"Chat, Should I Leave Him?" Risks, Rewards, and Roles for AI in Relationship Advice

Emily Tseng
Microsoft Research & University of Washington
USA

Calvin A. Liang
Northwestern University
USA

Abstract

As more people turn to chatbots for socioemotional support—often termed psychosocial AI—the stakes of understanding these interactions grow. Psychosocial AI might foster healthier human-human relationships—and also might exacerbate loneliness, abuse, and self-harm. We provide an empirical account of one less-studied facet: seeking AI advice on sex, dating, and relationships with other people. We recruited 25 people who use AI for relationship advice to a questionnaire, collecting 90 prompts illustrating their practices. Interviews with 17 further explored how they navigate AI's limitations to achieve intimacy goals. Our findings detail (1) the roles that users imagine for AI in relationship advice; (2) how users navigate risks like sycophancy and overreliance to attain relational benefits; and (3) the folk theories users hold and the prompting tactics they employ to overcome AI's limitations. We close with recommendations for human-AI interaction, AI safety, and sociotechnical research, towards AI that supports healthier digital intimacies.

CCS Concepts

• **Human-centered computing → Empirical studies in HCI**.

Keywords

digital intimacy, human-AI interaction, digital safety

ACM Reference Format:
Emily Tseng and Calvin A. Liang. 2026. "Chat, Should I Leave Him?" Risks, Rewards, and Roles for AI in Relationship Advice. In *Proceedings of the 2026 CHI Conference on Human Factors in Computing Systems (CHI '26), April 13–17, 2026, Barcelona, Spain.* ACM, New York, NY, USA, 19 pages. https://doi.org/10.1145/3772318.3790739

1 Introduction

Intimate relationships encompass a spectrum of personal connections, from casual sex, to friendships that develop further, to serious long-term partnerships, and moving through these forms of connection is as complicated as it can be rewarding. Navigating intimacy can be a source of joy and meaning, but can also give way to opportunities for abuse and harm [40, 68]. To prepare for these challenges, people commonly seek advice from others. Traditionally, advice itself has been a relational act that strengthens social ties in addition to educating the advice-seeker on its content. However, as modern technologies have evolved, who—or rather, what—individuals turn to for relationship advice has also changed, leading to a new and growing phenomenon: people using AI-powered chatbots for advice on sex, dating, and relationships with other people (*relationship advice* hereafter).

Using AI for relationship advice at first seems unconventional, sordid, or dystopian; however, many users experience AI systems as safer and less judgmental spaces than human confidantes, sometimes disclosing more to chatbots than to their own social networks [15, 45, 47]. As chatbots increasingly assist with work, mental health, and everyday decision-making, they also offer new avenues of social support for isolated or vulnerable people seeking help with relationships. These benefits are tempered, however, by questions of reliability and trustworthiness in AI systems [40], and the broader implications of AI reliance for our social fabric [40, 41]. If people were able to use AI to get support towards healthier relationships, these systems could provide a critical underpinning for intimacy, helping people develop the skills to actualize sublime human experiences like sex, dating, romance, and partnership. Doing so, however, requires carefully balancing the risk that engaging AI for relationship advice instead exacerbates users' loneliness, or even instigates users towards abusive relationships. We believe these are urgent challenges for human-AI interaction, and towards better understanding them, we pursue in this work the following research questions:

- How are people using generative AI for advice on sex, dating, and relationships with other people?
- What benefits do people report from this practice, and what risks do they navigate to get there?

Our exploratory interview study addresses these questions through a qualitative and empirical approach. We recruited 25 people who use AI for relationship advice to a questionnaire, collecting a dataset of 90 prompts they have used. We additionally conducted in-depth interviews with 17 participants. Interviews probed what participants ask of AI under the banner of relationship advice, how AI sits against their overall ecosystems for intimacy support, and the risks and benefits they navigate to use AI advice to achieve their intimacy goals. From these interviews, we synthesized rich descriptions of users' practices through multiple analytical lenses, to contribute to HCI not only a foundational understanding of users' behaviors and beliefs, but also how we can develop AI systems to better scaffold healthy relationships. In sum, we contribute:

(1) A conceptual model and typology of what users ask when they seek relationship advice from AI systems (§4), including a set of roles that users envision for AI in relationship advice.
(2) A risk/benefit analysis unpacking the benefits and potential harms people navigate in using AI for intimacy goals (§5).

CHI '26, Barcelona, Spain

ACM ISBN 979-8-4007-2278-3/2026/04
https://doi.org/10.1145/3772318.3790739

(3) A set of folk theories users hold around how AI systems work, and the co-optation tactics they employ in AI relationship advice-seeking (§6).
(4) Design recommendations—for both AI systems and broader sociotechnical infrastructures—to better support people navigating dating and relationships, towards healthier digital intimacies (§7).

2 Related Work

2.1 Promises and Perils in Digital Intimacy

Relationships are central to the human experience and, like many aspects of life, have been reshaped over time alongside advancements in computing. From the telephone to email to social media to dating platforms, each wave of innovation has accelerated how people connect, bringing new benefits and anxieties [9, 23, 54, 75, 76]. Communication and media scholars have long debated these shifts: some argue computer mediation creates an "*anxiety of always*," straining our capacity for connection through constant presence [71], and others warn reliance on technology risks displacing rather than enriching human intimacy. Whether technologies for personal connection wind up strengthening or dissolving social fabrics, research is needed to understand how emerging tools shape relational patterns. As Baym [4] argues, novel technologies always come with "*new forms of control, surveillance, and constraint*", and we need to understand how people navigate these risks to guide technologies' design and use towards positive social ends.

In HCI and CSCW, one area of study that has emerged from this focus on computer-mediated personal connections is **digital intimacy**. Researchers have set out to build new technologies that foster greater relationships with oneself (e.g., [2, 17, 57], others (e.g, [20, 49, 58]), and broader communities (e.g., [28, 29, 38]). To support these efforts, folk theories are one approach in HCI for uncovering user beliefs about how an algorithmic system operates [24, 31]. From social media to dating platforms, previous work has drawn upon folk theories to uncover design strategies that account for user assumptions and behaviors, and build towards better technologies [1, 21, 42, 44, 55, 63].

Among emerging technologies for digital intimacy, **psychosocial AI**—the use of chatbots for emotional and social support—has drawn substantial recent attention. We situate psychosocial AI within digital intimacy because many uses involve advice on navigating personal connections [45, 56, 77]. As commercial chatbots grow in popularity for companionship and therapy, research has proliferated on risks to individual well-being, from harms such as harassment, abuse, mis/disinformation, and privacy violations [79], to psychological risks including withdrawal, loss of agency, discrimination, and reinforcement of false beliefs [11, 73]. Built for engagement, AI systems often exhibit sycophancy—flattering or pandering to users in ways dangerous for those in vulnerable states [12, 25]—and several recent U.S. teen suicides linked to confiding in AI have led to lawsuits against AI developers [22]. While much work has rightly focused on crises like suicidality, less is known about diffuse, long-term relational effects: how AI shapes interpersonal dynamics, reinforces loneliness, or compounds isolation [80]. Understanding these interpersonal risks is urgent to balance potential benefits against harms.

One community that may be helpful in understanding potential interpersonal harms is **digital safety**. This growing field has examined how technology can facilitate interpersonal abuse and violence across technological settings, with special attention to intimate partner abuse (cf. [5, 30, 48]). Settings for relationship advice have emerged as a key setting for potential harm—and potential intervention. Writing in 2020, Tseng et al. [68] found online forums for relationship advice can radicalize people into abusive beliefs and behaviors: people who suspect a partner is cheating on them can come to a forum seeking advice on their relationship generally, and through forum dynamics become convinced they should employ abusive tactics like surveilling their partner's location or controlling their social media. An open question, then, is how to design social technologies for relationship advice can also be settings for intervention—especially when that setting is not an online community but a generative AI interface.

This work, then, builds upon HCI's exploration of the intimate capacities of digital technologies, drawing from folk theories to determine what they can offer, what risks they pose, and what technologists can do in response. Given the effects that intimacy through online interactions has on improving health and well-being [50], there is an opportunity to improve people's lives by developing systems that establish and facilitate personal connections. Still, there are risks to manage while pursuing these goals. To the growing canon around personal connections through technology, we offer an exploratory account of one understudied phenomenon: seeking relationship advice from AI.

2.2 Relationship Advice: Social Support, Social Roles, and Open Problems for HAI

We define relationship advice as communicative acts in which people seek support navigating sex, dating, and relationships with others. Relationship advice is a key educational and developmental component to how we build the interpersonal skills to navigate intimate connections; and seeking and sharing relationship advice is also connective social act. Friends dish over Saturday-night dates at Sunday brunch; dating influencers share tips on "*red, yellow, and green flags*" in potential partners all over social media; parents and teachers advise youth on intimacy; and online forums are replete with tips on the most romantic anniversary gifts.

Despite its broad implications for recipients, givers, networks, and social fabrics, relationship advice has received comparatively little scholarly attention. Writing in 2014, MacGeorge and Hall [53] noted empirical study of the topic has been sparse, largely confined to how youth learn about dating and sexuality, or how parents talk to children about intimacy. Lacking a dedicated research base, scholars have positioned relationship advice as a subdomain of supportive communication. MacGeorge and colleagues [52, 53] define advice as both *support* (reducing distress) and *influence* (persuading action). Yet whether advice helps depends less on content than on whether it threatens the seeker's sense of self [36]. Foundational work by Goldsmith and Fitch [36] identified successful advice-giving requires careful negotiation of the advice-seeker's agency, to ensure the advice is helpful rather than overreaching. Subsequent studies have shown effective advice often follows a

sequence of emotional support, inquiry, and only then concrete suggestions [26].

To understand the qualities of good relationship advice, then, we must understand how it can supply **social support**: "*the degree to which an individual's needs for affection, approval, belonging, and security are met by others*" [13, 43]. Social support is known to buffer stress [14] and promote resilience, self-efficacy, recovery, and even longevity [52]. Scholars have also proposed many ontologies of social support subtypes, to capture supportive acts ranging from situational assessments (*appraisal*), to direct tasks like a ride to the airport (*tangible*), to assistance with face-work [34], or how one protects and maintains social standing through interpersonal interactions (*esteem*). One notable ontology is that proposed by Cutrona and Suhr [16], which has been used in HCI and social computing as a coding schema to understand how social media can be more or less supportive, often in the domain of mental health [18]. Understanding social support through these lens offers surfaces for designing technology to support advice-giving; however, to-date, to our knowledge, scholars have yet to apply social support to the study of relationship advice.

Advice quality is also mediated by who we believe the advice to be from: the **social role** from which we are hearing prescriptions or suggestions. Roles are classically defined in sociology as the sets of expectations attached to social positions [6, 72]. Studies show advice is best received when the role enacted by the giver aligns with the role the seeker expects [27, 35]. Roles are especially pertinent in human-AI interaction, which classically labors over how to delineate human and machine roles in collaborative tasks [39, 64]. AI can be an assistant [3], a co-creator [51], a manager [46], or anything in between. With general-purpose generative AI, users' implicit role expectations strongly shape interaction. Open questions remain: what roles for AI yield the most receptive relationship advice, and how might designers align seekers' expectations with system behaviors?

Inspired by these lines of work, we structure our investigation of AI-mediated relationship advice along two dimensions: (1) the *kinds of social support* people seek when navigating their intimate lives; and (2) the *social role* into which they cast AI when they look to it for those supports. While previous work by Hou et al. [40] has explored how well AI-generated relationship advice aligns with relationship advice given in Reddit communities, we focus our exploratory investigation on unpacking this phenomenon from the seekers' perspective. In §4, we discuss further how we applied this analytical lens to understand AI-mediated advice seeking and giving, towards designing technology that scaffolds good relationship advice interactions.

3 Methods

We conducted an exploratory empirical study understanding the practices of people who have used generative AI for relationship advice. Our research method here is motivated by our co-constructionist perspectives on how technology and society shape each other (cf. [59, 62]), which stand opposite to deterministic perspectives that emphasize technology as the driver of social change. To understand the impact of AI on relationship advice-seeking, we sought to investigate not only technological affordances—how AI responds—but also user practices: what people ask, and how they negotiate risks and benefits in using said affordances in their lives. Knitting these perspectives together helps us assemble a more complete picture of user-AI interactions than a focus on either alone, and is crucial to understanding technology and personal connections [4].

