

How Would You Like Your AI to Respond? A Preliminary Study of Emotional Preferences for Chatbot Support Across Life Scenarios

Jiayin Huang, Dawei Xu, and Jonggi Hong

Stevens Institute of Technology, Hoboken, NJ 07307, USA

ABSTRACT

While much research has focused on detecting user emotions, far less is known about how chatbots should express emotion back to users. This paper explores user preferences for chatbot emotional intensity across everyday situations. We conducted a mixed-methods study with 51 participants who evaluated chatbot responses at three emotional levels, non-emotional, moderate, and deep emotional, across twelve realistic scenarios, complemented by surveys and interviews. Results suggest that preferences are highly context-dependent: deep empathy was often valued, but moderation was preferred in certain scenarios. We did not observe robust gender effects in these preference patterns. Interviews further revealed ambivalence, as participants appreciated empathetic support but expressed concerns about authenticity, dependency, and fairness. We offer preliminary empirical insights, design considerations for context-aware emotional adaptivity, and ethical reflections on emotionally responsive AI.

Keywords: Chatbot, Emotional response, Affective computing, Human-AI interaction

INTRODUCTION

Conversational AI systems are increasingly woven into our daily routines, from virtual assistants to customer support bots. As these systems evolve, they are no longer merely providers of information, they are also expected to respond in emotionally appropriate ways to users' needs. Yet, despite advances in affective computing, we still know relatively little about what kind of emotional style users prefer from chatbots across different situations. Affective computing has shown that emotion-aware systems can enhance user experience by increasing perceived empathy, trust, and satisfaction (Ghandeharioun et al., 2019; Svikhnushina and Pu, 2020). At the same time, implementation challenges remain: overly emotional responses risk manipulation, promote dependency, or perpetuate stereotypes assumptions (Wu et al.; Herath). Much prior work has focused on detecting user emotions (Kratzwald et al., 2018; Jaiswal et al., 2020; Binali et al., 2010; Majumder et al., 2019), but far less is known about how chatbots should express emotion back to users in context.

To explore this question, we conducted a mixed-methods study examining user preferences for chatbot emotional intensity across twelve realistic scenarios, spanning positive milestones, interpersonal conflicts, and health-related struggles. With 51 participants, we combined surveys, semi-structured interviews, and a preference-selection task in which participants chose among non-emotional, moderately emotional, and deeply emotional chatbot responses. Our findings suggest that preferences are highly context-dependent: while deeply emotional responses were often valued, moderation was preferred in certain scenarios. We did not observe robust gender effects in these preference patterns.

This work contributes: (1) exploratory empirical evidence on how preferences for chatbot emotional expression vary across everyday scenarios; (2) a comparison of users' attitudes toward chatbot emotional responses before and after experiencing example scenarios and responses; and (3) design considerations for emotionally adaptive chatbots that balance empathy, moderation, and user control.

RELATED WORK

Prior research on affective computing and conversational AI provides the foundation for our study. Much of this work has focused on detecting user emotions, identifying ethical risks, and examining patterns of emotional communication. However, less is known about how users actually prefer chatbots to express emotion, particularly across diverse everyday scenarios.

Emotion Recognition and Expression in Conversational AI. Affective computing has advanced rapidly in detecting emotions from text, speech, and multimodal signals (Kratzwald et al., 2018; Jaiswal et al., 2020; Binali et al., 2010; Majumder et al., 2019). These capabilities allow conversational agents to sense and respond to user affect, enabling personalization and more engaging interactions. Prior studies have shown that emotionally expressive chatbots can increase empathy, trust, and user satisfaction (Ghandeharioun et al., 2019; Svikhnushina and Pu, 2020). Yet implementation remains challenging: responses that are overly emotional may appear manipulative, artificial, or stereotype-reinforcing (Wu et al.; Herath). Moreover, prior work has concentrated primarily on emotion detection or on whether empathy is present at all, rather than on how much emotional expression users prefer from chatbots in different contexts.

