



อิทธิพลของประเภทตัวแทนในบทบาทของผู้ให้คำแนะนำและผู้ตัดสินใจต่อการกล่าวโทษผู้ตัดสินใจในนักศึกษาระดับปริญญาตรี

บุญยาพร มกรธวัช บุญธิชา เพียรดี พราราวรี ทองคำ พีพี สุวรรณสิงห์และ
ตระวิณ ชาลิร์กษตระกูล*

คณะศิลปศาสตร์ มหาวิทยาลัยธรรมศาสตร์ ประเทศไทย

The Influence of Advisor Agent Type and Decision Maker Agent Type on Blame Attribution toward Decision-makers in Undergraduate Students

Bunyaphon Makornatwat, Boonticha Piandee, Proundrawe Thongkam,
Pp Suwannasing, and Trawin Chaleeraktragoon*

Faculty of Liberal Arts, Thammasat University, Thailand

Article Info

Research article

Article History:

Received 29 September 2025

Revised 24 November 2025

Accepted 28 November 2025

คำสำคัญ

ปัญญาประดิษฐ์
การรับรู้ระดับจิตใจ
การกล่าวโทษ

* Corresponding author

E-mail address:

trawinc@tu.ac.th

บทคัดย่อ

ปัญญาประดิษฐ์ หรือ Artificial Intelligence (AI) ได้เข้ามามีบทบาทหลากหลายแง่มุมของสังคม ตั้งแต่ในด้านชีวิตประจำวันไปจนถึงบริบทสำคัญอย่างทางการแพทย์ แต่เมื่อเกิดข้อผิดพลาดขึ้นโดยมี AI เป็นส่วนหนึ่งในกระบวนการนั้นทำให้เกิดประเด็นสำคัญด้านการกล่าวโทษ งานวิจัยนี้จึงมีวัตถุประสงค์เพื่อศึกษาอิทธิพลของประเภทตัวแทนในบทบาทของผู้ให้คำแนะนำและผู้ตัดสินใจที่มีต่อการกล่าวโทษผู้ตัดสินใจในบริบททางการแพทย์ ในกรณีที่ผู้ตัดสินใจไม่ทำตามคำแนะนำและนำมาซึ่งผลลัพธ์ที่ไม่พึงประสงค์ โดยดำเนินการสุ่มกลุ่มตัวอย่างเป็นนักศึกษาระดับปริญญาตรีมหาวิทยาลัยธรรมศาสตร์ที่ศึกษาในรายวิชาจิตวิทยาแบบ Convenience Sampling จำนวน 97 คน ที่ได้อ่านสถานการณ์จำลองทางการแพทย์จำนวน 9 เรื่องไขที่ประกอบไปด้วยประเภทของตัวแทนในบทบาทผู้ให้คำแนะนำและผู้ตัดสินใจที่แตกต่างกันแบบ Within-subjects design จากนั้นประเมินการกล่าวโทษต่อผู้ตัดสินใจของแต่ละสถานการณ์ และทำการวิเคราะห์ผลด้วยสถิติแบบ two-way repeated measure ANOVA พบว่าเกิดอิทธิพลหลักของประเภทตัวแทนในบทบาทผู้ให้คำแนะนำ ($F(2,192) = 13.79, p < .001, \eta_p^2 = .126$) โดยผู้ตัดสินใจจะได้รับการกล่าวโทษมากที่สุดเมื่อปฏิเสธคำแนะนำจากมนุษย์ ซึ่งมากกว่ากรณีที่ปฏิเสธคำแนะนำจาก AI ที่มีจิตใจคล้ายมนุษย์ ($p = 0.018$) และ AI ปกติ ($p < 0.001$) อย่างมีนัยสำคัญทางสถิติ นอกจากนี้ ผู้ตัดสินใจยังได้รับการกล่าวโทษมากกว่าเมื่อปฏิเสธคำแนะนำจาก AI ที่มีจิตใจคล้ายมนุษย์เทียบกับ AI ปกติในระดับที่ใกล้เคียงกับนัยสำคัญทางสถิติ ในทางกลับกันประเภทของตัวแทนในบทบาทของผู้ตัดสินใจไม่มีอิทธิพลต่อการกล่าวโทษผู้ตัดสินใจและไม่พบอิทธิพลเชิงปฏิสัมพันธ์ระหว่างทั้งสองตัวแปรอิสระ

Keywords :

artificial intelligence,
mind perception,
blame attribution

Abstract

Artificial intelligence (AI) is playing an increasingly larger role in society, from aspects of everyday life to more significant ones such as those in the medical sphere. When AI-related systemic errors occur in advice or decision-making, blame can be ambiguous. Thus, this study investigates the influence of advisor and decision-maker agent types on attributing medical blame where decision-makers reject recommendations, leading to negative outcomes. The sample was obtained using convenience sampling. A total of 97 undergraduate students, enrolled in a psychology course at Thammasat University at the time, took part. Using a within-subjects design, participants read medical vignettes presented under nine experimental conditions featuring alternating advisor and decision-maker agent types. Participants were asked to evaluate degrees of blame attributed to the decision-maker in each scenario. The data were analyzed using a two-way repeated-measures ANOVA. Results indicated a main effect for advisor agent types ($F(2,192) = 13.79, p < .001, \eta_p^2 = .126$): decision-makers received the most blame for rejecting human, rather than human mind-like AI ($p = 0.018$) or generic AI ($p < 0.001$) advisors at a statistically significant level. Additionally, decision-makers received more blame for rejecting human mind-like AI advisors than for rejecting generic AI advisors at a marginally significant level. No main effect was found for decision-maker agent type. There was also no significant interaction between advisors and decision-maker agent types on blame attribution.

1. บทนำ

ปัญญาประดิษฐ์ หรือ Artificial intelligence (AI) คือเทคโนโลยีที่พัฒนาให้เครื่องจักรหรือระบบปฏิบัติการคอมพิวเตอร์สามารถคิด วิเคราะห์ และเรียนรู้ความสามารถบางอย่างได้อย่างชำนาญ อีกทั้งยังลอกเลียนแบบการตัดสินใจโดยใช้เหตุผลได้อย่างมนุษย์ (สำนักงานพัฒนาวิทยาศาสตร์และเทคโนโลยีแห่งชาติ, 2567) โดยในปัจจุบัน AI ได้เข้ามามีบทบาทสำคัญอย่างมากในหลากหลายแง่มุมของชีวิตมนุษย์ จากผลสำรวจพบว่าคนไทยกว่าร้อยละ 73.84 มีการใช้งานเทคโนโลยี AI ตั้งแต่ในด้านที่เกี่ยวข้องกับชีวิตประจำวันตลอดจนบริบทสำคัญอื่น ๆ (สำนักงานบริหารเทคโนโลยีสารสนเทศเพื่อการศึกษา, 2567) โดยเฉพาะในทางการแพทย์ของประเทศไทย ระบบ AI กำลังมีบทบาทสำคัญและถูกมองว่ามีศักยภาพสูงในการรักษาผู้ป่วย โดยเฉพาะในด้านการคัดกรองและวินิจฉัยโรค เช่น ช่วยในการวิเคราะห์ภาพเอกซเรย์เพื่อตรวจหาความผิดปกติของโรค ซึ่งสามารถช่วยเพิ่มความแม่นยำและลดระยะเวลาในกระบวนการบางอย่างลงได้ (โรงพยาบาลบำรุงราษฎร์, 2565) เป็นต้น นอกจากนี้ ทางสำนักงานพัฒนาวิทยาศาสตร์และเทคโนโลยีแห่งชาติ (สวทช.) ยังได้มีการพัฒนา Medical AI Data Platform หรือศูนย์ข้อมูลกลางทางการแพทย์ โดยมีเป้าหมายเพื่อลดภาระงานของบุคลากรทางการแพทย์และยกระดับคุณภาพทางการแพทย์ให้สูงขึ้นอีกในอนาคต (“เปิดตัว Medical AI”, 2568)

อย่างไรก็ตาม แม้ระบบอัจฉริยะนี้จะมีประโยชน์อย่างมากในสังคม แต่ในบางครั้ง การนำเข้ามาใช้งานกลับทำให้เกิดปัญหาที่ไม่คาดคิดขึ้น ยกตัวอย่าง จากกรณีที่ศูนย์การแพทย์ของมหาวิทยาลัย Stanford ได้มีการจัดลำดับการแจกจ่ายวัคซีนป้องกันโควิดโดยใช้ระบบ AI แต่ระบบดังกล่าวไม่ได้รวบรวมบุคลากรทางการแพทย์ด้านหน้าเข้าไปในฐานข้อมูลด้วยจึงทำให้เกิดการประท้วงเกิดขึ้น อย่างไรก็ตามทางผู้อำนวยการของศูนย์การแพทย์ได้แถลงการณ์ขอโทษต่อเหตุการณ์ที่เกิดขึ้นและให้เหตุผลว่าปัญหานี้เกิดจากความผิดพลาดของระบบการทำงานของ AI แต่ผู้คนกลับมองว่าเหตุผลนี้เป็นเพียงคำกล่าวอ้างเพื่อโยนความผิดให้กับ AI (Wamsley, 2020) กรณีนี้ชี้ให้เห็นถึงประเด็นที่ซับซ้อนว่า หากมีข้อผิดพลาดเกิดขึ้นโดยมี AI เข้าไปมีส่วนร่วมในกระบวนการทำงานนั้นแล้วการกล่าวโทษนี้ควรมุ่งเน้นไปที่บุคคลใด

ในช่วงไม่กี่ปีที่ผ่านมาแนวโน้มที่เพิ่มขึ้นของงานวิจัยที่มุ่งศึกษาบทบาทของ AI ในกระบวนการตัดสินใจโดยเฉพาะในบทบาทของผู้ให้คำแนะนำ โดยพบว่าผู้คนมักเชื่อมโยงความรับผิดชอบกับผู้ให้คำแนะนำที่เป็นมนุษย์มากกว่าผู้ให้คำแนะนำที่เป็น AI (Gazit et al., 2023, p. 5) และเมื่อเกิดผลลัพธ์ที่ไม่พึงประสงค์ขึ้น ผู้ที่เป็นมนุษย์มักถูกกล่าวโทษมากกว่า (Gómez-Sánchez et al., 2023, p. 262) อย่างไรก็ตามงานวิจัยจำนวนมากไม่น้อยยังคงมอง AI ในฐานะเครื่องมือที่ใช้สนับสนุนมนุษย์มากกว่าจะพิจารณาว่าเป็นตัวแทนที่มีเจตจำนง (agency) หรือมีจิตใจ (mind) เป็นของตนเอง (Stefanidis et al., 2022) แม้แต่ในงานวิจัยที่ศึกษาเรื่อง Explainable AI (XAI)¹ ได้แก่ งานของ David et al. (2021) ที่มีการเสนอรูปแบบ

¹ Explainable Artificial Intelligence (XAI) หมายถึง ระบบที่สามารถอธิบายหรือเปิดเผยเหตุผลเบื้องหลังการตัดสินใจของตนเองหรือของระบบอื่นได้ (Miller, 2019, p. 2)

การอธิบายเบื้องหลังการให้คำแนะนำของ AI ก็ยังเน้นไปที่การเพิ่มความโปร่งใสของกระบวนการคำนวณมากกว่าการสื่อสารว่า AI มีความคิดหรือความตั้งใจในเชิงจิตวิทยา อันเป็นมิติที่อาจมีความสำคัญต่อการประเมินความรับผิดชอบเชิงจริยธรรม เช่น การกล่าวโทษ ซึ่งเป็นกระบวนการอธิบายทางสังคมอย่างหนึ่งเมื่อเหตุการณ์นั้นมีผลเป็นลบ อีกทั้งยังเกี่ยวข้องกับการตัดสินใจอย่างมีเหตุผล และต้องการกำหนดความรับผิดชอบไปที่บุคคลใด-บุคคลหนึ่ง (Shaver, 1985, pp. 6-7)

