสดงการเคลื่อนไหวต่อเนื่อง

5. ผลการศึกษา

5.1 ลักษณะข้อมูลที่ได้

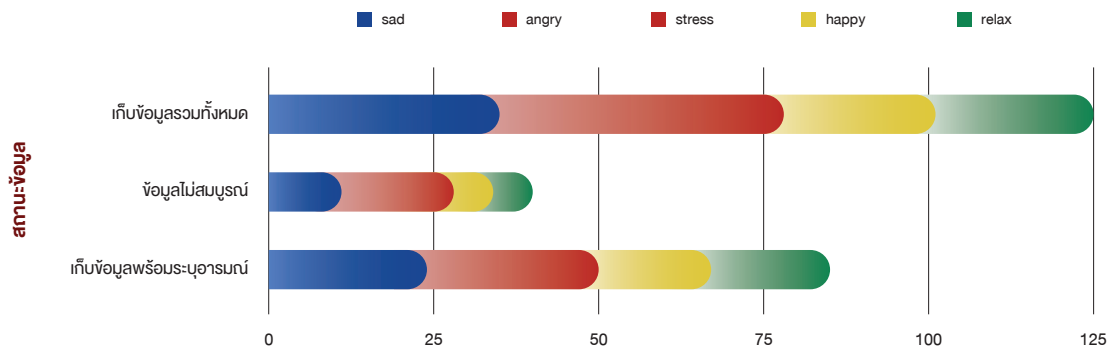
จากการเก็บข้อมูลตั้งแต่เดือนกันยายน พ.ศ. 2564 ถึงเดือนกุมภาพันธ์ พ.ศ. 2566 มีผู้สนใจและลงทะเบียนเข้าร่วมโครงการจำนวนทั้งสิ้น 587 ราย ทั้งนี้ มีผู้ที่คะแนนสุขภาพจิตผ่านเกณฑ์ที่กำหนด 286 ราย และดำเนินการเก็บข้อมูลเสร็จสิ้น 125 ราย โดยคลิปวิดีโอทั้งหมดสามารถนำไปใช้ระบุอารมณ์และระบุอารมณ์เสร็จสิ้นแล้วจำนวน 85 คลิป กลุ่มตัวอย่างเป็นเพศหญิงและชาย มีช่วงอายุตั้งแต่ 18-55 ปี ใช้สถิติเชิงพรรณนาแสดงรายละเอียดดังตารางที่ 4 และ 5 ดังนี้

ตารางที่ 4 จำนวนผู้เข้าร่วมโครงการทั้งหมด จำแนกเป็นกรณีเก็บข้อมูลได้และเก็บข้อมูลพร้อมระบุอารมณ์

ข้อมูลส่วนตัว	เก็บข้อมูลพร้อมระบุอารมณ์	ข้อมูลไม่สมบูรณ์	เก็บข้อมูลรวมทั้งหมด
เพศหญิง	52 (ร้อยละ 61.2)	30 (ร้อยละ 75)	82 (ร้อยละ 65.6)
เพศชาย	33 (ร้อยละ 38.8)	10 (ร้อยละ 25)	43 (ร้อยละ 34.4)
รวม	85 (ร้อยละ 100)	40 (ร้อยละ 100)	125 (ร้อยละ 100)
อายุ	อายุเฉลี่ย 29 SD 8 (18-55)	อายุเฉลี่ย 30 SD 10 (18-53)	อายุเฉลี่ย 29 SD 9 (18-55)

