The Role of Social Dialogue and Errors in Robots

Gale M. Lucas, Jill Boberg, David Traum, Ron Artstein, Jon Gratch, Alesia Gainer, Emmanuel Johnson, Anton Leuski
USC Institute for Creative Technologies, Los Angeles, USA
lucas@ict.usc.edu

Mikio Nakano
Honda Research Institute Japan Co., Ltd.
Wako, Saitama, Japan
nakano@jp.honda-ri.com

ABSTRACT
Social robots establish rapport with human users. This work explores the extent to which rapport-building can benefit (or harm) conversations with robots, and under what circumstances this occurs. For example, previous work has shown that agents that make conversational errors are less capable of influencing people than agents that do not make errors [1]. Some work has shown this effect with robots, but prior research has not considered additional factors such as the level of rapport between the person and the robot. We predicted that building rapport through a social dialogue (such as an ice-breaker) could mitigate the detrimental effect of a robot's errors on influence. Our study used a Nao robot programmed to persuade users to agree with its rankings on two "survival tasks" (e.g., lunar survival task). We manipulated both errors and social dialogue: the robot either exhibited errors in the second survival task or not, and users either engaged in an ice-breaker with the robot between the two survival tasks or completed a control task. Replicating previous research, errors tended to reduce the robot's influence in the second survival task. Contrary to our prediction, results revealed that the ice-breaker did not mitigate the effect of errors, and if anything, errors were more harmful after the ice-breaker (intended to build rapport) than in the control condition. This backfiring of attempted rapport-building may be due to a contrast effect, suggesting that the design of social robots should avoid introducing dialogues of incongruent quality.

Author Keywords
Social robots; influence; social dialogue; rapport; errors.

ACM Classification Keywords
I.2.9 [**Artificial Intelligence**]: Robotics
H.1.2 [**Models and Principles**]: User/Machine Systems – *Human factors*
H.5.2 [**Information Interfaces and Presentation**]: User Interfaces – *Natural language*

HAI '17, October 17–20, 2017, Bielefeld, Germany

ACM ISBN 978-1-4503-5113-3/17/10…$15.00
https://doi.org/10.1145/3125739.3132617

INTRODUCTION
Conversational robots and agents working in offices and at home are expected to play increasingly important roles. Such robots are expected to establish rapport with human users through repeated interaction and personalization [2]. Rapport can be created within a single interaction by engaging in relational or social dialogue. We seek to understand the extent to which such rapport-building can benefit (or harm) conversations with robots and agents, and under what circumstances this occurs.

RELATED WORK
Conversational robots and virtual agents can be used for both assistance on tasks as well as social interaction. For example, Siri includes a number of application services, such as weather and calendar functions, but also has some social responses and jokes. Systems like REA [3] and SASO [4] help users perform tasks, but also use relational or social dialogue to build a relationship with users.

While some assistive systems are built to assist users on information-seeking or service tasks (like Siri), others focus on influencing the user. Examples include systems to motivate people to exercise [5] or negotiate on a course of action [4,6]. Previous work with virtual agents has shown that errors in dialogue systems can reduce such social influence [1,7]. Evidence with robots is more mixed. Some work finds that errors have no impact on robots' influence [8], while other work indicates that errors in robots' dialogue systems negatively affect their influence [9–11]. However, research has yet to consider the impact of robots' social dialogue on mitigating the impact of errors.

PRESENT RESEARCH
We test the possibility that building rapport through a social dialogue (e.g., ice-breaker) could mitigate the detrimental effect of errors on influence. In contrast, other work suggests that trust is easily lost but not easy to regain [11]. If participants experience a loss of trust when the errors begin, rapport built in the ice-breaker may not be enough to mitigate the detrimental effect of errors. This experiment considers two factors: users either engaged in an ice-breaker with the robot or completed a control task, and afterwards, the robot either exhibited conversational errors or not. One hundred and fifteen participants were recruited from Craigslist for this experiment. The data for eight users were later removed due to technical issues experienced during the session.

Participants first completed the lunar survival task [12–14]: participants were asked to imagine that they are part of a space crew that crashed on the moon, and were asked to rank 10 items as to their importance for surviving long enough to be rescued. Participants ranked these items individually, and were then told that they should rank them again with the help of another crewmember, our robot. Participants then engaged in dialogue with the NAO robot; the robot was controlled by a human operator ("Wizard of Oz") and acted as a confederate [15], providing factual arguments for ranking the items in a specific order. Following the dialogue, participants re-ranked the items; the differences between initial rankings and final rankings served as a measure of influence [16].