ในขณะเดียวกัน งานวิจัยบางชิ้นเริ่มหันมาให้ความสนใจกับประเด็นด้านการรับรู้ ว่า AI มีจิตใจหรือไม่ และผลของการรับรู้ด้านจิตใจนี้ส่งผลอย่างไรต่อการกล่าวโทษ โดยการรับรู้จิตใจ (Mind perception) หมายถึง ความสามารถในการรับรู้ว่ามีจิตใจอยู่ในสิ่งมีชีวิตหรือไม่มีชีวิตอื่น ๆ ซึ่งเป็นการอนุมานว่าสิ่งนั้นมีจิตใจคล้ายมนุษย์ (Waytz et al., 2010, p. 386) เช่น Joo (2024) พบว่า AI ที่ถูกมองว่ามีจิตใจคล้ายมนุษย์มักถูกกล่าวโทษมากกว่า AI ปกติ ที่ไม่ได้มีจิตใจคล้ายมนุษย์ ในขณะที่งานของ Stuart and Kneer (2021, pp. 17-18) ชี้ให้เห็นว่าผู้คนมีแนวโน้มจะกล่าวโทษ AI มากขึ้นหากเชื่อว่า AI มีความคิดและเจตนา อย่างไรก็ตาม เมื่อต้องเผชิญกับผลเสียที่เกิดขึ้นจริง คนกลับลงเลที่จะกล่าวโทษ AI และมีแนวโน้มจะหาผู้รับผิดชอบที่เป็นมนุษย์แทน

แม้จะมีการศึกษาความสัมพันธ์ระหว่างระดับการรับรู้จิตใจกับการกล่าวโทษอยู่บ้าง อย่างไรก็ตาม งานวิจัยเหล่านี้ที่ศึกษาเกี่ยวกับระดับการรับรู้จิตใจของ AI และการกล่าวโทษส่วนใหญ่ยังคงเน้นศึกษา AI ในบทบาทของผู้ตัดสินใจเพียงอย่างเดียว โดยยังไม่เปรียบเทียบกับกรณีที่ AI ทำหน้าที่เป็นผู้ให้คำแนะนำ ซึ่งเป็นบทบาทที่พบได้บ่อยในระบบ AI ที่ถูกนำไปใช้จริง โดยเฉพาะอย่างยิ่งในบริบทที่มีความเสี่ยงสูง เช่น การแพทย์ นอกจากนี้ แม้จะมีงานวิจัยบางชิ้นที่เปรียบเทียบผลของบทบาทผู้ให้คำแนะนำกับผู้ตัดสินใจของ AI แล้ว (เช่น Lima et al., 2021) แต่ก็ยังขาดการพิจารณาประเด็นด้านลักษณะเชิงจิตใจของ AI ว่ามีผลต่อการประเมินความรับผิดชอบหรือไม่ ทั้งที่คุณลักษณะดังกล่าวอาจเป็นตัวแปรสำคัญที่กำหนดว่าผู้คนจะกล่าวโทษ AI ในลักษณะใด อีกทั้งงานวิจัยที่ผสมผสานทั้งบทบาทของ AI และระดับของการรับรู้จิตใจเข้าด้วยกัน โดยเฉพาะในบริบททางการแพทย์ซึ่งมีลักษณะเฉพาะและความเสี่ยงสูง กลับมีอยู่อย่างจำกัด

ดังนั้น งานวิจัยในครั้งนี้จึงมีวัตถุประสงค์เพื่อศึกษาผลของประเภทของตัวแทน (มนุษย์, AI ปกติ, และ AI ที่มีจิตใจคล้าย-มนุษย์) และบทบาทของตัวแทน (ผู้ให้คำแนะนำและผู้ตัดสินใจ) ต่อระดับของการกล่าวโทษผู้ตัดสินใจในบริบททางการแพทย์กรณีผู้ให้คำแนะนำ ให้คำแนะนำที่เหมาะสม แต่ผู้ตัดสินใจไม่ได้กระทำตามคำแนะนำ หรือกล่าวคือตัดสินใจกระทำในแนวทางที่ไม่เหมาะสม โดยก่อให้เกิดผลลัพธ์ที่ไม่พึงประสงค์และความร้ายแรงของผลลัพธ์ไม่แตกต่างกัน ซึ่งเป็นการเติมเต็มช่องว่างขององค์ความรู้ที่ยังขาดแคลนในงานวิจัยที่ผ่านมาและนำมาซึ่งคำถามว่าประเภทตัวแทนที่แตกต่างกันในบทบาทผู้ให้คำแนะนำและผู้ตัดสินใจจะมีอิทธิพลต่อการกล่าวโทษต่อผู้ตัดสินใจหรือไม่ โดยงานวิจัยครั้งนี้ทำการศึกษาในนักศึกษาระดับปริญญาตรีที่ลงทะเบียนเรียนในรายวิชาจิตวิทยา และทำการเลือกกลุ่มตัวอย่างด้วยวิธีการสุ่มตามสะดวก (convenience sampling) รวมถึงจะมีการอ่านสถานการณ์จำลองทางการแพทย์ที่ประกอบไปด้วยประเภทของตัวแทนในบทบาทผู้ให้คำแนะนำและผู้ตัดสินใจที่แตกต่างกัน จากนั้นจึง

ประเมินการกล่าวโทษต่อผู้ตัดสินใจของแต่ละสถานการณ์ อย่างไรก็ตาม กลุ่มนักศึกษาดังกล่าวถือว่ามีความเหมาะสมต่อการศึกษาระดับต้นด้านการอธิบายโทษ เนื่องจากอยู่ในช่วงวัยผู้ใหญ่ตอนต้นที่สามารถประเมินสถานการณ์ทางสังคมได้อย่างมีเหตุผล อีกทั้งนักศึกษายังเป็นกลุ่มที่นิยมใช้ในงานวิจัยลักษณะนี้ เนื่องจากเข้าถึงได้ง่ายและมีความเข้าใจในแนวคิดทางจิตวิทยาในระดับหนึ่ง

2. การทบทวนวรรณกรรม

2.1 การรับรู้ระดับจิตใจของ AI (Mind Perception of AI)

Mind perception หมายถึง ความสามารถในการรับรู้ว่ามีจิตใจอยู่ในสิ่งมีชีวิตอื่น ๆ หรือแม้แต่ในสิ่งไม่มีชีวิต เช่น ต้นไม้ สัตว์ หรือหุ่นยนต์ เป็นต้น โดยเฉพาะเมื่อบุคคลรับรู้ว่ามีระบบปัญญาประดิษฐ์ (AI) มีความสามารถทางจิตใจ ซึ่งสะท้อนถึงการอนุมานว่าสิ่งที่ไม่ใช่มนุษย์มีจิตใจ ปรากฏการณ์นี้เรียกว่า Anthropomorphism หรือการอุปมาว่าสิ่งไม่มีชีวิตมีลักษณะคล้ายมนุษย์ (McMurtrie, 2023, p. 54; Waytz et al., 2010, p. 386)

การรับรู้จิตใจสามารถจำแนกออกเป็นสองมิติหลัก มิติแรก คือ Experience หรือความสามารถในการรับรู้อารมณ์และความรู้สึก เช่น ความสุข ความเจ็บปวด ความกลัว ความหิว เป็นต้น มิติที่สอง คือ Agency หรือความสามารถในการกระทำอย่างมีเหตุผล การมีเจตนา การควบคุมตนเอง การวางแผน และการกระทำโดยมีเป้าหมายที่ชัดเจน โดยมนุษย์มักประเมินผู้อื่นแตกต่างกันตามระดับของ experience และ agency ที่รับรู้ได้ ตัวอย่างเช่น มนุษย์ผู้ใหญ่มักมองว่าทั้งสองมิติในระดับสูง ขณะที่หุ่นยนต์หรือ AI มักถูกมองว่ามี agency อยู่ในระดับปานกลางแต่มี experience ต่ำ (Gray et al., 2007, p. 619; McMurtrie, 2023, p. 55) อย่างไรก็ตาม ความสามารถในการรับรู้อารมณ์มีความสำคัญต่อการตัดสินใจเชิงจริยธรรม โดยเฉพาะในแง่ของความเข้าใจและความเห็นอกเห็นใจผู้อื่น ซึ่งเป็นพื้นฐานสำคัญในการตระหนักถึงผลกระทบของพฤติกรรมที่มีต่อผู้อื่น ความสามารถเหล่านี้ทำให้บุคคลสามารถเข้าใจผลของการกระทำของตนได้อย่างลึกซึ้ง ดังนั้น การส่งเสริมการรับรู้ในมิติของ experience จึงถือเป็นปัจจัยสำคัญที่ทำให้ AI ถูกมองว่ามีจิตใจ

ความก้าวหน้าอย่างรวดเร็วของเทคโนโลยี AI ได้นำไปสู่การพัฒนาระบบที่มีความซับซ้อนและสมจริงมากยิ่งขึ้น ซึ่งถูกนำไปใช้ประโยชน์ในหลากหลายสาขา ไม่ว่าจะเป็นการแพทย์ ระบบความปลอดภัย การศึกษา ตลอดจนระบบขนส่ง งานวิจัยของ Waytz et al. (2014, pp. 114-116) พบว่า เมื่อระบบ AI ในรถยนต์ไร้คนขับถูกออกแบบให้มีลักษณะคล้ายมนุษย์ ตัวอย่างเช่น การมีชื่อ เพศ และเสียงพูด เป็นต้น จะส่งผลให้ผู้ใช้งานแสดงระดับความไว้วางใจสูงขึ้น โดยประเมินว่าระบบดังกล่าว มีความสามารถทางจิตใจ และมีความน่าเชื่อถือมากกว่ารถยนต์ไร้คนขับที่มีเพียงระบบ AI สำหรับควบคุมพวงมาลัยและเร่งความเร็ว โดยไม่มีลักษณะคล้ายมนุษย์ นอกจากนี้ เมื่อเกิดเหตุการณ์อุบัติเหตุขึ้น ผู้ใช้งานรถยนต์ไร้คนขับที่มีระบบ AI ลักษณะคล้ายมนุษย์ ยังมีแนวโน้มรู้สึกผ่อนคลายมากกว่าเมื่อเทียบกับผู้ใช้งานรถยนต์ไร้คนขับที่มีระบบ AI ปกติโดยไม่มีลักษณะคล้ายมนุษย์ อย่างไรก็ตามรถยนต์ไร้คนขับที่ทำงานด้วยระบบ AI ลักษณะ

คล้ายมนุษย์กลับถูกกล่าวโทษมากกว่ารถยนต์ที่ผู้ขับควบคุมด้วยตนเอง โดยในงานวิจัยได้ให้เหตุผลไว้ว่า วัตถุหรือระบบที่ไม่มีความสามารถในการกระทำอย่างมีเหตุผล (agency) ไม่สามารถถือว่าเป็นผู้รับผิดชอบต่อการกระทำใด ๆ ได้

แม้ว่าการรับรู้จิตใจใน AI อาจก่อให้เกิดปฏิกิริยาเชิงบวกในบางสถานการณ์ แต่ Shank et al. (2019, p. 264) พบว่าเมื่อผู้เข้าร่วมรับรู้ว่ามี AI มีความสามารถด้านการคิดและอารมณ์ในบริบทของการปฏิสัมพันธ์ในชีวิตประจำวัน อาจนำไปสู่ปฏิกิริยาเชิงลบ โดยเฉพาะความรู้สึกไม่สบายใจหรือสับสนจากการที่ AI แสดงลักษณะคล้ายมนุษย์มากเกินไป ซึ่งสอดคล้องกับงานของ Kitchens and Meier (2025, pp. 5-6) ที่พบว่าผู้เข้าร่วมมีแนวโน้มรู้สึกหวาดกลัวและมองว่า AI เป็นภัยคุกคามมากขึ้น เมื่อ AI ถูกอธิบายว่ามีความสามารถด้านปัญญาและอารมณ์ในระดับสูงเมื่อเปรียบเทียบกับกรณีที่มี AI มีความสามารถในระดับต่ำ จากงานวิจัยข้างต้นสะท้อนให้เห็นว่า AI ที่มีความสามารถด้านการคิดและอารมณ์ จะทำให้บุคคลเกิดการรับรู้ความรู้สึกเชิงลบและมองว่ามีความอันตรายมากขึ้น ดังนั้นเมื่อ AI มีความสามารถในการกระทำอย่างมีเหตุผล (agency) และความสามารถในการรับรู้อารมณ์ความรู้สึก (experience) จึงถูกมองว่าสามารถเป็นตัวแทนเชิงจริยธรรม (moral agent) ที่กระทำหรือได้รับผลจากการกระทำไม่ต่างจากมนุษย์ (Floridi & Sanders, 2004, pp.14-16)