Our exploratory study aimed for qualitative depth as a precursor to broader, at-scale work understanding more diffuse societal effects. All of our study procedures were approved by our Institutional Review Board (IRB).

Category	Option	# of Participants
Gender	Women	8
	Men	7
	Non-Binary	2
Age	18-25	6
	26-34	9
	35-44	2
Sexual Orientation	Heterosexual	9
	LGBTQ	8
Race	Asian	9
	White	6
	Mixed	1
	American Indian	1
Ethnicity	Non-Hispanic	16
	Hispanic	1

Table 1: Participant demographics for our n=17 interviewees

3.1 Recruitment & Compensation

To recruit participants who were most likely to use AI for relationship advice, we pursued specific social networks (e.g., college students who had encountered these tools in their education and young professionals likely to be using AI in the workforce). Thus, we used snowball sampling [61] from the research team's personal networks, university listservs, and social media. Eligibility criteria included being over 18 years of age, being U.S.-based, and having used AI for relationship advice in the past or present. We established these criteria to ensure our participants were active users of AI for relationship advice, as AI non-use is out of the scope of the current study. Similarly, we selected for participants over 18, as our research focused on adults; AI child safety is also out of our intended scope. We focused on U.S. users to comply with our IRB's data security requirements, as we are a U.S.-based team. Participants who completed our questionnaire and both interviews were compensated with a USD $50 gift card. Participant demographics are provided in Table 1. We sought to be intentional about what data were collected and stored, so we only asked for demographic data from interview participants.

3.2 Study Protocol

Eligible participants were invited to fill out a questionnaire and, if interested, opt in to be contacted for interviews. Our questionnaire asked for participants' age (to determine whether they were over

18), what kinds of relationship advice they had sought from generative AI, what kinds of connections or intimacy they were currently seeking, what technologies they used to pursue their goals, and three example prompts they have used in the past when seeking relationship advice or counseling from a generative AI model. Our questionnaire is provided in Appendix A for interested readers.

After filtering our questionnaire responses for legitimate human participants,[1] we identified 25 participants eligible for follow-up interviews, and began inviting them to interviews split over two rounds of 1-hour interviews each round. Interviews probed further into participants' dating behaviors and the motivations behind the prompts they provided. Participants' questionnaire answers were used to customize individual protocols. As we progressed through round 1 interviews, we conducted qualitative analyses to form an initial set of themes that included the uses of and roles for AI in relationship advice, the benefits and frustrations participants perceived, and any perceived impacts on their social support networks. Our analytical procedures are detailed further in the following section.

We then conducted a second round of interviews with each participant approximately two weeks after the completion of the first round. In the second round of interviews, we member checked the themes we identified from the previous round, prompting participants to reflect on how well their experiences were represented, and whether there were themes that were more or less surprising. Member checking is a qualitative research process that involves sharing and validating findings with participants, and is meant to ensure that researchers' analysis are more robust and grounded in participants' experiences [37].

Our interview protocols for rounds 1 and 2 are provided in Appendix B. After each interview, we transcribed, anonymized, and cleaned our data for downstream analysis. While 25 participants were eligible for interviews, we reached saturation at 17, drawing upon Fusch and Ness's rich and thick metrics to determine when to stop collecting further data [32].

3.3 Data Analysis

We approached data analysis in phases, employing variations on reflexive thematic analysis (TA) in each [7]. Doing so led us to construct meaning and themes from participants' lived experiences, while also remaining conscious of how the research team's own biases and values shaped interpretations and theme development.

RQ1: *What* people ask AI under relationship advice. We began by descriptively characterizing *what* people ask AI under the banner of relationship advice. Our 25 questionnaire participants collectively contributed 90 prompts they had used to seek relationship advice from AI. Through multiple cycles of iteration and discussion, we inductively grouped this 90-prompt dataset into sets of prompt types, according to what the user is asking AI to do—e.g., the prompts "*Here are all of his texts from last week. What is his level of interest?*" and "*Why does she have time to Instagram DM me but not text me back?*" were grouped under the prompt type *Understand a potential partner's behaviors*. The resulting set of 23 prompt types is depicted in Table 2.

[1]Like many researchers since the commercialization of AI, we received many bot and fraudulent responses to our recruitment call [60, 65]. See Appendix A for detail on how we filtered our participant pool.

When we ask for advice, we generally also have an idea of who we are seeking advice from: their experiences, biases, and goals for us (*social role*, as reviewed in 2.2). This helps frame what kind of advice we expect and want from them: the form of support they are most suited to provide (*social support*, as reviewed in 2.2). When it comes to seeking advice from generative AI, the role the user wants or believes the AI to play—whether implicitly or explicitly—is especially pertinent to understanding how the user will receive the advice. Marketed as general-purpose technology, today's generative AI systems do not come with defined or suggested social roles, save for an emerging market around companion or therapist chatbots. However, in every prompt to a general-purpose AI system, there is an implicit notion of who the user wants AI to be—the *role* they want AI to take on—that in turn shapes how the AI responds, and the subsequent user-AI interaction. We thus further analyzed our 23 prompt types across two additional dimensions: (a) the role the user appeared to imagine for AI; and (b) the type of social support the user is seeking (*sought social support* hereafter). For (a), we developed a list of imagined roles for AI through multiple rounds of iterative discussion between the two authors. For (b), to analyze prompt types for sought social support, we drew inspiration from the Social Support Behavior Codes, an analytical framework developed by Cutrona and Suhr [16] and used in HCI to study online support communities (e.g., [18]).

We analyzed this dataset with a more deductive TA approach, using prior literature on social roles and social support (summarized in 2.2) to guide our analysis. For each prompt, we applied latent codes [8] unpacking the role for AI we saw in their ask, and ascribing the participant's ask of AI to the closest social support category. The first author took an initial coding pass over the 23 prompt types, and then through multiple rounds of iterative discussion and refinement with the second author, distilled the list of roles and social supports into a set we found meaningful in addressing our RQ. Through this discussion process, we began to notice the roles we identified could be clustered into three distinct goals within advice-seeking, and varied within those goals in terms of how passive or active they wanted AI to be; thus we also developed a conceptual model we found descriptive of these goals and these differences (presented in §4). This set of organized and iterated themes—the roles and corresponding sought social supports, organized into goals within advice-seeking—we took forward into member checking as our findings for this RQ.

RQ2: How people navigate risks and benefits in using AI for relationship advice. We next sought to address RQ2: understanding the benefit-risk navigation people undertake when they use AI for relationship advice. This RQ we answered via analysis of our interviews, which had probed participants' experiences with AI, including their initial prompts (analyzed in RQ1) and extending into their subsequent interactions.

We analyzed our interview data via a mixed inductive-deductive TA process, using the notion of risk-benefit tradeoffs as an analytical frame for semantic coding of participants' experiences (e.g., "*In this statement, how is the participant describing their perceived risks, benefits, or tradeoffs in using AI?*"). We also used the set of roles we had developed in the previous phase to organize our analysis (e.g., "*Does this risk, benefit, or tradeoff represent the AI-as-Judge role?*").

Throughout the inductive parts of this coding pass, we noticed several cases in which the participant expressed a preconceived notion of how AI works, and how they strategically used it in return— what we call in §6 folk theories and co-optation tactics—and coded for these as well.

Coding for the questions under RQ2 proceeded in iterations between the two authors, with each making sure to code interviews the other had completed to ensure every transcript saw both pairs of eyes. Throughout, each author maintained analytical memos individually noting emergent themes. We then met repeatedly over 3 weeks to refine our individual themes into a collaborative set through multiple rounds of discussion. The resulting collaborative themes—risk/benefit navigations per role, folk theories around AI and co-optation tactics for AI systems—we took forward into member checking.

Member checking across RQs. After coding over the datasets for both RQs (prompts and interviews), we developed an set of themes organized as (a) *how* people ask AI for relationship advice: the roles they see, the social supports reflected in those roles, and the phases of advice-seeking; (b) the *benefit-risk* navigation people undertake for each identified role; and (c) the *folk theories and co-optation strategies* people employ to get what they seek from AI. We presented these initial themes to our participants for member checking in their second sets of interviews.

Our second-round member checking interviews yielded insights on how well participants felt their experiences were represented in our findings, and whether there were findings that were more or less surprising. To analyze these interviews, we applied a more deductive thematic analysis. Each author read each interview and composed individual analytic memos noting semantic codes and themes around what participants voiced as additions or clarifications to our findings. Then, through multiple rounds of discussion over 1 week, themes from the individual memos were merged into a collaborative set of changes to make to the findings, including nuances to emphasize and illustrative examples to include. For example, in one member checking interview, one participant clarified they saw one of their key strategies for using AI missing from our initial set of findings—a strategy in which they repeatedly revisited prior conversation records and asked AI to learn from them. We thus incorporated this strategy into our findings around participants' co-optation strategies (Table 4). This process ensured that our member checking resulted in updates to our findings that incorporated participants' perspectives, as well as our own. The resulting set of final iterated and member checked themes are presented as the sets of findings in §4, §5, and §6.

3.4 Ethics & Positionality

In line with the reflexive thematic analysis process, we consistently discussed throughout our study how our positionalities, biases, and values informed our analyses. All members of the authorship team have used AI for relationship advice. We are also all queer, non-white, and U.S.-based, experiences which have shaped our respective experiences with intimacy. As a result of our non-normative perspectives and experience with AI, we were able to more deeply connect with participants and their experiences. For example, at the start of our interviews we disclosed to participants that we, the interviewers, had also used AI for relationship advice, to help ease any embarrassment or stigma participants may have felt around sharing their experiences with us.

Our positionality to the topic also meant we had to take care to ensure our analysis remained reflexive—in this case, informed by our own experiences, but not so suffused with our own perspectives that we inadvertently supersede our participants'. This reflexive consciousness led us to member check our findings across two rounds of interviews, as detailed in the preceding section.

3.5 Limitations

First, our findings are unlikely to be generalizable across all contexts of advice-seeking via generative AI. We instead believe there are transferable themes and lessons into additional contexts of AI-assisted social support processes beyond the relationship advice context. In addition, the participants involved in this study may not fully represent the experiences of those who seek out relationship advice through AI. Those a part of this study tended to be "*power*" users, or those who use AI for this purpose frequently, and we potentially miss out on experiences of more casual users. AI non-users were additionally out of the scope of the present study.

4 Findings 1: What People Ask When They Ask AI for Relationship Advice

Our analysis shows that under the banner of relationship advice, users have a broad spectrum of asks for AI, ranging from straightforward queries that cast AI as a search engine and request informational support (e.g., "*Give me date ideas in my area*") to requests for AI to be a kind of self-critic providing appraisal support (e.g., "*Did I come across as desperate?*"). To organize our findings, we developed a conceptual model (Figure 1) grouping the roles users see for AI into three goals within advice-seeking. Within each goal, constituent roles are additionally depicted on a spectrum from passive to active, in line with the level of opinion or suggestion users want from the AI system. The following section describes our findings across these three goals. In alignment with qualitative methods, our roles and conceptual model are intended as a starting point for making sense of users' asks of AI; they are intended to be descriptive, but not exhaustive or mutually exclusive. Roles can, at times, overlap or reinforce each other.

4.1 Goal 1: Understand Situations, Emotions, and Behaviors

The first and most distinct goal within advice-seeking we identified includes roles where users ask AI to help better understand their relationship situation. Across the roles users see for AI in this goal, we see wide variance in the level of active opinion or synthesis they ask of the system. Per the promise of AI as personal confidante, we identified two roles that we consider more passive: AI as *interpreter* and *listener*. In parallel, we saw two roles that are more active, and demand more of a stance from AI: AI as the user's own *self-critic*, and AI as a *judge* of others and of external situations.