ตารางที่ 5 จำนวนข้อมูล จำแนกตามอารมณ์ที่กระตุ้น

สถานะข้อมูล	Mood Induction					รวม
	sad (neg-de)	angry (neg-ac)	stress (neg-ac)	happy (pos-ac)	relax (pos-de)	
เก็บข้อมูล รวมทั้งหมด	35 (ร้อยละ 28)	21 (ร้อยละ 16.8)	22 (ร้อยละ 17.6)	23 (ร้อยละ 18.4)	24 (ร้อยละ 19.2)	125 (ร้อยละ 100)
ข้อมูลไม่สมบูรณ์	11 (ร้อยละ 27.5)	9 (ร้อยละ 22.5)	8 (ร้อยละ 20)	6 (ร้อยละ 15)	6 (ร้อยละ 15)	40 (ร้อยละ 100)
เก็บข้อมูล พร้อมระบุอารมณ์	24 (ร้อยละ 28.2)	12 (ร้อยละ 14.1)	14 (ร้อยละ 16.5)	17 (ร้อยละ 20)	18 (ร้อยละ 21.2)	85 (ร้อยละ 100)



ภาพที่ 2 สัดส่วนคลิปวิดีโอตามสถานะ: จำแนกตามอารมณ์ที่ระบุ

ข้อมูลที่เก็บได้ทั้งหมด ส่วนมากเป็นข้อมูลจากสถานการณ์กระตุ้นอารมณ์ทางลบที่มีพลังงาน ประกอบด้วยอารมณ์โกรธและเครียด (สีแดง ดังภาพที่ 2) โดยคิดเป็นร้อยละ 33.4 ของข้อมูลที่เก็บได้ทั้งหมด รองลงมา คือ กลุ่มอารมณ์ทางลบที่ไม่มีพลังงาน คือ อารมณ์เศร้า คิดเป็นร้อยละ 28 (สีน้ำเงิน ดังภาพที่ 2) ตามด้วยอารมณ์ทางบวกที่ไม่มีพลังงาน และอารมณ์ทางบวกที่มีพลังงานในสัดส่วนที่ใกล้เคียงกัน โดยข้อมูลในกลุ่มอารมณ์ทางบวกทั้ง 2 รูปแบบ มีสัดส่วนที่น้อยกว่าข้อมูลในกลุ่มอารมณ์ทางลบ (สีเหลืองและสีเขียว ดังภาพที่ 2)

ทั้งนี้ จากภาพที่ 2 ยังพบสัดส่วนที่คล้ายคลึงกันในกลุ่มข้อมูลที่ไม่สมบูรณ์ กล่าวคือ คลิปวิดีโอส่วนมากเป็นกลุ่มอารมณ์ทางลบที่มีพลังงาน ตามด้วยอารมณ์ทางลบที่ไม่มีพลังงาน โดยข้อมูลที่ไม่สมบูรณ์นี้เกิดจากปัญหาทางเทคนิคในระหว่างบันทึกข้อมูล ได้แก่ สัญญาณอินเทอร์เน็ตไม่คงที่ ทำให้คลิปวิดีโอกระตุกหรือค้างระหว่างให้ข้อมูล รวมถึงปัญหาด้านสิ่งแวดล้อมและสถานที่ของผู้ร่วมวิจัย เช่น มีเสียงรบกวนเกิดขึ้นในขณะที่ให้ข้อมูล หรือการปรับเปลี่ยนตำแหน่งที่นั่งในระหว่างพูดคุยจนทำให้องค์ประกอบของใบหน้าไม่สมบูรณ์

ส่วนข้อมูลที่น่ามาสร้างฐานข้อมูล เป็นข้อมูลที่ระบุอารมณ์ครบถ้วน มีจำนวนทั้งสิ้น 85 ชุด ซึ่งมีสัดส่วนเป็นอารมณ์ทางลบที่มีพลังงาน มากที่สุดโดยคิดเป็นร้อยละ 30.6 รองลงมา คือ อารมณ์ทางลบที่ไม่มีพลังงาน ร้อยละ 28.2 อารมณ์ทางบวกที่ไม่มีพลังงาน และอารมณ์ทางบวกที่มีพลังงาน ในสัดส่วนที่ใกล้เคียงกัน โดยข้อมูลในกลุ่มอารมณ์ทางบวกทั้ง 2 รูปแบบ มีสัดส่วนที่น้อยกว่าข้อมูลในกลุ่มอารมณ์ทางลบเช่นเดียวกัน

