Making Robot's Attitudes Predictable: A Stereotype Content Model for Human-Robot Interaction in Groups

Raquel Oliveira
Instituto Universitário de Lisboa (ISCTE-IUL), CIS-IUL
Lisbon, Portugal
rsaoa@iscte-iul.pt

Patrícia Arriaga
Instituto Universitário de Lisboa (ISCTE-IUL), CIS-IUL
Lisbon, Portugal
patricia.arriaga@iscte-iul.pt

Filipa Correia
Instituto Superior Técnico, Universidade de Lisboa & INESC-ID
Lisbon, Portugal
filipacorreia@tecnico.ulisboa.pt

Ana Paiva
Instituto Superior Técnico, Universidade de Lisboa & INESC-ID
Lisbon, Portugal
ana.paiva@inesc-id.pt

ABSTRACT

Stereotypes play a key role both in human and in robot perception. As such, these may play an important role in predicting behavior towards the targets of said stereotypes. In this paper, we argue for the usefulness of exploring how stable dimensions of stereotype content, more specifically warmth and competence, apply to HRI. We propose to do so by examining the impact of these characteristics, displayed by robots, on the emotions, behaviors and future intention of participants to interact and work with robots. We chose these two stereotypical dimensions given that research has demonstrated their underlying and ubiquitous influence on people perception and responses. Moreover, we decided to explore this issue in the context of small group interactions due to the ever-common nature of this type of social arrangements in people's everyday life.

KEYWORDS

Stereotype Content Model; Emotions; Intentions to Work, Human-Robot Interaction

1 INTRODUCTION

People's perception about themselves and others is influenced by a multitude of factors. One of those factors is the stereotypes individuals develop and apply to make sense of the world. Cognitive psychology theories argue that stereotypes are better categorized in terms of implicit personality theory, as they constitute heuristics for stimuli interpretation that due to their general acceptance can be considered as part of a process of social cognition [4]. Thus, these constitute part of a larger set of knowledge structures that guide interpretation of other's people behavior and are relevant predictors of own behavior [2], [3], [9]. Moreover, research in Human-Robot Interaction (henceforth, HRI) has demonstrated the presence of this type of social categorization (for example, based on gender [11]), drawing attention to the question of how different dimensions of stereotype content can impact the way robots are perceived, and thus, responded to. Answering this question can yield important findings to take into account during the robot conception design by contributing to the creation of more socially effective robots (by providing people a consistent form of social information about the robot that they can take into account to predict its behavior and thus, adjust their expectations about it and act accordingly). Thus, if we consider stereotypes can function as social cues in the context of HRI, this can spare the user from the effort of looking for further information during the interaction, thereby increasing the level of transparency and intuitiveness of this interaction.

1.1 Externalizing individual attitudes in groups of human-robot interactions

The present work is done under the AMIGOS project [1] which aims to contribute to the literature by considering the roles of emotions and adaptation in Human-Robot-Interaction, in the context of group interactions. Our aim is to create a data-driven model for group interactions that allows robots in that context, to adjust their behavior according to the situational characteristics and the preferences of each individual set of users. Thus, our goal is to endow the robot with the ability to generate context-adaptive responses using interactive machine learning techniques. In order to do so, we use a card-game scenario in which two participants engage in an entertaining task with two robots (for more details, see [1]), and analyze how mixed group behave towards one another in the role of partners and opponents. Thus, our goal is to endow the robot with the ability to generate context-adaptive responses as robots and humans interact in small groups. As such, making explicit certain types of attitudes and behaviors in robots is very important for a natural communication to emerge. Significant work has been done in making expressions of robots legible and predictable, in particular at the motion level [8] [13]. Other work has pointed out the need for exaggeration of actions and features in robots, to make them more natural[19]. Here we argue that attitudes can be made more salient in a robot by leveraging the power that stereotypes have in communication.

1.2 Stereotypes in Human and Robot Interaction

People are social creatures who tend to attribute human-like characteristics to a broad range of non-human elements. This might include random patterns [16], virtual agents [15] and robots [20]. In the case of robot perception, studies have already demonstrated the important role of stereotypes in robot's trait evaluation. For example, Eyssel and Hegel compared two robots displaying different

gender facial characteristics [10]. These authors found that the male robot was perceived as more agentic, whereas the female robot was perceived as being more communal. Furthermore, using a voice gender manipulation, that could be either synthesized robot-like or human-like, Eyssel and colleagues [12] also found that participants tended to evaluate the same-gender robot more positively across a large range of social dimensions, suggesting the existence of some sort of projection mechanism. Moreover, when robots' are stereotyped, it appears that the stereotypes that are associated with them are congruent with those that occur in Human to Human Interaction (HHI). For example, the aforementioned typical assumed gender-role stereotypes have been also been consistently verified in HRI [20], hinting at the existence of some level of extrapolation of people's stereotypes about other people to stereotypes about what other social actors, in this instance, robots, do or are supposed to do in a certain situation.

