

A Robot that Receives Moral Concern

Sachie Yamada

Abstract

Social robots, which were first introduced into human society more than 20 years ago, are rapidly becoming more widespread. Unlike industrial robots and AI used in factories, social robots are implemented to perform tasks (roles) on behalf of humans or to act as companions. For example, in the home, pet robots are used as companions, as learning companions for children, and as companions for the elderly, suggesting health activities and reminding them to take their medication. They have already been introduced into society as hotel staff and security guards in public places. The roles expected of such robots will undoubtedly increase in the future.

Social robots not only interact with users, but when used in public places, they encounter many people who happen to be there. The people who happen to be there may be potential future users. When robots are introduced into such public places, they may be helped when they encounter obstacles when performing tasks. On the other hand, they may also be teased, hit and/or kicked. This is a phenomenon known as robot abuse. There are no clear reports of this happening with household robots, but it is not difficult to imagine that something similar may occur.

In this paper, we started with bottom-up studies of the two types of human behavior towards robots (abuse and help) that occurred when robots were introduced into the real world. We then understood these behaviors in terms of the concept of moral concern for robots.

In this paper, we define moral concern for robots as “broad concern for robots based on moral reasons related to daily life that emerge from perceiving a robot as a moral patient and/or from an individual’s relationship with a robot.” From the perspective of moral concern for robots, we can understand that robot abuse is an action without moral concerns, and that

robot-helping behavior is an action with moral concerns. In order to empirically study moral concern for robots, we created a scale to measure moral concern for robots, and explored the factors of moral concern for robots.

In Study 1, we clarified the process of escalating robot abuse. We qualitatively analyzed the behavior of the nine children who escalated the robot abuse using the Trajectory Equifinality Model (TEM) to determine how the robot abuse escalated, based on the phenomenon of robot abuse observed in the field trial. From the analysis, we developed a multistage model: the abuse escalation model. The model has four stages: approach, mild abuse, physical abuse, and severe abuse. For each stage, we identified five social guides (SGs), which are influencing events that fuel the stage. We quantitatively verified the effects of SG1 to SG3, which are social guides from Stage 1 to Stage 3, based on data from 12 hours of observation of 522 children who came to the plaza. Next, we verified the effects of SG4 and SG5, which are Stage 4 SGs, based on observation data from 40 hours of observation of 363 children who physically abused robots. Our quantitative analysis confirmed the correlations of all five SGs and children's behaviors: the presence of other children related a new child to approach the robot (SG1); mild abuse by another child related a child to do mild abuse (SG2); physical abuse by another child related a child to conduct physical abuse (SG3); joint action of other child (SG4) and encouragement from others (SG5) related a child to escalate the abuse. The results of the study 1 showed the importance of understanding robot abuse as a process that involves the people around the abuser and the robot, rather than as a two-way relationship between the abuser and the robot. We proposed preventive measures for each stage of robot abuse.

During the previous field trial with a robot and its interaction with shopping mall visitors, we observed a surprising event during which a leaflet-distributing robot was abused, although it was subsequently helped by one of its previous abusers. We qualitatively analyzed 19 cases in which the robot was saved out of 47 cases in which the robot dropped the leaflets during the field trials using TEM. We identified the following interaction process that started with individuals who paid attention to the robot, whether they had abusive or non-abusive interactions with it, whether they noticed its failure, and finally whether they helped it. The presence of others encouraged the person to focus on the robot, and the interactions with it led to helping,

regardless whether the interaction was abusive. The absence of others when the robot dropped the leaflet encouraged helping. The findings of this study will motivate interaction designs for social robots that can leverage human help. In the future, it will be necessary to quantitatively verify this process.

In Study 3, we created a new scale to measure moral concern for robots which is necessary for empirically studying moral concerns for robots. In the Study 3-1, we created a 15-item scale with three factor structures: moral care, concern about abuse, and concern for welfare. In the Study 3-2, we confirmed the factor structure derived from previous study using confirmatory factor analysis. And the construct validity was confirmed by authenticating correlations with other measures that were predicted to be theoretically relevant. In the Study 3-3, we verified predictive validity that the MCRS positively influenced the willingness to engage in prosocial behavior and better predicted it than other relevant scales.

In Study 4, we used the scale created in Study 3 to examine the factors that influence moral concern for robots. We manipulated the characteristics of the robot, including appearance, intelligence and sociality, and the relationship between the robot and human, in terms of reciprocity. The results showed that robots with appearances more similar to humans were more likely to receive moral concern. When comparing intelligence using a semi-human-like robot, it was found that the robot with higher intelligence was more likely to receive moral concern. For the semi-human-like robot with high intelligence, it was shown that that with higher sociality was more likely to receive moral concern than that with lower sociality. In the case of the robot that is semi-human-like and have high intelligence and sociality, it was found that the robot that behaved reciprocally towards humans was more likely to receive moral concern. In the study 4, because the conditions could not be created with current robot, a vignette methodology was used. In the future, it will be necessary to confirm the effects of these four factors using actual robots. It will also be necessary to explore other factors that are considered necessary for the social implementation of robots.

One of the findings of this paper is the process of human-robot interaction that occurs in the real world. Understanding interaction as a process expands our knowledge of how people encounter and interact with robots in the real world. Furthermore, it is possible to consider measures to prevent robot abuse at each stage. Also, it is possible to propose measures that make it easier for

robots to be helped at each stage. In the world of the future, where robots are evolving, moral concern for robots will become a larger topic of discussion in robot ethics. A scale to measure moral concern for robots can provide empirical data on moral concern for robots. It will provide insight into the complex mechanisms of human morality. Furthermore, understanding what factors elicit moral concern will provide a basis for proposing designs for robots that can receive moral concern on a more task-specific basis.