Unveiling Gender Dynamics for Mental Health Posts in Social Media and Generative Artificial Intelligence

Tom Mattson
University of Richmond
tmattson@richmond.edu

Qin Weng
Baylor University
qin_weng@baylor.edu

Jie Ren
Fordham University
jren11@fordham.edu

Abstract

We investigate the level of empathy mental health posts receive on social media and generative artificial intelligence (GenAI). Specifically, we examine gender effects to determine if posts authored by self-identified men, women, or unknown (no self-identified gender discloser) receive varying levels of empathy across different technical platforms. Using a sample of mental health posts from Reddit, we find that self-identified women receive more empathy relative to men across all platforms. We further find that Inflection Pi, a GenAI tool specifically designed to be empathetic, provides the most empathy, but it still favors self-identified women over men. Self-identified men attempting to receive empathy for their prolonged emotional distress are disadvantaged relative to self-identified women.

Keywords: Mental health, Gender effects, Social media, Generative artificial intelligence, Empathy

1. Introduction

Due to a variety of factors, mental health issues such as anxiety, prolonged emotional distress, loneliness, and depression represent significant public health issues (Bommersbach et al., 2022; Chau et al., 2020; Twenge et al., 2021). These mental health issues, however, impact women more than men (WHO, 2021). Stepanikova et al. (2020), for instance, report a two-fold difference in depression between men and women due to unequal social and economic factors.

Women experiencing mental health challenges may be marginalized due to their gender and the social stigma of being labeled as mentally ill (Fernández et al., 2022; Stepanikova et al., 2020). Therefore, they must overcome at least two stigmas to conjure up the courage to seek help, which may result in hiding their mental struggles. Seemingly, there are more socio-cultural constraints for mentally challenged women relative to men to seek professional help or help from real-life family and friends.

Using technology to find an empathetic audience may help reduce some of the real-life structural inhibitors that emotionally distressed individuals (women in particular) face. Currently, two popular technical options are social media and generative artificial intelligence (GenAI) (Chau et al., 2020; Hussain et al., 2020; Xue et al., 2023). Both platforms have their pros and cons when used to seek empathy. Social media may provide access to other individuals who have had similar lived experiences (pro) or may be a playground for victim blaming and cyberbullying (con). GenAI may provide hallucinogenic responses that are abrasive (con) or may provide empathetic responses due to their algorithmic nature (pro).

The prior literature on gender, technology, and mental health is nascent. It is an open theoretical and empirical question whether self-identified men or women will receive varying levels of empathy across different technologies. It is possible that societal gender inequalities are manifest in the large language models (LLMs) associated with GenAI systems or the human labelers involved in creating the GenAI systems mitigate or exacerbate those societal gender inequities. On social media, it is possible the minimal group nature of those interactions might result in either more or less empathy for women relative to men. As a result, we address the following research question:

RQ: Do mental health posts from self-identified men or women receive different levels of empathy within and between social media and GenAI?

To address this question, we extracted a sample of mental health posts from multiple mental illness subreddits on Reddit where individuals post highly personal and emotional narratives about their mental challenges. We then inputted those posts into three GenAI systems (ChatGPT-3.5, Bard, and Inflection Pi) to determine GenAI responses and their level of empathy to those posts. We find that posts authored by

URI: https://hdl.handle.net/10125/109632
978-0-9981331-8-8

self-identified women received more empathy across all technical platforms.

2. Literature review

2.1. Gender

Gender is a social construction and categorization based on how one identifies themselves (Ely, 1995). As a result, gender does not have to be a binary social classification (e.g., woman, trans-woman, man, or trans-man). Contrarily, biological sex is defined based on reproductive organs (i.e., male or female). Consistent with this literature, we use the terms women and men to refer to individuals' social (gender) classification.

For almost a century, social scientists have studied gender effects (its antecedents and consequences) across different individual, organizational, and societal outcomes (Feldberg, 2022; Schmader et al., 2008). A large stream of gender research is related to women in the workforce investigating women entrepreneurs, leaders, compensation gaps, job roles, and career progression differences between men and women (Belingheri et al., 2021; Fischbacher et al., 2024). Gender inequities are deeply rooted in societies, cultures, and organizations, which create gender stereotypes that impact external and internal perceptions for women that perpetuate inequities (Eagly & Karau, 2002; Ridgeway, 2001). Gender stereotypes influence the gender gap in job performance and the formation of negative perceptions about abilities to perform job tasks (Bordalo et al., 2019; Spencer et al., 1999). Gender stereotypes and discriminatory practices may be explicit or implicit (Coffman et al., 2021; Koch et al., 2015).

In the mental health space, men and women experience different types and frequencies of mental health problems. Rosenfield and Mouzon (2013) note that women tend to have more internalized disorders while men tend to have more externalized disorders. By early adulthood, women are approximately twice as likely as men to have experienced depression, and this increased likelihood persists for the next four decades (Cyranowski et al., 2000; Rosenfield & Mouzon, 2013). These differences are due to a variety of societal and cultural inequities such as lower employment rates, lower pay, and fewer leadership opportunities (Cyranowski et al., 2000; Stepanikova et al., 2020).