การรับรู้ระดับจิตใจของ AI จึงมีความเกี่ยวข้องโดยตรงกับการตัดสินใจทางจริยธรรม โดยเฉพาะในบริบทที่มีความผิดพลาดซึ่งนำไปสู่ผลลัพธ์รุนแรง เช่น การเสียชีวิตของบุคคล แนวคิดข้างต้นได้รับการสนับสนุนเพิ่มเติมจากงานวิจัยของ Joo (2024, pp. 7-8) พบว่าเมื่อเกิดการกระทำผิดเชิงจริยธรรม ผู้เข้าร่วมมีแนวโน้มกล่าวโทษ AI ที่มีลักษณะคล้ายมนุษย์มากกว่า AI ปกติ โดยเฉพาะเมื่อระบบนั้นได้รับข้อมูลพื้นหลังเพิ่มเติม เช่น ชื่อ อายุ หรืองานอดิเรก เป็นต้น ซึ่งสะท้อนว่า การรับรู้ว่ามีจิตใจส่งผลโดยตรงต่อการประเมินความรับผิดชอบเชิงจริยธรรมของ AI

นอกจากนี้ ในงานวิจัยของ Leichtmann et al. (2024, pp. 32416-32417) ยังแสดงให้เห็นว่าระดับการกล่าวโทษต่อ AI ที่มีการให้เหตุผลประกอบการตัดสินใจ (Explainable AI: XAI) และการกล่าวโทษตนเอง ไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ ซึ่งสะท้อนว่าผู้เข้าร่วมมองว่า AI มีความรับผิดชอบเชิงจริยธรรมในระดับเดียวกับมนุษย์ อีกทั้งยังพบว่าผู้เข้าร่วมรับรู้คุณลักษณะเชิงจิตใจของ AI ผ่านพฤติกรรมการตัดสินใจ โดยการกล่าวโทษนี้ มีความสัมพันธ์กับการรับรู้ว่ามี AI มีความรับผิดชอบและสัมพันธ์กับการรับรู้ว่ามีคุณลักษณะทางจิตใจ ได้แก่ ความสามารถในการตัดสินใจ การมีจิตใจของตนเอง การมีเจตนา (intentionality) การมีอิสระในการตัดสินใจ (autonomy) ซึ่งชี้ให้เห็นว่า AI ถูกมองว่าเป็นผู้กระทำที่มีจิตใจในบริบทของการตัดสินใจทางจริยธรรม

ในทางตรงกันข้าม งานวิจัยของ Lima et al. (2021) ที่ศึกษาบริบทการตัดสินใจในกระบวนการกฎหมาย พบว่าผู้เข้าร่วมมีแนวโน้มกล่าวโทษ AI ปกติและมนุษย์ในบทบาทของผู้ให้คำแนะนำและผู้ตัดสินใจในระดับใกล้เคียงกัน โดยไม่มีความแตกต่างกันอย่างเด่นชัด ซึ่งแสดงให้เห็นว่า แม้งานนี้ไม่ได้เน้นลักษณะ

เชิงจิตใจของ AI แต่ระดับของการกล่าวโทษอาจขึ้นอยู่กับบริบทเฉพาะ ลักษณะของตัวแทน และบทบาทในสถานการณ์ มากกว่าความเป็นมนุษย์หรือไม่ของระบบเพียงอย่างเดียว

2.2 บทบาทของตัวแทนในการตัดสินใจทางจริยธรรม

ในการตัดสินใจเชิงจริยธรรม บทบาทของตัวแทนถือเป็นองค์ประกอบสำคัญ เนื่องจากแต่ละบทบาทสะท้อนระดับของการมีส่วนร่วมและอำนาจในการตัดสินใจที่แตกต่างกัน งานวิจัยที่ผ่านมาได้แสดงให้เห็นว่าความแตกต่างในบทบาทของตัวแทนมีผลต่อการประเมินทางจริยธรรมอย่างเห็นได้ชัด ตัวอย่างเช่น Hanson et al. (2024, p. 243) พบว่าผู้คนมีแนวโน้มที่จะกล่าวโทษมากขึ้นเมื่อหุ่นยนต์เพิกเฉยต่อคำแนะนำจากมนุษย์ ขณะเดียวกันในบริบทที่ระบบปัญญาประดิษฐ์ (AI) ทำหน้าที่ตัดสินใจโดยตรง แทนที่จะให้คำแนะนำ ผู้เข้าร่วมมักประเมินว่า AI มีความรับผิดชอบมากกว่า โดยเฉพาะในระบบที่มนุษย์ยังคงมีบทบาทกำกับดูแลในกระบวนการตัดสินใจ (human-on-the-loop) ซึ่งสะท้อนให้เห็นว่าระดับความเป็นอิสระในการตัดสินใจ (autonomy) มีผลต่อการรับรู้ทางจริยธรรมอย่างชัดเจน (Tolmeijer et al., 2022, p. 13)

นอกจากนี้ ตัวแทนที่ทำหน้าที่ตัดสินใจมักถูกมองว่ามีอำนาจควบคุมผลลัพธ์มากกว่า แม้จะมีการรับคำแนะนำจากบุคคลอื่นก็ตาม ในทางตรงกันข้าม ผู้ให้คำแนะนำที่ไม่ได้เป็นผู้ลงมือกระทำด้วยตนเอง มักได้รับการประเมินว่ามีความรับผิดชอบน้อยกว่า ซึ่งสะท้อนถึงลำดับชั้นทางจริยธรรมที่แตกต่างกันตามบทบาท

กรอบแนวคิดทางทฤษฎีหลายกรอบสามารถอธิบายปรากฏการณ์นี้ได้ โดยเฉพาะทฤษฎี Path of Blame (Malle et al., 2014, p. 151) ที่เสนอว่าการกล่าวโทษจะเกิดขึ้นเมื่อมีผลลัพธ์ที่ไม่พึงประสงค์ และผู้คนจะพิจารณาว่าผู้กระทำมีเจตนาในการทำให้เกิดผลนั้นหรือไม่ ซึ่งการมีอิสระในการตัดสินใจ (autonomy) ย่อมสัมพันธ์กับการมีเจตนา (intention) และส่งผลให้บทบาทของผู้ตัดสินใจถูกพิจารณาว่ามีความรับผิดชอบสูงกว่า

ขณะเดียวกัน ทฤษฎี Outcome Bias (Baron & Hershey, 1988, pp. 569-570) ซึ่งชี้ให้เห็นว่าผู้คนมักใช้ผลลัพธ์ในการประเมินความถูกต้องของการตัดสินใจ โดยเฉพาะเมื่อผลลัพธ์นั้นเลวร้าย ความผิดมักจะถูกโยนไปยังผู้ที่ตัดสินใจมากกว่าผู้ที่ให้คำแนะนำ ซึ่งแสดงให้เห็นว่า ในสถานการณ์ที่มีผลลัพธ์ไม่พึงประสงค์ ผู้ตัดสินใจจะถูกตั้งคำถามถึงการตัดสินใจที่อาจขาดการไตร่ตรองอย่างเหมาะสม

จากกรอบแนวคิดข้างต้นสามารถสรุปได้ว่า บทบาทของตัวแทน โดยเฉพาะผู้ที่ทำหน้าที่ตัดสินใจโดยตรง มีความสำคัญต่อการประเมินทางจริยธรรมอย่างมาก ทั้งในแง่ของการถูกมองว่ามีเจตนาในการกระทำและการเป็นผู้รับผิดชอบต่อผลลัพธ์โดยตรง ในทางกลับกันผู้ที่ไม่ได้ลงมือกระทำด้วยตนเอง เช่น ผู้ให้คำแนะนำ มักจะถูกลดทอนระดับความรับผิดชอบลง อย่างมีนัยสำคัญ ในบริบทที่มีความเสี่ยงสูง เช่น การตัดสินใจทางการแพทย์ บทบาทของผู้ตัดสินใจยิ่งได้รับความสนใจ เนื่องจากอาจหมายถึงผลลัพธ์ที่เกี่ยวข้องกับชีวิตของผู้ป่วย อย่างไรก็ตาม บทบาทของผู้ให้คำแนะนำก็มีความสำคัญเช่นกัน โดยเฉพาะเมื่อเกิดผลลัพธ์ที่ไม่พึงประสงค์ การประเมินทางจริยธรรมมักเข้มข้นขึ้นในลักษณะที่สะท้อนผ่านการกล่าวโทษ

ซึ่งมักถูกโยนไปยังที่ถูกมองว่ามีอำนาจและความรับผิดชอบต่อการตัดสินใจมากที่สุด (Baron & Hershey, 1988; Malle et al., 2014)

2.3 การกล่าวโทษ (Blame) ในการตัดสินใจทางการแพทย์

การกล่าวโทษ (Blame Attribution) เป็นการอธิบายทางสังคมรูปแบบหนึ่ง เป็นผลลัพธ์ของกระบวนการที่เริ่มต้นจากเหตุการณ์ที่มีผลลบ ซึ่งเกี่ยวข้องกับการตัดสินใจเกี่ยวกับความเป็นเหตุความรับผิดชอบส่วนบุคคลและปัจจัยที่อาจลดทอนความผิดได้ ผลลัพธ์สุดท้ายอาจเป็นการยืนยันหรือปฏิเสธความสมควรถูกตำหนิของบุคคลนั้น (Shaver, 1985, pp. 6-7) ซึ่งถือได้ว่าการกล่าวโทษเป็นการกำหนดความรับผิดชอบต่อการกระทำผิดนั้น

Malle et al. (2014, p. 5) ได้ศึกษาปัจจัยที่มีผลต่อการกล่าวโทษโดยอาศัยแบบจำลอง Path Model of Blame ซึ่งเสนอว่า เมื่อเกิดเหตุการณ์ที่ละเมิดบรรทัดฐานและมีผู้กระทำเป็นสาเหตุของเหตุการณ์นั้น การประเมินเจตนาของผู้กระทำมีบทบาทสำคัญ หากพบว่าผู้กระทำมีเจตนาในการก่อให้เกิดเหตุการณ์ดังกล่าว การกล่าวโทษจะอยู่ในระดับสูงมาก แต่หากไม่มีเจตนา ผู้ประเมินจะพิจารณาต่อว่าผู้กระทำนั้นมีหน้าที่หรือมีความสามารถที่จะป้องกันเหตุได้หรือไม่ ซึ่งหากทั้งสองเงื่อนไขนี้มีอยู่ ก็อาจนำไปสู่การกล่าวโทษได้

กระบวนการดังกล่าวกลายเป็นประเด็นที่ซับซ้อนยิ่งขึ้น เมื่อระบบปัญญาประดิษฐ์ (AI) เข้ามามีบทบาทสำคัญในการสนับสนุนกระบวนการวินิจฉัยทางการแพทย์ ไม่ว่าจะเป็นการตรวจพบมะเร็งเต้านม โรคเบาหวาน โรคหัวใจ โรคตับและไต (Kaur et al., 2020) ซึ่งเกี่ยวข้องกับความจริยธรรมและการตัดสินใจที่มีผลต่อชีวิต โดยการตัดสินใจที่ผิดพลาดอาจนำไปสู่ผลลัพธ์ที่ร้ายแรง เช่น การวินิจฉัยผิดหรือมีผู้เสียชีวิต เป็นต้น จึงนำไปสู่ประเด็นของการกล่าวโทษ โดยเฉพาะในกรณีที่ต้องตัดสินใจว่าใครควรรับผิดชอบต่อผลลัพธ์ที่เกิดขึ้น

Verdicchio and Perin (2022, p. 1) ได้เสนอว่า สำหรับการใช้ระบบ AI ในการวินิจฉัยทางการแพทย์ ความเป็นตัวการทางจริยธรรมควรถูกจำกัดไว้เฉพาะในมนุษย์เท่านั้น เนื่องจากมนุษย์มีความเป็นอิสระ (autonomy) และความรู้สึกหรือการตระหนักรู้ (sentience) ซึ่งระบบ AI ขาดคุณสมบัติเหล่านี้ ดังนั้น แม้ AI จะเป็นต้นเหตุของผลลัพธ์ที่ไม่พึงประสงค์ ก็ไม่ควรกล่าวโทษ AI เพราะไม่มีความหมายใดในเชิงจริยธรรมหรือสังคม การวิเคราะห์ความรับผิดชอบจึงควรพิจารณามนุษย์ในฐานะผู้ที่มีหน้าที่ควบคุมและตัดสินใจร่วมกับ AI เช่น นักพัฒนาระบบ หรือแพทย์ผู้ใช้งานที่ละเลยการตรวจสอบอย่างรอบคอบ เป็นต้น

