Author's copy. Appeared in: Nørskov, M., Seibt J., Quick O. 2020. Culturally Sustainable Social Robotics – Proceedings of Robophilosophy 2020. Series Frontiers of AI and Its Applications, IOS Press, Amsterdam.

Social Norms for Artificial Systems

Anna STRASSER [a,1]
[a] *Independent researcher*

Abstract. This paper investigates reasons to argue for social norms regulating our behavior towards artificial agents. By problematizing the assertion that moral agency is, in principle, a necessary prerequisite for any form of moral patiency, reasons are examined which are independent of attributing moral agency to artificial agents, but which speak for morally appropriate behavior towards artificial systems. Suggesting a consequentialist strategy, potential negative impacts of human-machine interactions are analyzed with a focus on factors that support a transfer of behavioral patterns from human-machine interactions to human-human interactions.

Keywords. Moral agency, moral patiency, social norms, social robotics

1. Introduction

Various kinds of artificial systems play a role in our social life already now. In the future, a growing number of products coming from social robotics will presumably be used in industry, transportation, healthcare, military, education, and children's toys. All these devices invade our social life, cause new behavioral patterns, and raise questions about responsibilities, duties, and rights. For example, it is a difficult question to decide how to judge people who cruelly abuse their humanoid robots. Issues about how humans should organize their co-existence with highly developed artificial systems are of great public interest.

Although there are controversial approaches concerning social norms that may regulate our behavior towards artificial systems, many positions seem to assume that artificial systems could only be treated as moral objects if they display moral agency. One could even go as far as Floridi and Sanders [1] and refer to positions taking moral agency as a necessary condition for any kind of moral patiency as standard positions. However, this paper aims to show that such a strong presupposition is problematic. Apart from the fact that we are far from agreeing on the status of artificial systems as moral agents, the well-known is/ought problem additionally questions the very necessity of concluding normative judgments from factual assertions [2]. Moreover, regarding our everyday practice, one can refer to several examples in which social norms already determine how specific agents should be treated, even though no moral agency is ascribed to them.

By taking moral agency as an indispensable precondition, discussions about social norms concerning artificial systems are postponed into a distant future until there is an agreement on the moral status of artificial systems. Up to now, positions concerning the status of artificial systems range from claims that artificial systems can only be regarded as tools [3-5] to positions arguing that artificial agents may qualify as social, moral agents [6, 7].

But even if we would agree on the status of artificial agents, considerations about the is/ought problem [2], as originally introduced by Hume [8], call the necessity of normative judgments into question. Indeed, I claim that a negative decision regarding the status as a moral agent can by no means exclude the possibility that artificial systems nevertheless qualify as objects of moral actions. In order to distinguish between cases of moral patiency, which do not necessarily presuppose moral agency, and cases where moral agency is presupposed, I refer to moral objects to describe cases in which objects are considered as worthy of moral consideration without requiring moral agency.

Instead of discussing both the is/ought problem and the infinite debates about the attribution of a certain status at length, this paper suggests a consequentialist strategy and examines reasons that are independent of factual claims about the status but speak for morally appropriate behavior towards artificial systems. For this purpose, negative impacts of human-machine interaction on human-human interaction, which one would prefer to avoid, are investigated. Starting from the assertion that if such negative transfers of behavior are likely, artificial systems should have a moral status in the sense that we should follow social norms governing our behavior towards them, human properties and characteristics of artificial systems are studied to identify factors contributing to undesired transfers of behavioral patterns learned in human-machine interactions. Based on the working hypothesis that the similarity between the two forms of interaction, in particular, increases the probability of transfers of behavioral patterns, I investigate factors contributing to this similarity. In other words, it will be shown that as soon as human-machine interactions evolve from mere tool-use to a kind of social interaction, differentiations become more complicated, and thus, the probability of transfers of behavioral patterns from one area to another increases. A high probability of transfers motivates to regulate human-machine interactions before they are transferred.