Much of the gender research in the information systems literature has focused on the IT workforce (or IS profession), gender stereotypes, and technology usage differences between genders (Aggarwal et al., 2023; Ahuja & Thatcher, 2005; Habib & Cornford, 2002; Oreglia & Srinivasam, 2016; Trauth & Connolly, 2021). Technologies are social products that are rarely gender neutral (Galyani Moghaddam, 2010). For instance, gender impacts perceptions of technology adoption and attitudes toward the riskiness of different technologies (Gefen & Straub, 1997; Venkatesh & Morri, 2000). Galyani Moghaddam (2010, p. 730) posits that "access and use of ICT (information communication technologies) are interwoven with socio-cultural issues and the gender gap is seen among all nations in the world." Chai et al. (2011) demonstrate that offline social issues including gender inequities persist in virtual environments.

2.2. Empathy

Empathy refers to the social, emotional, and cognitive processes that help individuals share, understand, and respond to the emotional states of others (Cuff et al., 2016; Decety, 2021). Empathy shapes how individuals interact with one another because individuals define themselves through their interactions and understandings of others (Jami et al., 2024). Empathetic individuals adapt their actions and emotional states to another individual. Empathy is a set of processes that involves the subjective understanding of the emotional states of another individual to act as if one were that person (Murphy et al., 2022).

Empathy involves the situational, contextual, and vicarious feelings (affect) and understandings (cognitions) that shape their actions toward others. In our paper, we define empathy as "a social activity performed at the intersection of individuals' constant interaction between internalized empathy-inducing situations and available social representations of the same phenomenon" (Jami et al., 2024, p. 2965-66). There is a behavioral element (or action component) to this definition in the sense that a responder (GenAI system or commenters on social media) chooses to write a response (action) to a mental health related post in a manner that may be interpreted as having greater, less, or no empathy. In this context, empathy is not necessarily a binary proposition (i.e., all-or-nothing versus varying levels of empathy) (Murphy et al., 2022).

Individuals exhibit empathy based on many situational, cultural, and contextual factors (Decety, 2021). The level of empathizing may be impacted by the following: 1) common group membership (Berendt, van Leeuwen, & Uhrich, 2024; Zaki, 2014), 2) perceptions of power and morality (Lammers et al., 2015), 3) interpersonal dynamics and the nature (direct versus indirect) of the social interactions (Beeney et al., 2011), and 4) general similarities between the target and empathizer (Batson et al., 1996).

On social media, there is a minimal group nature associated with interactions where individuals often must read between the lines of text-based inputs to infer another individual's emotions (Mattson, 2017). For some posts such as "I am grieving and want to cut myself," it is relatively easy to understand the emotional state of the poster. For other posts such as "I had an awkward encounter with my brother," it might be more challenging to understand the emotional state of the poster. On mental illness subreddits, most participating individuals have had prior direct or indirect experiences with mental health issues. Those similar life experiences may help them understand the plight of the poster, which can help commenters display varying levels of empathy in their responses. Common experiences are important antecedents for empathy (Preston & Hofelich, 2012).

On GenAI systems, sharing common experiences is not possible between a human (target) and a machine (empathizer). Boukricha and Wachsmuth (2011) argue that autonomous agents (machines) must have three elements to understand human's emotional situation to display human empathy: 1) an empathy mechanism (process whereby an empathetic emotion arises), 2) an empathy modulation (process to modulate and determine the degree of desired empathy), and 3) a response mechanism to be able to formulate an empathic response (how empathy is communicated to the target). GenAI systems have demonstrated that they can display empathy and sustain socioemotional relationships with humans (Ki et al., 2020).

2.3. Social media

Social media platforms allow individuals to make virtual connections with others, publish material (videos, images, and text), and share content with others on the platform (Kane et al., 2014). Popular social media platforms such as Facebook, Twitter (X), and Reddit have been extensively studied in the literature. Prior research has explained the variability in posting patterns, forms of expression, organizational performance, consumer engagement, public health, and misinformation (Chen et al., 2022; Dwivedi et al., 2023; Faraj & Johnson, 2011; Kim et al., 2019; Kitchens et al., 2020; Olan et al., 2024; Tajvidi & Karami, 2021). Individuals may create social and emotional bonds with others on social media platforms (Ren et al., 2012; Wang et al., 2021), but excessive social media use (left unmanaged) may be detrimental to individuals' mental well-being (Braghieri et al., 2022; Valkenburg, 2022). Emotional support, compassion, and empathy are relatively understudied outcome variables in the social media literature (Deuze, 2015).

Braghieri et al. (2022) reported an increased level of mental health issues related to the adoption of Facebook. Social media platforms are designed for engagement and are highly addictive. Social media often (not always) create a fictional reality that cannot be emulated, which may result in disappointment and emotional distress. Other related research has focused on developing algorithms to identify individual users on social media who might be at-risk for having or developing mental problems. For instance, Chau et al. (2020) developed a machine learning model to identify bloggers at risk of prolonged emotional distress. Similarly, Kumar et al. (2022) used advanced text analytics to help identify patterns related to depressive and suicidal thoughts.

Individuals often share deeply personal and emotional experiences on social media platforms in exchange for different forms of social support, compassion, empathy, and/or emotional support (Liu et al., 2020). In a mental health context, positive conversations are at the heart of good mental health (Fisher et al., 2012; Peterson, 2023). These positive conversations inevitably include elements of empathy (Mirzaei & Esmaeilzadeh, 2021; Yan & Tan, 2014). Tifferet (2020) reported that women on social media give and receive greater social support than men. Social support is an antecedent to emotional support, which includes empathy (Chen et al., 2019, 2020). Not all conversations on social media, however, are positive and supportive. Unmanaged forums and discussion threads may turn nasty and become overly negative. Those types of threads may have deleterious effects on mental well-being. Petter and Giddens (2023), for instance, note that women are often subjected to victim blaming on social media, which can lead to emotional distress and social exclusion.

