Mind Attribution: From Simple Shapes to Social Agents

Friederike Eyssel, Birte Schiffhauer, Fabio Dalla Libera, Yuichiro Yoshikawa, Jakub Złotowski, and Ricarda Wullenkord, and Hiroshi Ishiguro

Abstract— **We present an overview of recent empirical work that has examined mind attribution to simple animated shapes in Japan and Germany, respectively. In this work, we show that both Japanese and German participants differentially anthropomorphized geometric figures by endowing them with mind. Thereby, we not only replicate, but also extend classic findings on animism in nonliving entities. We discuss this notion in light of the universality of mind perception and anthropomorphization and reflect upon the impact of our findings on the perception of social agents, such as robots.**

I. INTRODUCTION

To identify dimensions of mind perception, [1] have shown that people seem to differentially attribute mind to a wide range of targets including animals (e.g., a frog), humans, and nonliving entities (e.g., god, a robot). Their findings suggest two core dimensions of mind perception, namely agency and experience. While agency entails the capability to plan and act intentionally, experience refers to the capability to experience emotions, inner life, and vitality. In previous work [1], humans were perceived as high on both dimensions. These results indicate that indeed, affect and cognition are core human capacities. However, once sociability, intentionality or other typically human qualities are attributed to nonliving entities, this could qualify as anthropomorphisation [2, 3]. Importantly, the process of attributing typically human characteristics to nonhuman entities is conceptually distinct from animism, mere lifelikeness or apparent naturalness of movement [2,3]. To illustrate, one may observe the following scenario: a big, triangle-shaped object is moving towards another small, triangle-shaped object. If one would infer from this scenario

F. Eyssel is with the Center of Excellence Cognitive Interaction Technology (CITEC) at Bielefeld University, Bielefeld, Germany (e-mail: feyssel@cit-ec.uni-bielefeld.de).

B. Schiffhauer is with CITEC, Bielefeld University, Germany (e-mail: birte.schiffhauer@cit-ec.uni-bielefeld.de).

F. Dalla Libera is with Graduate School of Engineering, Osaka University, Japan (e-mail: fabiodl@gmail.com).

Y. Yoshikawa is with Graduate School of Engineering, Osaka University, Japan, and JST ERATO Ishiguro Symbiotic Human-Robot Interaction Project, Japan (e-mail: yoshikawa@irl.sys.es.osaka-u.ac.jp)

J. Złotowski is with University of Canterbury, Christchurch, New Zealand (e-mail: jakub.zlotowski@pg.canterbury.ac.nz).

R. Wullenkord is with CITEC, Bielefeld University, Germany (e-mail: rwullenk@cit-ec.uni-bielefeld.de)

H. Ishiguro is with Graduate School of Engineering, Osaka University, and JST ERATO Ishiguro Symbiotic Human-Robot Interaction Project (e-mail: ishiguro@ams.eng.osakau.ac.jp).

that a big triangle would be *threatening* a small triangle, having negative intent to harm and the capacity to do so, we would describe this as an instance of an anthropomorphic inference. This means, in order to qualify as an instance of psychological anthropomorphism, judgments have to go beyond simple observations. By attributing mind, that is, through the process of anthropomorphization, we cognitively transform a simple a geometric figure into a social agent.

Classic research in psychology has already shown decades ago [4] that people are eager to attribute intentions, motivations, and emotions to animated geometric shapes. So far, however, evidence is based on qualitative data and,, experimental evidence is still sparse. More recently, [5] have demonstrated that preverbal 10-month-olds were displaying signs of sympathy towards animated shapes that were presented in a victim-versus-aggressor constellation, accompanied by a neutral 'bystander'. As illustrated in [5], these infants were even able to realize the assigned 'roles' of the animated figures and this determined subsequent object choice. We built on this previous work by presenting participants with the original materials used by [4] to explore baseline mind perception towards animated geometric shapes using self-reports. Specifically, we sought to examine whether participants would attribute agency and experience, as in [1] to nonliving entities, going beyond anecdotal, predominantly qualitative evidence. However, instead of exploring preverbal infants as in [5], we tested adult participants. Moreover, we replicated and extended findings cross-culturally by taking into account samples from Japan and Germany, respectively. We conducted three studies to shed light on mind perception processes as a function of target types and their respective 'behaviors' – thus, we explored what it takes to turn a nonliving entity into a 'social' agent that is perceived capable of thinking and even feeling [1,5].

II. RESEARCH GOALS

To replicate and extend previous work, we used variants of the animated clips by [4] and [5], and explored mind perception towards the two main 'protagonists' and a neutral third agent. We were particularly interested in examining mind perception in a triadic constellation, as previous HRI-related work has demonstrated the positive effects of additional robots on perceptions of social telepresence [6]. Analogously, it is plausible that the presence of a bystander agent could have an impact on the evaluation of all targets involved. By conducting basic research on mind perception in stimuli as abstract as geometric shapes, we can apply this knowledge to entities of more social nature, and learn about psychological processes involved in multi-party interactions between nonhuman agents. To explore this further, and to enrich the theoretical and empirical literature on mind

perception based on [1], [4], and [5], we conducted a set of three preliminary studies in Japan and Germany.