[1] Corresponding Author: Anna Strasser, Independent Researcher, Berlin, Germany; e-mail: annakatharinastrasser@gmail.com

2. How to argue for social norms

Social norms play an essential role in the organization of behavior in social groups. A social normative relation obtains when members of a group understand that they should do something because it is obligatory according to a norm-giving power. This power can be an institution with norm-giving legitimacy, a social group, or a society as a whole.

Social norms vary with respect to their normative level. For instance, that I should stop at a red light is rather a convention, whereas that I should not kill people is a moral norm. Social norms provide guidelines addressing both the active and passive side of potentially moral actions, namely the moral agent and the moral object. On the one hand, social norms determine what kind of behavior of a moral agent is acceptable. Thus, they define obligations and duties of moral agents. On the other hand, social norms also attribute rights and values to objects of action, indicating what kind of behavior towards such objects is morally appropriate.

We may agree that most moral agents are also moral objects. However, in contrast to the standard view claiming that moral agency is a necessary prerequisite for any kind of moral patiency, this paper argues for the claim that not all moral objects have to be moral agents. Already now, actual social norms concern a variety of moral objects. Among these, there are, of course, moral agents, but also other living beings and inanimate objects or specific subgroups of one of those.

It seems obvious that human societies agree on attributing values to a wide range of objects and agents. For example, even though non-human animals are often not regarded as proper moral agents, they are considered moral objects according to animal rights. Another example is our idea of nature conservation. The postulation of nature reserves explicitly demands a careful treatment of nature, such as plants or lakes. Neither lakes nor plants are regarded as moral agents, yet social norms regulate our behavior towards them.

Similarly, one can argue that certain human-machine interactions are worthy of moral consideration. In other words, one can argue for social norms regulating how to interact with artificial systems. If certain artificial systems are regarded as moral objects, this will constitute a new subset of inanimate but moral objects. Consequently, social norms regulating our behavior towards artificial agents can be introduced regardless of whether artificial agents have moral agency. To do so, one has to work out reasons why certain types of human-machine interactions require regulation.

With respect to social norms concerning artificial systems, I distinguish between norms that specify the behavior and properties of artificial systems and those that regulate how humans should behave as interaction partners of artificial systems. The former cases are not discussed in this paper. This would require to develop a framework of desirable features of artificial systems potentially involving the attribution of duties (obligations) that may be based on the claim that artificial agents are moral agents. However, such social norms can also concern the construction process and thereby determine the behavior of their creators. In this paper, I focus on the latter, namely on human behavior in human-machine interactions. Thereby, artificial systems are considered as moral objects.

Interestingly, even rules that aim at the behavior of artificial systems are often finally directed to the behavior of their creators and users. For example, the three laws on robots [9], which state that robots may not harm humans or other living beings, seem at first glance to consider robots as moral agents. However, in the end, the moral responsibility for complying with these laws lies more with the creators and users of such robots.

When it comes to the question of how humans should behave towards artificial systems, there are strong objections against attributing moral patiency to inanimate objects at all. For example, the lack of the capacity to suffer is repeatedly cited as a reason why destructive behavior towards artificial systems would be just as morally uncritical as towards other inanimate objects. In summary, it can be said that artificial systems are often denied both moral agency and moral patiency. However, this paper distinguishes between strong cases of moral patiency, in which moral agency is presupposed, from other cases where moral agency is not necessarily presupposed. The latter are cases in which objects are considered as worthy of moral consideration without requiring moral agency. This opens up the possibility to argue for considering artificial systems as moral objects without requiring the ability to suffer or moral agency.