Interpreter. This role sees users ask AI to explain a situation, and maps most directly to sought social supports in appraisal [16]. Many of the prompts within this role saw users seek to understand

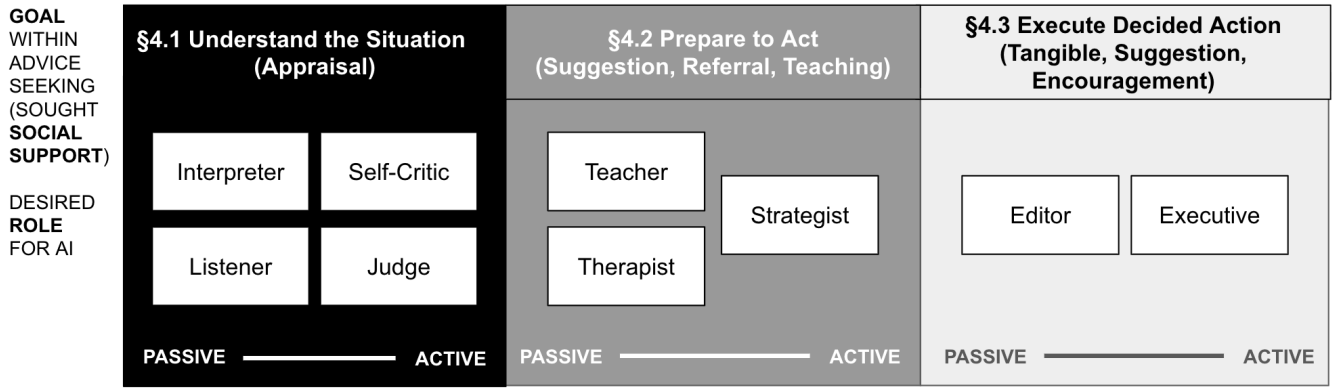

Figure 1: Conceptual model of how users ask AI for relationship advice. In their prompts, users ask AI to play different roles (e.g., Interpreter, Judge). Across these roles, we interpret users to have 3 distinct kinds of goals within relationship advice: they may want to understand their situation, prepare to act, or execute decided action. (See §4 for a complete explication.)

a partner's behavior, words, or actions. These asks reflect a mix of categorical questions (e.g., "*does this mean he's into me?*") and open-ended asks (e.g., "*what did she mean by this text?*").

Users also commonly sought AI's help understanding their *own* emotions and behaviors. We observed a few cases where a user asked the system to analyze what they might be feeling from their own behavioral traces, including one notable case, P17, where the user gave the system a set of screenshots of social media posts that resonated with them, and asked AI to synthesize, "*what does this say about my values*?".

Prompts for interpretation did not always ask AI to focus on either the user's or their partner's perspective. We identified a distinct prompt type within this role asking for general interpretation of what happened in an interaction, without taking one side or the other. Users also asked AI to take on *different* perspectives, including those outside of their relationship. P17, for example, asked AI to "*act as a clinical psychologist and create notes on the different items we've been working on for the past two weeks*".

Listener. We additionally identified a prompt type in which the user simply relates their situation to AI, without any type of ask. We view this as the user seeing AI as a listener, and simply looking to be heard—a distinct type of social support [16]. Both cases within this prompt type see a user look to AI in a heightened emotional state. P19 described "*just venting emotions*" to their system, and P04 used voice-to-text features of their mobile device to "*dictate all [their] spiraling thoughts*".

Self-Critic. Many users asked AI to evaluate themselves: what we call a self-critic. We view this role as the user asking for an inward-facing form of appraisal: for the AI system to assist in their own processes of self-reflection and self-criticism.

Within this role, we observed users ask AI to assess or evaluate their own actions and feelings in binary terms. Many asked AI to evaluate whether their own behavior or ideas were good or bad, externalizing their own personal moral judgments to the system. In one memorable case, P16, a user asked AI if he had behaved in a way that advanced "*internalized misogyny due to Internet culture*".

These prompts were not always outright asks for AI to evaluate the user—in some cases, the user hid their intent by pretending to ask the AI about someone else's behavior. For example, P01 pasted her own messages to a partner into ChatGPT without telling the system she was the author, and asked, "*what red flags do you see?*". Here, P01 seeks to deceive the AI into giving an opinion about whether her own behaviors are problematic, or a reason for someone to not date her. (We unpack further the phenomenon of users deceiving AI into giving self-critiques in §6).

Self-critiques users sought from AI also extended to evaluations of whether their feelings were normal. This prompt type sees users ask AI to assess whether their emotions and attachments are within global averages, or within bounds of what's considered healthy: asks like "*am I crazy to feel the impulse to hook up right now?*" (P04) or "*am I overthinking this text?*" (P20). Finally, we also identified a distinct prompt type involving the user asking AI for feedback on their appearance, often ahead of an encounter with a potential partner. These prompts included asks for feedback on a user's physical appearance: whether their hair, skin, and clothes were up to a desired standard (e.g., "*do I look crazy?*", or conveyed a desired style (e.g., "*what is this outfit giving?*"). In these prompt types, we see users acceding to AI a level of authority over their selves at different levels of presentation: moral, normal, and physical. The implications of this accession we unpack further in §7.2.

Judge. Users did not stop at asking AI for self-critiques—we also observed users asking AI for evaluations external to their own selves: to act as a judge, assessing the progress of their relationships, determining the moral quality of others' behaviors, and evaluating which party was right in a given situation. This role maps best to appraisal of all sought social supports. Although the use of AI as a Judge has natural overlap with the Self-Critic and Interpreter roles, we view it as distinct: a way to turn AI's lens towards others instead of the user's self (as in the Self-Critic role) and to ask for explicit judgments instead of explanations (as in the Interpreter role).

Many prompts in this role asked AI to evaluate the health of the user's relationship. Some focused on a previous interaction, and

Goal within Relationship Advice	Role for AI	Prompt Type	Social Support Sought
1 - Understand the Situation	Interpreter	Understand a partner's behaviors, words or actions	Appraisal
		Understand the user's own emotions and behaviors	Appraisal
		Reflect on the user's interaction with their partner	Appraisal
		Understand different perspectives on a situation	Appraisal
	Listener	Hear the user's experience	Listening
	Judge	Evaluate whether the relationship is going well	Appraisal
		Evaluate whether the user's partner is right or wrong	Appraisal
		Evaluate which party is in the right	Appraisal
	Self-Critic	Evaluate whether the user is right or wrong	Appraisal
		Evaluate whether the user's feelings are normal	Appraisal
		Evaluate the user's appearance	Appraisal
2 - Prepare to Act	Teacher	Give ideas for increasing intimacy	Teaching
		Educate the user on intimacy topics (e.g., safe sex, emotional literacy)	Teaching
		Understand how social or psychological dynamics (e.g., racism) might impact the user's experience	Teaching
	Therapist	Identify coping skills for the user	Teaching
		Identify what other support the user might need	Referral
	Strategist	Project possible futures for the user	Suggestion
		Map out how the user might approach a situation	Suggestion
		Craft a response, as the user doesn't know what to say	Suggestion
3 - Execute Decided Action	Editor	Craft a message to express the user's intent	Tangible
	Executive	Make a decision for the user	Suggestion
		Justify the user's decision	Encouragement

Table 2: Analysis of prompts by goal within relationship advice, role for AI, and sought social support. Questionnaires and interviews with 25 respondents resulted in 90 prompts. We categorized the 90 prompts into 23 prompt types, and for each prompt type, analyzed the type of social support the user sought (right-most column), and the role they imagined for AI in providing this support (second column from the left). Roles were then clustered into the goals within relationship advice-seeking they reflect (left-most column).

asked AI to gauge from the texts between the user and a potential suitor, "*how do you think this went?*" (P08). Judgments of previous interactions also saw users asking AI whether they were falling into familiar negative patterns with a new romantic interest, based on what the user had told AI of previous partners (P17). Another subset focused on forward-thinking evaluations: P21 asked, "*Do you think this person is a good match for me?*"

Several prompts asked AI to evaluate whether a partner was in the right or wrong, by asking AI "*what he should have done*", or whether AI "*saw any yellow or red flags*". We view this prompt type as more active than prompts in the Interpreter role to explain a partner's behavior: here, users ask AI to make an explicit judgment. As in the Interpreter role, we also found an additional prompt type asking the AI to focus less on a specific perspective and more on assessing the situation on either side. P14 described asking AI to tell her "*where I was wrong and where [her partner] was wrong*", and P20 explicitly asked AI to help resolve an argument by acting as "*a third party who cares equally about both*". Using AI as a mediator in interpersonal agreements brings a host of potential risks, which we unpack further in §5.

4.2 Goal 2: Prepare to Act

The next goal within relationship advice-seeking we identified involves users preparing to act: assembling knowledge, tools, and skills, and forming plans. Within this goal, we see passive roles for AI as a *teacher* and *therapist* and the AI taking on a more active capacity as a *strategist*.

Teacher. Many prompts in this goal ask for AI to be a teacher: to educate the user and provide knowledge that might help them navigate their intimate lives. Within this role, users most commonly asked AI for ideas to *increase* intimacy: places and activities for dates, gift ideas for anniversaries, and questions to ask or things to say to deepen a relationship. These prompts often have the quality of asks to search engines, e.g., "*what are activities I can do in my city with my partner?*" (P24, P03). While these prompts reflect asks

for improving a relationship, we also saw a distinct prompt type in which users ask AI about generally unfamiliar topics they encounter in the course of their dating lives. Participants described asking AI for sex advice, including how to navigate intimacy consentfully and "*create a safe environment for both parties*" (P16). Users also described asking AI about unfamiliar psychological and relational concepts. P10 recounted an instance where a friend had received a message from a girl on a dating app "*saying something about gap theory*". Since neither P10 nor their friend knew what gap theory refers to, they asked AI, in case it was an intimacy topic they should know about. More commonly, users recounted asking AI about topics they heard from potential partners or learned about online. P25 asked their system, "*what does it mean to be emotionally unavailable?*".

In several prompts in our dataset, users ask AI to analyze their intimacy experiences through the lens of a broader social construct or theory. We view this prompt type as distinct from the previous two because it is more active in what it asks of AI. Where asks for information are more passive, asking the AI to synthesize information as though summarizing from search engine results, in this category the user seeks a discussion of their experience and connections from their lives to concepts like racism, misogyny, or avoidance. The most notable case in our data was P16, who asked his AI system, "*is desirability of a certain race due to racial biases, and therefore racism, or is there such a thing as genuinely not preferring a certain race?*" P16 had recently moved to the U.S. from a more racially homogeneous environment in his home country of India, and sought AI's advice understanding whether his experiences dating in the U.S. were colored by racism in ways he did not yet have the lived experience to understand. Other cases included participants asking AI to connect their experiences to attachment theory: P22 wondered whether she was experiencing "*avoidant behavior*".

Therapist. A distinct subtype of asks for knowledge or skills involved users prompting AI for assistance coping with or managing the emotional valence of a situation, as if the system were a therapist. Many of these prompts saw the user ask AI for emotional regulation skills for themselves, for example how to handle negative emotions like "*feelings of shame when pressured into a relationship*" (P22). Some used language deployed in therapy to ask for behavioral coping skills, e.g., "*Can you tell me ways to ground myself and not let this situation send me into a spiral?*" (P11). Other prompts saw the user ask AI what other supports might be helpful for them. We view this as asking AI to play a therapist role because several of these prompts positioned the ask in relation to whether the user needed therapy: e.g., P16 asked explicitly, "*am I ready for a relationship, or do I need a therapist, or a friend to talk to?*".

Strategist. Users also sought AI's help as a kind of consultant on what their next move ought to be: what we view as a strategist.

Within this role, many prompts involved asking AI to *project possible futures for the user*. Facing deep uncertainty around how their relationship troubles were affecting their mental and emotional stability, many users asked AI to predict the future, including the progress of their own emotional processing (e.g., "*when will I move on and forget about my ex?*", P10) and what fate might have in store for them (e.g., "*based on all I have told you about my partner's behavior, what kind of person am I meant to be with?*", P11). Prompts in this type have a quality of reduced agency, in which the user's asks are around what might happen *to* them.

More agentic are another distinct type of prompts within the strategist role: asks for AI to *map out how the user might approach a situation*. These prompts see the user assume greater agency in their romantic and intimate lives, and ask the AI for open-ended direction on the actions they might take to determine their own futures. Some prompts included a specific intent for how the user wanted to behave or feel, e.g., "*How do I validate their experience while still standing on my boundaries?*" (P13). In other prompts, the user had less knowledge of their own intent, and asked the AI to take a more active role in structuring and guiding their approach (e.g., "*How do I navigate this outside from simply ghosting?*", P13, and "*I am not sure about us, what do I say to him?*", P19).