แนวคิดทางจริยธรรมสอดคล้องบางประการกับ Kempt et al. (2023, p. 1413) ที่ชี้ว่าในบริบทของการปฏิสัมพันธ์ระหว่างแพทย์กับ AI ในการวินิจฉัยทางการแพทย์ แม้แพทย์จะยังคงเป็นผู้มีอำนาจตัดสินใจหลักและต้องรับผิดชอบต่อความผิดพลาดที่เกิดขึ้น แต่ระดับของการถูกกล่าวโทษควรลดลงในกรณีที่แพทย์ใช้ระบบ AI ประกอบการตัดสินใจอย่างเหมาะสม โดยเฉพาะเมื่อเทียบกับแพทย์ที่ไม่ใช้ระบบดังกล่าวเลย เนื่องจากได้ปฏิบัติตามหน้าที่ในการพิจารณาทุกแง่มุมที่กระบวนการทางคลินิกอำนวยความสะดวกที่อีกคนละเลยเครื่องมือที่มีอยู่

อย่างไรก็ตาม จากงานวิจัยของ Krügel et al. (2024, p. 1485) ที่ศึกษาเปรียบเทียบกรณีแพทย์ได้รับคำแนะนำจาก AI กับกรณีที่ได้รับคำแนะนำจากเพื่อนร่วมงานมนุษย์ พบว่า คนมีแนวโน้มที่จะมอบความรับผิดชอบทางจริยธรรมให้กับ AI มากกว่า โดยเฉพาะในกลุ่มที่เชื่อว่า AI อาจมีจิตสำนึกได้ในอนาคต ขณะเดียวกันเมื่อเปรียบเทียบกับทีมที่เป็นมนุษย์ล้วน ความรับผิดชอบที่มอบให้กับมนุษย์กลับลดลง ในสถานการณ์ที่เป็นทีมวินิจฉัยแบบผสม (มนุษย์ร่วมกับ AI) สะท้อนให้เห็นถึงการเปลี่ยนแปลงในการรับรู้และการแจกแจงความรับผิดชอบทางศีลธรรมเมื่อ AI เข้ามามีบทบาทในกระบวนการตัดสินใจทางการแพทย์ แม้จะไม่มีคุณลักษณะของจิตใจ

นอกจากนี้ ในผลงานวิจัยของ Schoenherr and Thomson (2024, pp. 66-67) ช่วยชี้ให้เห็นถึงการแจกแจงความรับผิดชอบตามระดับความอิสระในการตัดสินใจของระบบและโครงสร้างของผู้เกี่ยวข้อง ซึ่งศึกษาการประเมินความรับผิดชอบเมื่อระบบ AI ทำงานผิดพลาด พบว่าระดับความอิสระในการตัดสินใจของระบบ AI มีอิทธิพลอย่างชัดเจนต่อการตัดสินใจว่าใครควรรับผิดชอบ โดยในสถานการณ์ที่ AI มีอิสระในการตัดสินใจ (full autonomy) ที่เกิดความผิดพลาดขึ้น ผู้ปฏิบัติงานมนุษย์ (operator) ซึ่งเป็นผู้สั่งเปิดใช้งาน AI จะถูกมองว่ามีความรับผิดชอบมากที่สุด เนื่องจากการมอบอำนาจตัดสินใจให้ AI ถูกตีความว่าเป็นการกระทำที่มีผลโดยตรงต่อผลลัพธ์และความน่าเชื่อถือของ AI ลดลงอย่างเห็นได้ชัดเมื่อระบบล้มเหลว ในขณะที่เมื่อระบบ AI ไม่มีอิสระในการตัดสินใจ (no autonomy) เกิดข้อผิดพลาดและอยู่ภายใต้การควบคุมของมนุษย์ตลอดเวลา ความรับผิดชอบจะถูกโยนไปที่นักพัฒนาระบบ (developer) แทน สะท้อนให้เห็นว่าผู้เข้าร่วมใช้เหตุผลในการตัดสินใจความรับผิดชอบโดยพิจารณาจากระดับความควบคุมของมนุษย์ต่อ AI และลักษณะของเทคโนโลยีที่เกี่ยวข้อง

ในมิติของความรู้สึก Bigman and Gray (2018, pp. 26-30) ได้สำรวจปฏิกริยาของมนุษย์เมื่อให้ AI เป็นผู้ตัดสินใจทางจริยธรรมในบริบททางการแพทย์ โดยเปรียบเทียบกรณีผลลัพธ์จากการตัดสินใจเป็นทั้งเชิงลบและเชิงบวก พบว่าผู้เข้าร่วมมีแนวโน้มไม่ยอมรับการตัดสินใจของ AI มากกว่าเมื่อเทียบกับมนุษย์ที่ทำการตัดสินใจแบบเดียวกันและให้ผลลัพธ์เหมือนกัน อย่างไรก็ตาม เมื่อเปลี่ยนบทบาทของ AI ให้ทำหน้าที่เพียงเป็นที่ปรึกษาโดยให้มนุษย์เป็นผู้ตัดสินใจขั้นสุดท้าย พบว่ามนุษย์ยอมรับการมีส่วนร่วมของ AI ได้มากขึ้น นอกจากนี้ ยังมีสาเหตุสำคัญจากการที่มนุษย์รับรู้ AI ไม่มีคุณลักษณะของจิตใจ แม้ว่า AI จะสามารถตัดสินใจได้อย่างถูกต้อง แต่ผู้คนยังคง มองว่า AI มีระดับของความสามารถในการกระทำอย่างมีเหตุผล (agency) และความสามารถในการรับรู้อารมณ์ความรู้สึก (experience) ที่ต่ำกว่ามนุษย์ ซึ่งไม่เพียงพอสำหรับการตัดสินใจในประเด็นทางจริยธรรมจึงส่งผลให้เกิดการไม่ยอมรับอย่างต่อเนื่อง และถึงแม้งานวิจัยนี้จะไม่ได้ศึกษาถึงการกล่าวโทษโดยตรง แต่อาจนำไปสู่การกล่าวโทษโดยอ้อมได้ตามทฤษฎี Path of blame (Malle et al., 2014) ในกรณีที่ AI เข้ามามีบทบาทในสถานการณ์ที่มีผลกระทบทางจริยธรรม แม้ว่า AI จะถูกรับรู้ว่ามี agency และ experience แต่กลับถูกมองว่ามีความสามารถและหน้าที่ในการป้องกันเหตุที่ก่อให้เกิดผลลัพธ์เชิงลบ ซึ่งอาจทำให้เกิดการกล่าวโทษต่อ AI ได้

นอกเหนือจากระดับการรับรู้จิตใจและความรู้สึกแล้ว อีกหนึ่งปัจจัยที่ส่งผลต่อการกล่าวโทษอย่างเห็นได้ชัดเจนนั่นคือ ระดับความเป็นอิสระของตัวแทน จากงานของ Kim and Hinds (2006, p. 83) กล่าวว่า

ในบริบทการทำงานร่วมกันระหว่างหุ่นยนต์กับมนุษย์ มนุษย์มีแนวโน้มที่จะกล่าวโทษหุ่นยนต์ที่มีระดับความเป็นอิสระ (autonomy) สูงมากกว่าหุ่นยนต์ที่มีอิสระต่ำ และกล่าวโทษตนเองน้อยลงเมื่อทำงานร่วมกับหุ่นยนต์ที่มีความอิสระสูง สะท้อนให้เห็นว่า เมื่อหุ่นยนต์แสดงพฤติกรรมที่ดูเหมือนมีเจตนา การกล่าวโทษของมนุษย์จะเปลี่ยนไป โดยมองว่าหุ่นยนต์สามารถเป็นผู้กระทำที่รับผิดชอบต่อผลลัพธ์ได้ สอดคล้องกับงานวิจัยของ Furlough et al. (2019, pp. 4-6) ที่พบว่าในสถานการณ์ที่เกิดผลลัพธ์ล้มเหลวเมื่อทำงานร่วมกัน หุ่นยนต์ที่มีระดับความเป็นอิสระสูง จะได้รับการกล่าวโทษสูงใกล้เคียงมนุษย์โดยยังมีความเป็นอิสระมากเท่าไรก็ยังมีแนวโน้มที่จะถูกกล่าวโทษมากขึ้น แสดงให้เห็นว่ามนุษย์มีแนวโน้มที่จะกล่าวโทษผู้กระทำที่มีความเป็นอิสระสูงมากกว่า เพราะสามารถควบคุมพฤติกรรมการตัดสินใจของตนเองได้ (Alicke, 2000, pp. 599-560)

2.4 ช่องว่างของงานวิจัยและความจำเป็นของการศึกษา

ในช่วงไม่กี่ปีที่ผ่านมา มีแนวโน้มที่เพิ่มขึ้นของงานวิจัยที่มุ่งศึกษาบทบาทของระบบปัญญาประดิษฐ์ (AI) ในกระบวนการตัดสินใจโดยเฉพาะในฐานะของผู้ให้คำแนะนำ พบว่า ผู้คนมักเชื่อมโยงความรับผิดชอบกับตัวแทนที่เป็นมนุษย์มากกว่าตัวแทนที่เป็น AI (Gazit et al., 2023, p. 5) และเมื่อเกิดผลลัพธ์ที่ไม่พึงประสงค์ขึ้น ผู้ที่เป็นมนุษย์มักถูกกล่าวโทษมากกว่า (Gómez-Sánchez et al., 2023, p. 262) อย่างไรก็ตาม งานวิจัยจำนวนไม่น้อยยังคงมอง AI ในฐานะเครื่องมือที่ใช้สนับสนุนมนุษย์มากกว่าจะพิจารณาว่าเป็นตัวแทนที่มีเจตจำนง (agency) หรือมีจิตใจ (mind) เป็นของตนเอง (Stefanidis et al., 2022) แม้แต่ในงานวิจัยที่ศึกษาเรื่อง Explainable AI เช่น David et al. (2021) ซึ่งมีการเสนอรูปแบบการอธิบายเบื้องหลังคำแนะนำของ AI ก็ยังเน้นไปที่การเพิ่มความโปร่งใสของกระบวนการคำนวณ มากกว่าการสื่อสารว่า AI มีความคิดหรือความตั้งใจในเชิงจิตวิทยา ซึ่งเป็นมิติที่อาจมีความสำคัญต่อการประเมินความรับผิดชอบเชิงจริยธรรม เช่น การกล่าวโทษ เป็นต้น

ในขณะเดียวกัน งานวิจัยบางชิ้นเริ่มหันมาให้ความสนใจกับประเด็นด้านการรับรู้ ว่า AI มีจิตใจหรือไม่ และผลของการรับรู้ที่มีต่อการกล่าวโทษ เช่น Joo (2024, p. 6) พบว่า AI ที่ถูกมองว่ามีจิตใจคล้ายมนุษย์มักถูกกล่าวโทษมากกว่า ในขณะที่งานของ Stuart and Kneer (2021, pp. 17-18) ชี้ให้เห็นว่า ผู้คนมีแนวโน้มจะกล่าวโทษ AI มากขึ้นหากเชื่อว่า AI มีความคิดและเจตนา อย่างไรก็ตาม เมื่อต้องเผชิญกับผลเสียที่เกิดขึ้นจริง คนกลับลังเลที่จะกล่าวโทษ AI และมีแนวโน้มที่จะหาผู้รับผิดชอบที่เป็นมนุษย์แทน