2.1. Varieties of human-machine interactions

Of course, promoting social norms with respect to artificial systems cannot apply to all types of human-machine interactions. Many human-machine interactions are best characterized as pure tool-use and do not require any further social norms. However, in view of the rise of social robotics, there are more and more human-machine interactions that cannot that easily be reduced to mere tool-use. Especially social robots are intentionally designed as companions of humans and not as tools. In sum, it is obvious that our way of interacting with social robotics' devices is fundamentally different from the way we use other machines, such as simple calculators. Due to the increasing number of social robotics' devices, I predict that we must be prepared that a lot of human-machine interactions will not only be part of our social life but may also shape our social interactions with each other.

The special nature of our attitude towards social robotics products becomes clear when we look, for example, at social norms aiming to avoid destructive behavior towards interaction partners. At first glance, destructive behavior may be, in principle, negatively connoted by social norms, but fundamental differences are made between inanimate objects and living beings. However, it is precisely this seemingly clear line between living and inanimate objects that may become somewhat blurred by social robots and other technical devices that enter our social life. We are obviously able to distinguish between machines that are considered as mere tools from those that play a role in social human-machine interactions. The latter move closer to living creatures in our evaluation, so that we judge destructive behavior towards simple objects, such as toasters, quite differently from destructive behavior towards lifelike artificial systems [10].

2.2. The human side

The human attitude towards artificial agents plays an essential role in how human-machine interactions are evaluated. Analyzing emotional attitudes towards certain artificial agents, one can even report moral concerns on how to treat artificial agents. Studies showed that people hesitate to practice destructive behavior towards social robots [11]. It is not surprising that particularly toy-robots are capable of triggering some kind of mothering behavior. Kate Darling [10] showed that even a brief interaction with the dinosaur toy-robot called Pleo leads to moral concerns when it comes to destructive behavior. She reports that after the participants had time to interact with Pleo, they were asked to tie up, beat, and "kill" their Pleo. All participants refused to "hurt" their robots. Also, in the military sphere, where one might suspect less emotional bonds, one can refer to reports illustrating how behavioral patterns are developed that one might expect to be more likely to be found amongst humans. For example, soldiers develop relationships with their robot comrades, which has led to the fact that, for example, mine-sweeping robots have been buried with honors [12].

When interacting with certain artificial systems, humans already have a tendency to act according to social norms, which are quite similar to the ones used in human-human interactions. They do this even though they know that their counterparts are only machines. However, relying exclusively on the fact that humans tend to act according to social norms is not yet sufficient to justify social norms.

3. Impact on human-human interactions

Arguing in favor of social norms regulating human behavior in interactions with artificial agents can be motivated by the consequences behavioral patterns practiced in human-machine interactions have on human-human interactions. This idea follows a consequentialist strategy.

The ability to generalize is one of the most important cognitive abilities of humans. Humans are able to apply behavioral patterns learned in one context to many other situations. However, concerning artificial systems, this ability can lead to both desirable and disastrous transfers. For instance, transferring skills learned in flight simulators to the real world is rather advantageous for pilots. Further positively evaluated transfers can be found in therapeutic settings using Virtual Reality (VR) for the treatments of phobias, post-traumatic stress disorders, and schizophrenia. The success of such interventions is based on the fact that behavioral patterns practiced in VR are put into practice in the real world. For instance, virtual reality exposure therapy has been established as an efficient treatment of phobias [13].

However, not every transfer is welcome. There are behavioral patterns practiced in human-machine interactions for which one can only hope that they will not be transferred to human-human interactions. The following examples illustrate that the spectrum of questionable transfers can range from decreasing politeness to really cruel behavior.

For example, anecdotes about virtual assistants like ALEXA nicely illustrate potential transfers of behavioral patterns to human-human interactions. Due to frequent practice to interact with virtual assistants exclusively in a rude commanding tone, even adults may lose their polite attitude towards other people. In particular, regarding children, there are legitimate fears that they are in danger of unlearning polite expressions. This is, for example, reflected in the advertisement of a new Echo Dot Kids Edition that emphasizes that those artificial agents give the child positive reinforcement when they say "please" and "thank you."