Emotional Communication Frameworks. A substantial body of work on supportive communication provides a theoretically grounded account of why different emotional tones succeed or fail across contexts. We draw explicitly on Burleson's person-centered support (PCS) framework (Burleson, 2008), a central model in this literature. PCS distinguishes messages by the degree to which they acknowledge, legitimize, and elaborate on the recipient's feelings while also helping with meaning-making or coping. Low PCS messages tend to minimize or redirect feelings and move quickly to information or advice; moderate PCS messages provide some recognition or sympathy but limited elaboration; high PCS messages validate emotions, reflect the recipient's perspective, and offer more contextualized support. Prior work has already

begun to operationalize emotional tone in chatbot communication. For example, Liao et al. (Liao and Yan, 2022) distinguish between empathetic and non-empathetic chatbots using validated emotional acknowledgment strategies. Building on this binary line of work, our study examines three levels of emotional expression: non-emotional, moderate, and deeply emotional, to enable a more fine-grained analysis of user preferences across scenarios.

People also pursue different emotion-regulation goals depending on appraisal and context (Lumen Learning, n.d.) (e.g., soothing, reappraisal, meaning-making). In identity/achievement episodes, language that amplifies positive affect and recognition is functional (Gable and Reis, 2010); in rumination-prone settings (e.g., regret, sleep anxiety), measured, reappraisal-oriented support is preferred, which is consistent with process models in affect regulation (Gross, 1998; Lazarus, 1984; Scherer, 2009). These theories anticipate our finding that scenario semantics shape preferences for emotional intensity.

Gender and Ethical Concerns in Emotionally Responsive AI. Prior work in interpersonal supportive communication suggests that gender differences in emotional expression and support preferences are generally modest and highly context-dependent (Burlison, 2003; MacGeorge et al., 2011; Goldsmith, 2004; Chaplin and Aldao, 2013; Christov-Moore et al., 2014). In contrast, HCI and conversational agent research has rarely modeled user gender as a moderator of responses to emotional tone, even though many studies manipulate empathic versus non-empathic styles in chatbots (e.g., Liu and Sundar, 2018; Seitz, 2024; Liu et al., 2025; Jiang et al., 2025). Existing work typically tests main effects of empathic wording on warmth, trust, or perceived support, without systematically examining whether users of different genders prefer different levels of emotional intensity. As emotionally capable AI becomes embedded in everyday life, scholars have warned about its potential for manipulation and dependency (Bakir et al., 2024; De Freitas et al., 2025; Lee and See, 2004; Glikson and Woolley, 2020).

METHOD

We adopted a mixed methods design to investigate how user preferences for chatbot emotional intensity vary across everyday scenarios and whether these patterns differed by gender. The study consisted of four components: (1) a pre-task survey capturing demographics and prior AI use, (2) a semi-structured interview probing participants' attitudes toward emotional AI, (3) a preference-selection task in which participants evaluated chatbot responses of varying emotional intensity across 12 real-life scenarios, and (4) a post-task survey assessing how these experiences shaped participants' perceptions, as shown in Figure 1.

Participants. Fifty-one participants (25 men, 26 women; $M = 27.47$, $SD = 5.07$, range = 19–49) participated in the study. Participants were recruited through university mailing lists, social media postings, and word-of-mouth. All participants provided informed written consent, and the study protocol was approved by the university's institutional review board (IRB).

The study lasted approximately 54.84 minutes ($SD = 13.04$). Participants were compensated at a rate of \$20/hour via Amazon gift card.

Study Design. To construct the study scenarios, we drew on prior work on basic emotions in everyday life (Ekman, 1992; Fehr and Russell, 1984; Shaver et al., 1987) and created 12 emotionally salient scenarios spanning both positive and negative domains. These scenarios were based on real-life stories collected from diverse individuals and included experiences such as receiving a dream job offer, relocating to a new city, earning a promotion, navigating conflicts with friends or parents, coping with insomnia or illness, and grieving a loved one.

