

HUMAN-ROBOT INTERACTION

Teaching robots the art of human social synchrony

Rachael E. Jack*

Humanoid robots can now learn the art of social synchrony using neural networks.

Copyright © 2024 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works

In the digital age, social robots are fast becoming part of mainstream society, from training doctors and educating children to providing talk therapies and customer service. Now a global multibillion-dollar industry, the increasing demand for robots with human-like social intelligence marks a significant milestone in our technological history. Thanks to rapid developments in artificial intelligence (AI), robots—once primarily confined to dull, dirty, and dangerous work, such as stocking shelves, cleaning floors, and deactivating bombs—are now elevated to join the human social world, with immense transformative potential for society (1). Furthermore, as human digital twins, such robots can serve as invaluable tools for scientific inquiry, enabling researchers to simulate, study, and better understand complex human social behaviors and cognitive processes (2).

However, if you have interacted with a social robot, you probably quickly realized the limits of this potential. Although their physical appearance is increasingly impressive, their interactions are often clunky, stilted, and awkward. Something feels off. Why? One critical limitation is that current social robots lack the art of social synchrony, where nods, smiles, gestures, and speech are carefully orchestrated across conversation partners (3). Although such turn-taking is seemingly effortless for most of us, it is a highly complex skill that involves rapidly processing others' speech, vocal tone, facial expressions, and gestures and planning precisely when and how to respond (4). Present in every language, spoken and signed, these sophisticated exchange dynamics are widely considered to be the fundamental “engine” for successful social interaction (5). Therefore, for

social robots to engage in human social interactions, such social synchrony skills are essential. Yet, as with many other human social behaviors, equipping social robots with a sophisticated human-like social intelligence is challenging (6).

In their recent paper, Hu *et al.* (7) addressed the art of social synchrony by endowing one such humanoid social robot, Emo, with it. Emo is a soft-skinned anthropomorphic facial robot that can display a wide range of nuanced facial expressions using 26 magnet-controlled facial actuators. It also has high-resolution cameras in its eye sockets to detect different types of facial expressions. Although Emo can mimic the human facial expressions that it detects, engaging in social synchrony involves a more refined planning and execution of responses. To achieve this, Hu and colleagues trained Emo using neural networks to

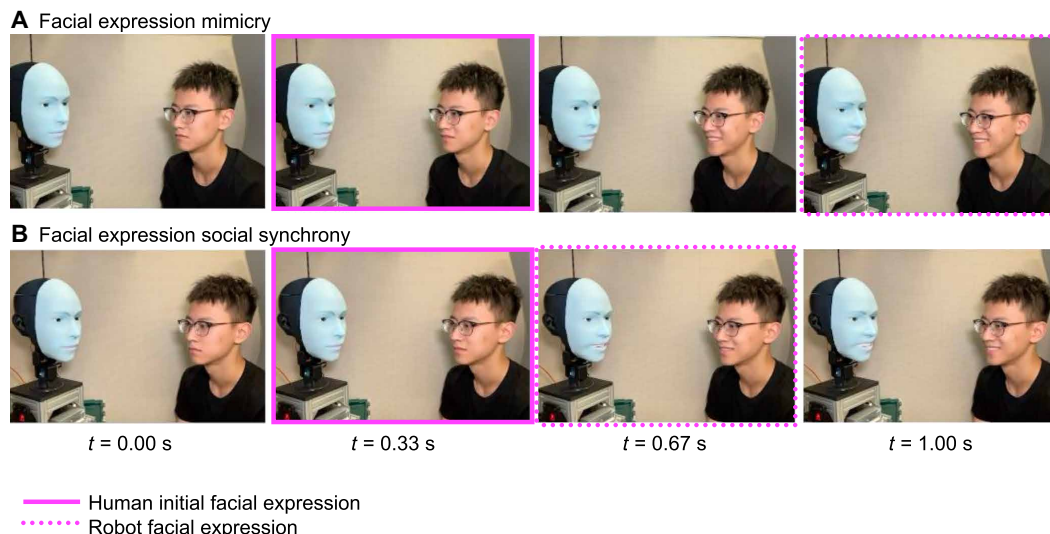


Fig. 1. Facial expression mimicry versus facial expression social synchrony in robot-human interaction. (A) In facial expression mimicry, the social robot copies the human's facial expression, which inevitably involves a noticeable delay and subsequent asynchrony. For example, at time point 3, while the human is displaying their facial expression, the robot remains neutral, creating a mismatch between the robot and human. (B) In facial expression social synchrony, the robot can predict the human's facial expression from early facial movements (see time point 2). This enables the robot to synchronize its facial expressions with the human's. For example, at time point 3, the robot and the human each display the same facial expression at the same time.

School of Psychology and Neuroscience, University of Glasgow, Glasgow, Scotland, UK.