Another example comes from a therapeutic setting and deals with the effects of interactions with sex-robots. Questions of whether sex-robots could help to normalize deviant behavior are still controversially discussed. Based on ideas of substituting actions, some positions assume a possible positive effect by arguing that sex-robots could help to reduce sexual crimes. However, especially outside the therapeutic settings, there is increasing evidence of critical consequences such interactions can have for interpersonal behavior. According to such views, it is claimed that sex robots promote the ubiquitous idea that even living women are sex objects and thus increase physical and sexual violence against women [14]. A particularly sensitive area concerns the effects of child sex-robots, so-called paedobots. Although there are no conclusive studies yet, there are indications that dealing with paedobots is likely to lower inhibition thresholds and thus put children to higher risk.

The common denominator of these examples is that, although the concrete behavior towards the artificial agent is not yet morally judged, it does take on a moral dimension in view of potential transfers. As soon as we realize that behaviors towards social robots can influence behaviors among humans, we start to think about whether these behaviors should be regulated in order to avoid undesirable transfers. As Kate Darling [10] states: "One reason why humans might want to prevent the 'abuse' of robot companions is to protect social values."

The critical point is the question of how likely it is that behavioral patterns are transferred from one domain to another. When analyzing the probability of such transfers, it is important to examine to what extent such transfers may occur automatically. If human beings could intentionally suppress transfers to human-human interactions, one would only need social norms that mark unwanted transfers as unacceptable. However, as long as we cannot exclude that behavioral patterns acquired through interactions with human-like artificial systems are particularly predestined to be reflected in human-human interactions, we have reasons to regulate such behavioral patterns before they can be transferred.

4. Factors contributing to the transfer

To identify relevant factors increasing the probability of transfers, the role of both human counterparts and artificial agents is investigated. Assuming that both sides make a significant contribution, I argue that the interplay between the characteristics of human and artificial interaction partners promotes the likeliness of transfers of behavioral patterns.

4.1. Anthropomorphizing

There is no doubt that human attitudes towards artificial systems shape how people interact with artificial systems. First and foremost, the human tendency to anthropomorphize plays an extraordinary role. It contributes significantly to the fact that human-machine interactions often appear somehow similar to social human-human interactions.

Anthropomorphizing, understood as assigning 'as-if' attributions of social characteristics, does not imply ontological claims. That means such attributions remain neutral regarding the question of what kind of implementation causes the observed behavior. Thereby, 'as-if' attributions differ from justified attributions we make concerning other human beings.

However, humans are not very selective when it comes to anthropomorphization. Indeed, not only objects with seemingly social characteristics are anthropomorphized. People interpret all kinds of non-social data as if they were social data. For example, merely moving geometric objects are described by social characteristics. This was shown by the famous experiment by Heider and Simmel [15]: Even though participants knew that geometric objects have no social characteristics, they used social narratives to describe their movements. Assigning social characteristics makes it easier to understand, describe, and anticipate the behavior of many entities in our world. Due to this explanatory value, humans are well-advised to use social ascriptions instead of relying on more technical descriptions. In other words, taking the intentional stance [16] helps to make sense of the world and improves our abilities to understand and anticipate events in the world.

However, there is something special about anthropomorphizing in human-machine interactions. Instead of being limited to an observational explicative perspective, such 'as-if' attributions of social characteristics have an impact on the way we interact with artificial systems. Although 'as-if' ascriptions make no ontological claims, they motivate us to orient ourselves on social norms in subsequent interactions. In this way, a change from an observational perspective to a more interactive perspective is achieved. Through anthropomorphization, we ascribe values to artificial systems only in an 'as-if' mode, which nevertheless allows us to regard them as social interaction partners.

Moreover, our tendency to anthropomorphize puts us in situations where we automatically switch from an observational perspective to an interacting perspective. Although our philosophical ideas may tell us that our counterparts are not really social agents, we cannot help it and react socially. Without a doubt, anthropomorphizing contributes a lot to making human-machine interactions more similar to human-human interactions. Due to this similarity, distinguishing between the two becomes more difficult, which in turn may foster possible transfers of behavioral patterns.