2.4. Generative Artificial Intelligence (GenAI)

GenAI systems such as ChatGPT and Bard are a unique form of artificial intelligence that generates new content from a variety of inputs using pre-trained transformers and LLMs (Euchner, 2023). Different GenAI systems have been used to write computer code, craft legal briefs, and develop investing strategies. The academic literature has reported pros and cons associated with GenAI for knowledge work (Benbya et al., 2024), coding (Dohmke et al., 2023; Peng et al., 2023), employee efficiency (Brynjolfsson et al., 2023; Noy & Zhang, 2023), creative writing (Doshi & Hauser, 2023), and perceptual analyses (Li et al., 2024).

GenAI systems have a propensity to hallucinate. That is, they may provide responses that are non-sensical, against the rules, or factually inaccurate. When used to provide empathy, however, factual accuracy and even the occasional hallucination are not as important as understanding the context, plight, feelings, and emotions from the different inputs. GenAI responses may be empathetic even if the responses represent half-truths or factual inaccuracies. Like many complex information systems, GenAI systems are currently being used in ways that are different from what the original designers intended. For instance, ChatGPT and Bard were not designed to provide empathy or emotional support, but individuals (particularly younger individuals) are nevertheless using them for that purpose.

Ki et al. (2020) and Brandtzaeg and Følstad (2018) argue that individuals may get social support from GenAI systems due to their human-like competencies. Many technical systems have been theorized to be social actors, but this is even more applicable to GenAI systems. Conceptualizing GenAI systems as social actors means that the communication, reasoning, and social processes embedded in these systems are interwoven together (Guzman & Lewis, 2020). The ability for a GenAI system to adjust their responses to subtle cues makes the interactions mimic human-to-human communication. Individuals often perceive the interactions with GenAI systems as human-like (Brandtzaeg & Følstad, 2018; Ki et al., 2020). Individuals often mindlessly apply offline social rules to non-human entities such that they overuse offline social categories and gender stereotypes to guide their online interactions (Nass & Moon, 2002). Gender impacts how individuals perceive the trustworthiness and likability of the non-human entity (Nass & Moon, 2002). Therefore, scholars may use social theories to explain their socio-technical attributes (Peter & Kühne, 2018).

3. Research model

Figure 1 displays our research model. We hypothesize about differing levels of empathy across posts authored by different genders (self-disclosed or self-identified women versus men) and technologies (social media and GenAI systems). Our research model is *not relevant* for "information seeking" posts such as "is this a good therapist" or "can you recommend a good website?" For our hypotheses to be relevant, a post must be personal and emotional. The post must include a personal narrative about their own mental health challenges, or the challenges associated with an acquaintance, family member, or friend. Additionally, relevant posts must convey affect or an emotional state such that the responders (social media commenters or GenAI systems) can display empathy. The empathy outcome variable is related to matching and understanding the emotions of others. Therefore, posts must have an emotional element for empathy to be a relevant response. These personal and affective posts about their mental health struggles (e.g., "I want to cut myself" or "My brother just committed suicide") often do not ask the community any question. They simply present the narrative without any type of question to the community.

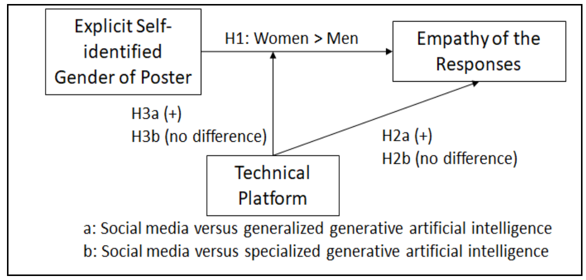

Figure 1. Research model

We argue that the main effect of gender on the level of empathy will be greater for self-identified women relative to men for two primary reasons. First, societal norms make it more appropriate for women to outwardly display their emotions more than men. Prentice and Carranza (2002) note that there is prescriptive stereotype that men should be strong while women should be caring. Based on these gender stereotypes, feminine characteristics include being affectionate, compassionate, soothing, gentle, and warm while masculine characteristics include being aggressive, competitive, ambitious, assertive, dominant, and forceful (Prentice & Carranza, 2002).

To be clear, we are not implying that these stereotypes are good. Instead, we are saying that they are deeply rooted in many societies and cultures, which means that they cannot be ignored when evaluating gender effects (Eagly & Karau, 2002; Koch et al., 2015; Ridgeway, 2001). As a result, when men outwardly display emotions or discuss their weaknesses, it goes against societal stereotypes for gender roles, which makes it hard for communities to determine an appropriate emotional (empathic) response. Furthermore, women compared with men are seen more positively, relative to societal standards when discussing their vulnerabilities (Nass & Moon, 2002; Prentice & Carranza, 2002). Mental health posts discussing the root causes of a women's emotional distress are more normative relative to men. Therefore, they will have a higher propensity to receive more empathy from the community.

Second, women and men have different communication styles due to socialization differences across societies and cultures. In general, men are expected to communicate with language that expresses independence and competitiveness while women are expected to communicate with language that expresses connections, caring, relationships, and emotions (Tannen, 1995). Mental health posts focused more on independence and competitiveness might come across as individualistic and cold. As a result, they might receive less empathetic responses, possibly resulting in more analytically driven responses. Contrarily, posts that are more affectionate and focused on relationships will probably result in less analytical and more empathetic responses due to the emotional and caring language. The LLMs associated with GenAI systems have similar patterns favoring women over men because social media responses are inevitably contained in the sample of text used to train and validate them. As a result, we hypothesize about the following main effect:

H1: Mental health related posts authored by self-identified women relative to men will receive greater empathy across both social media and GenAI.