1.3 The Stereotype Content Model: Why does Competence Matter

The term stereotypes implies a gestalt view of people perception, suggesting the notion that some traits can be more central than others, in organizing our perception of other people [20]. According to the Stereotype Content Model (hereinafter, SCM) [14]; [7] warmth and competence are the two main stable content dimensions of stereotypes. These are central to group stereotypes and have been linked to specific emotional and behavioral outcomes [7] (see fig 1). SCM considers two levels of competence and warmth (low and high). The combination of those two, can be associated with a set of traits that are considered more or less socially desirable and, thus can elicit different emotional and behavioral responses [14]. Social robots have not yet been perfected to the point where a certain degree of incompetence (i.e. failure to adjust and choose a proper response) is not to be expected. In fact, social situations, especially those involving more than one human, can present very challenging environments that make it hard for a robot to identify the relevant characteristics to take into account to form a response that is adequate. Furthermore, a high level of competence, might not always be desirable as it can evoke machine-like associated stereotypes (i.e.*computers/machines don't fail*) or because a high level of task-orientation might be perceived as a threat [17]. Research in HHI demonstrates consistent group stereotypes associated with perceptions of different levels of competence and warmth [14], that can result in different behavioral approaches.

Goals and Hypothesis. Although some studies have already focused on the competence and warmth dimensions in robot perception (e.g.: [5], we believe that a more systematic understanding of how different levels of competence and warmth displayed by the robot intertwine to mold an overall impression of the robot, is still missing. An examination of how these differences in perception affect the emotional and behavioral responses towards the robot is also lacking. In this context, we expect participants to display similar emotional responses to those observed in HHI, towards robots displaying different levels of warmth and competence. (see fig.1). We also aim to examine the moderating role of trait congruency in the future intention to work with robots. Furthermore, we also expect participants to display a higher level of future intention to work with the robot that displays similar characteristics to him/herself.

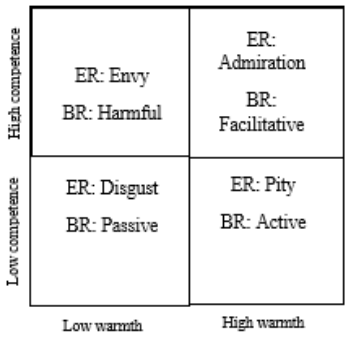

Figure 1: Content dimensions of SCM: ER stands for expected emotional response, whereas BR stands for expected behavioral response [14]

2 ONGOING WORK

Social interactions are complex phenomena, that might include different forms of interpersonal communication and social messaging. These social interactions become increasingly more complex if one looks at them from a group interaction standpoint, as this adds complexity to the analysis.

However, as groups of humans and robots might emerge in a near future it becomes relevant to consider how they act and communicate in groups. Answering the question we present in this paper will allow for the development of robots that can interact both according to the situation and to the type of emotional or behavioral response it wishes to evoke. This also helps the development of robots that can evoke consistent mental models and social categorizations from humans and that, in turn, will make robots' behaviors and intentions appear more predictable from the user perspective.

To address these questions, we are currently working on an entertaining interaction card-game scenario , where two human participants are required to play with two robots [6], displaying high and low levels of warmth and competence. This task has been used in previous studies of the AMIGOS project and details regarding the experimental design and task can be found elsewhere [1], [17] . Warmth will be manipulated through the utterances spoken by the robot whereas competence will be manipulated through the game-solving algorithm implemented in each robot (for more details, see [18]). Thus, our main goal is to understand the behavioral and emotional implications of different judgments of warmth and competence levels and how these might affect future intention to work with robots.

REFERENCES

[1] Patricia Arriaga, Raquel Oliveira, Patrícia Alves-Oliveira, Ana Paiva, Sofia Petisca, and Filipa Correia. 2017. AMIGOS: Affect Modeling for robots In GrOup Social interactions. (Nov 2017). DOI : http://dx.doi.org/10.17605/OSF.IO/ZP34V

[2] Richard D Ashmore and Frances K Del Boca. 1979. Sex stereotypes and implicit personality theory: Toward a cognitive—Social psychological conceptualization.

