Who to Blame? Understanding Distribution of Responsibility in Human-Robot Interaction

Stephen Gaschignard
Affiliation
Address
e-mail address

Mert Oguz
Affiliation
Address
e-mail address

Xiang Zhi Tan
Affiliation
Address
e-mail address

ABSTRACT

In this paper we describe the formatting requirements for SIGCHI Conference Proceedings, and this sample file offers recommendations on writing for the worldwide SIGCHI readership. Please review this document even if you have submitted to SIGCHI conferences before, some format details have changed relative to previous years.

Author Keywords

Guides; instructions; author's kit; conference publications; keywords should be separated by a semi-colon.
Optional section to be included in your final version, but strongly encouraged.

ACM Classification Keywords

H.5.m. Information Interfaces and Presentation (e.g. HCI): Miscellaneous

See: `http://www.acm.org/about/class/1998/` for more information and the full list of ACM classifiers and descriptors. Optional section to be included in your final version, but strongly encouraged. On the submission page only the classifiers letter-number combination will need to be entered.

INTRODUCTION

With the enhancement of robotics and artificial intelligence studies, robots have entered our lives in numerous ways. From industrial robots used in factories to everyday electronics like cleaning robots, humans take advantage of the cutting-edge technology that is designed for them. While studies about robotics deal with development and optimization of such technologies, the scope of human robot interaction studies expands through understanding psychological processes which humans come across while interacting with robots. Many notions which are vastly studied for human-human interaction such as trust, harm and perception of mistakes have been extended into human-robot interaction. [1]. It is clear that robots making mistakes during the interaction, therefore the performance of the robot is one of the most important reasons for trust to diminish. [2]. Whether the other body is a human or robot, people tend to trust if there is consistency in the behavior of the other. Making mistakes on the other hand, is a natural outcome of poor analysis of the enclosing environment, sometimes caused by different deficiencies. Humans have a high try and fail learning capability, besides many creative mistake recovery strategies. With today's technology, robots are still not as talented as humans in terms of communication with other humans or direct interaction. Therefore, in the upcoming years, we predict that a number of studies will be conducted about the robot-made mistakes. As it is interesting to study why robots make mistake, studying how people perceive those mistakes and how they react in such situations is also crucial for designing successful mistake recovery systems.

Our research question focuses on the human perception of the possible robot mistakes in a restaurant setting. We have decided on this setting since we believe that robot waiters might be conventional in the near future. Restaurants are the places where at least a number of interactions occur between the customer and the waiter. These interactions include physical ones where the waiter brings the menu, dish or the bill. Conventionally, a waiter talks to a customer for getting the order or accepting the payment. These interactions have a psychological effect on the customer. Many customers decide on the amount of tip they wish to give based on the behavior of the waiter. Therefore, we categorized the interactions into different types to be studied. We analyzed 3 different types of mistakes with 2 different severity levels. Our research is a within-participant study conducted with 49 participants (26 of them were females) where each participant is a U.S. resident. We hired participants through Amazon Mechanical Turk, paying them for completing our online questionnaire. In order to reach a more random participant population, we used video-based research instead of lab-based live settings. [3] showed that video-based methods have reasonable results for new innovative studies. This approach enabled us to conduct our study with much smaller budget and time spending. Instead, we could focus on the scenarios and the quality of the questionnaire. 6 scenarios (3 types of mistakes: Physical, Psychological, Financial and 2 severity levels: Severe, Non severe) are designed for the study. After watching every video scenario, we asked participants how they attribute responsibility to each of the stakeholders (The robot, programmer of the robot, owner of the restaurant, manager of the restaurant,

Type	Non-Severe	Severe	p Value
Financial	3.633	4.918	0.0003
Psychological	2.633	5.020	< 0.0001
Psychological	2.816	5.020	< 0.0001

Table 1. Mean Difference between perceive severity

manufacturer). Then, we asked them what would be their reaction in such situations. We hypothesized that both the type and severity of the mistake will influence participants' distribution of blame. We further predicted that people will blame the robot more for psychological abuse. Our last hypothesis was people will blame the robot for non-severe situations, whereas they will blame other stakeholders for severe situations, if the mistake is a financial exploitation or physical abuse.

METHODOLOGY

RESULT

Manipulation Check

We conducted a manipulation check to verify the success manipulation of our condition. After watching each video clip, we ask participants to rate the severity for each scenario. We found in all scenarios, the participants rated the severe condition significantly higher than the non-severe condition.

DISCUSSION

CONCLUSION

REFERENCES FORMAT

References must be the same font size as other body text.

REFERENCES

1. Freedy, E., DeVisser, E., Weltman, G., and Coeyman, N. Measurement of trust in human-robot collaboration. In *Collaborative Technologies and Systems, 2007. CTS 2007. International Symposium on*, IEEE (2007), 106–114.
2. Hancock, P. A., Billings, D. R., Schaefer, K. E., Chen, J. Y., De Visser, E. J., and Parasuraman, R. A meta-analysis of factors affecting trust in human-robot interaction. *Human Factors: The Journal of the Human Factors and Ergonomics Society 53*, 5 (2011), 517–527.
3. Woods, S., Walters, M., Koay, K. L., and Dautenhahn, K. Comparing human robot interaction scenarios using live and video based methods: towards a novel methodological approach. In *Advanced Motion Control, 2006. 9th IEEE International Workshop on*, IEEE (2006), 750–755.