Our next set of hypotheses are related to the main effect of the technical platform on the level of empathy. Empathy is a distinct human emotion. Therefore, GenAI systems cannot feel empathy, but they can learn how to respond empathically based on the text used to train their LLMs and pre-trained transformers. Here, we distinguish between *generalized GenAI systems* such as Claude, ChatGPT, and Bard that have hundreds of use cases versus *specialized GenAI systems* such as Woebot and Inflection Pi that have fewer use cases. By their nature, generalized GenAI systems are a "jack of all trades but a master of none." In their training data, generalized GenAI systems will inevitably include responses that vary significantly in their conveyed empathy. Therefore, the language used in the responses to mental health related inputs might display an average level of empathy.

Specialized GenAI systems that are designed to be sympathetic, compassionate, and empathetic will have a different pattern of responses relative to their generalized GenAI counterparts because their human labelers are focused on maximizing the level of empathy in their responses. As a result, we expect the responses from specialized GenAI systems such as Inflection Pi, Woebot, and Xiaoice to be more empathetic (and human-like) than generalized GenAI systems such as ChatGPT and Bard. They are specifically focused on constructing systems and responses that are empathetic. Therefore, they are almost certainly removing text that is non-empathetic during their training processes. In this manner, these specialized GenAI systems are more like humans in terms of providing empathy. As a result, we propose the following platform main effects:

H2a: Regardless of self-identified gender, social media will provide more empathy relative to generalized GenAI.

H2b: Regardless of self-identified gender, social media will provide no difference in empathy relative to specialized GenAI.

Our final set of hypotheses is related to the interaction of gender and technical platform. That is, we expect the gender effect will vary across social media, generalized GenAI, and specialized GenAI. Our argument for generalized GenAI is similar to the argument we previously made except now we consider the gender variable. Inevitably, generalized GenAI systems will use text that favors men over women and women over men in their training samples. Doing so will flatten out the gender differences between men and women. Humans are generally more empathic to women than to men due to stereotypes, which will get coded into specialized GenAI systems. Therefore, we expect specialized GenAI systems to have a similar effect as actual humans interacting on social media. Therefore, we hypothesize the following qualifying effects:

H3a: The effect of self-identified gender will be stronger on social media relative to generalized GenAI.

H3b: The effect of self-identified gender will be not different between social media and specialized GenAI.

4. Research design & methods

4.1. Research context

We tested our research model using a sample of posts and comments from Reddit. Reddit is a social media platform containing topic-specific discussion forums (subreddits). We sampled posts and associated comments from the "Mental Health", "Mental Illness", "Self-Harm" and "Suicide" subreddits along with those same terms as keywords. We then inputted those Reddit posts to three GenAI systems – ChatGPT-3.5, Bard, and Inflection Pi. Neither ChatGPT nor Bard were designed to provide empathetic responses or emotional support. However, individuals are using them in that manner. We confirmed this conjecture via a discussion with a sample of college students. Inflection Pi is different. Per its mission, Inflection Pi has been constructed specifically to be empathetic, useful, compassionate, and safe. Inflection Pi incorporates empathetic tuning to give the specialized GenAI system a high emotional quotient, which is directly related to empathy.

ChatGPT-3.5, Bard, and Inflection PI are open (i.e., they are not rule-based) and volatile systems, which means they may provide different or hallucinogenic responses to the same inputs. In our study, we did not observe hallucinogenic responses probably due to the richness of the mental narratives that were inputted. We also entered a sample of posts into the three GenAI systems multiple times to determine how volatile the responses were. The responses were virtually identical. However, as the GenAI systems get updated over time, there is no guarantee that the responses will be the same.

All three GenAI systems in our study may respond to particularly disturbing inputs with responses such as “I am not qualified to help you.” We removed those responses from our study. We saw very few (five) of these reductionist responses in our data. The three GenAI systems offered a diversity of responses with varying levels of empathy to our sample of Reddit posts.

4.2. Data collection methods

We used a combination of the Reddit application programming interface (API) and the Pushshift API to download our sample. After downloading our sample of posts and their respective comments, we manually entered each Reddit post in ChatGPT-3.5, Bard, and Inflection Pi to capture their responses. After gathering all the responses and Reddit comments, we used the linguistic inquiry and word count (LIWC-22) software to determine the linguistic characteristics of the posts and comments. LIWC is increasingly being used in top-level information systems research (Gnewuch, Morana, Hinz, Kellner, & Maedche, 2023). The LIWC-22 dictionary is organized hierarchically into categories and subcategories. We used these LIWC categories primarily as controls and to contextualize the posts, responses, and comments. LIWC provides a percentage score between 0 and 100 for each category based on the overall word count of the inputted text. We converted those percentages to counts because the percentages were biased against longer posts.

4.3. Gender of the poster

We manually coded the gender of each poster based on their explicit self-identification of the poster’s gender. Certain posters explicitly self-disclosed their gender in the content of their mental health post, so there was no ambiguity in their self-identified gender. For instance, “I am a 19F” for a 19-year-old woman or “I am a late 50s man” for a man. We used that self-identification or self-disclosure as our operational definition of their gender, which has nothing to do with the type of language (e.g., masculinity or femininity) used in the post. In fact, using the masculinity versus femininity text analyzer (https://app.readable.com/text/gender/), we found no difference in the masculinity or femininity of posts authored by self-identified women and men.