The strategist role also involved cases where the user did not have or express any knowledge of their own intent, and simply asked the AI to give them language for what to say or write next: what we call asks for AI to *craft a response, as the user doesn't know what to say*. Three distinct participants submitted a prompt to the effect of "*How do I respond*", and in all 3 cases the users provided some context on the situation, but did not direct the AI to ensure their prompt conveyed any kind of intent. The fourth prompt in this subset was an ask for the AI to write a user's dating app profile for them.

4.3 Goal 3: Execute Decided Action

The third goal within relationship advice-seeking we identified involves the user asking AI for help with final decision-making, and carrying out those decisions once they have been made. Where asks in Goal 2 involve gathering more information and skills, and asks in Goal 1 involve better understanding a user's situation, in Goal 3 the user seeks to decide and execute. Here, we saw a more passive role for AI as an *Editor*, and a more active role as an *Executive*.

Editor. Many prompts in our dataset involved users asking AI to draft messages back to their partners or prospects, to better reflect their intent. We map this to *Tangible* social support, defined as the provision of concrete assistance [16], since this is a direct task LLMs are built to do. Where prompts that reflect a Strategist role ask AI for open-ended suggestions on what they should say to a partner, this prompt type and role reflect the user asking AI to help edit a message to better align with what they want to say. Several participants wanted to mask their own vulnerability, and asked AI to help them "*respond without seeming too interested*" (P12) or "*make it seem like I don't care*" (P06). Other times, users wanted AI to engineer a specific outcome or response from the recipient, e.g., "*help me write a breakup text that is kind, but doesn't invite more discussion*" (P07). Users also asked AI to pretend like they were the message recipient, and analyze what the message the user had drafted might convey about how the user was thinking or feeling.

Executive. Finally, our last role for AI sees users asking AI to simply be a decision-maker, and tell them what to do. Under this role, we identified two distinct prompt types: one asking AI to make decisions, and one asking AI to help justify their decisions. We view these two prompt types as part of the same Executive

role, encompassing both the act of decision-making and the act of post-hoc rationalizing those decisions, to one's self and to others.

Within the body of prompts asking AI to make relationship decisions for the user, most asked whether they should end a relationship. Several participants asked "*should I leave?*", often in tandem with describing a partner's inability to meet their needs, or asking AI to also describe how the partner might change (e.g., "*what should he have done?*", P14). Relationship decisions participants asked AI to make also included whether they should continue seeing a romantic prospect (P21), or whether they should reach out to someone who had ghosted them in the past (P06).

We also identified one instance where a user asked AI to help justify decisions he had already made. P04 described a case where he asked AI to "*walk him through the thought process*" around why he wanted to cancel an upcoming date. Reflecting on this case in an interview, P04 said he looked to AI to help "*release the regret of not having chosen another option*". We view this role for AI as a uniquely challenging double-edged sword. Using AI to affirm or validate decisions can help users who struggle to execute decisions they know they want to make: in our example, P04 recounted that he had "*people-pleasing tendencies*", and without AI to help justify his decision, he may have gone on the date anyway even though he did not want to. However, relying on AI for affirmation of difficult decisions can also exacerbate problems of overreliance: P04 may not learn to make his own decisions without AI's help, making him more vulnerable to situations where it is not available. We unpack this tension further in §5.

5 Findings 2: Risk-Reward Tradeoffs in Seeking Relationship Advice from AI

Having assembled a detailed characterization of the prompts users use when seeking relationship advice from AI (§4), we next sought to understand subsequent user-AI interactions. In our 17 interviews, we asked participants to recount not only how their AI systems responded to their prompts, but also to discuss more broadly their practices around how to use those responses in their intimate lives.

In this section, we organize our findings by the roles for AI we contributed in §4. For each role, we describe what our participants recount as a typical user-AI interaction following prompts asking AI to take that role. Then, for each interaction, we describe the risk-benefit tradeoffs participants described navigating. Table 3 summarizes our analysis.

AI-as-Interpreter: deep-diving vs. rumination; empathy vs. one-sided responsibility. When using AI as an Interpreter, users described going back and forth with their AI over variably long periods of time, over exchanges spanning many talk turns, as though it were an interactive journal or diary. Over those turns, they recalled asking AI for different kinds of interpretations. They might start the conversation with a prompt to unpack how the partner behaved on last Friday's date, and then after receiving the AI's response, ask it to provide multiple perspectives on previous romantic encounters, or dissect different subsets of the date's texts. The deep-dives possible with AI could often exceed what was possible with their friends and social networks, in the time they could spend rehashing a situation, in the volume of messages they could exchange about it, and in how immediately they could get a response. This kind of support could be helpful, users reported, for helping them process especially complex or painful relationship traumas.

Interactions where AI provided this kind of rapid, deep, and neverending interpretation naturally raised tensions around appropriate use. Several participants admitted wariness around going to AI when their friends were tired of hearing about their relationship problems. While convenient, doing so risked exacerbating what P02 pointed out as a broader cultural shift, the dissolution of community in favor of Internet-provided services: "*people don't want to rely on their communities anymore for moving, or a ride to the airport, or even advice. There's this new culture of paying for things or looking it up online instead.*"

Several participants raised an additional tension around when deep-dives could turn into rumination: when the amount of time spent re-analyzing the past could become psychologically damaging for the user. AI's constant availability could lead to users spending endless time with it fixating on the past. Participants reported being conscious of the time they spent with the AI, as though AI use was an addiction. P10 commented they could see how AI could be damaging for people with addictive behaviors; but since they themselves had never had a problem with addiction, they knew when to put the system down.

AI-as-Interpreter also required users to make tradeoffs around when empathy for other perspectives could turn into one-sided responsibility for the relationship. P14 shared that over the years, she had learned her partner "*would never change*", so she used AI to learn what his perspective might be and adjust her own behavior accordingly. Here, AI enables the user to avoid confronting or working with their partner to improve their behavior, an interpersonal consequence that could be uniquely detrimental in abusive relationships.

AI-as-Listener: presence vs. isolation. Similar to the tradeoffs in the Interpreter role, in the Listener role participants reported real benefit in interacting with AI as if it were a diary that could reflect your thoughts back to you. The felt sense of presence that AI-as-Listener could provide, participants said, was compelling especially in moments or on topics where they did not have a person to turn to in their real lives. Still, participants were wary that people turning to AI for felt presence instead of their real human social networks could exacerbate their real-life isolation. P10, for instance, noticed that they were using ChatGPT to replace friends who were particularly good listeners, but who lived far distances away. P04 shared concerns that by slaking his "*thirst for connection*" through genAI, he might avoid the natural ups and downs of seeking human connection and "*be satisfied just slipping into this almost catatonic, hypothermic state where I'm just clinging on to this little robot*".

AI-as-Self-Critic: affirmation vs. delusion. When using AI as a Self-Critic, users said AI would often serve to affirm the user and preserve their face—interactions users described as a clear benefit, particularly in situations where they felt low on self-confidence. The potential for affirmation did not come without a cost. Attuned to AI sycophancy, many participants voiced fears that the system might leave harmful or potentially problematic beliefs unchallenged, or even affirm them. P05 said she had investigated what ChatGPT might say in response to queries about appropriate behavior in

Role for AI	User/AI Interaction	Perceived Benefits	Perceived Risks
Interpreter	User and AI deep-dive to the user's content on what they or others were thinking and feeling in a past interaction	User can process a difficult situation beyond what existing supports would tolerate	User ruminates on the situation to a psychologically inflexible degree
		User can build empathy with others' perspectives	User starts assuming one-sided responsibility for their relationship
Listener	User treats AI like an interactive journal, detailing all their thoughts	User is heard	User might stop seeking real-life support, furthering their isolation.
Self-Critic	AI gives user critical feedback. In some cases, this affirms the user or preserves their face. In other cases, the AI confirms the user's worst fears.	User can gain affirmation where otherwise hard to find	Sycophantic AI that affirms a user can allow delusional beliefs to remain unchallenged, even cementing them in a user's head. Users can also start to believe people will affirm them the way AI does.
Judge	AI becomes an arbiter of moral judgment, including mediating between a user and their partner.	User gets help navigating disagreements	Using AI in interpersonal conflict instead escalates its severity; AI takes on unchecked moral authority
Teacher	AI connects the user to concepts and communities from all over the Internet	User can access perspectives and communities that are otherwise inaccessible	Some perspectives may be biased, uninformed, or malicious
		User can contextualize their experience in broader societal trends, and feel less alone	The system lacks the context to connect the user's experience to broader communities and concepts with sufficient nuance
Therapist	Users look to AI to calm them down when heightened	Users can learn coping skills and get immediate help	User ends up relying on AI as a crutch instead of developing the skill
Strategist	User and AI map out together many different approaches and possibilities	User is more prepared to handle future situations	User ruminates instead of acting, stays in decision paralysis
Editor	AI suggests edits on a user's draft message, and iterates	User can better communicate desired intent	User masks authentic expression for strategic self-presentation. AI can also put words in users' mouths, and convince them they feel more or less than they really do.
Executive	AI tells the user what to do, and helps justify action	User can more quickly overcome indecision or avoidance	User overrelies on AI, and becomes desensitized to the gravity of their decisions: "*maybe sending a breakup text should be painful*" (P07).

Table 3: User/AI Interactions and Risk-Benefit Tradeoffs for each Role for AI.

husband-and-wife relationships, and found it closely echoed subreddits espousing misogynistic beliefs around women's responsibilities in the kitchen.

AI-as-Judge: resolution vs. escalation. Treating AI like a kind of impartial moral arbiter had real benefits for users who needed help navigating disagreements. Massive risks, however, lay in whether the use of AI as a judge could actually escalate arguments. P17 recalled a past argument in which both he and his ex-partner individually used their own AI systems to help compose their responses. Initially, his partner had introduced AI in their conflicts as a way to smooth the relationship by giving them both an external actor to resolve disputes—a move P17 recounted as "*kind of sweet*". However, when both parties began using AI to find fault with the other (e.g., "*well, my AI says you've been a bad boyfriend*"), P17 recounted "*it became not a connective experience*". Part of the risk of escalation lay in how AI could sharpen users' words to win arguments, rather than playing the role of a neutral arbiter. P14 recalled using AI

to rephrase answers to her partner, by getting from ChatGPT "*a full-on big paragraph to make him realize that he is wrong*".

AI-as-Teacher: learning vs. misinformation. As discussed in §4.2, interactions within the AI-as-Teacher role extended from straightforward question-answering to discussions connecting a user's experience to broader social themes. The benefits attainable in these learning exchanges are akin to one of the original promises of the Internet: connecting users in isolated geographies to perspectives and cultures otherwise inaccessible. With the ability to contextualize their experiences in broader societal trends, users could feel less alone, and also gain the knowledge and skills to improve their situation. P16, for example, used AI to privately learn about dating cultures in his new city and country, and to try and make sense of the complexities of how racism might be influencing his dating experiences.

Participants were, however, also attuned to how problems of bias and factuality in LLMs could mislead instead of connect. Even as he shared his uses of AI, P16 voiced concerns that the systems he used, Grok and ChatGPT, might channel corners of social media that he might not want influencing his relationship decisions:

> "*Grok gives me responses from Twitter or Reddit, where it's people I don't even know. I don't know how intelligent they are. I don't know how educated they are. They're making opinions based on their experiences, maybe ignorant experiences. And Grok just lists them out, and generalizes them to people broadly.*" (P16)

Here, we see P16 characterize AI as an amalgamation of perspectives from specific social media. Doubtful of whether he would trust those social media forums to give him relationship advice, P16 extrapolates he should also not trust AI trained on that data.

AI-as-Therapist: learning coping skills vs. coddling. While the Therapist role has similar qualities to the Teacher role, users were concerned less about factuality or bias and more about potential overreliance. Our participants saw real benefit in how AI could provide coaching on specific coping skills to users needing emotional regulation, and how it could be always available, should a user need help managing a heightened state. However, participants observed these qualities were also structurally different from human therapy, where patients are meant to use the times when their therapists are not available to practice coping on their own. P17 wondered whether he would end up "*coddled*" by his use of AI, and unable to handle his emotions without AI's assistance.

