What Robots Can Teach Us About Ourselves

Cognitive and psychological researchers Emily Cross and Arvid Kappas have devoted decades to understanding how people make sense of other social agents. To someone outside of their field, the term “agent” may seem odd, but it’s no mistake.

At first, it was understanding interactions between humans, but this wording doesn’t exclude the non-living. “Somehow, machines crept in,” Kappas quipped. The increasing complexity of robots created a demand for their expertise and the creation of the field known as social robotics.

The field's main focus is to create robots that can meaningfully integrate into the nuanced and chaotic social world of humans. Cross and Kappas have shown that achieving this takes more than just good engineering and smart code. It will take a deep understanding of concepts like emotions, group behavior and language, to name a few.

Cross, a professor at ETH Zürich, and Kappas, a professor at Constructor University, published their new January paper, “Social Robotics Is Not (Just) About Machines, It Is About People: Psychology's Role in Developing Social Machines,” in Annual Review of Psychology.

However, underneath the efforts to cultivate human-robot interaction (coined HRI in the field), there is an equally interesting development — robots are teaching us about ourselves.

“There is a very romantic undertow to thinking about how humans relate to one another,” Kappas noted. While human connection feels like a uniquely special thing, it might not be as selective as we think. A 2025 Harvard Business Review Report found that ‘therapy/companionship’ was the #1 use of generative AI.

Kappas compared our ability to form social connections to tentacles moving around the environment and grabbing onto things. “I run around in the world with an evolutionarily adapted sense that there are sentient beings around me,” he said. Our tentacles grab onto a lot, from talking to our pets and treasuring family heirlooms to engaging with robots.

It’s not only that we have a drive to connect with other agents (living or not). There is also a sense that we misidentify what we seek in connection. “I think there is a real big difference between what people think they want and what they actually respond to,” Cross added. It’s normal to think we want connections that revolve around genuineness and honesty, and that the alternative is unacceptable.

Yet, the research leans toward the opposite. Whether it is a physical robot of limited capability, generative AI, or even a robot knowingly controlled by a human, people are ready to divulge personal details and put concepts like genuineness on the shelf. “They want the real deal. But when it comes down to it, the kind of ‘fake deal’ isn't quite as off-putting as people think it will be,” Cross said.

Kappas adds to this with the observation that we simply don't always want deep connection. He described a quick chat on the plane, lobbing thoughts about the Super Bowl and Bad Bunny's performance, and in the end, "it was not so important what that person thought." These little passing and pleasing moments are fundamental to social well-being. They allow us to think aloud and self-express. In the context of interesting experiments and new AI, there is an element of 'why not?" that fits neatly into our more casual tentacles.

Cross also shared the sometimes too-true fact that "people like to talk about themselves." This is something that should not be ignored when discussing motivators for connection.

Kappas thinks the greatest takeaway from what robots teach us about humans is a shift in perspective, rather than a specific finding or headline; Psychology has historically focused on humans' reactions to stimuli. And the robots of today are functional in that when they think, they compute, and when they don't think, they do nothing.

While the behaviors of robots and humans might align, there still seems to be the unshakable feeling that machines do not truly understand or feel. Kappas calls us to think about the background 'noise' of our brains and behaviors. This 'in between activity' we generate apart from notable events might not only be the fundamental difference between humans and robots, but an underappreciated element in what makes us human.