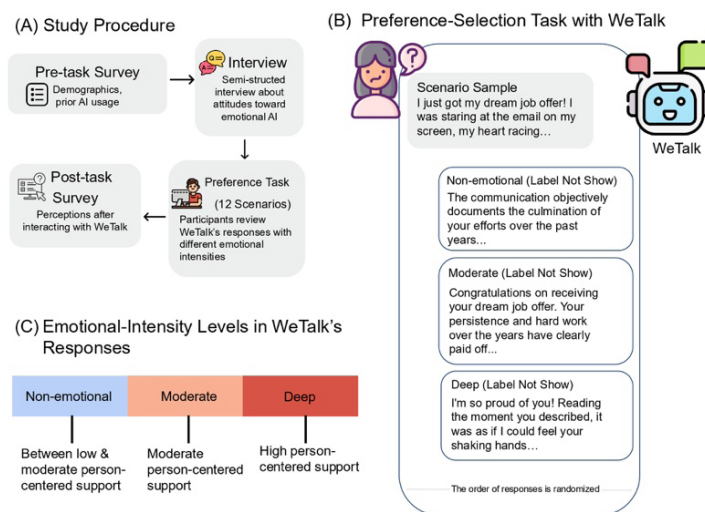


Figure 1: Overview of the study design and emotional-intensity manipulation in WeTalk.

For each scenario, we used two large language models, ChatGPT o3 and DeepSeek R1, to generate candidate chatbot responses. The prompt included the full scenario text and a rubric defining three emotional-intensity levels. Motivated by Burleson's person-centered support (PCS) framework (Burleson, 2008) and the Circumplex Model of Affect (Posner et al., 2005), we classified chatbot responses into three categories: non-emotional, moderate emotional, and deep emotional. Rather than adopting the original low, moderate, and high PCS categories verbatim, we adapted the framework to contemporary human–AI interaction contexts, in which chatbot responses are rarely designed to explicitly deny or invalidate users' emotional experiences.

Our three response styles were mapped as follows. Non-emotional responses were conceptually positioned between low and moderate PCS. They prioritized factual and situation-focused interpretation, using relatively neutral and analytical language while avoiding explicit emotional acknowledgment. Importantly, they did not challenge or invalidate users' feelings. Moderate emotional responses aligned with moderate PCS by maintaining clarity and objectivity while including a limited emotional tone, such as brief sympathy

or emotionally relevant framing. Deep emotional responses corresponded to high PCS by expressing stronger empathy, perspective-taking, and emotional engagement with the situation.

We then curated and standardized the model outputs to ensure clarity and comparability across intensity levels. The three responses within each scenario were designed to vary primarily in emotional style, while keeping response length, overall factual grounding, and advice density as comparable as possible. During the study, participants were presented with the scenario text and three pre-written replies shown in randomized order. They were not informed which model produced each reply, and the interface did not allow open-ended dialogue. This design reduced reputational bias and ensured that preference judgments were based on the perceived tone of the responses rather than model identity. At the beginning of the study, participants were also informed that any emotional responses from the chatbot were algorithmically generated rather than genuine empathy.

After completing the task, all participants were asked whether the three response styles formed a clear gradient from non-emotional to moderate to deep. All participants reported perceiving the intended ordering. Because this check was collected post hoc and relied on self-report, we report it as supportive validation of the stimulus design rather than as a formal manipulation check.

Measures and Analysis. The study employed both quantitative and qualitative measures to examine how preferences for chatbot emotional responses varied across scenario contexts, and whether these patterns differed by gender. Quantitative data consisted of participants' selections among the three response styles across the 12 scenarios. For exploratory analysis, we treated these response styles as an ordered approximation of emotional intensity and analyzed the data using Aligned Rank Transform (ART) ANOVA (Wobbrock et al., 2011), testing the effects of Scenario, Gender, and their interaction. The pre-task survey captured demographic information, AI usage habits, and familiarity with emotional AI, while the post-task survey assessed perceived emotional impact, authenticity of responses, trust in AI, and ethical concerns.

We analyzed the qualitative data using thematic analysis (Clarke and Braun, 2014). To assess coding reliability, two researchers independently coded the same interview responses using a multi-label coding scheme, since participants' responses often reflected multiple co-occurring themes. We calculated Krippendorff's α on a nominal scale and obtained $\alpha = 0.88$, indicating satisfactory agreement (Marzi et al., 2024). Remaining disagreements were resolved through discussion, after which the primary author coded the full dataset.