4.4. Empathy

We measured empathy using machine learning with a separate sample of comments related to mental health on Reddit. We started with a sample of 1,300 comments. After filtering out duplicates and incomplete comments, we were left with 1,275 comments to train and validate our machine learning models. We manually labeled each comment in a binary manner as either providing empathy or not providing empathy. A few examples of empathetic comments are as follows: 1) “This sucks, I am sorry you experienced that...Seeing someone die can be traumatic and leave a lasting scar” 2) “It's concerning that you're experiencing triggers and feelings of disgust toward yourself. It's essential to address these issues and seek support” and 3) “I understand that you're struggling with negative thoughts about your self-worth and that this is interfering with your ability to achieve your goals.”

We used two coders to label our data. The two coders first coded a sample of comments together to ensure consistency. We then coded the same sample of 50 separately to determine inter-rater reliability, which was 88%. The discrepancies were discussed and resolved collectively. We labeled the data together (not separately) during multiple coding sessions. The final labeled dataset that we used to train and validate the machine learning models had 492 comments coded as providing empathy and 783 as providing no empathy.

Using the labeled observations, we pre-processed our data using standard text analytic techniques (i.e., removed punctuation, stop words, and lemmatized the responses). We then constructed a bag of words matrix with our cleaned data. We tried different sizes for the bag of words from one to ten. In our data, a bag of five words performed the best. We randomly split our data 90% for training and 10% for testing. We tried the standard set of machine learning algorithms used in the information systems discipline. The support vector machine (SVM) with a radial-basis function (RBF) kernel had the best balance between bias and variance in our data. Our model had an accuracy score of 80% in the unseen testing data.

We then applied our trained machine learning model to our original sample of Reddit comments and GenAI responses. We used the SVM's estimated probability scores to estimate the range of empathy values for each comment and response. As a result, our final operationalization had empathy values between zero and one with lower values representing less empathy and higher values representing more empathy. In our regression models, we converted these probability scores to the log odds consistent with a sigmoid function. To validate that our SVM machine learning model worked effectively, we manually sampled a subset of observations with high empathy probabilities (>80%) and those with low empathy probabilities (<20%). We used two coders to validate the model's estimations. Both coders agreed with the machine learning model's predictions 88% and 92% of the time.

4.5. Control variables

We used several control variables captured using LIWC-22: 1) word count of the post, 2) word count of the comment, 3) authenticity of the post, 4) emotion (affect) of the post, and 5) personal pronouns used in the post. The longer the post (#1) generally means that the post contained a richer narrative, which might logically result in greater empathy. The longer the response (#2) generally means that there is a greater likelihood that at least part of the reply contained empathetic text. The more authentic (#3), emotional (#4), and personal (#5) the post, then the greater the likelihood that the post would receive an empathetic response, regardless of the self-identified gender of the poster.

4.6. Final sample for analysis

For this paper, we randomly selected 1135 out of the 1896 downloaded posts. Of those 1135, 136 were posted by self-identified men and 195 were posted by self-identified women. The remaining 809 were unknown. The unknown group acts like a control sample because these posts did not explicitly identify the gender of the poster. Given the unequal sample between men, women, and unknowns, we randomly selected 165 of the unknown posts. After filtering posts that were incomplete, had no emotion, and were not personal, we were left with 128 posts posted by self-identified men, 187 posts posted by self-identified women, and 119 posts posted by unknown gender. Table 1 displays the descriptive statistics about our sample. The sample size represents comments from Reddit and the responses from the three GenAI systems. A post may receive more than one comment from Reddit but only one from each of the GenAI systems.

Table 1. Descriptive statistics for all platforms

Variable	N	Mean	SD	Min	Max
Empathy	11924	0.41	0.25	0.003	0.99
Word Count (p)	11924	321.29	286.87	22.00	1909.00
Word Count (c)	11924	75.63	98.20	0.00	1206.00
Authenticity (p)	11924	199.29	169.57	0.43	1517.67
Affect (p)	11924	17.33	12.85	1.00	114.92
Pronoun (p)	11924	56.41	43.14	2.00	402.99

Note. (p) represents post, (c) represents comment.

5. Results

Table 2 displays the regression results for all platforms combined. Based on the main effects model (Model 1), the effect of self-identified gender is significant. Posts from self-identified women receive more empathy than those from self-identified men and unknown genders across all platforms, lending support to hypothesis H1. The effect of platform is also significant. Inflection Pi demonstrates higher empathy than Reddit, while both Bard and ChatGPT demonstrate less empathy than Reddit. Inflection Pi emulated human empathy, but it did it better than humans interacting on Reddit. Therefore, H2a is supported while H2b is not.

Table 2. Regression results for all platforms

	Model 1	Model 2
	Main	**Interaction**
Gender (Men)	-0.238 *	-0.254*
Gender (Unknown)	-0.305 *	-0.293*
Word Count (p)	-0.001*	-0.001*
Word Count (c)	0.007*	0.007*
Authenticity (p)	-0.001*	-0.001*
Affect (p)	0.00002	0.0002
Pronoun (p)	0.007*	0.007*
Intercept	-0.597*	-0.592*
Platform		
Bard	-0.840 *	-1.018*
ChatGPT	-1.605 *	-1.461*
Inflection Pi	0.802 *	0.750*
Gender Platform*		
Men*Bard		0.410*
Men*ChatGPT		-0.001
Men*InflectionPI		0.047
Unknown*Bard		0.202
Unknown*ChatGPT		-0.532*
Unknown*InflectionPI		0.13
Adjusted R2	0.28	0.29

Notes: 1) ** p value < 0.05 2) Reference group for gender is women, and 3) Reference group for platform is Reddit.*

Next, based on the interaction effects model (Model 2), we found a significantly different self-identified gender effect on Reddit relative to Bard and ChatGPT. However, we found no significant difference in the self-identified gender effect on

Reddit and Inflection Pi. Figure 2 displays the interaction for self-identified gender by platform to demonstrate the differential self-identified gender effects across platforms. Thus, both H3a and H3b are supported.