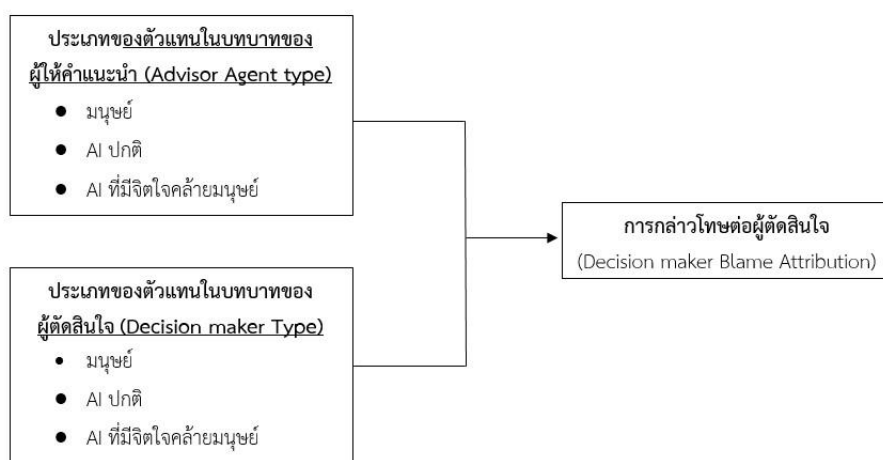
แม้จะมีการศึกษาความสัมพันธ์ระหว่างระดับการรับรู้จิตใจ (mind perception) กับการกล่าวโทษอยู่บ้าง อย่างไรก็ตาม งานวิจัยเหล่านี้ส่วนใหญ่มักเน้นศึกษา AI ในบทบาทของ “ผู้ตัดสินใจ” เพียงอย่างเดียว โดยยังไม่เปรียบเทียบกับกรณีที่ AI ทำหน้าที่เป็น “ผู้ให้คำแนะนำ” ซึ่งเป็นบทบาทที่พบได้บ่อยในระบบ AI ที่ถูกนำไปใช้จริง โดยเฉพาะอย่างยิ่งในบริบทที่มีความเสี่ยงสูง เช่น การแพทย์ เป็นต้น นอกจากนี้ แม้จะมีงานวิจัยบางชิ้นที่เปรียบเทียบผลของบทบาทผู้ให้คำแนะนำกับผู้ตัดสินใจของ AI แล้ว (เช่น Lima et al., 2021) แต่ก็ยังขาดการพิจารณาประเด็นด้านลักษณะเชิงจิตวิทยาของ AI ว่ามีผลต่อการประเมินความรับผิดชอบหรือไม่ ทั้งที่คุณลักษณะดังกล่าวอาจเป็นตัวแปรสำคัญที่กำหนดว่าผู้คนจะกล่าวโทษ AI ในลักษณะใด อีกทั้ง

งานวิจัยที่ผสมผสานทั้งบทบาทของ AI และระดับของ mind perception เข้าไว้ด้วยกัน โดยเฉพาะในบริบททางการแพทย์ซึ่งมีลักษณะเฉพาะและความเสี่ยงสูง กลับมีอยู่อย่างจำกัด

ดังนั้น งานวิจัยในครั้งนี้จึงมีวัตถุประสงค์เพื่อศึกษาผลของประเภทของตัวแทน (มนุษย์, AI ปกติ, และ AI ที่มีจิตใจคล้ายมนุษย์) และบทบาทของตัวแทน (ผู้ให้คำแนะนำ และผู้ตัดสินใจ) ต่อระดับของการกล่าวโทษผู้ตัดสินใจในบริบททางการแพทย์ ซึ่งเป็นการเติมเต็มช่องว่างขององค์ความรู้ที่ยังขาดแคลนในงานวิจัยที่ผ่านมา

แผนภาพที่ 1

กรอบแนวคิดการวิจัย



3. สมมติฐานการวิจัย

3.1 ประเภทของตัวแทนในบทบาทของผู้ให้คำแนะนำมีอิทธิพลต่อการกล่าวโทษอย่างมีนัยสำคัญทางสถิติ

3.2 ประเภทของตัวแทนในบทบาทของผู้ตัดสินใจมีอิทธิพลต่อการกล่าวโทษอย่างมีนัยสำคัญทางสถิติ

3.3 อิทธิพลเชิงปฏิสัมพันธ์ระหว่างประเภทของตัวแทนในบทบาทของผู้ให้คำแนะนำและประเภทของตัวแทนในบทบาทของผู้ตัดสินใจที่มีต่อการกล่าวโทษมีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ

4. วิธีการศึกษาวิจัย

4.1 วิธีการศึกษา

งานวิจัยนี้เป็นงานวิจัยกึ่งทดลอง (quasi-experimental design) เพื่อศึกษาอิทธิพลของประเภทของตัวแทนในบทบาทของผู้ให้คำแนะนำ (Advisor Agent type) และประเภทของตัวแทนในบทบาท

ผู้ตัดสินใจ (Decision maker agent type) ที่มีต่อการกล่าวโทษผู้ตัดสินใจในบริบททางการแพทย์ โดยผู้เข้าร่วมวิจัยทุกคนทำทุกเงื่อนไขการทดลอง (Within-subjects design)

4.2 ประชากรและกลุ่มตัวอย่าง

ประชากรในงานวิจัยครั้งนี้ คือนักศึกษาระดับปริญญาตรีที่ลงทะเบียนเรียนในรายวิชาจิตวิทยา เลือกกลุ่มตัวอย่างด้วยวิธีการสุ่มตามสะดวก (convenience sampling) ผ่านการลงทะเบียนใน Google Form ด้วยความสมัครใจ โดยผู้วิจัยได้ดำเนินการคำนวณหาขนาดกลุ่มตัวอย่างจากโปรแกรม G*Power สำหรับ two-way repeated-measures ANOVA ผลการคำนวณระบุว่าต้องมีผู้เข้าร่วมอย่างน้อย $N = 22$ คน ซึ่งการวิจัยในครั้งนี้ได้กลุ่มตัวอย่างเบื้องต้นมาจำนวน 115 คน มีผู้เข้าร่วมวิจัยที่ถอนตัวและไม่แสดงตัวทั้งสิ้น 18 คน และไม่พบผู้เข้าร่วมที่ไม่ผ่านการตรวจสอบความใส่ใจ (attention check) จึงทำให้มีข้อมูลกลุ่มตัวอย่างรวมทั้งสิ้น 97 คน ที่นำมาใช้ในการวิเคราะห์ผล

4.3 เครื่องมือที่ใช้ในการวิจัย

4.3.1 สถานการณ์จำลองทางการแพทย์

คณะผู้วิจัยสร้างสถานการณ์จำลองทางการแพทย์ขึ้นมาด้วย AI จำนวน 27 สถานการณ์ โดยกำหนดสถานการณ์ให้เป็นไปตามเงื่อนไขของงานวิจัย ดังนี้

- 1) เป็นสถานการณ์ที่ผู้ให้คำแนะนำ ให้คำแนะนำที่เหมาะสม แต่ผู้ตัดสินใจปฏิเสธที่จะทำตามคำแนะนำหรือกล่าวคือ ตัดสินใจกระทำในแนวทางที่ไม่เหมาะสม
- 2) เป็นสถานการณ์ที่นำมาซึ่งผลลัพธ์ที่ไม่พึงประสงค์ โดยที่ความร้ายแรงของผลลัพธ์ในทุกสถานการณ์มีความไม่แตกต่างกันจนเกินไป

หลังจากนั้นนำสถานการณ์ทั้งหมดให้ผู้มีความรู้ทางการแพทย์จำนวน 2 ท่าน ประเมินและตรวจสอบความถูกต้องของเนื้อหา ลำดับถัดมาจึงนำสถานการณ์ที่ได้หลังจากผ่านการตรวจสอบมาทำการศึกษา นำร่อง (pilot study) โดยการให้ผู้เข้าร่วมการศึกษาแบบนำร่องประเมินการกล่าวโทษต่อผู้ตัดสินใจในแต่ละสถานการณ์ และคัดเลือกสถานการณ์ที่ขึ้นให้เกิดการกล่าวโทษเฉลี่ยที่ใกล้เคียงกันที่สุดเพียง 9 สถานการณ์ตามเงื่อนไขที่ได้กล่าวไปข้างต้น เพื่อลดอคติในการกล่าวโทษที่อาจเกิดจากการชักจูงของสถานการณ์

4.3.2 มาตรการวัดการกล่าวโทษ (Blame Attribution)

แบบวัด Blame Attribution ที่ใช้ในงานวิจัยนี้พัฒนามาจากงานของ Kim and Hinds (2006, p. 82) ซึ่งได้พัฒนาและตรวจสอบความเชื่อมั่นของเครื่องมือด้วยการให้ผู้เข้าร่วมตอบคำถามสองข้อและหาค่าสัมประสิทธิ์แอลฟาของครอนบาค (Cronbach's Alpha) ในข้อคำถามเกี่ยวกับการกล่าวโทษต่อตัวหุ่นยนต์ มีค่าเฉลี่ยอยู่ที่ 0.808

สำหรับงานวิจัยปัจจุบันเนื่องจากคำถามที่นำมา มีลักษณะที่ตรงไปตรงมาและไม่ซับซ้อน ผู้วิจัยจึงได้แปลแบบทดสอบนี้และได้ดำเนินการแปลย้อน (back-translate) จากนั้นให้ผู้เชี่ยวชาญประเมินความตรงเชิงเนื้อหา (face validity) และนำปรับใช้จำนวนทั้งหมด 2 ข้อ วัดต่อ 1 สถานการณ์ โดยใช้เป็น

มาตรวัดแบบลิเคิร์ต (Likert scale) 7 ระดับในการประเมิน ตั้งแต่ 1 คะแนน (ไม่เห็นด้วยอย่างยิ่ง) จนถึง 7 คะแนน (เห็นด้วยอย่างยิ่ง) ซึ่งปรับคำในข้อคำถามจากคำว่า “robot” เป็น “ผู้ตัดสินใจ” ในแต่ละสถานการณ์ เพื่อความสอดคล้องกับจุดประสงค์ของงานวิจัย

4.4 กระบวนการวิจัยและเก็บรวบรวมข้อมูล

คณะผู้วิจัยทำการนัดหมายเวลากับผู้เข้าร่วมวิจัย เมื่อถึงเวลาให้ผู้เข้าร่วมวิจัยทำแบบสอบถามผ่านเว็บไซต์ Qualtrics โดยเริ่มจากการให้ผู้เข้าร่วมวิจัยอ่านหนังสือขอความยินยอมที่มีการชี้แจงถึงจุดประสงค์ ความเสี่ยงที่อาจได้รับ การเก็บรักษาความลับ และรวมถึงสิทธิในการถอนตัวของผู้เข้าร่วมวิจัยได้อย่างเสรีก่อนเริ่มทำการทดลอง หลังจากแสดงความยินยอมแล้วผู้เข้าร่วมวิจัยจึงจะได้อ่านสถานการณ์จำลองทางการแพทย์ที่กำหนดให้ผู้เข้าร่วมอย่างสุ่มในลักษณะที่ทุกคนทำทุกเงื่อนไขการทดลอง (Within-subjects design) ทั้งหมด 9 สถานการณ์ ซึ่งประกอบไปด้วย 3 (ประเภทของตัวแทนในบทบาทผู้ให้คำแนะนำ) $\times$ 3 (ประเภทของตัวแทนในบทบาทผู้ตัดสินใจ) ที่แตกต่างกัน ดังนี้

1. สถานการณ์ที่มีผู้ให้คำแนะนำเป็นมนุษย์ และผู้ตัดสินใจเป็นมนุษย์
2. สถานการณ์ที่มีผู้ให้คำแนะนำเป็นมนุษย์ และผู้ตัดสินใจเป็นปัญญาประดิษฐ์
3. สถานการณ์ที่มีผู้ให้คำแนะนำเป็นมนุษย์และผู้ตัดสินใจเป็นปัญญาประดิษฐ์ที่มีจิตใจคล้ายมนุษย์
4. สถานการณ์ที่มีผู้ให้คำแนะนำเป็นปัญญาประดิษฐ์ และผู้ตัดสินใจเป็นมนุษย์
5. สถานการณ์ที่มีผู้ให้คำแนะนำเป็นปัญญาประดิษฐ์และผู้ตัดสินใจเป็นปัญญาประดิษฐ์
6. สถานการณ์ที่มีผู้ให้คำแนะนำเป็นปัญญาประดิษฐ์และผู้ตัดสินใจเป็นปัญญาประดิษฐ์ที่มีจิตใจคล้ายมนุษย์
7. สถานการณ์ที่มีผู้ให้คำแนะนำเป็นปัญญาประดิษฐ์ที่มีจิตใจคล้ายมนุษย์และผู้ตัดสินใจเป็นมนุษย์
8. สถานการณ์ที่มีผู้ให้คำแนะนำเป็นปัญญาประดิษฐ์ที่มีจิตใจคล้ายมนุษย์และผู้ตัดสินใจเป็นปัญญาประดิษฐ์
9. สถานการณ์ที่มีผู้ให้คำแนะนำเป็นปัญญาประดิษฐ์ที่มีจิตใจคล้ายมนุษย์และผู้ตัดสินใจเป็นปัญญาประดิษฐ์ที่มีจิตใจคล้ายมนุษย์

โดยสถานการณ์ทั้งหมดที่ผู้เข้าร่วมวิจัยได้อ่าน คณะผู้วิจัยได้ทำการถ่วงดุลตัวแปรในรูปแบบ counterbalancing เพื่อป้องกันอิทธิพลอันเนื่องมาจากการนำเสนอ หลังจากนั้น เมื่ออ่านสถานการณ์จบ 1 สถานการณ์ ผู้เข้าร่วมวิจัยจะได้อบบแบบประเมินการกล่าวโทษและแบบสอบถามวัดความใส่ใจ (attention check) กระทำซ้ำจนครบทั้งหมด 9 เงื่อนไขการทดลอง ลำดับสุดท้ายจึงเป็นการเก็บข้อมูลประชากรศาสตร์และอ่านเอกสารชี้แจงข้อมูลหลังจบการวิจัย (Debrief)