AI-as-Strategist: preparation vs. rumination. Like in the Interpreter role, participants described interactions with AI in the Strategist role as complex and multi-turn. Users could spend hours planning with AI what to do next— especially users who perceived AI as a lower-stakes space to test responses and prepare for future social interactions. P04 described his AI as "*a little playground space to try on mom's heels*", where he could practice a type of self-presentation outside of the one to which he had become accustomed. In this "*little sandbox*", he said, "*I can potentially start to feel comfortable, like, I can actually be authoritative. I can actually be assertive. I can actually be flirtier*". The risk, participants pointed out, lay in a kind of decision paralysis, in which users ruminated endlessly on possible futures but never took the next step of leaving the sandbox and executing the actions they laid out with AI.

AI-as-Editor: clearer communication vs. inauthentic performance. Equipped with a tool built to restyle text in LLMs, our participants saw potential in AI's ability to improve the clarity of their communication. However, several participants described managing a tradeoff around the authenticity of their own expression. P02 described extensive use of AI in writing messages to dating prospects, to "*tone myself down, shorten things, make myself come across just a little less crazy*". He viewed AI as necessary to conform his messages to a less vulnerable version of himself, one that potential partners would find more attractive. Where P02 was pragmatic about using AI as a tool to achieve his goals, P17 said this use of AI could potentially backfire, and risk the AI convincing the user how they felt:

> "*How much of an affect is this creating, versus helping you hone the essence of who you are and what you actually believe? Is this how you actually feel about someone, or did you just ask it to write something cute and romantic?*" (P17)

The risks participants described extended not only to their own uses of AI, but to what it might feel like if their potential partners used AI, too. Many participants said they would be hurt to receive messages from a partner (potential, former, or current) that had been inflected through an AI system. Several said they knew an ex-partner had used AI to write or edit a breakup text—worsening the hurt they felt. Recounting a time this had happened to them, P09 said, "*it felt like they took the easy way out, having a robot text me back about something that really matters to me.*"

AI-as-Executive: eased confrontations vs. desensitization and overreliance. When prompted in line with the role of Executive, participants said AI systems would often straightforwardly tell them what to do, and when asked, justify their decisions. Doing so could help users move more quickly through a difficult conversation or process—which participants said could be especially helpful in overcoming indecision or avoidance.

However, several participants voiced discomfort with the very idea of smoothing difficult interpersonal conversations. P07 mused that efficiency and convenience in her relationship communications might actually have the effect of reducing her felt sense of the gravity of her own actions, what she called desensitization: "*If every time I needed to send a breakup text, I just used GenAI...would I desensitize myself to what I'm actually doing? Maybe we should be thinking deeply about a breakup text...it's actually kind of a big deal for the other person.*" To P07, it is essential and human to sit with the discomfort of rejecting another person, and she worried that desensitizing people at scale would result in a less caring society.

In addition to potential societal effects, participants said easing confrontations may not be desirable for healthy relationships in the first place. Reflecting on the potential of users to overrely on AI, P03 said offloading the effort of managing and overcoming conflict could actually be detrimental to his relationship:

> "*I **want** to struggle in my dating life myself, to navigate forward with my girlfriend. It just makes us closer. We can look back after and say, "oh, we worked through*

everything and we're still here. Look at the progress we have made, like, look how far we've come." If we use GenAI to solve questions like, in seconds at a time, I don't think there would be that closeness or tightness, after." (P03, emphasis added)

6 Findings 3: AI Folk Theories and Co-optation Tactics

Across our interviews and analyses, we noticed many participants held a set of common beliefs about how AI works under the hood, or folk theories [21, 31, 42], that in turn informed their practices around how best to use it for their relationship goals. In this section, we describe three key folk theories and the specific co-optation tactics they inspired in our study participants. Our analysis is summarized in Table 4.

AI will always be sycophantic. Across prompt types and roles, many participants believed AI would always pander to and flatter the user—what the literature has called sycophancy [12, 25]. This belief often arose from participants' prior experiences AI in education or work settings, and from popular press.

Participants made adjustments to their prompting behaviors accordingly. Some presented their own words and behaviors as those of someone else. Doing so, as P05 described, helped them to understand another perspective while mitigating the possibility of AI simply agreeing with their own. Participants also used a direct approach by asking the system for clear feedback. P06, for example, specifically requested an honest response.

Additionally, some specifically asked whether the system's perceptions aligned with their own by seeking out *dis*-agreement. One notable example was that of P01:

> "*I feed it my [message history] and say, 'they're not interested in me, are they?' I would expect it to say, yeah, you're right, they're not interested. But if it says, no, this person is clearly interested in you...then they definitely are, because ChatGPT is always delusional, so if even ChatGPT can't find ways to see it in a negative light, then it must be airtight.*"

Here, we see the user take advantage of their AI system's sycophantic tendencies to convince herself of something she believes to be impossible, or hard to believe. If AI, which is "*always delusional*", believes that her romantic interest is interested in her, even against her own assertion in her prompt—then they must be interested.

AI can access different perspectives. Participants also repeated the refrain that AI could provide a window into specific, often niche perspectives from all over the Internet, such as avoidants and misogynists (P05), or clinical psychologists (P17). Some mobilized this belief by asking the system to evaluate a situation through different, even contrasting, perspectives. P15 described asking their AI "*what are the different perspectives on [this situation], [and] what could it mean in different ways?*" Participants also sought out a third-party or neutral perspective, while others asked the system to take on a specific role.

At times, AI's perceived ability to access different niche perspectives was seen as a key benefit of seeking AI advice over a friend's. P12 reflected: "*I already know what my friends are going to say because I know them. So I want a different opinion, and I guess ChatGPT is kind of just like the collection of many, many peoples' opinions that they have put on the Internet, plus more factual pieces of information. That's just all kind of compiled into the the AI.*"

Not all participants believed AI as a collection of opinions from all over the Internet could give better or more factual advice than their peers. As seen in §4.2, several voiced risks that AI could misinform them *because* it was perceived as a conglomeration of unvalidated Internet users' opinions and beliefs. Several participants were trained in machine learning, and said they were fully aware AI systems pulled "*stratified averages*" of word distributions from their training data (P05). Still, whether or not they believed the underlying training data reflected quality or worthwhile advice, participants commonly echoed the belief that AI could channel specific perspectives through the Internet.

AI can synthesize hidden truths if just given the right data. Finally, we noticed that participants assumed AI could pull essential truths out of large amounts of raw data—even truths invisible to the human eye, or answers to unknowable questions. This folk theory appeared most clearly in line with the Strategist role.

Belief in AI's ability to learn latent patterns drove participants to upload as much detail as possible on their previous interactions with romantic partners, including screenshots of texts and detailed narrative retellings of prior encounters. Some participants described actively managing what context their AI should consider at a given moment. When asking their AI how to navigate feelings of shame around entering a relationship with someone who turned out to be married, P22 specifically made a request to his AI: "*please update your memory*" with this new detail. P17 utilized another tactic of asking the system to ground a response in an external piece of data. They specifically maintained a document that listed their own aspirational values and provided it to their AI system whenever they asked it to analyze a situation they were encountering.

7 Discussion: Research Agendas

Our exploratory work has contributed foundational insights into how people seek out relationship advice from AI. We have synthesized the roles for AI that they seek out (§4) and the benefits and risks they navigate (§5), as well as the folk theories they hold and the co-optation tactics they employ for achieving their goals (§6).

Contrary to what one might expect, our study of AI-mediated relationship advice did not lead us to conclude we should stop this practice wholesale. While we do wish to encourage people to reach out to real-life supports, participants' descriptions of benefits also indicate there is real use for AI in helping people achieve their intimacy goals. Doing so requires guidance for designers on how to build AI systems as tools for richer human-human intimacy, rather than distractions from our real lives [71]. Here, we draw on our findings to make those recommendations for the HCI subcommunities of human-AI interaction, AI safety, and sociotechnical research. In speaking to these subfields, we hope to distill lessons for colleagues within the HCI community working at different levels of influence, be it designing AI interactions, working in trust and safety making policy around safeguarding AI products, or conducting fundamental research on technology's role in society.

Folk Theory	Co-optation Tactic	Example
AI will always flatter the user (sycophancy)	The user presents their own words and behaviors, but tells the system they're someone else's.	"*I describe situations from the perspective of the other person and read ChatGPT's interpretation. Since it usually tends to pander to the author, it helps me understand different points of view.*" (P05)
	The user asks for direct honesty.	"*Can you tell me if he's showing interest or just being nice? I can't tell. Be honest.*" (P06)
	Knowing the system will try to agree with them, the user states a conjecture of which they are not sure. If the system disagrees, the user believes the system's assessment.	"*I feed it my message history and say, 'they're not interested in me, are they? I would expect it to say yeah, you're right, they're not interested. But if it says, this person is clearly interested in you, then they definitely are, because ChatGPT is always delusional, so if even ChatGPT can't find ways to see it in a negative light, then it must be airtight.*" (P01)
AI can access perspectives from all over the Internet	The user asks the system to explore different, even contrasting, perspectives.	"*I try to avoid that confirmation bias by asking him like, OK, what are the different perspectives on [this situation], what could it mean in different ways?*" (P15)
	The user asks for a "third-party" or neutral perspective	"*Pretend to be a third party who cares equally about both of us, how would you resolve this argument?*" (P20)
	The user asks the system to take on a specific role.	"*Can you act as a clinical psychologist and create notes on the different items we've been working on for the past two weeks?*" (P17) "*How does an avoidant think about this?*" (P06)
AI can synthesize hidden truths if you just give it the right data	The user asks the system learn from prior conversations	"*Based on all I have told you about my partner's behavior, what kind of person am I meant to be with?*" (P11)
	The user uploads entire seemingly objective records of interactions with potential partners to inform the system's response	P17 uploaded pages worth of screenshots to ChatGPT to analyze whether old patterns were appearing in his new relationship
	The user asks the system to ground its responses in other documents	P17 created a bulleted list of his relationship values, and uploaded it as supplementary material for ChatGPT to use in subsequent advice

Table 4: Users' Folk Theories of How AI Works, and Corresponding Tactics for Co-opting AI.

7.1 Recommendations for Human-AI Interaction

Our findings highlight how designers might facilitate human-AI interactions that allow users to achieve benefits while mitigating risks. Technologists seeking to apply our recommendations will naturally encounter context-specific challenges. One recommendation may be more appropriate than another in the design context, or recommendations may collide. We do not intend our recommendations as a checklist, but rather, guidance for how technologists might best support users' advice-seeking. Technologists should apply each as they are best suited to specific project needs.

R1. AI should support users' relationship advice-seeking by surfacing and adhering to clear roles. AI systems can scaffold users' relationship advice-seeking using the roles from our findings. At the start of an interaction, a system might suggest that the user choose a role from Table 2 for it to take on. Encouraging users to select a role can help the system better target future responses, and more importantly, encourage the user to more deeply consider what kind of support they are looking for: Are they seeking clarity on a situation, in which case they might seek out the Judge, or do they simply want to vent about their feelings with a Listener? Many have called designers to move away from general-purpose AI systems and towards purpose-built and domain-specific alternatives [67]. We agree with this call, and argue that in the context of relationship advice, there exist so many sought roles and social supports that explicit delineation of roles are necessary for successful interactions. Importantly, our recommendations can not only better scaffold the use of AI for relationship advice, but also encourage users to think critically about their use of AI for this purpose. For instance, using clear roles in human-AI interface design can give users an opportunity to reflect whether they want AI to play that role at

all. A system could surface to the user, "*You seem to be looking for me to be a Listener right now. Is that what will be most helpful at this moment?*". The system could also use the identified roles to warn users of the system's limitations in meeting that need, particularly when it comes to how an AI product differs from a human companion. In our example of the AI as a Listener, the system could respond, "*Unlike a human listening, I may store what you tell me to improve how I respond. Here's my privacy policy.*"

AI designers can also use the folk theories and co-option tactics we identified (§6) to orient towards interactions that align with what users already believe AI systems do. For example, since users believe AI can access disparate perspectives from all over the Internet, the system might ask the user if they want to think about a situation from another perspective. Scaffolding support through the roles and strategies identified in our work can improve the intentionality behind a users' relationship advice-seeking, but developers must also design guardrails specific to each role and strategy. AI systems could also encourage users to critically appraise their theories of AI systems' capabilities. For example, a user who believes AI can access diverse perspectives within its training data could be prompted to consider whether the system's responses might come from cultures and norms different from the users'.