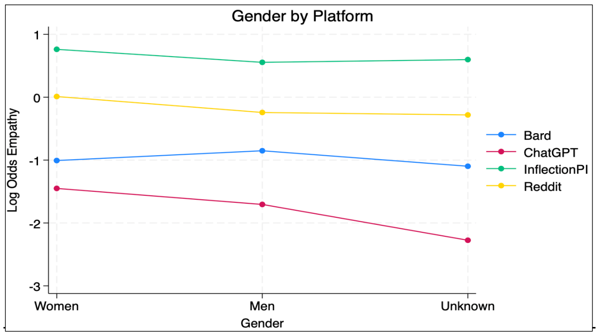

Figure 2. Empathy interaction

To further demonstrate the platform differences, we ran regressions for each platform separately (see Table 3). On Reddit, we find that self-identified women receive more empathy than both self-identified men and unknowns. On Inflection Pi and ChatGPT, self-identified women receive more empathy relative to unknowns but not self-identified men. We see no significant self-identified gender effect on Bard.

Table 3. Regression results by platforms

	Model 3 Bard	Model 4 Chat GPT	Model 5 Inflection Pi	Model 6 Reddit
Gender (Men)	0.373	-0.228	-0.131	-0.232*
Gender (Unknown)	-0.163	-0.865*	-0.359*	-0.270*
Word Count (p)	-0.003	-0.003	-0.005*	-0.001*
Word Count (c)	0.014*	0.003*	0.004*	0.008*
Authenticity (p)	0.0004	-0.002*	0.001	-0.001*
Affect (p)	-0.008	0.013	0.030*	0.0003
Pronoun (p)	0.01	0.019*	0.013*	0.008*
Intercept	-2.194*	-1.149*	1.111*	-0.678*
N	434	434	434	10,622
Adjusted R2	0.49	0.12	0.27	0.22

Notes: 1) * *p value < 0.05 2), Reference group for gender is women, and 3) Reference group for platform is Reddit.*

6. Discussion & conclusion

This paper is among the first attempts to study the effect of self-identified gender in the context of mental health and technologies (social media vs GenAI). Below we discuss a couple of limitations of this paper and suggest the directions for future research. It is important to note that none of us are medical doctors or healthcare professionals. The individual posts used in our study are primarily from mental health, mental illness, suicide, and self-harm subreddits. Many of the posters explicitly state that they are in therapy or are taking medication for a mental condition, but others are just using the terms depression and mentally ill in a colloquial manner. We have no basis or qualifications for clinically diagnosing them. Therefore, caution should be taken when comparing our empathy results with those results reported in mental health GenAI studies published in medical journals by trained medical doctors.

The empathy portrayed by the textual responses from GenAI systems and social media is clearly different from how trained medical professionals provide empathy in face-to-face settings. However, responding digitally with a heart emoji or kind words may offer a struggling individual temporary emotional comfort even if the technologies do so differently from trained mental therapists.

Gender as a socially constructed construct does not have to be binary. For instance, trans-women or trans-men might be even more marginalized than women, which might impact the pattern of empathy. In our data, we only had four entries from self-identified trans-men or trans-women, so we could not do any analyses. A fruitful area of future research could build off our results by investigating non-binary self-identified gender categorizations and empathy differences across the different technologies. Other social media platforms such as Weibo or Twitter may have different norms and moderating rules that might impact the level of empathy. Other GenAI systems such as Claude or ChatGPT-4.0 have different LLMs, which will impact their level of empathy in their responses. Our results, however, provide a first step toward understanding patterns of empathy for mental health posts across GenAI systems and social media.

7. References

Aggarwal, R., Midha, V., & Sullivan, N. (2023). The effect of gender expectations and physical attractiveness on discussion of weakness in online professional recommendations. *Information Systems Research, 35*(1), 87-103.

Ahuja, M., and Thatcher, J. B. 2005. Moving beyond intentions and toward the theory of trying: Effects of work environment and gender on post-adoption information technology use. *MIS Quarterly, 29*(3), 427-459.

Batson, C. D., Sympson, S. C., Hindman, J. L., Decruz, P., Todd, R. M., Weeks, J. L., Jennings, G., & Burns, C. T. (1996). "I've been there, too": Effect on empathy of prior experience with a need. *Personality and Social Psychology Bulletin, 22*(5), 474-482.

Beeney, J. E., Franklin Jr, R. G., Levy, K. N., & Adams Jr, R. B. (2011). I feel your pain: Emotional closeness modulates neural responses to empathically experienced rejection. *Social Neuroscience, 6*(4), 369-376.

Benbya, H., Strich, F., & Tamm, T. (2024). Navigating generative artificial intelligence promises and perils for knowledge and creative work. *Journal of the Association for Information Systems, 25*(1), 37-181.

Belingheri, P., Chiarello, F., Fronzetti Colladon, A., & Rovelli, P. (2021). Twenty years of gender equality research: A scoping review based on a new semantic indicator. *Plos One, 16*(9), e0256474.