RESULTS

Our findings suggest that preferences for emotionally adaptive chatbot responses were strongly context-dependent across the 12 scenarios. Although deep emotional responses were frequently selected overall, the preferred level of emotional intensity varied by situation. In contrast, we did not observe

robust gender effects, indicating that variation in emotional preference was driven primarily by scenario rather than by gender.

Baseline Attitudes Toward Emotional AI. Participants entered the study with substantial prior experience using chatbots and generally positive expectations about their emotional capabilities. In the pre-task survey, moderate emotionality was the most frequently preferred response style (26/51), followed by deep emotionality (19/51), with only a small minority favoring non-emotional responses (6/51). Moderate responses were often described as balancing empathy and objectivity, whereas deep responses were seen as more immersive but sometimes excessive.

Participants were also broadly open to the possibility of emotional connection with AI. In the pre-task survey, 38 of 51 participants expressed some level of agreement that they could form an emotional connection with a chatbot, while 7 disagreed and 5 remained neutral. Interview responses showed a similarly mixed stance. Participants reported using AI for information search and content generation (34/51), academic support (12/51), everyday personal tasks (13/51), coding assistance (4/51), business and organizational analysis (2/51), and emotional support (12/51). A majority (36/51) said they would consult a chatbot during emotional outbursts because of its immediacy and non-judgmental presence, while 8 preferred human support and 2 said it would depend on the situation.

Most participants (46/51) recognized that AI emotions are simulated, yet this did not preclude meaningful effects. Over half (29/51) felt that empathetic AI made them more empathetic toward others, while 12/51 reported becoming more indifferent and 9/51 reported mixed effects. Looking ahead, participants anticipated both benefits and risks for society. Perceived benefits included efficiency and productivity (22/51), mental and emotional support (12/51), and improved communication and information access (9/51). Reported concerns centered on societal or relational transformation (18/51), emotional dependence (7/51), privacy (5/51), and fraud or security risks (3/51). Taken together, these baseline findings suggest that participants were receptive to emotionally adaptive chatbots while remaining cautious about authenticity, dependency, and broader social consequences.

Scenario-Based Variations in Emotional Response Preferences. To account for the ordinal structure of emotional preference ratings and the repeated-measures design, we conducted an Aligned Rank Transform (ART) ANOVA, treating emotional preference as an ordered outcome (Non-emotional = 1, Moderate = 2, Deep = 3). The analysis revealed a significant main effect of Scenario on emotional preference, $F(11, 539) = 2.75, p = .002$, indicating that preferred levels of emotional intensity varied systematically across situations. In contrast, the main effect of Gender was not significant, $F(1, 49) = 2.08, p = .156$, nor was the Gender \times Scenario interaction, $F(11, 539) = 1.50, p = .126$. These results suggest that emotional response preferences were strongly context-dependent, while the overall pattern of variation was broadly similar across genders.

Figure 2 presents descriptive preference distributions across scenarios. These patterns are intended to illustrate general tendencies rather than scenario-level

inferential contrasts, as post hoc pairwise comparisons between scenarios did not remain significant after Holm correction. Descriptively, achievement and transition scenarios such as Promotion to Senior Designer, Dream Job Offer, and New Apartment, New Life more often elicited preferences for deep emotional responses. In reflective contexts such as Regret Over Past Mistakes, participants more often selected moderate emotional support. Patterns also differed across negatively valenced scenarios: The Death of a Loved One more often elicited deep and moderate selections, whereas Breakup After Three Years showed relatively more moderate and non-emotional preferences. Work–Life Imbalance showed a more heterogeneous pattern, with participants distributed across emotional styles rather than converging on a single preferred response type.

Satisfaction and Preference Shifts After Exposure. In the post-task survey, participants rated non-emotional responses least favorably ($M = 4.96$, $SD = 2.41$), often describing them as practical but emotionally insufficient.

	Deep emotional connection	Moderate emotional	Non-emotional
Breakup after three years	15	23	13
The death of a loved one	29	18	4
The pressure of lower back pain	24	21	6
Dream job offer	27	19	5
New apartment, new life	29	17	5
Promotion to senior designer	30	12	9
Communication with parents	24	21	6
Sleep deprivation concerns	18	25	8
Work-life imbalance	29	9	13
Confusion about not replying to messages	24	23	4
Product launch failure	33	12	6
Regret over past mistakes	16	29	6

■ Managing grief & health
■ Achievement and milestone contexts
■ Conflicting coping styles
■ Reflective contexts

Figure 2: Total preference counts across the 12 scenarios.