5.2 ข้อมูลภาพและเสียง

จากข้อมูลที่ได้จากกลุ่มตัวอย่างจำนวน 85 คน ภายใต้การจำลองสถานการณ์การให้คำปรึกษาทางจิตวิทยา ใน 4 กลุ่มอารมณ์ โดยมีสัดส่วนที่ใกล้เคียงกัน ทำให้การศึกษานี้ได้ชุดข้อมูลภาพและเสียงจากกลุ่มตัวอย่างที่แตกต่างกันทั้งหมด 85 ชุดข้อมูล ที่มีลักษณะการแสดงออกทางอารมณ์ผ่านทางอวัจนภาษาตามจริงและมีพลวัต ผ่านการสนทนาถึงประสบการณ์ทางอารมณ์ของกลุ่มตัวอย่างแต่ละคนซึ่งมีลักษณะเฉพาะตัวที่แตกต่างกัน ซึ่งข้อมูลทั้ง 85 ชุด ครอบคลุมอารมณ์ที่พบได้บ่อยในบริบทการให้คำปรึกษาทางจิตวิทยา เช่น อารมณ์เศร้า เสียใจ โกรธ เครียด กังวล มีความสุข และผ่อนคลาย โดยข้อมูลในแต่ละส่วนจะมีลักษณะ ดังนี้

5.2.1 ข้อมูลภาพ

จากการตัดภาพที่ได้จากคลิปวิดีโอความยาว 15-25 นาที ข้อมูลที่ได้ 1 ชุด จะมีภาพใบหน้าของกลุ่มตัวอย่างซึ่งเป็นภาพต่อเนื่องจำนวน 3,000-8,000 ภาพ โดยรวมถึงภาพในขณะที่กำลังพูด หรือมีการเคลื่อนไหว อวัยวะบนใบหน้า และภาพระหว่างการเปลี่ยนอารมณ์ และในระหว่างชุดข้อมูลแต่ละชุด จะได้รับการระบุอารมณ์ตามการแบ่งกลุ่มอารมณ์ ซึ่งสัดส่วนอารมณ์จะแตกต่างกันไป ขึ้นอยู่กับเงื่อนไขการกระตุ้นอารมณ์ที่ได้รับ ตัวอย่างในชุดข้อมูลจากกลุ่มตัวอย่างที่ได้รับการกระตุ้นอารมณ์เศร้า อาจพบข้อมูลระบุอารมณ์ทางลบที่ไม่มีพลังงานมากกว่าอารมณ์ในกลุ่มอื่น แต่ยังสามารถพบอารมณ์อื่น ๆ ในข้อมูลชุดนี้ได้ เช่น อารมณ์ทางลบที่มีพลังงานหรืออารมณ์กลุ่มทางบวกได้ แต่ในสัดส่วนที่น้อยกว่า เนื่องจากระหว่างการสนทนาอารมณ์ที่เกิดขึ้นมีความเป็นพลวัตระหว่างนักจิตวิทยาและกลุ่มตัวอย่าง

5.2.2 ข้อมูลเสียง

จากการตัดแบ่งช่วงเสียงตามการเว้นระยะการพูด ข้อมูลที่ได้ 1 ชุด จะมีจำนวนคลิปเสียงของกลุ่มตัวอย่างประมาณ 40-300 ไฟล์ โดยแต่ละไฟล์จะเป็นเสียงของกลุ่มตัวอย่างที่กำลังโต้ตอบกับนักจิตวิทยา ซึ่งอาจเป็นได้ทั้งเสียงตอบรับหรือเสียงที่เปล่งออกมาสั้น ๆ โดยไม่มีความหมายในเชิงภาษา หรือเป็นวลีที่มีใจความไม่ครบถ้วน หรือเป็นประโยคสมบูรณ์ที่มีความยาวหลายวินาที ชุดข้อมูลแต่ละชุดจะได้รับการระบุอารมณ์ตามการแบ่งกลุ่มอารมณ์ ซึ่งสัดส่วนอารมณ์จะแตกต่างกันไป ขึ้นอยู่กับเงื่อนไขการกระตุ้นอารมณ์ที่ได้รับ เช่นเดียวกันกับข้อมูลภาพ