Berendt, J., van Leeuwen, E., & Uhrich, S. (2024). Can't live with them, can't live without them: The ambivalent effects of existential outgroup threat on helping behavior. *Personality and Social Psychology Bulletin, 50*(6), 971-984.

Bommersbach, T. J., Rosenheck, R. A., & Rhee, T. G. (2022). National trends of mental health care among US adults who attempted suicide in the past 12 months. *JAMA Psychiatry, 79*(3), 219-231.

Bordalo, P., Coffman, K., Gennaioli, N., & Shleifer, A. (2019). Beliefs about gender. *American Economic Review, 109*(3), 739-773.

Boukricha, H., & Wachsmuth, I. (2011). Empathy-based emotional alignment for a virtual human: A three-step approach. *KI-Künstliche Intelligenz, (25)*3, 195–204.

Braghieri, L., Levy, R., & Makarin, A. (2022). Social media and mental health. *American Economic Review, 112*(11), 3660-3693.

Brandtzaeg, P. B., & Følstad, A. (2018). Chatbots: Changing user needs and motivations. *Interactions, 25*(5), 38-43.

Brynjolfsson, E., Li, D., & Raymond, L. R. (2023). Generative AI at work. *National Bureau of Economic Research*.

Chai, S., Das, S., & Rao, H. R. (2011). Factors affecting bloggers' knowledge sharing: An investigation across gender. *Journal of Management Information Systems, 28*(3), 309-342.

Chau, M., Li, T. M. H., Wong, P. W. C., Xu, J. J., Yip, P. S. F., & Chen, H. (2020). Finding people with emotional distress in online social media: A design combining machine learning and rule-based classification. *MIS Quarterly, 44*(2), 933-955.

Chen, L., Baird, A., & Straub, D. (2019). Fostering participant health knowledge and attitudes: An econometric study of a chronic disease-focused online health community. *Journal of Management Information Systems, 36*(1), 194-229.

Chen, L., Baird, A., & Straub, D. (2020). A linguistic signaling model of social support exchange in online health communities. *Decision Support Systems, 130*, 113233.

Coffman, K. B., Exley, C. L., & Niederle, M. (2021). The role of beliefs in driving gender discrimination. *Management Science, 67*(6), 3551-3569.

Cuff, B. M., Brown, S. J., Taylor, L., & Howat, D. J. (2016). Empathy: A review of the concept. *Emotion Review, 8*(2), 144-153.

Cyranowski, J. M., Frank, E., Young, E., & Shear, M. K. (2000). Adolescent onset of the gender difference in lifetime rates of major depression: a theoretical model. *Archives of General Psychiatry, 57*(1), 21-27.

Decety, J. (2021). Why empathy is not a reliable source of information in moral decision making. *Current Directions in Psychological Science, 30*(5), 425-430.

Deuze, M. (2015). A call for compassion in social media studies. *Social Media + Society, 1*, 1-2.

Dohmke, T., Iansiti, M., & Richards, G. (2023). *Sea change in software development: Economic and productivity analysis of the ai-powered developer lifecycle*. arXiv preprint arXiv:2306.15033.

Doshi, A. R., & Hauser, O. P. (2023). *Generative artificial intelligence enhances individual creativity but reduces the collective diversity of novel content*. arXiv preprint arXiv: 2312.00506

Dwivedi, Y. K., Ismagilova, E., Rana, N. P., & Raman, R. (2023). Social media adoption, usage and impact in business-to-business (B2B) context: A state-of-the-art literature review. *Information Systems Frontiers, 25*(3), 971-993.

Eagly, A. H., & S. J. Karau (2002). Role congruity theory of prejudice toward female leaders. *Psychological Review, 109*, 573–598.

Euchner, J. (2023). Generative AI. *Research-Technology Management, 66*(3), 71-74.

Ely, R. J. (1995). The power in demography: Women's social constructions of gender identity at work. *Academy of Management journal, 38*(3), 589-634.

Faraj, S., & Johnson, S. L. (2011). Network exchange patterns in online communities. *Organization Science, 22*(6), 1464-1480.

Feldberg, A. C. (2022). The task bind: Explaining gender differences in managerial tasks and performance. *Administrative Science Quarterly, 67*(4), 1049-1092.

Fernández, D., Grandón, P., López-Angulo, Y., Vielma-Aguilera, A. V., & Peñate, W. (2022). Systematic review of explanatory models of internalized stigma in people diagnosed with a mental disorder. *International Journal of Mental Health and Addiction, 20*(6), 3315-3338.

Fischbacher, U., Kübler, D., & Stüber, R. (2024). Betting on diversity—occupational segregation and gender stereotypes. *Management Science, 70*(8), 5502-5516.

Fisher, C. L., Goldsmith, D., Harrison, K., Hoffner, C. A., Segrin, C., Wright, K., & Miller, K. (2012). Communication and mental health: A conversation from the CM café. *Communication Monographs, 79*(4), 539-550.

Galyani Moghaddam, G. (2010). Information technology and gender gap: Toward a global view. *The Electronic Library, 28*(5), 722-733.

Gefen, D., & Straub, D.W. (1997), Gender differences in the perception and use of e-mail: An extension to the technology acceptance model. *MIS Quarterly, 21*(4), 389-400.

Gnewuch, U., Morana, S., Hinz, O., Kellner, R., & Maedche, A. (2023). More than a bot? The impact of disclosing human involvement on customer interactions with hybrid service agents. *Information Systems Research*, *Article in Advance*, 1-20.

Guzman, A. L., & Lewis, S. C. (2020). Artificial intelligence and communication: A human–machine communication research agenda. *New Media & Society, 22*(1), 70-86.