Both emotional response styles were rated more positively. Moderate emotional responses received the highest and most consistent ratings ($M = 7.98$, $SD = 1.32$), whereas deep emotional responses showed a similar mean ($M = 7.96$, $SD = 2.18$) but greater variability, suggesting stronger polarization. A Friedman test indicated significant differences across the three response styles, $\chi^2(2) = 39.99$, $p < .001$. Post hoc Wilcoxon tests showed that both moderate and deep responses were rated significantly more satisfying than non-emotional responses (both $p < .001$), whereas moderate and deep did not differ significantly from each other ($\text{padj} > .99$). Participants also reported a stronger perceived potential for emotional connection with chatbots after completing the study: agreement increased from 38 to 44 participants, while disagreement dropped from 7 to 5 and neutral responses from 6 to 2. Pre- and post-task comparisons further revealed both stability and change in preferred response style. Most participants maintained their original preference, especially for moderate (18/51) and deep (10/51) responses. Fourteen participants shifted toward deeper emotional responses, whereas nine shifted from deep to moderate, often citing concerns that highly emotional replies could feel exaggerated or artificial. No participants

shifted toward non-emotional responses. The Stuart–Maxwell test indicated a significant overall shift in the distribution of preferences across the study, $\chi^2(2) = 6.05, p = .048$.

DISCUSSION

Our results suggest that preferred emotional intensity varies systematically by scenario, whereas gender does not operate as a stable determinant of emotional preference. The ART ANOVA revealed a significant main effect of scenario, but neither a significant main effect of gender nor a Gender \times Scenario interaction. Taken together, these findings suggest that situational meaning plays a more important role than demographic grouping in shaping expectations for chatbot emotional engagement. For design, this means that emotional adaptivity should be grounded primarily in what is happening to the user, rather than in assumptions about who the user is.

System Design Guidance for Emotionally Adaptive Chatbots. For industry practice, our findings support a scenario-conditioned design strategy rather than a user-profile-driven one. Emotional intensity should be treated as a controllable system parameter shaped by situational context, interaction history, and user feedback. Systems should avoid fixed gender-based emotion policies, as our results do not support using gender as a default basis for emotional style. Instead, they should adapt tone to scenario meaning: achievement and milestone contexts more often favored deeper emotional validation, whereas reflective or problem-solving contexts more often favored moderate emotional engagement. Moderate emotionality is therefore a practical default, as it was rated highly and consistently, while deep emotional responses were more polarizing. Deep emotional language may be beneficial in some scenarios, but it should be used more selectively. Users' preferences may shift with repeated exposure, as shown by the changes between pre- and post-task preferences. A useful deployment strategy would therefore combine an initial scenario-based default with lightweight calibration from user behavior, explicit feedback, or repeated interaction patterns. In system terms, emotional style should be updated as part of an ongoing preference model rather than stored as a fixed profile attribute. Finally, systems should make emotional control visible to users and include safeguards in sensitive contexts, especially where stronger emotional language may risk overreach or emotional dependency.

Ethical Reflections on Emotional Adaptivity. While emotionally adaptive chatbots may offer more tailored and satisfying support, they also raise ethical concerns that extend beyond usability. Adjusting emotional tone in response to situational cues may increase the risk of emotional overreach, perceived manipulation, or user discomfort, especially when users are vulnerable. Participants' concerns about authenticity and emotional dependence suggest that emotional design should support users without encouraging confusion about the nature of the system. These findings also reinforce the need to avoid demographic shortcuts. Since gender did not provide robust explanatory value in our study, using it to guide emotional expression would introduce ethical risk without clear design benefit. A more responsible approach is

to prioritize contextual interpretation, user control, and transparency over demographic inference.