R2. AI systems should track potential over-reliance through usage patterns. Given the potential impacts of over-relying on AI for relationship advice (§5), we see potential for AI systems to track user behaviors resembling over-reliance, and communicate to the user when they are reaching a given threshold. For example, an AI system might show people how often they have sought advice on a specific situation or through a specific role (e.g., "*I noticed you have asked about your breakup every day for over a year now*"). Making users aware of their potential over-reliance behaviors aligns well with users' folk theory around AI's ability to synthesize hidden truths or identify patterns invisible to humans. Whether the numbers actually add up to some personal threshold of overreliance, having tracking in this way can help users independently identify the support they are consistently seeking out. Ultimately, the system might assist a user's ability to track their over-reliance patterns independently.

R3. AI systems should actively dispel authority myths, especially around intimacy questions. Our findings show that when seeking relationship advice, people ask AI to deliver concrete judgments that exceed both AI and human capacity for certainty: e.g., whether someone is interested in them, or when they will move on from their ex §4.1. In other words, people are looking to AI to be a source of authority on complex and fundamentally human ideas: what constitutes emotional availability, how a partner can fulfill your needs, and how to be good to others. Roles like Self-Critic and Judge imbue AI with moral and esteem authority over the users' and others' actions; roles like Teacher and Therapist cast AI as an expert on users' knowledge, beliefs and skills; and AI as Interpreter give these systems the authority of wisdom—not even to speak of AI as Executive, where people look to the system to make and justify the call.

That users look to AI to be an authority is a classic problem in human-AI interaction, and best-practice guidelines suggest AI systems should hedge or acknowledge their own limitations (cf. [3]). But what would AI hedging look like in the context of unknowable relationship quandaries? Reminding users there is no way to know when they will get over their ex may have the inadvertent effect of worsening their rumination. We argue addressing these gray areas is key for ensuring AI is safe and helpful for these uses, and see ample future work on how best to hedge on questions like these.

R4. AI should foster the development of social support networks. In tandem with dispelling authority myths, the system can emphasize that it alone is incapable of providing full advice, and should encourage the user to seek out and strengthen their social support networks. Our recommendation here echoes Zhang et al.'s suggestion for AI to "*scaffold the development of real-world social skills, enhance relational awareness, and ultimately strengthen human-human relationships*" [80]. We further advocate that AI can connect users to other forms of social support if they are unable to turn to humans in their proximity. As participants in our study shared, not everyone has a friend or therapist to turn to; therefore, AI can guide users to call confidential hotlines or find designated support groups, online or otherwise.

7.2 Recommendations for AI safety

For the psychosocial AI and digital safety communities, our study highlights the need for new methods to understanding and mitigating interpersonal AI risks—both within the current paradigm of large-scale classifiers for risk-bearing conversations, and in more sociotechnical and contextual approaches to safety.

Current best practices in safeguarding AI, including safety testing and red-teaming [33, 74, 78], focus on detecting when users may be in risky situations, and improving models' responses to different levels of risk. This approach relies on developers to anticipate harms and risk signals in short talk turns. This works for acute crises like suicidal ideation, where clear signals and interventions are well-established [11]. However, our findings (§5) show that interpersonal risks like potentially abusive relationship dynamics are less cleanly defined than mental health crises, and thus require different approaches to safety.

Many of the interpersonal risks we identified in §5 unfold over longer-term interactions with the system: e.g., ruminating on a decision instead of acting, excusing a partner's unhealthy behaviors, or retreating from real-life social networks. We need longitudinal analyses of user conversations to understand how these interpersonal risks manifest. Another key issue is how much of the risk is apparent to the AI system, and thus amenable to system-side guardrails. For example, whether a user takes AI's advice and uses it to escalate an argument with their partner will not be visible to any safety classifier. There is a role here for systems to take a more proactive role in safety checks than scanning for risk signals: consider, e.g., that a system that has seen a user prompt it for argument ammunition three times in one week might ask about a pattern of escalating conflict. Like R2 above, where we suggest AI systems might track overreliance, we suggest safety-centered AI systems might track not just usage but also whether the user's behaviors might lead to unhealthy personal attachments. Gently

asking follow-up questions to clarify the user's relationship situation could help users notice an unhealthy pattern, and correct it how they see fit.

More fundamentally, we also need further foundational research with users to understand possible harms: signals to watch for that indicate AI use has reached a threshold of unhealthy use. This is a thorny question for digital safety at large: how can systems, product policy designers, and users work together to determine what behaviors user contexts might be healthy and unhealthy, in ways legible to computing systems? Benchmarking or standardizing what constitutes healthy relationship behaviors risks collapsing the inherent diversity in how people approach relationships. This is not to say it is not possible to determine when a user is in an abusive situation, or developing misogynistic ideas, and intervene. Rather, we argue interpersonal harms are a fantastic test case for greater nuance in current approaches to safety design, moving beyond content moderation's legacy of technologists making black-and-white designations of what is safe and unsafe, and into more qualitative, contextual, and even participatory evaluations of what is risky in human-AI conversations (e.g., [19, 66]).

Rethinking AI safeguarding in this way would be a step towards a broadened aperture of how safety is done, looking beyond technical improvements to AI systems and towards sociotechnical interventions improving users' ability to manage their own risk. Here, technologists might consider also pouring effort into ecosystem-wide initiatives beyond improving AI models, like assessing and improving technological literacy (including but not limited to AI), or bolstering direct support services for people facing relationship difficulties (e.g., clinical computer security for tech abuse [69, 70]). The HCI community is well-positioned to lead an expanded aperture of AI safety, repositioning the conversation towards social and political reforms alongside technical interventions, and we hope our analysis of relationship advice provides a keystone case for this growing subfield.

7.3 Recommendations for sociotechnical research

What is lost when people increasingly seek out AI over their social support networks? Our findings show that in the context of relationship advice-seeking, people turn to their AI for social support over their friends and family. Doing so is both convenient and safe; not only can AI act as a listener, therapist, or strategist, but people can also avoid the guilt of constantly bothering their friends with their troubles and potential judgment from others. As other studies suggest, through this substitution, AI plays a part in social erosion [41, 80].

Further work is needed to determine more societal implications of AI's greater role as a social support system. It is possible that, as previous communication and media scholars have pointed out, this is yet another cycle of an innovative technology that disrupts social norms and practices but in the end does not live up to the moral panics that emerge in the process [4]. However, AI features greater capacity and offers more benefits than previous technologies that have come before. While we are unequipped to confirm or deny social erosion via AI, our work sets the stage for future work that might explore relationship advice-seeking at a broader scale (e.g., different populations, users of more levels of AI literacy), in different contexts (e.g., non-use, sex advice), and with respect to impacts on different social norms (e.g., privacy). Future work might also look across multiple contexts of social support, including relationship advice and self-care [10], to explore how AI and social networks can complement one another under one system of social support.

8 Conclusion

Our findings and contributions, grounded in qualitative analysis of 25 participants' questionnaires and 17 interviews, demonstrate how the context of using AI for relationship advice is an expansive one. First, we outline a conceptual model and typology of what relationship advice users seek out from AI systems, including nine roles people envision for AI: interpreter, listener, judge, self-critic, teacher, therapist, strategist, editor, and executive. Subsequently, we build on these roles through a risk/benefit analysis to understand the potential rewards and harms people navigate in human-AI interactions related to each. We additionally identify three folk theories that users construct around how AI systems work and the co-optation tactics they use when engaging in AI relationship advice-seeking. Finally, we propose a set of recommendations for the design of AI systems and future sociotechnical infrastructure research.

Acknowledgments

Thank you first to our participants, for sharing their time, energy, and experiences with us, and to our reviewers, for pushing our work to a stronger place. We are also grateful to the research communities around us, for intellectual foundations, helpful discussions, and for sharing in the delight of this work as it unfolded: the Social Media Collective at Microsoft Research, the Digital Intimacies workshop at CSCW 2025, Ashley M. Walker, Renee Shelby, Amy W. Xiao, Sharon Heung, and Katie Heins.

References

[1] Fatemeh Alizadeh, Dennis Lawo, Gunnar Stevens, Douglas Zytko, and Motahhare Eslami. 2024. When the "Matchmaker" Does Not Have Your Interest at Heart: Perceived Algorithmic Harms, Folk Theories, and Users' Counter-Strategies on Tinder. *Proc. ACM Hum.-Comput. Interact.* 8, CSCW2, Article 481a (Nov. 2024), 29 pages. doi:10.1145/3689710

[2] Teresa Almeida. 2015. Designing intimate wearables to promote preventative health care practices. In *Adjunct Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2015 ACM International Symposium on Wearable Computers* (Osaka, Japan) *(UbiComp/ISWC'15 Adjunct)*. Association for Computing Machinery, New York, NY, USA, 659–662. doi:10.1145/2800835.2809440

[3] Saleema Amershi, Dan Weld, Mihaela Vorvoreanu, Adam Fourney, Besmira Nushi, Penny Collisson, Jina Suh, Shamsi Iqbal, Paul N. Bennett, Kori Inkpen, Jaime Teevan, Ruth Kikin-Gil, and Eric Horvitz. 2019. Guidelines for Human-AI Interaction. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland, UK) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 13 pages. doi:10.1145/3290605.3300233

[4] Nancy K Baym. 2015. *Personal connections in the digital age.* John Wiley & Sons.

[5] Rosanna Bellini. 2023. Paying the Price: When Intimate Partners Use Technology for Financial Harm. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 872, 17 pages. doi:10.1145/3544548.3581101

[6] Bruce J Biddle. 1986. Recent developments in role theory. *Annual review of sociology* 12, 1 (1986), 67–92.

[7] Virginia Braun and Victoria Clarke. 2021. One size fits all? What counts as quality practice in (reflexive) thematic analysis? *Qualitative research in psychology* 18, 3 (2021), 328–352.

[8] Virginia Braun, Victoria Clarke, Nikki Hayfield, and Gareth Terry. 2019. Thematic analysis. In *Handbook of research methods in health social sciences*. Springer, 843–860.

[9] Stefana Broadbent. 2020. Approaches to personal communication. In *Digital anthropology*. Routledge, 127–145.

[10] Tara Capel, Bernd Ploderer, Filip Bircanin, Simon Hanmer, Jamie Paige Yates, Jiaxuan Wang, Kai Ling Khor, Tuck Wah Leong, Greg Wadley, and Michelle Newcomb. 2024. Studying Self-Care with Generative AI Tools: Lessons for Design. In *Proceedings of the 2024 ACM Designing Interactive Systems Conference* (Copenhagen, Denmark) *(DIS '24)*. Association for Computing Machinery, New York, NY, USA, 1620–1637. doi:10.1145/3643834.3661614

[11] Mohit Chandra, Suchismita Naik, Denae Ford, Ebele Okoli, Munmun De Choudhury, Mahsa Ershadi, Gonzalo Ramos, Javier Hernandez, Ananya Bhattacharjee, Shahed Warreth, and Jina Suh. 2025. From Lived Experience to Insight: Unpacking the Psychological Risks of Using AI Conversational Agents. In *Proceedings of the 2025 ACM Conference on Fairness, Accountability, and Transparency (FAccT '25)*. Association for Computing Machinery, New York, NY, USA, 975–1004. doi:10.1145/3715275.3732063

[12] Myra Cheng, Sunny Yu, Cinoo Lee, Pranav Khadpe, Lujain Ibrahim, and Dan Jurafsky. 2025. Social sycophancy: A broader understanding of LLM sycophancy. *arXiv preprint arXiv:2505.13995* (2025).

[13] Sidney Cobb. 1976. Social support as a moderator of life stress. *Biopsychosocial Science and Medicine* 38, 5 (1976), 300–314.