Habib, L., and Cornford, T. (2002). Computers in the home: Domestication and gender. *Information Technology & People, 15*(2), 159-174.

Hussain, J., Satti, F. A., Afzal, M., Khan, W. A., Bilal, H. S. M., Ansaar, M. Z., Ahmad, H. F., Hur, T., Bang, J., & Kim, J. I. (2020). Exploring the dominant features of social media for depression detection. *Journal of Information Science, 46*(6), 739-759.

Jami, P. Y., Walker, D. I., & Mansouri, B. (2024). Interaction of empathy and culture: A review. Current Psychology, 43(4), 2965-2980.

Kane, G. C., Alavi, M., Labianca, G. J., & Borgatti, S. P. (2014). What's different about social media networks? A framework and research agenda. *MIS Quarterly, 38*(1), 275-304.

Ki, C. W. C., Cho, E., & Lee, J. E. (2020). Can an intelligent personal assistant (IPA) be your friend? Para-friendship development mechanism between IPAs and their users. *Computers in Human Behavior, 111*, 1-10.

Kim, A., Moravec, P. L., & Dennis, A. R. (2019). Combating fake news on social media with source ratings: The effects of user and expert reputation ratings. *Journal of Management Information Systems, 36*(3), 931-968.

Kitchens, B., Johnson, S. L., & Gray, P. (2020). Understanding echo chambers and filter bubbles: The impact of social media on diversification and partisan shifts in news consumption. *MIS Quarterly, 44*(4), 1619-1649.

Koch, A. J., D'Mello, S. D., & Sackett, P. R. (2015). A meta-analysis of gender stereotypes and bias in experimental simulations of employment decision making. *Journal of Applied Psychology, 100*(1), 128-161.

Kumar, R., Mukherjee, S., Choi, T. M., & Dhamotharan, L. (2022). Mining voices from self-expressed messages on social-media: Diagnostics of mental distress during COVID-19. *Decision Support Systems, 162*, 113792.

Lammers, J., Galinsky, A. D., Dubois, D., & Rucker, D. D. (2015). Power and morality. *Current Opinion in Psychology, 6*, 15-19.

Li, P., Castelo, N., Katona, Z., & Sarvary, M. (2024). Frontiers: Determining the validity of large language models for automated perceptual analysis. *Marketing Science, 43*(2)

Liu, N., Tong, Y., & Chan, H. C. (2020). Dual effects of social support seeking in patient-centric online healthcare communities: A longitudinal study. *Information & Management, 57(*8).

Mattson, T. (2017). Noise or quality? Cross-nested hierarchical effects of culture on online ratings. *Communication of the Association of Information Systems, 40*(1), 25.

Mirzaei, T., & Esmaeilzadeh, P. (2021). Engagement in online health communities: Channel expansion and social exchanges. *Information & Management, 58*(1).

Murphy, B. A., Lilienfeld, S. O., & Algoe, S. B. (2022). Why we should reject the restrictive isomorphic matching definition of empathy. *Emotion Review, 14*(3), 167-181.

Nass, C., & Moon, Y. (2000) Machines and mindlessness: Social responses to computers. *Journal of Social Issues*, 81-103

Noy, S., & Zhang, W. (2023). Experimental evidence on the productivity effects of generative artificial intelligence. *Science, 381*(6654), 187-192.

Olan, F., Jayawickrama, U., Arakpogun, E. O., Suklan, J., & Liu, S. (2024). Fake news on social media: The impact on society. *Information Systems Frontiers, 26*, 443-458.

Oreglia, E., and Srinivasan, J. (2016). ICT, intermediaries, and the transformation of gendered power structures. *MIS Quarterly, 40*(2), 501-510.

Peng, S., Kalliamvakou, E., Cihon, P., & Demirer, M. (2023). *The impact of AI on developer productivity: Evidence from GitHub Copilot*. arXiv preprint arXiv:2302.06590.

Peter, J., & Kühne, R. (2018). The new frontier in communication research: Why we should study social robots. *Media and Communication, 6*(3), 73-76.

Peterson, C. C. (2023). Conversations that warm the heart: Positive and negative links to the mental health challenges of the Covid-19 pandemic. *Journal of Public Health, 45*(2), 499-504.

Petter, S., & Giddens, L. 2023. Is it your fault? Framing social media inclusion and exclusion using just world theory. *Journal of the Association for Information Systems, 24*(5), 1248-1270

Prentice, D. A., & Carranza, E. (2002). What women and men should be, shouldn't be, are allowed to be, and don't have to be: The contents of prescriptive gender stereotypes. *Psychology of Women Quarterly, 26*(4), 269-281.

Preston, S. D., & Hofelich, A. J. (2012). The many faces of empathy: Parsing empathic phenomena through a proximate, dynamic-systems view of representing the other in the self. *Emotion Review, 4*(1), 24-33.

Ren, Y., Harper, F. M., Drenner, S., Terveen, L., Kiesler, S., Riedl, J., & Kraut, R. E. (2012). Building member attachment in online communities: Applying theories of group identity and interpersonal bonds. *MIS Quarterly, 36*(3), 841-864.

Ridgeway, C. L. (2001). Gender, status, and leadership. *Journal of Social Issues, 57*, 637–655.

Rosenfield, S., & Mouzon, D. (2013). Gender and mental health. In C. S. Aneshensel, J. C. Phelan, & A. Bierman (Eds.), *Handbook of the Sociology of Mental Health* (pp. 277-296). Springer.