6. การอภิปรายผล

ผลการศึกษาครั้งนี้ ผู้วิจัยได้ฐานข้อมูลการแสดงออกทางอารมณ์ผ่านทางอวัจนภาษา จำแนกออกเป็นข้อมูลภาพใบหน้าและข้อมูลเสียงที่เก็บบันทึกได้จากสถานการณ์จำลองกระบวนการให้คำปรึกษาทางจิตวิทยา ซึ่งการแสดงออกของกลุ่มตัวอย่างทั้งหมดเป็นการแสดงอารมณ์ที่เกิดขึ้นตามจริง และเป็นข้อมูลที่มีความละเอียดในระดับที่เห็นการเปลี่ยนแปลงทางอารมณ์ที่เกิดขึ้นในระหว่างการพูดคุย อีกทั้งยังมีความซับซ้อน เนื่องจากการแสดงออกอาจมีความคลุมเครือ ไม่ชัดเจน ซึ่งเป็นธรรมชาติของการแสดงออกทางอารมณ์ของคนทั่วไป รวมถึงในบริบทการให้บริการทางจิตวิทยา หากเปรียบเทียบกับฐานข้อมูลการแสดงออกทางใบหน้าที่มีอยู่ในปัจจุบัน เช่น AffectNet จะพบว่า ข้อมูลที่มีอยู่มากเป็นภาพใบหน้าที่ตัวแบบกำลังแสดงสีหน้าอารมณ์ใดอารมณ์หนึ่งอยู่ ไม่ใช่ภาพในระหว่างการเคลื่อนไหวหรือภาพต่อเนื่อง ทำให้การวิเคราะห์อารมณ์จากภาพแต่ละภาพทำได้ง่ายกว่า ต่างจากชุดข้อมูลภาพที่ได้จากการศึกษานี้ ที่ข้อมูลเป็นลักษณะภาพที่มีความต่อเนื่องและอาจพบการแสดงออกผ่านสีหน้าที่ชัดเจนได้น้อย ฐานข้อมูลของการศึกษานี้จึงช่วยให้โมเดลวิเคราะห์อารมณ์ได้เรียนรู้อารมณ์ที่ละเอียดกว่า จนถึงในระดับที่เกิดการเปลี่ยนแปลงจากอารมณ์หนึ่งสู่อารมณ์หนึ่ง

ทั้งนี้ เมื่อเทียบกับฐานข้อมูลเสียงที่มีอยู่ในปัจจุบัน เช่น Thai-SER ที่ได้ข้อมูลการแสดงออกทางอารมณ์จากนักแสดงที่แสดงอารมณ์ผ่านน้ำเสียงตามโจทย์ที่กำหนด ซึ่งข้อมูลที่ได้จะมีความชัดเจนตามลักษณะการแสดง ในขณะที่ข้อมูลเสียงที่ได้จากการศึกษานี้ ได้จากกลุ่มตัวอย่างที่เป็นบุคคลทั่วไป ผ่านกระบวนการให้คำปรึกษาทางจิตวิทยา ความแตกต่างระหว่างตัวแบบที่ให้ข้อมูลและกระบวนการให้ได้มาซึ่งข้อมูล แสดงให้เห็นว่าข้อมูลที่ได้จากการศึกษาค้างนี้สอดคล้องกับวัตถุประสงค์ที่จะใช้ในการวิเคราะห์อารมณ์จากน้ำเสียงในระหว่างการให้บริการทางจิตวิทยาอย่างมาก เนื่องจากการแสดงอารมณ์ผ่านน้ำเสียงในแง่ของการแสดงและการให้คำปรึกษาจริงจะมีระดับของอารมณ์ที่แตกต่างกัน