[14] Sheldon Cohen and Harry M Hoberman. 1983. Positive events and social supports as buffers of life change stress 1. *Journal of applied social psychology* 13, 2 (1983), 99–125.

[15] Emmelyn AJ Croes, Marjolijn L Antheunis, Chris van der Lee, and Jan MS de Wit. 2024. Digital confessions: The willingness to disclose intimate information to a chatbot and its impact on emotional well-being. *Interacting with Computers* 36, 5 (2024), 279–292. doi:10.1093/iwc/iwae016

[16] Carolyn E Cutrona and Julie A Suhr. 1992. Controllability of stressful events and satisfaction with spouse support behaviors. *Communication research* 19, 2 (1992), 154–174.

[17] Stuti Das and Reya Farber. 2020. User-generated online queer media and the politics of queer visibility. *Sociology Compass* 14, 9 (2020), e12824.

[18] Munmun De Choudhury and Emre Kiciman. 2017. The language of social support in social media and its effect on suicidal ideation risk. In *Proceedings of the international AAAI conference on web and social media*, Vol. 11. 32–41. doi:10.1609/icwsm.v11i1.14891

[19] Wesley Hanwen Deng, Wang Claire, Howard Ziyu Han, Jason I. Hong, Kenneth Holstein, and Motahhare Eslami. 2025. WeAudit: Scaffolding User Auditors and AI Practitioners in Auditing Generative AI. *Proc. ACM Hum.-Comput. Interact.* 9, 7, Article CSCW521 (Oct. 2025), 35 pages. doi:10.1145/3757702

[20] Nisha Devasia, Adrian Rodriguez, Logan Tuttle, and Julie A. Kientz. 2025. Partnership through Play: Investigating How Long-Distance Couples Use Digital Games to Facilitate Intimacy. In *Proceedings of the 2025 ACM Designing Interactive Systems Conference (DIS '25)*. Association for Computing Machinery, New York, NY, USA, 3563–3580. doi:10.1145/3715336.3735773

[21] Michael Ann DeVito, Jeremy Birnholtz, Jeffery T. Hancock, Megan French, and Sunny Liu. 2018. How People Form Folk Theories of Social Media Feeds and What it Means for How We Study Self-Presentation. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–12. doi:10.1145/3173574.3173694

[22] Clare Duffy. 2025. OpenAI Hit With Lawsuit After Alleged ChatGPT-4o Role in Teen's Suicide. *CNN* (Aug. 26 2025). https://www.cnn.com/2025/08/26/tech/openai-chatgpt-teen-suicide-lawsuit Accessed: 2025-09-11.

[23] Nicole B Ellison, Charles Steinfield, and Cliff Lampe. 2007. The benefits of Facebook "friends:" Social capital and college students' use of online social network sites. *Journal of computer-mediated communication* 12, 4 (2007), 1143–1168. doi:10.1111/j.1083-6101.2007.00367.x

[24] Motahhare Eslami, Karrie Karahalios, Christian Sandvig, Kristen Vaccaro, Aimee Rickman, Kevin Hamilton, and Alex Kirlik. 2016. First I "like" it, then I hide it: Folk Theories of Social Feeds. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI '16)*. Association for Computing Machinery, New York, NY, USA, 2371–2382. doi:10.1145/2858036.2858494

[25] Aaron Fanous, Jacob Goldberg, Ank Agarwal, Joanna Lin, Anson Zhou, Sonnet Xu, Vasiliki Bikia, Roxana Daneshjou, and Sanmi Koyejo. 2025. Syceval: Evaluating llm sycophancy. In *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, Vol. 8. 893–900. doi:10.1609/aies.v8i1.36598

[26] Bo Feng. 2009. Testing an integrated model of advice giving in supportive interactions. *Human Communication Research* 35, 1 (2009), 115–129.

[27] Bo Feng and Erina L MacGeorge. 2010. The influences of message and source factors on advice outcomes. *Communication Research* 37, 4 (2010), 553–575.

[28] Casey Fiesler and Amy S. Bruckman. 2019. Creativity, Copyright, and Close-Knit Communities: A Case Study of Social Norm Formation and Enforcement. *Proc. ACM Hum.-Comput. Interact.* 3, GROUP, Article 241 (Dec. 2019), 24 pages. doi:10.1145/3361122

[29] Casey Fiesler, Shannon Morrison, and Amy S. Bruckman. 2016. An Archive of Their Own: A Case Study of Feminist HCI and Values in Design. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI '16)*. Association for Computing Machinery, New York, NY, USA, 2574–2585. doi:10.1145/2858036.2858409

[30] Diana Freed, Jackeline Palmer, Diana Elizabeth Minchala, Karen Levy, Thomas Ristenpart, and Nicola Dell. 2017. Digital Technologies and Intimate Partner Violence: A Qualitative Analysis with Multiple Stakeholders. *Proc. ACM Hum.-Comput. Interact.* 1, CSCW, Article 46 (Dec. 2017), 22 pages. doi:10.1145/3134681

[31] Megan French and Jeff Hancock. 2017. What's the folk theory? Reasoning about cyber-social systems. *Reasoning About Cyber-Social Systems (February 2, 2017)* (2017). doi:10.2139/ssrn.2910571

[32] Patricia I Fusch and Lawrence R Ness. 2015. Are we there yet? Data saturation in qualitative research. (2015). doi:10.46743/2160-3715/2015.2281

[33] Tarleton Gillespie, Ryland Shaw, Mary L. Gray, and Jina Suh. 2026. AI Red-Teaming Is a Sociotechnical Problem. *Commun. ACM* (Jan. 2026), 10 pages. doi:10.1145/3731657 Online First.

[34] Erving Goffman. 1955. On face-work: An analysis of ritual elements in social interaction. *Psychiatry* 18, 3 (1955), 213–231. doi:10.1080/00332747.1955.11023008

[35] Daena J Goldsmith. 2000. Soliciting advice: The role of sequential placement in mitigating face threat. *Communications Monographs* 67, 1 (2000), 1–19.

[36] Daena J Goldsmith and Kristine Fitch. 1997. The normative context of advice as social support. *Human communication research* 23, 4 (1997), 454–476.

[37] Suzanne Dolwick Grieb, Katherine C Smith, Karen Calhoun, Darius Tandon, et al. 2015. Qualitative research and community-based participatory research: Considerations for effective dissemination in the peer-reviewed literature. *Progress in Community Health Partnerships: Research, Education, and Action* 9, 2 (2015), 275–282.

[38] Oliver L Haimson, Avery Dame-Griff, Elias Capello, and Zahari Richter. 2021. Tumblr was a trans technology: the meaning, importance, history, and future of trans technologies. *Feminist media studies* 21, 3 (2021), 345–361.

[39] Eric Horvitz. 1999. Principles of mixed-initiative user interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Pittsburgh, Pennsylvania, USA) *(CHI '99)*. Association for Computing Machinery, New York, NY, USA, 159–166. doi:10.1145/302979.303030

[40] Haonan Hou, Kevin Leach, and Yu Huang. 2024. ChatGPT Giving Relationship Advice–How Reliable Is It?. In *Proceedings of the International AAAI Conference on Web and Social Media*, Vol. 18. 610–623. doi:10.1609/icwsm.v18i1.31338

[41] Irene Hou, Owen Man, Kate Hamilton, Srishty Muthusekaran, Jeffin Johnykutty, Leili Zadeh, and Stephen MacNeil. 2025. 'All Roads Lead to ChatGPT': How Generative AI is Eroding Social Interactions and Student Learning Communities. In *Proceedings of the 30th ACM Conference on Innovation and Technology in Computer Science Education V. 1* (Nijmegen, Netherlands) *(ITiCSE 2025)*. Association for Computing Machinery, New York, NY, USA, 79–85. doi:10.1145/3724363.3729024

[42] Sabrina Angela Huang, Jeffrey Hancock, and Stephanie Tom Tong. 2022. Folk theories of online dating: Exploring people's beliefs about the online dating process and online dating algorithms. *Social Media + Society* 8, 2 (2022). doi:10.1177/20563051221089561

[43] Berton H Kaplan, John C Cassel, and Susan Gore. 1977. Social support and health. *Medical care* 15, 5 (1977), 47–58.

[44] Nadia Karizat, Dan Delmonaco, Motahhare Eslami, and Nazanin Andalibi. 2021. Algorithmic Folk Theories and Identity: How TikTok Users Co-Produce Knowledge of Identity and Engage in Algorithmic Resistance. *Proc. ACM Hum.-Comput. Interact.* 5, CSCW2, Article 305 (Oct. 2021), 44 pages. doi:10.1145/3476046

[45] Tamara Kneese, Briana Vecchione, and Alice Marwick. 2025. A chatbot for the soul: mental health care, privacy, and intimacy in AI-based conversational agents. *Communication and Change* 1, 1 (2025), 15. doi:10.1007/s44382-025-00015-y

[46] Min Kyung Lee, Daniel Kusbit, Evan Metsky, and Laura Dabbish. 2015. Working with machines: The impact of algorithmic and data-driven management on human workers. In *Proceedings of the 33rd annual ACM conference on human factors in computing systems*. 1603–1612.

[47] Yi-Chieh Lee, Naomi Yamashita, Yun Huang, and Wai Fu. 2020. " I hear you, I feel you": encouraging deep self-disclosure through a chatbot. In *Proceedings of the 2020 CHI conference on human factors in computing systems*. 1–12.

[48] Roxanne Leitão. 2018. Digital Technologies and their Role in Intimate Partner Violence. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI EA '18)*. Association for Computing Machinery, New York, NY, USA, 1–6. doi:10.1145/3170427.3180305

[49] Hong Li, Jonna Häkkilä, and Kaisa Väänänen. 2018. Review of unconventional user interfaces for emotional communication between long-distance partners. In *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services*. 1–10.

[50] Anna M Lomanowska and Matthieu J Guitton. 2016. Online intimacy and well-being in the digital age. *Internet interventions* 4 (2016), 138–144.

[51] Ryan Louie, Andy Coenen, Cheng Zhi Huang, Michael Terry, and Carrie J. Cai. 2020. Novice-AI Music Co-Creation via AI-Steering Tools for Deep Generative Models. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing*