4.2. Characteristics of artificial systems

However, the human tendency towards anthropomorphizing cannot sufficiently explain why certain human-machine interactions are particularly easily confused with human-human interactions. An analysis of the specific characteristics of artificial systems of social robotics can show to what extent factors stemming from the artificial agents promote a higher probability of a transfer of behavioral patterns. Thereby, I focus on properties that contribute to a similarity with human-machine interactions and thus increase the difficulty of distinguishing between artificial and human interaction partners. In short, if we cannot take clearly apart humans from artificial interaction partners, the probability of transferring behavioral patterns increases. How challenging it can be to distinguish between reality and fantasy becomes obvious with respect to experiences one can make in virtual reality (VR).

Taking into account that the idea of social robotics is to design artificial agents that enter the human space of social interactions, the similarity to human-human interactions is a desired objective. In addition, specific human-human interactions serve as models for many human-machine interactions; thus, we are dealing with a duplication of already known interaction types.

Since reciprocal exchanges of social cues constitute a distinguishing feature of human-human interactions, the ability to process social cues is of particularly high relevance for social robotics. Consequently, artificial agents are designed to interpret social cues presented by their human counterparts as well as to send social clues back to make their ‚minds' visible.

To succeed in serving as a social interaction partner, artificial systems need specific detection systems, reasoning mechanisms, and abilities to express social cues. Detection systems enable artificial systems, for example, to recognize social cues, such as emotional states. The reasoning mechanisms make sure that social cues are interpreted, and appropriate responses are selected. Last but not least, the ability to express social cues in a way that can be interpreted by humans can establish a reciprocal exchange of social cues. Concerning affective expressions, Höök [17] defined an interactive process as an affective loop in which the human user first expresses her emotions; in a second step, the system responds with appropriate affective expressions, which again affects the user.

The ability to process social cues is an essential factor in social interactions. Although not yet unified in one system, one can refer to various abilities of artificial agents in this respect. Regarding detection systems, for example, one can point to deep-learning emotion-detection networks [18]. Interpretation abilities can be found in the so-called artificial retrieval of information assistants (ARIAS), which are able to handle multimodal social interactions by processing verbal and non-verbal social cues [19].

A special feature of social interactions, which is also based on the processing of social cues, is the ability to anticipate future actions of others. For example, joint actions are based on successful mindreading and intention-reading. This could be a particular challenge for artificial systems. However, it has been shown that artificial systems can also achieve certain forms of mindreading. For example, artificial agents can model human mental states in relation to the perspective of the human counterpart [20]. These artificial agents cannot only infer from their perception of the physical world to what a human counterpart can or cannot see, but they can also take into account that the fact of whether human agents can see an object will guide their future actions. In this respect, it can be argued that artificial agents are successful in some cases of mindreading.

Intentions are rather unobservable; they are normally derived from the context. However, there are also perceivable factors that contribute to successful attributions of intentions. For example, several studies showed that humans make use of information about movement kinematics to predict future actions [21, 22]. Of particular interest concerning artificial systems is a study by Cavallo and colleagues [23]. This study shows that humans are able to predict whether an observed agent grasps a bottle with the intention of either pouring water into a glass or drinking water from the bottle. Within the framework of this study, a classification and regression tree model (CART) was trained and tested. This CART model used information about the kinematic movement to predict the participant's intentions. Humans may not recognize that they implicitly use information about kinematic movement, but they would probably remark if this kind of information is not processed.

The ability to process social cues contributes to an increased similarity of human-machine interactions with human-human interactions. Of course, one has to admit that there is not only an immense richness of social cues used by humans but also that this richness is so far largely unexplored. Nevertheless, it seems as if even the ability to process some social cues can already make interactions with artificial systems confusing similar to social interactions among humans. Apart from the fact that increasing similarities make the transmission of behavioral patterns more likely, such characteristics trigger, in particular, our tendency to anthropomorphize, which in turn contributes to an even greater similarity.