7. ข้อสรุป

ผลจากการศึกษา ทำให้ได้ข้อมูลการแสดงออกทางอารมณ์ผ่านอวัจนภาษาของคนไทยผ่านใบหน้าและน้ำเสียงในบริบทการให้คำปรึกษาทางจิตวิทยาจากกลุ่มตัวอย่างจำนวน 85 คน ซึ่งข้อมูลที่มีอยู่จะได้รับการระบุอารมณ์ 5 กลุ่ม ได้แก่ อารมณ์เชิงลบที่มีพลังงาน อารมณ์เชิงลบที่ไม่มีพลังงาน อารมณ์เชิงบวกที่มีพลังงาน อารมณ์เชิงบวกที่ไม่มีพลังงาน และอารมณ์เฉย ๆ เพื่อใช้เป็นฐานข้อมูลสำหรับพัฒนาปัญญาประดิษฐ์เกี่ยวกับการแสดงออกทางอารมณ์ผ่านทางอวัจนภาษาของคนไทย ซึ่งผู้มีความประสงค์จะขอเข้าถึงฐานข้อมูลต้องทำเรื่องขออนุญาตตามที่ศูนย์เทคโนโลยีและนวัตกรรมสุขภาพจิตประเทศไทย คณะวิศวกรรมศาสตร์ มหาวิทยาลัยมหิดล โดยคณะผู้วิจัยจะพิจารณาให้ใช้งานตามความเหมาะสม เพื่อเป็นการปกป้องสิทธิและความเป็นส่วนตัวของผู้เข้าร่วมวิจัย

8. ข้อเสนอแนะ

8.1 ข้อเสนอแนะสำหรับการศึกษาในอนาคต

ฐานข้อมูลที่ได้จากการศึกษานี้ แม้จะมีการกำหนดลักษณะการแสดงอารมณ์อ้างอิงตามระบบ FACS ที่คัดเลือกมาแล้วว่าเป็นรูปแบบการแสดงออกที่พบได้ในหลายวัฒนธรรม แต่ในการศึกษานี้ยังไม่มีมีการกล่าวถึงลักษณะการแสดงอารมณ์ที่พบในกลุ่มตัวอย่างคนไทยว่าแตกต่างหรือสอดคล้องกับเกณฑ์ที่คัดเลือกไว้หรือไม่อย่างไร หากมีการศึกษาถึงลักษณะการแสดงออกที่พบเพิ่มเติม จะเป็นประโยชน์ต่อการพัฒนาความเข้าใจเกี่ยวกับการแสดงออกทางอารมณ์ผ่านทางสีหน้าและน้ำเสียงในวัฒนธรรมไทยมากยิ่งขึ้น

นอกจากนี้ ควรมีการศึกษาวเคราะห์ทางสถิติเพื่อเปรียบเทียบฐานข้อมูลในการศึกษานี้กับฐานข้อมูลที่มีอยู่ในปัจจุบัน เช่น AffectNet หรือ Thai-SER ว่าแต่ละงานมีข้อดี-ข้อด้อยต่อการพัฒนาโมเดลวิเคราะห์อารมณ์ในคนไทยอย่างไร

8.2 ข้อเสนอแนะเชิงนโยบายสำหรับกิจการสื่อสารดิจิทัล

ในอนาคต การได้รับการสนับสนุนอินเทอร์เน็ตความเร็วสูงอาจมีส่วนช่วยในการพัฒนาคุณภาพของข้อมูลหรือลดความผิดพลาดจากการเก็บข้อมูลในรูปแบบออนไลน์ได้ดีขึ้น