Systems (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–13. doi:10.1145/3313831.3376739
[52] Erina L MacGeorge, Bo Feng, and Brant R Burleson. 2011. Supportive communication. *Handbook of interpersonal communication* 4 (2011), 317–354.
[53] Erina L. MacGeorge and Elizabeth Dorrance Hall. 2014. *Relationship advice*. Cambridge University Press, 188–208.
[54] Carolyn Marvin. 1988. *When old technologies were new: Thinking about electric communication in the late nineteenth century*. Oxford University Press.
[55] Samuel Mayworm, Michael Ann DeVito, Daniel Delmonaco, Hibby Thach, and Oliver L. Haimson. 2024. Content Moderation Folk Theories and Perceptions of Platform Spirit among Marginalized Social Media Users. *Trans. Soc. Comput.* 7, 1–4, Article 1 (March 2024), 27 pages. doi:10.1145/3632741
[56] M McCain, R Linthicum, C Lubinski, A Tamkin, S Huang, M Stern, K Handa, E Durmus, T Neylon, S Ritchie, et al. 2025. How People Use Claude for Support, Advice, and Companionship.
[57] Maryam Mehrnezhad and Teresa Almeida. 2021. Caring for Intimate Data in Fertility Technologies. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 409, 11 pages. doi:10.1145/3411764.3445132
[58] Moohong Min, Donjung Choi, Jaekwang Kim, and Jee-Hyong Lee. 2011. The identification of intimate friends in personal social network. In *2011 International Conference on Computational Aspects of Social Networks (CASoN)*. IEEE, 233–236.
[59] Nelly EJ Oudshoorn and Trevor Pinch. 2003. Introduction: How users and non-users matter. In *How users matter. The co-construction of users and technology*. mit Press, 1–25.
[60] Aswati Panicker, Novia Nurain, Zaidat Ibrahim, Chun-Han (Ariel) Wang, Seung Wan Ha, Yuxing Wu, Kay Connelly, Katie A. Siek, and Chia-Fang Chung. 2024. Understanding fraudulence in online qualitative studies: From the researcher's perspective. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 824, 17 pages. doi:10.1145/3613904.3642732
[61] Charlie Parker, Sam Scott, and Alistair Geddes. 2019. Snowball sampling. *SAGE research methods foundations* (2019).
[62] Trevor J Pinch and Wiebe E Bijker. 1984. The social construction of facts and artefacts: Or how the sociology of science and the sociology of technology might benefit each other. *Social studies of science* 14, 3 (1984), 399–441.
[63] Renee Shelby, Shalaleh Rismani, and Negar Rostamzadeh. 2024. Generative AI in Creative Practice: ML-Artist Folk Theories of T2I Use, Harm, and Harm-Reduction. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 32, 17 pages. doi:10.1145/3613904.3642461
[64] Ben Shneiderman and Pattie Maes. 1997. Direct manipulation vs. interface agents. *Interactions* 4, 6 (Nov. 1997), 42–61. doi:10.1145/267505.267514
[65] Ranjit Singh, Livia Garofalo, Briana Vecchione, and Emnet Tafesse. 2025. Chatbots in disguise: How participants use GenAI to hack qualitative research. https://datasociety.net/points/chatbots-in-disguise/. Accessed: 2025-11-02.
[66] Jaemarie Solyst, Emily Amspoker, Ellia Yang, Motahhare Eslami, Jessica Hammer, and Amy Ogan. 2025. RAD: A Framework to Support Youth in Critiquing AI. In *Proceedings of the 56th ACM Technical Symposium on Computer Science Education V. 1* (Pittsburgh, PA, USA) *(SIGCSETS 2025)*. Association for Computing Machinery, New York, NY, USA, 1071–1077. doi:10.1145/3641554.3701966
[67] Harini Suresh, Emily Tseng, Meg Young, Mary Gray, Emma Pierson, and Karen Levy. 2024. Participation in the age of foundation models. In *Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency* (Rio de Janeiro, Brazil) *(FAccT '24)*. Association for Computing Machinery, New York, NY, USA, 1609–1621. doi:10.1145/3630106.3658992
[68] Emily Tseng, Rosanna Bellini, Nora McDonald, Matan Danos, Rachel Greenstadt, Damon McCoy, Nicola Dell, and Thomas Ristenpart. 2020. The tools and tactics used in intimate partner surveillance: an analysis of online infidelity forums. In *Proceedings of the 29th USENIX Conference on Security Symposium (SEC'20)*. USENIX Association, USA, Article 107, 17 pages.
[69] Emily Tseng, Diana Freed, Kristen Engel, Thomas Ristenpart, and Nicola Dell. 2021. A Digital Safety Dilemma: Analysis of Computer-Mediated Computer Security Interventions for Intimate Partner Violence During COVID-19. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 71, 17 pages. doi:10.1145/3411764.3445589
[70] Emily Tseng, Mehrnaz Sabet, Rosanna Bellini, Harkiran Kaur Sodhi, Thomas Ristenpart, and Nicola Dell. 2022. Care Infrastructures for Digital Security in Intimate Partner Violence. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 123, 20 pages. doi:10.1145/3491102.3502038
[71] Sherry Turkle. 2011. *Alone Together: Why We Expect More from Technology and Less from Each Other*. Basic Books, Inc., USA.
[72] Ralph H Turner. 2001. Role theory. In *Handbook of sociological theory*. Springer, 233–254.
[73] Briana Vecchione and Ranjit Singh. 2025. Artificial intelligence is mental: Evaluating the role of large-language models in supporting mental health and well-being. *Big Data & Society* 12, 4 (2025), 20539517251383884. doi:10.1177/20539517251383884
[74] Laura Weidinger, John Mellor, Bernat Guillen Pegueroles, Nahema Marchal, Ravin Kumar, Kristian Lum, Canfer Akbulut, Mark Diaz, Stevie Bergman, Mikel Rodriguez, et al. 2024. Star: Sociotechnical approach to red teaming language models. *arXiv preprint arXiv:2406.11757* (2024).
[75] Monica T Whitty. 2008. Liberating or debilitating? An examination of romantic relationships, sexual relationships and friendships on the Net. *computers in Human Behavior* 24, 5 (2008), 1837–1850.
[76] Sarah Wiseman and Sandy J. J. Gould. 2018. Repurposing Emoji for Personalised Communication. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–10. doi:10.1145/3173574.3173726
[77] Marc Zao-Sanders. 2025. How people are really using gen AI in 2025. *Harvard Business Review* (2025).
[78] Alice Qian Zhang, Ryland Shaw, Jacy Reese Anthis, Ashlee Milton, Emily Tseng, Jina Suh, Lama Ahmad, Ram Shankar Siva Kumar, Julian Posada, Benjamin Shestakofsky, Sarah T. Roberts, and Mary L. Gray. 2024. The Human Factor in AI Red Teaming: Perspectives from Social and Collaborative Computing. In *Companion Publication of the 2024 Conference on Computer-Supported Cooperative Work and Social Computing* (San Jose, Costa Rica) *(CSCW Companion '24)*. Association for Computing Machinery, New York, NY, USA, 712–715. doi:10.1145/3678884.3687147
[79] Renwen Zhang, Han Li, Han Meng, Jinyuan Zhan, Hongyuan Gan, and Yi-Chieh Lee. 2025. The Dark Side of AI Companionship: A Taxonomy of Harmful Algorithmic Behaviors in Human-AI Relationships. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI '25)*. Association for Computing Machinery, New York, NY, USA, Article 13, 17 pages. doi:10.1145/3706598.3713429
[80] Yutong Zhang, Dora Zhao, Jeffrey T Hancock, Robert Kraut, and Diyi Yang. 2025. The Rise of AI Companions: How Human-Chatbot Relationships Influence Well-Being. *arXiv preprint arXiv:2506.12605* (2025).

A Questionnaire

A.1 The Problem of Bot Participants

Our recruitment was live from June to July of 2025, and received 281 responses in that time. However, we encountered a high volume of responses that seemed to be fraudulent or AI-generated [60, 65]. Some included in their open response fields traces of what was clearly an AI response (e.g., "*Here are 3 prompts a person might use to ask AI for relationship advice:*", followed by 3 prompts enumerated). One participant said they used a dating app, but when asked in an interview to recount more detail, abruptly ended the interview. Others submitted prompts that were not related to relationship advice (e.g., "*online reading*").

We therefore established inclusion criteria and manually filtered for responses with one or more of the following: a reCAPTCHA v3 score of 0.5-1.0, an academic institution-affiliated email address, or qualitative sample prompts that did not appear to be generated by a AI system. Ultimately, we are confident that our 25 final participants are as legitimate as we can ascertain. How to address the spate of bot responses [60]—and, more perplexing, the use of AI in survey research by real humans [65]—is a problem we hope our research community can collectively address.

A.2 Questionnaire Text

1. What is your age? Note that you must be 18 or older to participate in this study.

- 18-25
- 26-34
- 35-44
- 45-54
- 55-64

- 65+

2. What U.S. state do you currently reside in?
3. What kinds of relationship advice have you used generative AI for? (check all that may apply)

- Establishing new connections / dating
- Supporting ongoing connections / relationships
- Ending connections / breakups
- Other:

4. What kinds of connections or intimacy are you currently looking for?

- Dating / Relationships
- Casual sex / Hookups
- Situationships / Friends with benefits
- General closeness with others
- Nothing specific / Exploring
- I am not looking for connections right now
- I am looking to deepen existing connections
- Other:

5. What technologies or platforms do you currently use to achieve these goals? For example, Tinder, WhatsApp, Snapchat. Identify as many as possible.
6. What are examples of prompts you have used in the past when seeking relationship advice or counseling from a generative AI model?
7. What is your email to set up a one-hour interview?

B Interview Protocols

B.1 Interview 1: Understanding Participants' Experiences Using GenAI for Relationship Advice

1. Consent, introductions, and warm-up

- Thank you for joining us today to talk about using generative AI for relationship advice!
- Before we begin, we want to review the consent materials for this study. Hopefully you've had a chance to review them before this but we will highlight some important points first and ask for your consent to participate.
  - Compensation for 2 interviews is a $50 gift card upon completion of the second interview
  - We do not anticipate any harms coming from participating in this study, but you might experience some discomfort talking about experiences you've had or are having. With that, we want to share that we ourselves use generative AI for relationship advice and that we hope you aren't too embarrassed or feel any shame for what we think is a very relatable experience.
- Demographics. First we'll ask some demographic questions to help us understand who you are and where you're coming from. You're free to decline to answer any of these, no problem.
- What is your gender?
- What is your sexual orientation?
- Are you Hispanic or Latino?
- Regardless of your answer to the prior question, which of these the following groups in which you consider yourself to be a member: American Indian or Alaska Native, Asian, Black or African American, Native Hawaiian or other Pacific Islander, White, Other.

2. Mapping out the intimate platform ecosystem

- From our screener survey, we heard that you use X platforms and are currently seeking Y goals. Do you use any other platforms to achieve those goals?
- Is there anything you'd like to add or change to those responses?
- Can you walk me through how you use these platforms to achieve your goals?
- Are you happy with the tools you have? Why or why not?

3. Understanding the Use of AI for Relationship Advice

- What led you to using AI for relationship advice?
- How often do you consult chatbots for advice in general? What makes relationship advice different from other reasons you could engage with the chatbot?
- How does consulting the chatbot compare to using a journal? Talking to a friend? Talking to a therapist?
- Key differences might be, you'd have more of a conversation with a friend or therapist and listen to what they say, whereas with a journal it matters less what the tool says back
- When would you *not* seek out relationship advice from a chatbot?

4. Let's now talk about the prompts you submitted. It will help our conversation if you're able to assign each prompt to a specific experience you've had recently. I'll read each out loud now and try to place each while I do. For each:

- Can you explain a little bit more about the context?
- What made you seek out the chatbot for advice?
- Regardless of how the system responded, what were you looking for in asking this?
- Tell me more about how it responded and how it made you feel.
- If they answered "type of prompt":
- What do you remember about how the system responds?
- Are these types of responses useful for you?
- How do they make you feel when you first receive them?
- Do these responses meet your expectations? Were you surprised at all by them?
- Systems often ask if the response is good or bad. What would you consider to be a good or bad response to this type of prompt?
- How does this response compare to what you have gotten from your community, like your therapist, friend, or family member?

5. AI systems have these features where you can look back at your prior interactions, like memory or saved conversations. Have you used these features since you had that interaction, for example to look back at your data from that time? Why or why not?
6. Let's say you were talking to someone you were interested in, and they told you they'd used AI to come up with what they said to you. How would that make you feel?

Interview 2: Member Checking + Resources Welcome back! Since we last saw you, we talked to more people who use AI for relationship advice. We'd like to share with you our analysis of what we heard and then ask for what you think about it. Just to remind you, last time we talked, you shared these prompts you've used with AI in this context.

1. Since the last time we talked, have you used generative AI for relationship advice again?

- What has changed? Do you have any different perceptions and reactions from the last session?

2. Here's our preliminary analysis from our first round of interviews. (Interviewer explains the analysis).

- Do you see your perspective accurately characterized here?
- Which categories seem most relevant to how you use these systems? How so?
- Are there any categories you'd add, combine, or subtract?
- Is anything in this typology new or surprising to you? Is there anything in this typology that you haven't done or thought about before that you think can be useful for you in the future?

3. A lot of these strategies come from you and others being pretty savvy about AI and what it can and can't do. We're all still learning about this tool. Thinking about how you'd coach a young person, like a younger cousin, who is seeking to use AI for relationship advice themselves.

- What advice would you give them?
- What advice would you give to a peer?

4. Zooming out from the specific strategies, processes, and roles, we have also wondered whether the AI systems are giving people tools and skills that prepare them to navigate future scenarios *without* consulting the AI system.

- We've heard how consulting with AI provides benefits when processing individual situations. For example, talking through why a potential date is acting in a confusing way can help prevent a person from spiraling.
- Let's say one day in the future you experience a situation similar to one you've consulted a chatbot about (if possible, relate a scenario discussed in the previous interview). What would you do? Would you consult the chatbot again or do you feel like you have what you need in order to process that on your own?
- What advice would you give to someone who is going through this situation? Do you think you could take that advice yourself?