Next, participants answered a series of personal questions. In the ice-breaker condition, the participant had an interactive dialogue with the robot. The robot shared its own stories, in first person, while eliciting stories from the participant. In the control condition, the subject participated in a non-interactive oral survey, with the same personal questions being asked by a female (non-robot) voice.

Participants then completed a second ranking task, the Save-the-Art Task [16]. Participants were asked to imagine that they were a manager at a museum that was on fire. They were asked to rank 10 pieces of art as to their importance in being saved. Again, they ranked the items individually, had an interactive dialog with the robot about which items should be saved, and then re-ranked the items. Differences between rankings again indexed influence.

In the error condition, the robot made a series of errors while interacting with the participant on the Save-the-Art Task. When users asked questions about the robot's rankings, errors were introduced in one of several ways: asking users to repeat themselves, answering a different question, repeating the answer to the previous question, answering with a question or non sequitur, or not answering at all. Errors were introduced into the dialogue according to a set order at a rate of about one of these errors per two utterances.

RESULTS AND DISCUSSION

A 2 (ice-breaker) x 2 (error) ANOVA was run on change in influence from task 1 (lunar) to 2 (art). Because initial agreement with the robot limits influence, initial agreement with the robot on each task was entered as a co-variate to statistically equate participants on this factor.

There was a main effect such that participants were marginally influenced more in the absence of errors ($F(1,95) = 3.27$, $p = .07$). While there did not seem to be an effect of ice-breaker $F(1,96) = 1.082$, $p = .30$), the interaction between error and ice-breaker conditions approached a trend ($F(1,96) = 1.54$, $p = .22$). The effect of errors, if anything, tended to be driven more by the ice-breaker condition than the control condition (Figure 1). That is, the effect of errors tended to be worse after the ice-breaker.

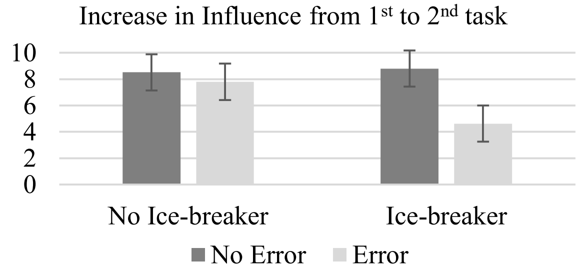

Figure 1. Effects of ice-breaker and error on influence.

As in previous research, errors reduced influence in the second survival task. However, contrary to our prediction, errors were more harmful, if anything, after the ice-breaker.

This work contributes further evidence to an area that has mixed findings: some work finds that errors have no impact on the robot's influence [8], but our work and others' shows that errors in robot's dialogue systems negatively affect their influence [9–11]. In all of these cases, however, it seems to depend on when the errors occur: errors after a period of good performance were much more harmful to influence than those that occur earlier [9–11]. In Desai et al. [9–10], drops in reliability after a period of good performance were much more harmful to influence than early failures. Weigmann et al. [11] likewise found that robots that shifted from 100% reliable to 80% reliable had less influence than those that were 80% reliable from the start.

The present work also finds that errors are more harmful after good performance during an ice-breaker conversation than without such a conversation. Indeed, this may result from the timing of the ice-breaker. As in the above research, the ice-breaker made the robot appear to perform well for longer (compared to the control condition). Indeed, a contrast effect may have occurred, whereby errors stand out more after an ice-breaker interaction. The contrast effect is a common concept in social psychology, which notes that if someone has experienced a positive interaction, their response scale is anchored in the positive, and subsequent negative experiences are judged against that positive scale [17].

These results have implications for HRI and robot design. We have shown that conversational errors hinder users from taking the advice of the robot, undermining the robot's persuasiveness. While it seems that errors are particularly damaging when they suddenly appear after good performance (here during a social dialogue), more research is needed to isolate the precise impact of errors and possible interventions. For example, the role of errors versus the contrast effect could be explored by introducing consistent errors in all tasks, or just introducing errors before the ice-breaker. Research could also explore specific error-mitigation rapport-building, such as incorporating apologies, explanations, and negative self-disclosure.