8.3 ข้อเสนอแนะสำหรับการนำฐานข้อมูลไปใช้ประโยชน์

ฐานข้อมูลการแสดงออกทางอารมณ์ผ่านอวัจนภาษาที่ได้จากการศึกษาครั้งนี้ ถือเป็นข้อมูลการแสดงออกทางอารมณ์ที่มีความเฉพาะเจาะจงในบริบทของการบำบัดหรือการให้คำปรึกษาทางจิตวิทยาในกลุ่มประชากรไทย เหมาะจะใช้เป็นฐานข้อมูลสำหรับพัฒนาโมเดลวิเคราะห์อารมณ์จากภาพใบหน้า และน้ำเสียงเพื่อใช้ในงานบริการด้านสุขภาพจิตในประเทศไทย เนื่องมาจากกระบวนการได้มาซึ่งข้อมูลเป็นการจำลองกระบวนการให้คำปรึกษามีการกระตุ้นให้กลุ่มตัวอย่างเกิดอารมณ์ต่าง ๆ ครอบคลุมปัญหาทางอารมณ์ที่สามารถพบได้ในการบำบัดจริงในกลุ่มตัวอย่างที่เป็นบุคคลทั่วไป ข้อมูลทั้งเสียงและภาพได้จากการแสดงออกตามธรรมชาติ ซึ่งมีทั้งข้อมูลที่ชัดเจนและคลุมเครือซับซ้อนตามลักษณะการแสดงออกทางอารมณ์ที่พบได้ในบริบทของวัฒนธรรมไทย ทั้งยังเป็นข้อมูลที่ละเอียดในระดับการเปลี่ยนอารมณ์ ต่างจากฐานข้อมูลที่มีในปัจจุบัน เช่น Thai-SER ที่แม้ว่าจะมีข้อมูลน้ำเสียงที่สื่อถึงอารมณ์ต่าง ๆ แต่ข้อมูลเหล่านั้นมีความชัดเจนระดับหนึ่งจากการแสดงบทบาทสมมติตามอารมณ์ที่กำหนดให้ และ AffectNet ซึ่งเป็นฐานข้อมูลภาพนิ่งที่มีลักษณะการแสดงออกที่ชัดเจน อีกทั้งยังเป็นภาพที่รวบรวมจากอินเทอร์เน็ตและมีความหลากหลายทางเชื้อชาติ วัฒนธรรม ซึ่งในความเป็นจริง บุคคลทั่วไปที่อยู่ในภาวะอารมณ์ต่าง ๆ ตามบริบทการบำบัดหรือให้คำปรึกษาทางจิตวิทยาในวัฒนธรรมไทย อาจมีการแสดงออกที่แตกต่างจากการแสดงตามบทบาท

กิตติกรรมประกาศ

โครงการนี้ได้รับการสนับสนุนงบประมาณวิจัยจากกองทุนวิจัยและพัฒนากิจการกระจายเสียง กิจการโทรทัศน์ และกิจการโทรคมนาคม เพื่อประโยชน์สาธารณะ (กทปส.)