4.5 การวิเคราะห์ผลและการแปลผลข้อมูล

ผู้วิจัยได้นำข้อมูลที่เก็บรวบรวมได้จากแบบสอบถามมาวิเคราะห์ข้อมูลโดยใช้โปรแกรม Statistical Package for the Social Science: SPSS และวิเคราะห์ข้อมูล ดังนี้

1. การวิเคราะห์สถิติเชิงพรรณนาจากข้อมูลลักษณะประชากร ได้แก่ เพศ และชั้นปีการศึกษา
2. การทดสอบสมมติฐานเพื่อทดสอบอิทธิพลของประเภทของตัวแทนในบทบาทของผู้ให้คำแนะนำและประเภทของตัวแทนในบทบาทผู้ตัดสินใจที่มีต่อการกล่าวโทษด้วยการวิเคราะห์ความแปรปรวนสองทางแบบวัดซ้ำ (two-way repeated measures ANOVA)

5. ผลการวิจัย

5.1 การตรวจสอบความใส่ใจ (attention check)

การตรวจสอบความใส่ใจของผู้เข้าร่วมวิจัยกระทำด้วยข้อคำถามวัดความใส่ใจจำนวน 4 ข้อ ผลพบว่าผู้เข้าร่วมวิจัยทั้งหมด 97 คน ไม่มีผู้ที่ไม่ผ่านเกณฑ์ ดังนั้น จึงเลือกกลุ่มตัวอย่างที่นำมาวิเคราะห์ข้อมูลต่อได้ทั้งสิ้นจำนวน 97 คน คิดเป็นร้อยละ 100

5.2 การวิเคราะห์สถิติเชิงพรรณนา

ตารางที่ 1

จำนวนและร้อยละของลักษณะทางประชากรศาสตร์ของกลุ่มตัวอย่าง

ข้อมูลกลุ่มตัวอย่าง	จำนวน (คน)	ร้อยละ
เพศ		
ชาย	30	30.93
หญิง	61	62.89
LGBTQIA+	5	5.15
ไม่ระบุ	1	1.03
ชั้นปีการศึกษา		
ปี 1	46	47.42
ปี 2	19	19.59
ปี 3	21	21.65
ปี 4	10	10.31
สูงกว่าชั้นปีที่ 4	1	1.03

จากตารางที่ 1 แสดงให้เห็นว่ากลุ่มตัวอย่างโดยส่วนใหญ่เป็นเพศหญิง จำนวน 62 คน (ร้อยละ 63.92) และโดยส่วนมากศึกษาอยู่ในชั้นปีที่ 1 จำนวน 46 คน (ร้อยละ 47.42)

5.3 การวิเคราะห์ค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐาน

ค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของตัวแปรที่ศึกษา คือ การกล่าวโทษต่อผู้ตัดสินใจ โดยวัดจากมาตรวัดการกล่าวโทษ จำนวน 2 ข้อต่อ 1 สถานการณ์ ประเมินผ่าน Likert scale 7 ระดับ ตั้งแต่ 1

(เห็นน้อยที่สุด) ถึง 7 (เห็นด้วยอย่างยิ่ง) มีค่าเฉลี่ยเลขคณิตและค่าส่วนเบี่ยงเบนมาตรฐาน ดังแสดงในตารางที่ 2 ต่อไปนี้

ตารางที่ 2

ตารางแสดงค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของการกล่าวโทษผู้ตัดสินใจ (Decision-maker blame attribution) ในแต่ละเงื่อนไขการทดลอง

เงื่อนไข	M	SD
ผู้ให้คำแนะนำมนุษย์ * ผู้ตัดสินใจมนุษย์	5.20	1.59
ผู้ให้คำแนะนำมนุษย์ * ผู้ตัดสินใจปัญญาประดิษฐ์	4.93	1.70
ผู้ให้คำแนะนำมนุษย์ * ผู้ตัดสินใจปัญญาประดิษฐ์ที่มีจิตใจคล้ายมนุษย์	5.11	1.78
ผู้ให้คำแนะนำปัญญาประดิษฐ์ * ผู้ตัดสินใจมนุษย์	4.45	1.70
ผู้ให้คำแนะนำปัญญาประดิษฐ์ * ผู้ตัดสินใจปัญญาประดิษฐ์	4.64	1.78
ผู้ให้คำแนะนำปัญญาประดิษฐ์ * ผู้ตัดสินใจปัญญาประดิษฐ์ที่มีจิตใจคล้ายมนุษย์	4.56	1.94
ผู้ให้คำแนะนำปัญญาประดิษฐ์ที่มีจิตใจคล้ายมนุษย์ * ผู้ตัดสินใจมนุษย์	4.82	1.71
ผู้ให้คำแนะนำปัญญาประดิษฐ์ที่มีจิตใจคล้ายมนุษย์ * ผู้ตัดสินใจปัญญาประดิษฐ์	4.78	1.84
ผู้ให้คำแนะนำปัญญาประดิษฐ์ที่มีจิตใจคล้ายมนุษย์ * ผู้ตัดสินใจปัญญาประดิษฐ์ที่มีจิตใจคล้ายมนุษย์	4.70	1.84

n = 97

5.3 การทดสอบสมมติฐาน

สมมติฐานงานวิจัย

1. ประเภทของตัวแทนในบทบาทของผู้ให้คำแนะนำมีอิทธิพลต่อการกล่าวโทษ
2. ประเภทของตัวแทนในบทบาทผู้ตัดสินใจมีอิทธิพลต่อการกล่าวโทษ
3. ประเภทของตัวแทนในบทบาทของผู้ให้คำแนะนำและประเภทของตัวแทนในบทบาทผู้ตัดสินใจมีอิทธิพลเชิงปฏิสัมพันธ์ต่อการกล่าวโทษ

ตารางที่ 3

การทดสอบ Mauchly's Test of Sphericity

	χ^2	p	df
ประเภทของตัวแทนในบทบาทของผู้ให้คำแนะนำ	4.47	.107	2
ประเภทของตัวแทนในบทบาทผู้ตัดสินใจ	21.34	< .001*	2
ประเภทของตัวแทนในบทบาทของผู้ให้คำแนะนำ * ประเภทของตัวแทนในบทบาทผู้ตัดสินใจ	18.31	.032*	9

n = 97

* หมายถึง $p < .05$

ตารางที่ 4

การวิเคราะห์ความแปรปรวนสองทางแบบวัดซ้ำ

	<i>F</i>	<i>p</i>	η_p^2
ประเภทของตัวแทนในบทบาทของผู้ให้คำแนะนำ	13.79	< .001*	.126
ประเภทของตัวแทนในบทบาทผู้ตัดสินใจ	.07	.908	.001
ประเภทของตัวแทนในบทบาทของผู้ให้คำแนะนำ*ประเภทของตัวแทนในบทบาทผู้ตัดสินใจ	.94	.437	.010

n = 97

* หมายถึง $p < .05$

จากการทดสอบสมมติฐานที่ 1 ด้วย Mauchly's test พบว่าเงื่อนไขของ sphericity เป็นไปตามสมมติฐาน, $\chi^2 (2) = 4.47, p = .107$ ดังนั้นจึงใช้ค่าความเป็นอิสระ (degrees of freedom) แบบไม่ปรับแก้ ซึ่งพบว่าประเภทของตัวแทนในบทบาทของผู้ให้คำแนะนำมีอิทธิพลต่อการกล่าวโทษผู้ตัดสินใจอย่างมีนัยสำคัญทางสถิติที่ระดับ .05, $F (2,192) = 13.79, p < .001, \eta_p^2 = .126$ (แสดงในตารางที่ 3 และ 4)

จากนั้นจึงทำการวิเคราะห์เปรียบเทียบรายคู่ (Pairwise Comparisons) ตัวแปรประเภทของตัวแทนในบทบาทของผู้ให้คำแนะนำด้วยวิธี Bonferroni พบว่าการกล่าวโทษในบริบทที่ผู้ให้คำแนะนำเป็นมนุษย์ ($M = 5.08, SD = .13$) นั้นมีค่าเฉลี่ยสูงที่สุดและมีความแตกต่างอย่างมีนัยสำคัญทางสถิติกับผู้ให้คำแนะนำที่เป็นปัญญาประดิษฐ์ ($M = 4.55, SD = .15$), $p < .001$ และผู้ให้คำแนะนำที่เป็นปัญญาประดิษฐ์ที่มีจิตใจคล้ายมนุษย์ ($M = 4.77, SD = .14$), $p = .018$ นอกจากนั้นไม่พบความแตกต่างอย่างมีนัยสำคัญทางสถิติระหว่างผู้ให้คำแนะนำที่เป็นปัญญาประดิษฐ์และผู้ให้คำแนะนำที่เป็นปัญญาประดิษฐ์ที่มีจิตใจคล้ายมนุษย์ $p = .059$ (แสดงในตารางที่ 5)

ตารางที่ 5

การวิเคราะห์เปรียบเทียบรายคู่ (Pairwise Comparisons) ของตัวแปรประเภทของตัวแทนในบทบาทของผู้ให้คำแนะนำ

Measure		Mean	Std.	<i>p</i>	95% CI	
Advisor (I)	Advisor (J)	Diff. (I-J)	Error		Lower	Upper
H	AI	.529*	.100	< .001	.285	.774
	HAI	.311*	.111	.018	.041	.581
AI	H	-.529*	.100	< .001	-.774	.285
	HAI	-.218	.092	.059	-.442	.006
HAI	H	-.311*	.111	.018	-.581	.041
	AI	.218	.092	.059	-.006	.440

n = 97

* หมายถึง $p < .05$

H และ HAI หมายถึง Human และ Human mind-liked AI ตามลำดับ

ต่อมาเมื่อทดสอบสมมติฐานที่ 2 ด้วย Mauchly's test พบว่าเงื่อนไขของ sphericity นั้นถูกละเมิด, $\chi^2 (2) = 21.34, p < .001$ ดังนั้นจึงทำการปรับค่าความเป็นอิสระ (degrees of freedom) โดยใช้การปรับด้วย Huynh-Feldt ($\epsilon = .845$) ซึ่งหลังจากทำการปรับค่าแล้ว พบว่าประเภของตัวแทนในบทบาทผู้ตัดสินใจไม่มีอิทธิพลต่อการกล่าวโทษอย่างมีนัยสำคัญทางสถิติที่ระดับ .05, $F (1.69, 2.88) = .07, p = .908, \eta_p^2 = .001$ (แสดงในตารางที่ 3 และ 4)

และลำดับสุดท้ายทำการทดสอบสมมติฐานที่ 3 ด้วย Mauchly's test พบว่าเงื่อนไขของ sphericity นั้นถูกละเมิด, $\chi^2 (9) = 18.31, p = .032$ ดังนั้นจึงได้ทำการปรับค่าความเป็นอิสระ (degrees of freedom) โดยใช้การปรับด้วย Huynh-Feldt ($\epsilon = .962$) ซึ่งหลังจากทำการปรับค่าแล้ว พบว่าผลของอิทธิพลเชิงปฏิสัมพันธ์ระหว่างประเภของตัวแทนในบทบาทของผู้ให้คำแนะนำและประเภของตัวแทนในบทบาทผู้ตัดสินใจต่อการกล่าวโทษ ไม่มีความแตกต่างอย่างมีนัยสำคัญทางสถิติที่ระดับ .05, $F (3.85, 369.25) = .94, p = .437, \eta_p^2 = .010$ (แสดงในตารางที่ 3 และ 4)

6. อภิปรายผล

ผลการวิจัยพบว่าประเภของตัวแทนในบทบาทผู้ให้คำแนะนำมีอิทธิพลต่อการกล่าวโทษผู้ตัดสินใจอย่างมีนัยสำคัญทางสถิติ โดยเมื่อผู้ให้คำแนะนำอยู่ในรูปแบบของมนุษย์จะทำให้เกิดการกล่าวโทษต่อผู้ตัดสินใจมากกว่าผู้ให้คำแนะนำรูปแบบ AI หรือ AI ที่มีจิตใจคล้ายมนุษย์ นอกจากนี้ AI ที่มีจิตใจคล้ายมนุษย์ยังมีแนวโน้มที่จะทำให้ผู้ตัดสินใจถูกกล่าวโทษสูงกว่า AI โดยทั่วไปเช่นกัน แต่ถึงอย่างไรก็ตาม ไม่พบอิทธิพลของประเภของตัวแทนในบทบาทผู้ตัดสินใจต่อการกล่าวโทษผู้ตัดสินใจอย่างมีนัยสำคัญทางสถิติ