Limitations and Future Directions. This study used a single-session design with curated, non-interactive responses. That structure helped isolate emotional intensity, but it does not capture the full complexity of real-world chatbot use, where emotional preferences may evolve over longer periods and across multi-turn interaction. Future work could extend this study in more interactive and longitudinal settings, examine how emotional preferences shift over time, and test how scenario-based emotional adaptation performs in deployed systems.

CONCLUSION

Across 12 everyday scenarios, preferences for chatbot emotional intensity were strongly context dependent. Moderate responses produced the most consistent satisfaction, while deeper emotional responses were preferred in some situations but were also more polarizing. No robust gender effects were observed. These findings suggest that emotionally adaptive chatbots should rely primarily on scenario context rather than fixed demographic assumptions. Such systems should support adjustable and transparent emotional expression while guarding against overreach and emotional dependence. Overall, this study offers preliminary empirical evidence and practical design guidance for context-sensitive emotional adaptivity in conversational AI.

REFERENCES

- Bakir, V. Laffer, A. McStay, A. Miranda, D. Urquhart, L. (2024) On Manipulation by Emotional AI: UK Adults' Views and Governance Implications, *Frontiers in Sociology*, Volume 9, 1339834.
- Binali, H. Wu, C. Potdar, V. (2010) Computational Approaches for Emotion Detection in Text, proceedings of the 4th IEEE International Conference on Digital Ecosystems and Technologies, IEEE, pp. 172–177.
- Burleson, B.R. (2003) The Experience and Effects of Emotional Support: What the Study of Cultural and Gender Differences Can Tell Us About Close Relationships, Emotion, and Interpersonal Communication, *Personal Relationships*, Volume 10, pp. 1–23.
- Burleson, B.R. (2008) What Counts as Effective Emotional Support? Explorations of Individual and Situational Differences, in *Studies in Applied Interpersonal Communication*, SAGE Publications, Inc., pp. 207–228.
- Chaplin, T.M. Aldao, A. (2013) Gender Differences in Emotion Expression in Children: A Meta-Analytic Review, *Psychological Bulletin*, Volume 139, 735.
- Christov-Moore, L. Simpson, E.A. Coudé, G. Grigaityte, K. Iacoboni, M. Ferrari, P.F. (2014) Empathy: Gender Effects in Brain and Behavior, *Neuroscience & Biobehavioral Reviews*, Volume 46, pp. 604–627.
- Clarke, V. Braun, V. (2014) Thematic Analysis, in *Encyclopedia of Critical Psychology*, Springer, pp. 1947–1952.
- De Freitas, J. Oğuz-Uğuralp, Z. Kaan-Uğuralp, A. (2025) Emotional Manipulation by AI Companions, arXiv preprint arXiv:2508.19258.
- Ekman, P. (1992) An Argument for Basic Emotions, *Cognition & Emotion*, Volume 6, pp. 169–200.