รายการเอกสารอ้างอิง

- นันทวัช สิริธีรภัทร์, กมลเนตร วรณเสวก, กมลพร วรณฤทธิ, ปณต ผู้กฤตยาคามี, สุพร อภินันทเวช, และ พนม เกตุมาน. (2558). *จิตเวช ศิริราช DSM-5*. ภาควิชาจิตเวชศาสตร์ คณะแพทยศาสตร์ศิริราชพยาบาล มหาวิทยาลัยมหิดล.
- วรางคณา โสมะนันท์, คาลอส บุญสุภา, และ พลอยไพลิน กมลนาวัน. (2564). การให้บริการการปรึกษาเชิงจิตวิทยาแบบออนไลน์: มิติใหม่ของการให้บริการปรึกษาเชิงจิตวิทยา. *วารสารบัณฑิตศึกษามหาวิทยาลัยราชภัฏวไลยอลงกรณ์ ในพระบรมราชูปถัมภ์*, 15(1), 247-260. <https://opac02.rbru.ac.th/cgi-bin/koha/opac-detail.pl?biblionumber=4465>
- Borges, V., Duarte, R. P., Cunha, C. A., & Mota, D. B. (2019). Are you lost? Using facial recognition to detect customer emotions in retail stores. *Advances in Human-oriented and Personalized Mechanisms, Technologies, and Services. CENTRIC 2019* (pp. 49-54). Valencia: Spain. https://www.researchgate.net/publication/338019900_Are_you_Lost_Using_Facial_Recognition_to_Detect_Customer_Emotions_in_Retail_Stores
- Busso, C., Bulut, M., Lee, C. -C., Kazemzadeh, A., Mower, E., Kim, S., Chang, J. N., Lee, S., & Narayanan, S. (2008). IEMOCAP: Interactive emotional dyadic motion capture database. *Language Resources and Evaluation*, 42, 335-359. <https://doi.org/10.1007/s10579-008-9076-6>
- Cordaro, D. T., Sun, R., Keltner, D., Kamble, S., Huddar, N., & McNeil, G. (2018). Universals and cultural variations in 22 emotional expressions across five cultures. *Emotion*, 18(1), 75-93. <https://doi.org/10.1037/emo0000302>
- Ekman, P., & Friesen, W. V. (1978). *Facial Action Coding System (FACS)* [Database record]. APA PsycTests. <https://doi.org/10.1037/t27734-000>
- Gross, J. J. (2015). *Handbook of Emotion Regulation Second Edition*. Guilford.
- Kasuriya, S., Theeramunkong, T., Wutiwiwatchai, C., & Sukhummek, P. (2019). Developing a Thai emotional speech corpus from Lakorn (EMOLA). *Language Resources and Evaluation*, 53, 17-55. <https://doi.org/10.1007/s10579-018-9428-9>
- Keltner, D., Sauter, D., Tracy, J., & Cowen, A. (2019). Emotional expression: Advances in basic emotion theory. *Journal of Nonverbal Behavior*, 43(2), 133-160. <https://doi.org/10.1007/s10919-019-00293-3>
- Lapakko, D. (2007). Communication is 93% Nonverbal: An Urban Legend Proliferates. *Communication and Theater Association of Minnesota Journal*, 34, 7-19. <https://cornerstone.lib.mnsu.edu/ctamj/vol34/iss1/2/>

- Luna-Jiménez, C., Griol, D., Callejas, Z., Kleinlein, R., Montero, J. M., & Fernández-Martínez, F. (2021). Multimodal emotion recognition on RAVDESS dataset using transfer learning. *Sensors*, 21(22), 7665. <https://doi.org/10.3390/s21227665>
- Matsumoto, D., & Willingham, B. (2009). Spontaneous facial expressions of emotion of congenitally and noncongenitally blind individuals. *Journal of Personality and Social Psychology*, 96(1), 1-10. <https://doi.org/10.1037/a0014037>
- Mollahosseini, A., Hasani, B., & Mahoor, M. H. (2019). AffectNet: A Database for Facial Expression, Valence, and Arousal Computing in the Wild. *IEEE Transactions on Affective Computing*, 10(1), 18-31. <https://doi.org/10.1109/TAFFC.2017.2740923>
- Russell, J. A. (1980). A circumplex model of affect. *Journal of personality and Social Psychology*, 39(6), 1161-1178. <https://doi.org/10.1037/h0077714>
- Siedlecka, E., & Denson, T. F. (2019). Experimental methods for inducing basic emotions: A qualitative review. *Emotion Review*, 11(1), 87-97. <https://doi.org/10.1177/1754073917749016>
- Sobin, C., & Alpert, M. (1999). Emotion in speech: The acoustic attributes of fear, anger, sadness, and joy. *Journal of psycholinguistic research*, 28(4), 347-365. <https://doi.org/10.1023/a:1023237014909>
- Zeren, S. G., Erus, S. M., Amanvermez, Y., Genc, A. B., Yilmaz, M. B., & Duy, B. (2020). The Effectiveness of Online Counseling for University Students in Turkey: A Non-Randomized Controlled Trial. *European Journal of Educational Research*, 9(2), 825-834. <https://doi.org/10.12973/eu-jer.9.2.825>
- Zhang, L., Walter, S., Ma, X., Werner, P., Al-Hamadi, A., Traue, H. C., & Gruss, S. (2016). "BioVid Emo DB": A multimodal database for emotion analyses validated by subjective ratings. *2016 IEEE Symposium Series on Computational Intelligence. SSCI* (pp. 1-6). IEEE. <https://doi.org/10.1109/SSCI.2016.7849931>