III. Preliminary Findings

In Study 1, participants watched a video sequence from the original clip by [4]. The video depicted animated geometric shapes (i.e., a big triangle, a small triangle, and a circle). The big triangle appeared as if it would be 'attacking' the small triangle, while the circle would take on the role of a passive 'bystander'. Japanese participants perceived the characters as equally high in experience. Unexpectedly, the big triangle was perceived as least mindful in terms of agency, followed by the small triangle, and the circle. To investigate this effect further, we conducted an extended replication in Germany. Importantly, however, in Study 2, we manipulated the endings of the video clips to disambiguate them. This was done because the video used in Study 1 could have been interpreted as showcasing an 'unhappy' ending, given the plausible interpretation that the 'attacker' could have allegedly defeated the small triangle. In Study 2, participants were thus presented with the very same clip as in Study 1 or they viewed a version with a manipulated 'happy' ending. Results conveyed that the big triangle was attributed more 'default' agency in the 'attack' video condition compared to the 'happy end' condition. This might be due to the fact that intentions to harm are attributed to an alleged 'attacker', whereas participants could interpret the 'happy end' as some sort of reconciliation between the three characters where no further negative outcomes would be expected. In that condition, the big triangle was perceived lower in experience and agency compared to the other two targets. Additionally, we explored a sub-component of agency, namely moral agency. As predicted, the big triangle was also perceived as having lesser moral agency than the circle and the small triangle. This seems plausible in light of the fact that an attacker is commonly perceived as acting immorally. Findings in this regard appeared reasonable and promising, and could be researched in future studies that investigate mind attribution to robots or other social agents. To further clarify the roles of attacker vs. victim vs. bystander, we decided to conduct a third study with adapted materials proposed by [5]. The up-to-date clips developed by these authors feature three geometric shapes that can clearly be distinguished by color, making subsequent within-subjects ratings less confusing to participants. In the control video, the blue circle chases a yellow square and crashes into it. The red cylinder moves far away from the other two figures throughout the video. Video 2 features an ending in which the red cylinder moves close to the yellow square. In Video 3, however, the red cylinder moves close to the blue circle. Since video type did not have an effect, Figure 1 depicts the results regarding mind attribution to the three colored shapes across conditions. Taken together, all three studies revealed that both Japanese and German participants were prone to attribute mind to minimalist agents, such as animated geometric shapes. In three empirical studies, participants did not differentiate

much between targets in terms of experience attribution, and effects turned out more pronounced regarding agency perception.
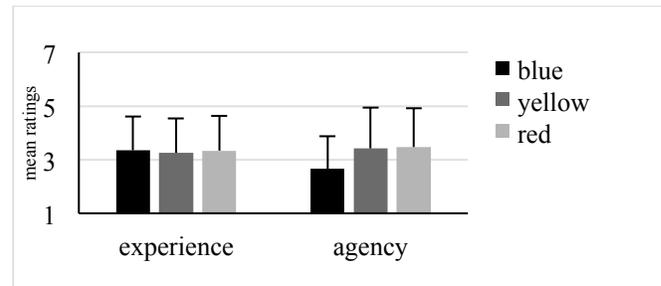

Fig. 1. Mean attribution of agency and experience to the characters in Study 3.

It is worth noting that across studies, no floor effects were observed, so that the utilized self-report scales for mind perception can likewise be recommended for future research. More research is needed to validate the moral agency subscale and to establish further the psychometric quality of all self-report measures used in this research. However, taking into account the fact that mean levels of mind perception were relatively high, even higher mean values might be expected once agents that indeed come close to actual social interactants are to be evaluated (e.g., humanoid social robots). Follow-up studies need to focus even more on the role of bystanders in social robotics, to examine further whether a bystander potentially changes evaluations of a core protagonists, as indicated by preliminary results on robot bystander effects on social telepresence experiences [6]. Further research needs to demonstrate the impact of the presence of bystander robots on anthropomorphic inferences about a given target robot of interest. Therefore, we suggest validating the present findings by examining basic robot platforms or small-scale humanoid systems like NAO (Aldebaran Robotics) or M^3Synchy2 [6] in triadic interaction constellations. Clearly, having participants experience immersion in an actual interaction with a robot would rule out problematic issues associated with video materials.

References

[1] H. M. Gray, K. Gray, and D. M. Wegner, "Dimensions of mind perception," *Science*, vol. 315, no. 5812, p. 619, 2007.

[2] C. Bartneck, E. Croft, and D. Kulic, "Measuring the anthropomorphism, animacy, likeability, perceived intelligence and safety of robots," in *Proc. of the Metrics of Human-Robot Interaction Workshop*, Technical Report 471, pp. 37-41, 2008.

[3] N. Epley, A. Waytz, and J. T. Cacioppo, "On Seeing Human: A three-factor theory of anthropomorphism," *Psychological Review*, vol. 114, pp. 864-886, 2007.

[4] F. Heider, and M. Simmel, "An experimental study of apparent behavior," *The American Journal of Psychology*, pp. 243-259, 1944.

[5] Y. Kanakogi, Y. Okumura, Y. Inoue, M. Kitazaki, M., and S. Itakura, "Rudimentary sympathy in preverbal infants: preference for others in distress," PLOS One, vol. 8, no. 6, e65292, 2013.

[6] T. Arimoto, Y.Yoshikawa, and H. Ishiguro, "Nodding responses by collective proxy robots for enhancing social telepresence", in *Proceedings of the 2nd International Conference on Human-agent Interaction* (HAI '14). ACM, New York, NY, USA, pp.97-102.