4.3. Automaticity

In addition to the ability to generalize behavior, humans can also recognize situations in which transfers are not welcome. For example, many behavioral rules in human societies are specific to distinct social roles. It seems natural that people are able to behave according to specific social roles of their counterparts. Typically, family members are treated differently from work colleagues. This means that learned patterns of behavior, such as how to greet someone, are not generally applied to all greeting situations. Pronounced characteristics help people in a human society not to use inappropriate behavioral patterns towards particular role holders. However, it is important to note that people usually do not pretend that their role is confusingly similar to another social role. In contrast, artificial agents are designed to appear as similar as possible to human counterparts.

Because of their built-in similarities, artificial systems behave as if they were human counterparts. There are no pronounced characteristics that would support fine-grained differentiations. Therefore, certain human-machine interactions are confusing similar to already well-known human-human interactions. As described above, this similarity is a result of the fact that human-machine interactions are designed to exhibit social features we know from human-human interactions. This kind of intended similarity makes it particularly challenging to distinguish human and non-human addressees in interactions clearly. Furthermore, artificial systems are constructed to trigger our tendency to anthropomorphize to an extraordinary degree. I, therefore, assume that possible mix-ups are almost pre-programmed. Nevertheless, one may ask whether unwelcome transfers of behavioral patterns practiced in human-machine interactions could be reduced once artificial systems are given a specific, distinguishable social role. But to this day, no clear role has been attributed to artificial systems. Therefore, we are confronted with both positive and negative transfer effects.

5. Conclusion

The above reflections about undesirable impacts behavior towards artificial agents may have on human-human interactions motivate arguing for social norms regulating interactions with artificial agents. Consequently, arguing for social norms is dependent on the presupposition that behavioral patterns are often transferred. Nevertheless, it would be a rather radical claim to demand that social norms comparable to those governing human-human interactions should be applied to all kinds of human-machine interactions that exhibit a minimum of social characteristics. I assume that this would go too far. With regard to social norms concerning politeness and other less dramatic topics, one can probably accept the risk that people develop bad habits through frequent interactions with artificial social systems. Moreover, one may even hope that people could learn to suppress unwanted transfers in the long run, despite their tendency to anthropomorphize and the quasi-automatic reactions to social cues. Furthermore, future research in social robotics could also contribute to making differentiation between artificial and human interaction partners easier by developing pronounced characteristics.

However, taking severe violations of social norms into account, even a just probable negative impact becomes decisive. Therefore, I suggest establishing social norms that exclude any kind of behavior towards artificial agents that would be classified as downright criminal when acted out with another human. If we treat artificial systems along with social norms that prevent abuse and other cruelties, we can at least avoid such unwelcome transfers.

On the one hand, this limited proposal goes too far for those who argue that machines should always be treated as rightless tools due to the lack of moral agency. On the other hand, this proposal may go not far enough for others since it focuses exclusively on the consequences for humans and does not discuss arguments as to whether artificial systems deserve to be treated according to social standards. Nevertheless, this proposal can serve as a promising starting point for future research.

References

[1] Floridi L, Sanders J. On the Morality of Artificial Agents. Minds and Machine 2004 Aug;14(3):349-79. doi:10.1023/B:MIND.0000035461.63578.9d

[2] Gunkel, D. The other question: can and should robots have rights? Ethics and Information Technology 2018 Jun;20(2):87-99. doi:10.1007/s10676-017-9442-4.

[3] Feenberg A. Critical theory of technology. Oxford: Oxford University Press; 1991.

[4] Johnson DG. Computer systems: Moral entities but not moral agents. Ethics and Information Technology 2006 Nov;8(4):195-204. doi:10.1007/s10676-006-9111-5.