*Corresponding author. Email: rachael.jack@glasgow.ac.uk

predict the facial expressions of its human interlocutors based on their early facial movements. For example, detecting raised corners of the mouth suggests that a happy facial expression will be displayed. Emo's predictive ability enables it to plan and execute its own facial expressions in response, achieving a more human-like social synchrony. Hu and colleagues also upgraded Emo's processing capacity to run on lightweight computing facilities, freeing up processing power for the development of other functions, such as speech and listening.

Using this simple and elegant approach, Hu and colleagues elevated Emo's social interaction skills from mere mimicry (Fig. 1A) to the art of social synchrony (Fig. 1B). Such a development has profound implications for the future of social robots. For example, even in this nascent form based solely on facial expressions, these turn-taking skills could radically improve trust and rapport building in human-robot interactions, bringing social robots one step closer to reaching their potential. Importantly, the success of Emo's new skills will depend on what its human conversation partners make of them. Do these new skills improve social exchanges with its human users, engendering trust and empathy? Or is something still off? If so, what is it? Given that Emo's facial expressions can be precisely controlled, a fruitful approach to addressing this question would be to use data-driven social psychophysics

methods from the human behavioral sciences (8). Specifically, different features of Emo's expressions, such as the exact timing of when they start or their specific facial movement components, would be experimentally manipulated to test how they influence human user behaviors, such as how much participants trust or engage with Emo. By pinpointing the specific features that facilitate or hinder human user engagement, such feedback could then be used to improve Emo's social communication skills for both general use and bespoke applications, including cross-cultural interactions (9), thereby enhancing its utility, accessibility, and marketability (10).

Last, human social interactions are inherently multimodal, involving complex combinations of visual and auditory signals, such as nodding, "uhms" and "ahhs," raised eyebrows, averted gazes, long blinks, and hand gestures. Here, Hu and colleagues focus on a single modality—facial expressions—but their results pave the way to developing social synchrony skills with more complex multimodal signals. Such a feat, although a complex interdisciplinary endeavor, could truly enable social robots to join the human social world.

REFERENCES AND NOTES

1. B. Lugin, C. Pelachaud, D. Traum, *The Handbook on Socially Interactive Agents: 20 Years of Research on*

Embodied Conversational Agents, Intelligent Virtual Agents, and Social Robotics Volume 2: Interactivity, Platforms, Application (ACM, 2022).

2. F. Barachini, C. Stary, *From Digital Twins to Digital Selves and Beyond: Engineering and Social Models for a Trans-Humanist World* (Springer Nature, 2022).
3. S. C. Levinson, Turn-taking in human communication—origins and implications for language processing. *Trends Cogn. Sci.* **20**, 6–14 (2016).
4. M. J. Pickering, S. Garrod, The use of prediction to drive alignment in dialogue, in *Grounding Sociality: Neurons, Mind and Culture*, G. R. Semin, G. Echterhoff, Eds. (Psychology Press, 2011), pp. 175–191.
5. S. C. Levinson, in *Roots of Human Sociality* (Routledge, 2020), pp. 39–69.
6. C. Breazeal, K. Dautenhahn, T. Kanda, Social robotics, in *Springer Handbook of Robotics*, B. Siciliano, O. Khatib, Eds. (Springer, 2016), pp. 1935–1972.
7. Y. Hu, B. Lin, J. Wang, Y. Wang, C. Mehlman, H. Lipson, Human-robot facial coexpression. *Sci. Robot.* **9**, eadi4724 (2024).
8. R. E. Jack, P. G. Schyns, Toward a social psychophysics of face communication. *Annu. Rev. Psychol.* **68**, 269–297 (2017).
9. C. Chen, L. B. Hensel, Y. Duan, R. A. A. Ince, O. G. B. Garrod, J. Beskow, R. E. Jack, P. G. Schyns, Equipping social robots with culturally-sensitive facial expressions of emotion using data-driven methods, in *2019 14th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2019)* (IEEE, 2019), pp. 1–8.
10. A. Akinade, Y. Haile, N. Mutangana, C. Tucker, D. Vernon, Culturally competent social robots target inclusion in Africa. *Sci. Robot.* **8**, eadm7012 (2023).

Acknowledgments

Funding: I thank the European Research Council (ERC) for funding provided under the European Union's Horizon 2020 research and innovation programme (grant agreement no. 759796) awarded to R.E.J.

10.1126/scirobotics.ado5755

Teaching robots the art of human social synchrony

Rachael E. Jack

Sci. Robot. **9** (88), eado5755. DOI: 10.1126/scirobotics.ado5755

View the article online

<https://www.science.org/doi/10.1126/scirobotics.ado5755>

Permissions

<https://www.science.org/help/reprints-and-permissions>

Use of this article is subject to the [Terms of service](#)

Science Robotics (ISSN 2470-9476) is published by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. The title *Science Robotics* is a registered trademark of AAAS.

Copyright © 2024 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works