Sex roles 5, 2 (1979), 219–248.

[3] Richard D Ashmore and Frances K Del Boca. 1981. Conceptual approaches to stereotypes and stereotyping. *Cognitive processes in stereotyping and intergroup behavior* 1 (1981), 35.

[4] Richard D Ashmore and Frances K Del Boca. 2013. *The social psychology of female-male relations: A critical analysis of central concepts.* Elsevier.

[5] Colleen M Carpinella, Alisa B Wyman, Michael A Perez, and Steven J Stroessner. 2017. The robotic social attributes scale (RoSAS): development and validation. In *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction.* ACM, 254–262.

[6] Filipa Correia, Sofia Petisca, Patrıcia Alves-Oliveira, Tiago Ribeiro, Francisco S Melo, and Ana Paiva. Groups of humans and robots: Understanding membership preferences and team formation. (????).

[7] Amy JC Cuddy, Susan T Fiske, and Peter Glick. 2008. Warmth and competence as universal dimensions of social perception: The stereotype content model and the BIAS map. *Advances in experimental social psychology* 40 (2008), 61–149.

[8] Anca D Dragan, Kenton CT Lee, and Siddhartha S Srinivasa. 2013. Legibility and predictability of robot motion. In *Human-Robot Interaction (HRI), 2013 8th ACM/IEEE International Conference on.* IEEE, 301–308.

[9] Alice H Eagly, Richard D Ashmore, Mona G Makhijani, and Laura C Longo. 1991. What is beautiful is good, but…: A meta-analytic review of research on the physical attractiveness stereotype. *Psychological bulletin* 110, 1 (1991), 109.

[10] Friederike Eyssel and Frank Hegel. 2012. (S) he's got the look: gender stereotyping of robots. *Journal of Applied Social Psychology* 42, 9 (2012), 2213–2230.

[11] Friederike Eyssel and Dieta Kuchenbrandt. 2012. Social categorization of social robots: Anthropomorphism as a function of robot group membership. *British Journal of Social Psychology* 51, 4 (2012), 724–731.

[12] Friederike Eyssel, Dieta Kuchenbrandt, Simon Bobinger, Laura de Ruiter, and Frank Hegel. 2012. 'If you sound like me, you must be more human': On the interplay of robot and user features on human-robot acceptance and anthropomorphism. In *Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction.* ACM, 125–126.

[13] Miguel Faria, Rui Silva, Patrıcia Alves-Oliveira, Francisco S Melo, and Ana Paiva. 2017. "Me and You Together" Movement Impact in Multi-user Collaboration Tasks. (2017).

[14] Susan T Fiske, Amy JC Cuddy, Peter Glick, and Jun Xu. 2002. A model of (often mixed) stereotype content: competence and warmth respectively follow from perceived status and competition. *Journal of personality and social psychology* 82, 6 (2002), 878.

[15] Magnus Haake and Agneta Gulz. 2008. Visual stereotypes and virtual pedagogical agents. *Journal of Educational Technology & Society* 11, 4 (2008), 1.

[16] Jiangang Liu, Jun Li, Lu Feng, Ling Li, Jie Tian, and Kang Lee. 2014. Seeing Jesus in toast: neural and behavioral correlates of face pareidolia. *Cortex* 53 (2014), 60–77.

[17] Raquel Oliveira, Patrícia Arriaga, Patrícia Alves-Oliveira, Filipa Correia, Sofia Petisca, and Ana Paiva. 2018. Friends Or Foes? Socioemotional Support and Gaze Behavior In Mixed Groups of Humans And Robots. (Mar 2018).

[18] Raquel Oliveira, Ana Paiva, Sofia Petisca, Patricia Arriaga, Patrícia Alves-Oliveira, and Filipa Correia. 2018. Stereotype Content Model applied to Human and Robot Interaction. (Feb 2018). osf.io/waj28

[19] Tiago Ribeiro and Ana Paiva. 2012. The illusion of robotic life: principles and practices of animation for robots. In *Human-Robot Interaction (HRI), 2012 7th ACM/IEEE International Conference on.* IEEE, 383–390.

[20] Benedict Tay, Younbo Jung, and Taezoon Park. 2014. When stereotypes meet robots: the double-edge sword of robot gender and personality in human–robot interaction. *Computers in Human Behavior* 38 (2014), 75–84.