[5] Levy D. Robots unlimited: Life in a virtual age. Boca Raton, FL: CRC Press; 2005. doi:10.1201/b10697.

[6] Coeckelbergh M. Growing moral relations: Critique of moral status ascription. New York: Palgrave MacMillan; 2012. doi:10.1057/9781137025968.

[7] Sparrow R. The turing triage test. Ethics and Information Technology 2004 Dec;6(4):203-13. doi:10.1007/s10676-004-6491-2.

[8] Hume D. A treatise of human nature. New York: Oxford University Press; 1980.

[9] Asimov I. I, Robot. Gnome Press; 1950.

[10] Darling K. Extending legal protection to social robots: The effects of anthropomorphism, empathy, and violent behavior toward robotic objects. In: Calo R, Froomkin AM, Kerr I, editors. Robot law. Northampton, MA: Edward Elgar; 2016. p. 213-31. doi:10.2139/ssrn.2044797.

[11] Jacobsson M. Play, Belief and Stories about Robots: A Case Study of a Pleo Blogging Community. RO-MAN 2009 - The 18th IEEE International Symposium on Robot and Human Interactive Communication; 2009;232-37. doi:10.1109/ROMAN.2009.5326213

[12] Carpenter J. Culture and human–robot interaction in militarized spaces: a war story. London: Routledge; 2016. doi:10.4324/9781315562698

[13] Peperkorn HM, Diemer J, Alpers GW, Mühlberger A. Representation of Patients' Hand Modulates Fear Reactions of Patients with Spider Phobia in Virtual Reality. Frontiers in Psychology 2016 Feb;7:268. doi:10.3389/fpsyg.2016.00268.

[14] Cox-George C, Bewley SI. Sex Robot: the health implications of the sex robot industry. BMJ Sexual & Reproductive Health 2018 Jul;44(3):161-64. doi:10.1136/bmjsrh-2017-200012.

[15] Heider F, Simmel M. An experimental study of apparent behavior. American Journal of Psychology 1944 Apr;57(2), 243-59. doi:10.2307/1416950.

[16] Dennett D. The Intentional Stance. MIT Press; 1987.

[17] Höök K. Affective loop experiences: designing for interactional embodiment. Phil. Trans. R. Soc. B 2009 Dec;364(1535): 3585-95. doi:10.1098/rstb.2009.0202.

[18] Mossbridge J, Monroe E. Team Hanson-Lia-SingularityNet: Deep-learning Assessment of Emotional Dynamics Predicts Self-Transcendent Feelings During Constrained Brief Interactions with Emotionally Responsive AI Embedded in Android Technology. Unpublished XPrize Submission 2018.

[19] Baur T, Mehlmann G, Damian I, Gebhard P, Lingenfelser F, Wagner J, Lugrin B, André E. Context-aware automated analysis and annotation of social human-agent interactions. ACM Transactions on Interactive Intelligent Systems 2015 Jun;5(2):11. doi:10.1145/2764921.

[20] Gray J, Breazeal C. Manipulating Mental States Through Physical Action – A Self-as-Simulator Approach to Choosing Physical Actions Based on Mental State Outcomes. International Journal of Social Robotics 2014 Aug;6(3), 315-27.

[21] Manera V, Becchio C, Cavallo A, Sartori L, Castiello U. Cooperation or competition? Discriminating between social intentions by observing prehensile movements. Experimental Brain Research 2011 Jun;211(3-4): 547-56. doi:10.1007/s00221-011-2649-4.

[22] Sartori L, Becchio C, Castiello U. Cues to intention: The role of movement information. Cognition 2011 May;119(2): 242-52. doi:10.1016/j.cognition.2011.01.014.

[23] Cavallo A, Koul A, Ansuini C, Capozzi F, Becchio C. Decoding intentions from movement kinematics. Scientific reports 2016 Nov;6:37036. doi:10.1038/srep37036.