หนึ่งในเหตุผลที่อาจอธิบายผลของการศึกษาในครั้งนี้ที่พบว่าความแตกต่างในการกล่าวโทษผู้ตัดสินใจตามประเภของผู้ให้คำแนะนำ โดยเฉพาะอย่างยิ่งในกรณีที่ผู้ให้คำแนะนำเป็นมนุษย์ คือ การที่มนุษย์มักรับรู้ว่ามีระดับความสามารถสูงกว่า AI อย่างนัยสำคัญทางสถิติ ดังนั้น เมื่อผู้ตัดสินใจปฏิเสธที่จะทำตามผู้ให้คำแนะนำที่เป็นมนุษย์ การกล่าวโทษที่มีต่อผู้ตัดสินใจจึงมีแนวโน้มสูงกว่า เมื่อเทียบกับกรณีที่ผู้ให้คำแนะนำที่เป็น AI หรือประเภอื่น ๆ ข้อสังเกตนี้สอดคล้องกับงานวิจัยของ Frischknecht (2021, p. 5) ที่พบว่า แม้ AI จะมีระดับความเป็นอิสระทางเทคนิคสูง แต่ก็ยังคงถูกมองว่ามีระดับความสามารถต่ำกว่ามนุษย์อย่างมีนัยสำคัญ ซึ่งอาจส่งผลต่อการรับรู้และการกล่าวโทษในสถานการณ์ที่เกี่ยวข้องกับการตัดสินใจ โดยข้อสังเกตนี้ยังสนับสนุนแนวโน้มที่ว่าผู้ตัดสินใจจะถูกกล่าวโทษมากขึ้นหาก AI ที่ให้คำแนะนำถูกบรรยายว่ามีคุณลักษณะบางประการที่เกี่ยวข้องกับมนุษย์ ซึ่งอาจทำให้มนุษย์มองว่า AI เหล่านี้มีความคล้ายคลึงมนุษย์มากขึ้น จึงรับรู้ความสามารถของ AI ลักษณะนี้สูงกว่า AI ทั่วไปและนำไปสู่การประเมินความรับผิดชอบต่อความผิดทางจริยธรรมที่มากกว่า โดยเฉพาะเมื่อเกิดการรับรู้ถึงระดับจิตใจของ AI (Joo, 2024, pp. 7-8)

อย่างไรก็ตาม ผลการวิจัยส่วนหนึ่งที่พบว่าประเภของผู้ตัดสินใจไม่ได้ส่งผลอย่างมีนัยสำคัญทางสถิติต่อการกล่าวโทษผู้ตัดสินใจ ซึ่งอาจสะท้อนได้ว่าผู้คนไม่ได้ให้ความสำคัญกับลักษณะของประเภของตัวแทนในการตัดสินใจ หากแต่ให้ความสำคัญกับผลลัพธ์ที่ไม่พึงประสงค์เพียงอย่างเดียว ปรากฏการณ์นี้สามารถอธิบายได้โดยทฤษฎี Outcome bias ซึ่งเสนอว่า ผู้คนมักประเมินคุณภาพของ

การตัดสินใจจากผลลัพธ์ที่เกิดขึ้น มากกว่าการพิจารณาข้อมูลที่ผู้ตัดสินใจมีในขณะนั้น นำไปสู่การตัดสินใจที่อาจไม่ยุติธรรม โดยเฉพาะอย่างยิ่ง เมื่อผลลัพธ์ออกมาดี การตัดสินใจจะถูกมองในแง่บวกมากกว่ากรณีที่ผลลัพธ์ล้มเหลว (Baron & Hershey, 1988, p. 569)

รวมไปถึงไม่พบอิทธิพลเชิงปฏิสัมพันธ์ระหว่างประเภทของตัวแทนในบทบาทผู้ให้คำแนะนำและประเภทของตัวแทนในบทบาทผู้ตัดสินใจ กล่าวคือ ประเภทของตัวแทนในบทบาทผู้ให้คำแนะนำและผู้ตัดสินใจส่งผลต่อระดับการกล่าวโทษอย่างเป็นอิสระต่อกัน

ดังนั้น ผลการวิจัยในครั้งนี้ผู้เข้าร่วมการวิจัยจะได้รับประโยชน์โดยตรงจากการได้มีโอกาสเรียนรู้และทำความเข้าใจเกี่ยวกับการตัดสินใจในบริบททางการแพทย์ภายใต้สถานการณ์ที่เกี่ยวข้องกับการให้คำแนะนำและการตัดสินใจของทั้งมนุษย์และ AI ซึ่งช่วยให้ผู้เข้าร่วมได้ตระหนักถึงความซับซ้อนของการตัดสินใจในสภาพแวดล้อมที่มีเทคโนโลยีเข้ามามีส่วนร่วม นอกจากนี้ การศึกษานี้มีส่วนสนับสนุนต่อทั้งแนวคิดเชิงทฤษฎีและการประยุกต์ใช้

ในเชิงปฏิบัติ กล่าวคือ งานวิจัยได้แสดงให้เห็นว่าการรับรู้ต่อผู้ให้คำแนะนำในรูปแบบต่าง ๆ โดยเฉพาะมนุษย์ส่งผลต่อระดับการกล่าวโทษที่มีต่อผู้ตัดสินใจ ซึ่งสะท้อนให้เห็นถึงบทบาทของความคาดหวังทางสังคม โดยเฉพาะอย่างยิ่งในกลุ่มผู้เข้าร่วมที่เป็นนักศึกษาที่มีความรู้หรือพื้นฐานด้านจิตวิทยาและเทคโนโลยี จะมีแนวโน้มให้การประเมินและการกล่าวโทษที่สะท้อนถึงทัศนคติและการรับรู้ต่อความรับผิดชอบของ AI เทียบกับมนุษย์ได้อย่างชัดเจน ซึ่งช่วยให้เข้าใจทัศนคติของคณาจารย์ใหม่ต่อการใช้ AI ในการตัดสินใจทางการแพทย์ได้ดียิ่งขึ้น

และการรับรู้ถึงความสามารถของตัวแทนในการตัดสินใจในเชิงทฤษฎี ผลลัพธ์นี้ช่วยเสริมสร้างความเข้าใจต่อกลไกการกล่าวโทษในบริบทของการตัดสินใจร่วมระหว่างมนุษย์และ AI โดยเฉพาะอย่างยิ่งในสถานการณ์ที่มีผลลัพธ์เชิงลบ ดังนั้นความสำคัญของการกำหนดบทบาทและความสัมพันธ์ระหว่างตัวแทนต่าง ๆ ในระบบตัดสินใจร่วมระหว่างมนุษย์กับ AI จึงเป็นสิ่งที่ควรพิจารณา โดยเฉพาะอย่างยิ่งในกรณีที่ผู้ให้คำแนะนำและผู้ตัดสินใจแยกจากกัน การออกแบบระบบควรคำนึงถึงการสื่อสารโครงสร้างอำนาจการตัดสินใจให้ชัดเจน เพื่อป้องกันความเข้าใจคลาดเคลื่อนเกี่ยวกับความรับผิดชอบ โดยเฉพาะอย่างยิ่งในกรณีที่มนุษย์เป็นผู้ให้คำแนะนำ เนื่องจากผู้ใช้ระบบอาจมีแนวโน้มคาดหวังว่าคำแนะนำของมนุษย์จะมีคุณภาพหรือความน่าเชื่อถือที่สูงกว่า AI ซึ่งการคาดหวังนี้อาจส่งผลโดยอ้อมต่อการกล่าวโทษต่อผู้ตัดสินใจหากไม่ปฏิบัติตามคำแนะนำนั้น รวมไปถึงในแง่มุมมองของการออกแบบ AI ที่ผลการศึกษาในครั้งนี้ชี้ว่า เมื่อผู้คนรับรู้ว่ามีคุณลักษณะทางจิตใจใกล้เคียงกับมนุษย์มากขึ้น จะมีแนวโน้มในการกล่าวโทษผู้ตัดสินใจที่มากกว่า AI ธรรมดา ดังนั้นการลดองค์ประกอบที่ทำให้ AI ดูคล้ายมนุษย์และการรับรู้ระดับจิตใจของ AI จึงยังเป็นอีกหนึ่งประเด็นที่ควรนำมาพิจารณา โดยเฉพาะการนำมาใช้งานในระบบการตัดสินใจเชิงจริยธรรม

7. ข้อจำกัดและข้อเสนอแนะ

งานวิจัยนี้ยังมีข้อจำกัดบางประการที่อาจส่งผลต่อการตีความผลลัพธ์ กล่าวคืองานวิจัยนี้ยังไม่ได้มีการพิจารณาอย่างละเอียดเกี่ยวกับการรับรู้ความสามารถของตัวแทนแต่ละประเภทในบทบาทต่าง ๆ และยังไม่ได้ทำการศึกษาในเรื่องของการแยกระดับความเสี่ยงของสถานการณ์ทางการแพทย์ ดังนั้น

งานวิจัยในอนาคตอาจต่อยอดประเด็นนี้ โดยการสำรวจเพิ่มเติมถึงระดับความสามารถที่แตกต่างกันของตัวแทนและระดับความเสี่ยงที่แตกต่างกันในแต่ละบริบทว่าจะส่งผลอย่างไรต่อการรับรู้การกล่าวโทษ

ประการต่อมา การศึกษาครั้งนี้ยังไม่ได้มุ่งเน้นกับการให้เหตุผลของผู้ตัดสินใจ จึงควรมีการศึกษาต่อยอดที่เน้นการอธิบายเหตุผลประกอบการตัดสินใจของผู้กระทำ โดยเฉพาะในบริบทที่การตัดสินใจนำไปสู่ผลลัพธ์ที่ไม่พึงประสงค์งานวิจัยที่เกี่ยวข้องกับ Explainable AI (XAI) พบว่าการศึกษาโดยเฉพาะอย่างยิ่งในกลุ่มผู้ที่ไม่มีความเชี่ยวชาญแสดงให้เห็นว่า เมื่อ AI เกิดความผิดพลาด ผู้ใช้มีแนวโน้มให้ความสำคัญกับคำอธิบายเชิงกลไกหรือคำอธิบายที่มีรายละเอียดมากขึ้น (Schoenherr & Thomson, 2024, p. 9) เช่นเดียวกับกับในทฤษฎี Path of Blame (Malle et al., 2014, p. 151) ที่เชื่อว่าผู้คนจะพิจารณาเหตุผลประกอบการตัดสินใจเมื่อรู้ว่าผู้กระทำมีเจตนาในการทำให้เกิดผลลัพธ์ที่ไม่พึงประสงค์ ดังนั้นรูปแบบและเนื้อหาของคำอธิบายอาจส่งผลต่อการประเมิน การกล่าวโทษของทั้งมนุษย์และ AI ในกรณีที่เกิดผลลัพธ์ที่ไม่พึงประสงค์ ซึ่งเป็นประเด็นที่ควรได้รับการพิจารณาอย่างจริงจังในการวิจัยในอนาคต

นอกจากนี้ เนื่องจากงานวิจัยนี้ศึกษาในกรณีที่ผู้ตัดสินใจได้รับคำแนะนำที่เหมาะสม แต่ปฏิเสธที่จะปฏิบัติตามคำแนะนำ จึงไม่มีความจำเป็นในการประเมินการกล่าวโทษผู้ให้คำแนะนำ การศึกษาในอนาคตอาจพิจารณาการสำรวจสถานการณ์ที่ผู้ตัดสินใจเลือกปฏิบัติตามคำแนะนำ แต่ผลลัพธ์ยังคงเป็นลบเพื่อเปรียบเทียบกับกรณีที่ไม่เชื่อฟังคำแนะนำว่ามีผลต่อระดับการกล่าวโทษในลักษณะเดียวกันหรือไม่