- Fehr, B. Russell, J.A. (1984) Concept of Emotion Viewed from a Prototype Perspective, *Journal of Experimental Psychology: General*, Volume 113, 464.
- Gable, S.L. Reis, H.T. (2010) Good News! Capitalizing on Positive Events in an Interpersonal Context, in *Advances in Experimental Social Psychology*, Elsevier, Volume 42, pp. 195–257.
- Ghandeharioun, A. McDuff, D. Czerwinski, M. Rowan, K. (2019) Emma: An Emotion-Aware Wellbeing Chatbot, proceedings of the 2019 8th International Conference on Affective Computing and Intelligent Interaction (ACII), IEEE, pp. 1–7.
- Glikson, E. Woolley, A.W. (2020) Human Trust in Artificial Intelligence: Review of Empirical Research, *Academy of Management Annals*, Volume 14, pp. 627–660.
- Goldsmith, D.J. (2004) *Communicating Social Support*, Cambridge University Press.
- Gross, J.J. (1998) Antecedent- and Response-Focused Emotion Regulation: Divergent Consequences for Experience, Expression, and Physiology, *Journal of Personality and Social Psychology*, Volume 74, 224.
- Herath, R. Emotionally Intelligent Chatbots in Mental Health: A Review of Psychological, Ethical, and Developmental Impacts, *International Journal of Computer Applications* 975, 8887.
- Jaiswal, A. Raju, A.K. Deb, S. (2020) Facial Emotion Detection Using Deep Learning, proceedings of the 2020 International Conference for Emerging Technology (INCET), IEEE, pp. 1–5.
- Jiang, T. Huang, C. Xu, Y. Zheng, H. (2025) Cognitive vs. Emotional Empathy: Exploring Their Impact on User Outcomes in Health-Assistant Chatbots, *Behaviour & Information Technology*, pp. 1–16.
- Kratzwald, B. Ilić, S. Kraus, M. Feuerriegel, S. Prendinger, H. (2018) Deep Learning for Affective Computing: Text-Based Emotion Recognition in Decision Support, *Decision Support Systems*, Volume 115, pp. 24–35.
- Lazarus, R.S. (1984) *Stress, Appraisal, and Coping*, Springer, Volume 445.
- Lee, J.D. See, K.A. (2004) Trust in Automation: Designing for Appropriate Reliance, *Human Factors*, Volume 46, pp. 50–80.
- Liao, T. Yan, B. (2022) Are You Feeling Happy? The Effect of Emotions on People's Interaction Experience Toward Empathetic Chatbots, proceedings of the International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, American Society of Mechanical Engineers, p. V03BT03A022.
- Liu, B. Sundar, S.S. (2018) Should Machines Express Sympathy and Empathy? Experiments with a Health Advice Chatbot, *Cyberpsychology, Behavior, and Social Networking*, Volume 21, pp. 625–636.
- Liu, T. Giorgi, S. Aich, A. Lahnala, A. Curtis, B. Ungar, L. Sedoc, J. (2025) The Illusion of Empathy: How AI Chatbots Shape Conversation Perception, proceedings of the AAAI Conference on Artificial Intelligence, pp. 14327–14335.
- Lumen Learning. *The Coping Process*. Health Psychology (SUNY HVCC) Website: <https://courses.lumenlearning.com/suny-hvcc-healthpsychology/chapter/the-coping-process/>
- MacGeorge, E.L. Feng, B. Burleson, B.R. (2011) Supportive Communication, *Handbook of Interpersonal Communication*, Volume 4, pp. 317–354.
- Majumder, N. Poria, S. Hazarika, D. Mihalcea, R. Gelbukh, A. Cambria, E. (2019) DialogueRNN: An Attentive RNN for Emotion Detection in Conversations, proceedings of the AAAI Conference on Artificial Intelligence, pp. 6818–6825.

- Marzi, G. Balzano, M. Marchiori, D. (2024) K-Alpha Calculator – Krippendorff's Alpha Calculator: A User-Friendly Tool for Computing Krippendorff's Alpha Inter-Rater Reliability Coefficient, *MethodsX*, Volume 12, 102545.
- Posner, J. Russell, J.A. Peterson, B.S. (2005) The Circumplex Model of Affect: An Integrative Approach to Affective Neuroscience, *Cognitive Development, and Psychopathology, Development and Psychopathology*, Volume 17, pp. 715–734.
- Scherer, K.R. (2009) Emotions Are Emergent Processes: They Require a Dynamic Computational Architecture, *Philosophical Transactions of the Royal Society B: Biological Sciences*, Volume 364, pp. 3459–3474.
- Seitz, L. (2024) Artificial Empathy in Healthcare Chatbots: Does It Feel Authentic? *Computers in Human Behavior: Artificial Humans*, Volume 2, 100067.
- Shaver, P. Schwartz, J. Kirson, D. O'Connor, C. (1987) Emotion Knowledge: Further Exploration of a Prototype Approach, *Journal of Personality and Social Psychology*, Volume 52, 1061.
- Svikhnushina, E. Pu, P. (2020) Social and Emotional Etiquette of Chatbots: A Qualitative Approach to Understanding User Needs and Expectations, *arXiv preprint arXiv:2006.13883*.
- Wobbrock, J.O. Findlater, L. Gergle, D. Higgins, J.J. (2011) The Aligned Rank Transform for Nonparametric Factorial Analyses Using Only ANOVA Procedures, *proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 143–146.
- Wu, H. Strelnikov, K. Montag, C. Potential Dangers of Emotional AI Agents in Society, Available at SSRN 5397038.