

Relational Norms for Human-AI Interaction: Some Preliminary Empirical Findings

Madeline G. Reinecke, Andreas Kappes, Ilina Singh, Julian Savulescu, & Brian D. Earp

Keywords: relationships, norms, cognitive science

Abstract

Artificial intelligence (AI) systems are increasingly filling social roles traditionally filled by humans, raising questions about whether people apply the same relational norms to human-AI interactions as they do to human-human ones. To investigate this, we surveyed a gender-representative sample of U.S. adults ($N = 432$) on their expectations regarding four categories of relational norms—care, hierarchy, transaction, and mating—across seven types of relationships (i.e., fellow workers, supervisor-assistant, teacher-student, mental health provider-patient, customer-seller, long-term romantic partners, close friends). Depending on condition, we described these relationships as consisting of either of two humans or one human and a superintelligent AI. Participants then rated the extent to which each party in the relationship should—or should not—behave in a manner consistent with care, hierarchy, transaction, and mating.

Our results revealed a complex picture: While norms related to care and mating were strongly endorsed in certain kinds of human-human relationships (e.g., between romantic partners or close friends), people often resisted applying these same norms to analogous human-AI interactions. Hierarchical norms were likewise applied differently to human-AI contexts, with participants emphasizing that humans should maintain equivalent or greater authority than AI in nearly all relationship types (e.g., such that a human student should have equivalent authority to an AI teacher). These and other findings, described below, suggest that relational norms developed for human-human interactions may not transfer to human-AI relationships, highlighting the need for behavioral science to inform the design of socially acceptable AI systems.

Background: The integration of AI into the human social world—via chatbots like Replika or Xiaoice—has transformed how individuals interact with machines. These AI systems can mimic roles traditionally reserved for humans, such as personal assistants, therapists, or even romantic partners. However, the psychological and normative dimensions of these relationships remain underexplored. Do people expect to treat AI "partners" the same way they treat human partners? Can norms developed for human relationships adequately capture human-AI social dynamics?

Prior research in psychology identifies four key relational functions—care, hierarchy, transaction, and mating—as central to coordinating human-human interactions (Earp, 2021; Earp et al., 2021). These norms emerged across human societies to resolve recurrent coordination challenges faced by our species (Curry et al., 2019); but they may not seamlessly extend to interactions with AI, which lack the same biological and psychological underpinnings. Consider the following “translation” problems, highlighting key differences between humans and AI that might impact the application of each relational norm:

1. **Care:** Norms of care involve non-contingent support for another’s welfare needs, as seen in parent-child or close friendship dynamics. But AI systems do not have welfare needs on most accounts. In human-AI contexts, might participants see care as unidirectional, with the expectation that AIs should behave in a caring manner toward humans but not vice versa?
2. **Hierarchy:** Human relationships often involve hierarchies, as in manager-employee or teacher-student interactions. Yet there is reason to think humans may be resistant to ceding any type of authority to an AI (e.g., due to fear, Cugurullo &

Relational Norms for Human-AI Interaction: Some Preliminary Empirical Findings

Acheampong, 2024), even if the latter is in a traditionally hierarchical role, such as a supervisor.

3. **Transaction:** Transactional norms, based on reciprocity and exchange, are central to peer relationships. Human-AI relationships, however, are often mediated by third parties, such as AI developers or companies. This transactional framing may influence how users perceive even non-transactional roles, such as AI "friends" or "romantic partners."
4. **Mating:** Romantic and sexual relationships are underpinned by norms of commitment, affection, and, in some cases, reproduction. While such norms may seem natural and appropriate within human-human relationships, might they be viewed as unnatural or inauthentic in human-AI romantic interactions?

In this work, we empirically investigate whether participants apply relational norms from human-human relationships to human-AI relationships, providing a foundation for future work on human-AI relational norms in both psychology and ethics.

Methodology: Participants were asked to imagine a world 10 years into the future, in "a world that... includes artificially intelligent (AI) agents." We prompted them to evaluate relational norms across seven kinds of relationships, from which we could compare human-human to human-AI dynamics. Additional measures assessed participants' fears about AI assuming social roles, their perceptions of the "realness" of AI relationships, and their attributions of sentience and autonomy to AI systems.

Key Findings and Implications: On the whole, human-AI relational norms varied systematically from human-human relational norms. There was no single category of norms (e.g., hierarchy) that was entirely aligned or misaligned between human-human and human-AI cases. Even in the domain of mating, where norms varied most widely between human-human and human-AI relationships, we observed similarity in some instances (e.g., in denying romantic behavior as appropriate for the teacher-student relationship; $K-S = .08$, Bonferroni-corrected $p = 1.00$). Moreover, no single human-AI relationship entirely aligned (or misaligned) with its human-human equivalent. Even for romantic partnerships, where participants distinguished most strongly between human-human and human-AI pairs (e.g., in terms of how appropriate romantic behavior would be), there was still marked similarity in some areas (e.g., transaction, $K-S = 0.15$, Bonferroni-corrected $p = 1.00$). Overall, however, we observed a trend towards resisting care and mating in human-AI relationships, which rely on mutual emotional investment. This suggests a potential barrier to the social acceptance of AI in intimate or emotionally significant roles. Next, we also observed a tendency for participants to endorse norms that would maintain human authority over AI — even in collaborative or peer-like relationships, and even in relationships in which the human occupied the traditionally subordinate role (e.g., a human student and an AI teacher). Again, this suggests that human-human norms may not successfully transfer to human-AI relationships.

Ethical Design Considerations: As AI increasingly occupies social roles, understanding how relational norms translate—or fail to translate—can inform the development of AI systems that align with human values and expectations.

Conclusion: Our findings underscore the complexities of applying human relational norms to human-AI interaction. As AI continues to evolve, it is imperative to consider these dynamics to foster ethical, socially appropriate integration into human society.

References

Earp, B. D. (2021). *Relational Morality* (Doctoral dissertation, Yale University).

Earp, B. D., McLoughlin, K. L., Monrad, J. T., Clark, M. S., & Crockett, M. J. (2021). How social relationships shape moral wrongness judgments. *Nature Communications*, 12(1), 5776.

Relational Norms for Human-AI Interaction: Some Preliminary Empirical Findings

Cugurullo, F., & Acheampong, R. A. (2024). Fear of AI: an inquiry into the adoption of autonomous cars in spite of fear, and a theoretical framework for the study of artificial intelligence technology acceptance. *AI & Society*, 39(4), 1569-1584.

Curry, O. S., Mullins, D. A., & Whitehouse, H. (2019). Is it good to cooperate? Testing the theory of morality-as-cooperation in 60 societies. *Current Anthropology*, 60(1), 47-69.