อีกทั้ง กลุ่มตัวอย่างในการศึกษาครั้งนี้เป็นนักศึกษาระดับปริญญาตรี ซึ่งอาจไม่มีประสบการณ์โดยตรงในบริบททางการแพทย์ ทำให้การรับรู้เกี่ยวกับความซับซ้อนของกระบวนการตัดสินใจทางการแพทย์อาจมีข้อจำกัด ส่งผลให้ระดับการกล่าวโทษผู้ตัดสินใจอาจแตกต่างจากกลุ่มตัวอย่างที่มีประสบการณ์จริงในวิชาชีพดังกล่าว นอกจากนี้ การขาดประสบการณ์ทางคลินิกอาจทำให้ผู้เข้าร่วมประเมินการกระทำของ AI หรือมนุษย์บนพื้นฐานของความเหมือนมนุษย์ (anthropomorphic cues) มากกว่าการพิจารณาเหตุผลเชิงวิชาชีพ ส่งผลต่อการรับรู้จิตใจของ AI และอาจเป็นปัจจัยหนึ่งที่ทำให้ผลการศึกษามีแนวโน้มแตกต่างจากกรณีของผู้เชี่ยวชาญทางการแพทย์จริง ดังนั้น งานวิจัยในอนาคตอาจพิจารณาศึกษาเพิ่มเติมโดยเปรียบเทียบระดับการกล่าวโทษผู้ตัดสินใจระหว่างกลุ่มที่มีและไม่มีประสบการณ์ทางการแพทย์โดยตรง รวมถึงขยายกลุ่มตัวอย่างให้มีความหลากหลายมากยิ่งขึ้น ทั้งในด้านช่วงอายุและภูมิหลังทางสังคมอื่น ๆ เพื่อให้เข้าใจความแตกต่างของการรับรู้และการตัดสินใจจริยธรรมต่อบทบาทของ AI ในบริบทเสี่ยงสูงได้อย่างรอบด้านยิ่งขึ้น

ประการสุดท้ายการศึกษาในครั้งนี้ต้องการศึกษาเฉพาะสถานการณ์ที่ผู้ตัดสินใจปฏิเสธที่จะทำตามผู้ให้คำแนะนำ อย่างไรก็ตาม ในสถานการณ์จริงมักประกอบด้วยทางเลือกมากกว่าสองทาง การวิจัยในอนาคตจึงอาจพิจารณาศึกษาระดับการกล่าวโทษในบริบทที่ผู้ให้คำแนะนำเสนอทางเลือกที่ไม่เหมาะสมและผู้ตัดสินใจเลือกทางเลือกอื่นซึ่งแม้จะไม่ตรงกับคำแนะนำ แต่ยังคงเป็นทางเลือกที่ไม่เหมาะสมเช่นกัน โดยที่ผลลัพธ์ที่เกิดขึ้นยังคงเป็นผลลัพธ์ที่ไม่พึงประสงค์ เพื่อสำรวจว่าปัจจัยเหล่านี้ส่งผลต่อระดับการกล่าวโทษอย่างไร

เอกสารอ้างอิง/References

- เปิดตัว medical AI data platform หวังใช้ AI ช่วยคัดกรอง-วินิจฉัยโรค. (21 เมษายน 2568). เดลินิวส์.
<https://www.dailynews.co.th/news/4629471/>
- โรงพยาบาลบำรุงราษฎร์. (23 พฤษภาคม 2565). Ai เทคโนโลยีปัญญาประดิษฐ์ในการวินิจฉัยภาวะ
 ความผิดปกติของปอดและมะเร็งเต้านม. <https://www.bumrungrad.com/th/health-blog/may-2022/artificial-intelligence-cancer-diagnosis>
- สำนักงานบริหารเทคโนโลยีสารสนเทศเพื่อการศึกษา. (25 กรกฎาคม 2567). คนไทย 73.84% ใช้ AI ใน
 ชีวิตประจำวัน ยิ่งรายได้สูง ยิ่งใช้ AI เยอะ. <https://www.uni.net.th/index.php/uncategorized/8012/>
- สำนักงานพัฒนาวิทยาศาสตร์และเทคโนโลยีแห่งชาติ. (1 กันยายน 2567). ปัญญาประดิษฐ์ (AI) คืออะไร.
https://www.nstda.or.th/home/knowledge_post/ปัญญาประดิษฐ์-ai-คืออะไร/
- Alicke, M. D. (2000). Culpable control and the psychology of blame. *Psychological bulletin*, 126(4),
 556-574. <https://doi.org/10.1037/0033-2909.126.4.556>
- Baron, J., & Hershey, J. C. (1988). Outcome bias in decision evaluation. *Journal of Personality and
 Social Psychology*, 54(4), 569-579. <https://doi.org/10.1037//0022-3514.54.4.569>
- Bigman, Y. E., & Gray, K. (2018). People are averse to machines making moral decisions.
Cognition, 181, 21-34. <https://doi.org/10.1016/j.cognition.2018.08.003>
- Floridi, L., & Sanders, J. (2004). On the morality of artificial agents. *Minds and Machines*, 14,
 349-379. <https://doi.org/10.1023/B:MIND.0000035461.63578.9d>
- David, B. D., Resheff, Y. S., & Tron, T. (2021). Explainable AI and adoption of financial algorithmic
 advisors. In *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society*
 (pp. 390-400). <https://doi.org/10.1145/3461702.3462565>
- Frischknecht, R. (2021). A social cognition perspective on autonomous technology. *Computers in
 Human Behavior*, 122, Article 106815. <https://doi.org/10.1016/j.chb.2021.106815>
- Furlough, C., Stokes, T., & Gillan, D. J. (2019). Attributing blame to robots: I. the influence of robot
 autonomy. *Human Factors: The Journal of the Human Factors and Ergonomics Society*,
 63(4), 592-602. <https://doi.org/10.1177/0018720819880641>
- Gazit, L., Arazy, O., & Hertz, U. (2023). Choosing between human and algorithmic advisors:
 The role of responsibility sharing. *Computers in Human Behavior: Artificial Humans*, 1(2),
 Article 100009. <https://doi.org/10.1016/j.chbah.2023.100009>
- Gómez-Sánchez, J., Gordo, C., Franklin, M., Fernandez-Basso, C., & Lagnado, D. (2023). Who is
 to blame? Responsibility attribution in AI systems vs human agents in the field of air
 crashes. In *Lecture Notes in Computer Science* (pp. 256-264). https://doi.org/10.1007/978-3-031-42935-4_21

- Gray, H. M., Gray, K., & Wegner, D. M. (2007). Dimensions of mind perception. *Science*, 315(5812), 619-619. <https://doi.org/10.1126/science.1134475>
- Hanson, A., Starr, N. D., Emnett, C., Wen, R., Malle, B. F., & Williams, T. (2024). The power of advice: Differential blame for human and robot advisors and deciders in a moral advising context. In *Proceedings of the 2024 ACM/IEEE International Conference on Human-Robot Interaction* (pp. 240-249). <https://doi.org/10.1145/3610977.3634942>
- Joo, M. (2024). It's the AI's fault, not mine: Mind perception increases blame attribution to AI. *PLOS ONE*, 19(12), Article e0314559. <https://doi.org/10.1371/journal.pone.0314559>
- Kaur, S., Singla, J., Nkenyereye, L., Jha, S., Prashar, D., Joshi, G. P., El-Sappagh, S., Islam, Md. S., & Islam, S. M. (2020). Medical diagnostic systems using artificial intelligence (AI) algorithms: principles and perspectives. *IEEE Access*, 8, 228049-228069. <https://doi.org/10.1109/access.2020.3042273>
- Kempton, H., Heilinger, J. C., & Nagel, S. K. (2023). "I'm afraid I can't let you do that, doctor": Meaningful disagreements with AI in medical contexts. *AI & Society*, 38, 1407-1414. <https://doi.org/10.1007/s00146-022-01418-x>
- Kim, T., & Hinds, P. (2006). Who should I blame? Effects of autonomy and transparency on attributions in human-robot interaction. In *ROMAN 2006: The 15th IEEE International Symposium on Robot and Human Interactive Communication* (pp. 80-85). <https://doi.org/10.1109/ROMAN.2006.314398>
- Kitchens, M. B., & Meier, B. P. (2025). The fearful mind of Artificial Intelligence: Fear and perceived existential threat of artificial intelligence as a function of its cognitive and emotional capabilities. *The Journal of Social Psychology*, 166(4), 423-436. <https://doi.org/10.1080/00224545.2025.2503006>
- Krügel, S., Ammeling, J., Aubreville, M., Fritz, A., Kießig, A., & Uhl, M. (2024). Perceived responsibility in AI-supported medicine. *AI & SOCIETY*, 40(3), 1485-1495. <https://doi.org/10.1007/s00146-024-01972-6>
- Leichtmann, B., Hinterreiter, A., Humer, C., Ventura, A., Streit, M., & Mara, M. (2024). Moral reasoning in a digital age: Blaming artificial intelligence for incorrect high-risk decisions. *Current Psychology*, 43(41), 32412-32421. <https://doi.org/10.1007/s12144-024-06658-2>
- Lima, G., Grgić-Hlača, N., & Cha, M. (2021). Human perceptions on moral responsibility of AI: A case study in AI-assisted bail decision-making. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (pp. 1-17). <https://doi.org/10.1145/3411764.3445260>
- Malle, B. F., Guglielmo, S., & Monroe, A. E. (2014). A theory of blame. *Psychological Inquiry*, 25(2), 147-186. <https://doi.org/10.1080/1047840x.2014.877340>
- McMurtrie, B. L. (2023). Investigating the dimensions of mind perception. *Journal of European Psychology Students*, 14(1), 54-68. <https://doi.org/10.59477/jeps.594>

- Miller, T. (2019). Explanation in artificial intelligence: Insights from the social sciences. *Artificial Intelligence*, 267, 1-38. <https://doi.org/10.1016/j.artint.2018.07.007>
- Schoenherr, J. R., & Thomson, R. (2024). When AI fails, who do we blame? Attributing responsibility in human–AI interactions. *IEEE Transactions on Technology and Society*, 5(1), 61-70. <https://doi.org/10.1109/TTS.2024.3370095>
- Shank, D. B., Graves, C., Gott, A., Gamez, P., & Rodriguez, S. (2019). Feeling our way to machine minds: People's emotions when perceiving mind in artificial intelligence. *Computers in Human Behavior*, 98, 256-266. <https://doi.org/10.1016/j.chb.2019.04.001>
- Shaver, K. G. (1985). *The attribution of blame*. Springer. <https://doi.org/10.1007/978-1-4612-5094-4>
- Stefanidis, K., Tsatsou, D., Konstantinidis, D., Gymnopoulos, L., Daras, P., Wilson-Barnes, S., Hart, K., Cornelissen, V., Decorte, E., Lalama, E., Pfeiffer, A., Hassapidou, M., Pagkalos, I., Argiriou, A., Rouskas, K., Hadjimitriou, S., Charisis, V., Dias, S. B., Diniz, J. A., ... Dimitropoulos, K. (2022). Protein AI advisor: A knowledge-based recommendation framework using expert-validated meals for healthy diets. *Nutrients*, 14(20), Article 4435. <https://doi.org/10.3390/nu14204435>
- Stuart, M. T., & Kneer, M. (2021). Guilty artificial minds: Folk attributions of mens rea and culpability to artificially intelligent agents. *Proceedings of the ACM on Human-Computer Interaction*, 5(CSCW2), Article 363. <https://doi.org/10.1145/3479507>
- Tolmeijer, S., Christen, M., Kandul, S., Kneer, M., & Bernstein, A. (2022). Capable but amoral? Comparing AI and human expert collaboration in ethical decision making. In *CHI Conference on Human Factors in Computing Systems* (pp. 1-17). <https://doi.org/10.1145/3491102.3517732>
- Verdicchio, M., & Perin, A. (2022). When doctors and AI interact: On human responsibility for artificial risks. *Philosophy & Technology*, 35(1), Article 16. <https://doi.org/10.1007/s13347-022-00506-6>
- Wamsley, L. (2020, December 19). *Stanford apologizes after vaccine allocation leaves out nearly all medical residents*. NPR. <https://www.npr.org/sections/coronavirus-live-updates/2020/12/18/948176807/stanford-apologizes-after-vaccine-allocation-leaves-out-nearly-all>
- Waytz, A., Gray, K., Epley, N., & Wegner, D. M. (2010). Causes and consequences of mind perception. *Trends in Cognitive Sciences*, 14(8), 383-388. <https://doi.org/10.1016/j.tics.2010.05.006>
- Waytz, A., Heafner, J., & Epley, N. (2014). The mind in the machine: Anthropomorphism increases trust in an autonomous vehicle. *Journal of Experimental Social Psychology*, 52, 113-117. <https://doi.org/10.1016/j.jesp.2014.01.005>