In sum, our work – combined with prior research – highlights the risk of errors in robot's dialogue. Some null findings [8] notwithstanding, errors appear to reduce robots' influence. Design could still focus on other ways of mitigating errors, but merely placing a social dialogue before the errors appears to be a poor option.

ACKNOWLEDGMENTS
This work was supported in part by the U.S. Army. Any opinion, content or information presented does not necessarily reflect the position or the policy of the United States Government, and no official endorsement should be inferred.

REFERENCES
1. Wang, Y., P. Khooshabeh, et al. (2013). Looking Real and Making Mistakes. Intelligent Virtual Agents: 13th International Conference, IVA 2013, Edinburgh, UK, August 29–31, 2013. Proceedings. R. Aylett, B. Krenn, C. Pelachaud and H. Shimodaira. Berlin, Heidelberg: Springer: 339–348.
2. Kanda, T., Shiomi, M., Miyashita, Z., Ishiguro, H., & Hagita, N. (2010). A communication robot in a shopping mall. IEEE Transactions on Robotics, 26(5), 897–913.
3. Cassell, J., & Bickmore, T. (2003). Negotiated collusion: Modeling social language and its relationship effects in intelligent agents. User modeling and user-adapted interaction 13(1): 89–132.
4. Traum, D., Swartout, W., Marsella, S., & Gratch, J. (2005). Fight, flight, or negotiate: Believable strategies for conversing under crisis. In Proceedings of Intelligent Virtual Agents, 52–64.
5. Manuvinakurike, R., Velicer, W. F., & Bickmore, T. W. (2014). Automated indexing of Internet stories for health behavior change: weight loss attitude pilot study. Journal of medical Internet research, 16(12).
6. Hiraoka, T., Neubig, G., Sakti, S., Toda, T., & Nakamura, S. (2014, August). Reinforcement Learning of Cooperative Persuasive Dialogue Policies using Framing. In COLING (pp. 1706–1717).
7. Blascovich, J., & McCall, C. (2013). Social influence in virtual environments. In K. Dill (Ed.), The Oxford handbook of media psychology (pp. 305–315). New York, NY: Oxford University Press.
8. Salem, M., Lakatos, G., Amirabdollahian, F., & Dautenhahn, K. (2015, March). Would you trust a (faulty) robot?: Effects of error, task type and personality on human-robot cooperation and trust. In Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction (pp. 141–148). ACM.
9. Desai, M., Medvedev, M., Vázquez, M., McSheehy, S., Gadea-Omelchenko, S., Bruggeman, C., ... & Yanco, H. (2012, March). Effects of changing reliability on trust of robot systems. In Human-Robot Interaction (HRI), 2012 7th ACM/IEEE International Conference on (pp. 73–80). IEEE.
10. Desai, M., Kaniarasu, P., Medvedev, M., Steinfeld, A., & Yanco, H. (2013, March). Impact of robot failures and feedback on real-time trust. In Proceedings of the 8th ACM/IEEE international conference on Human-robot interaction (pp. 251–258). IEEE.
11. Wiegmann, D. A., Rich, A., & Zhang, H. (2001). Automated diagnostic aids: The effects of aid reliability on users' trust and reliance. Theoretical Issues in Ergonomics Science, 2(4), 352–367.
12. Moon, Y. (1998, January). The effects of distance in local versus remote human-computer interaction. In Proceedings of the SIGCHI conference on Human factors in computing systems (pp. 103–108). ACM Press/Addison-Wesley Publishing Co.
13. Khooshabeh, P., McCall, C., Gandhe, S., Gratch, J., & Blascovich, J. (2011, May). Does it matter if a computer jokes. In CHI'11 Extended Abstracts on Human Factors in Computing Systems (pp. 77–86). ACM.
14. Ramachandran, D., & J. Canny (2008). The persuasive power of human-machine dialogue. International Conference on Persuasive Technology, Springer.
15. Kelley, J. F. (1984). An iterative design methodology for user-friendly natural language office information applications. ACM Transactions on Office Information Systems, 2, 26–41.
16. Artstein, R., Traum, D., Boberg, J., Gainer, A., Gratch, J., Johnson, E., Leuski, A., & Nakano, M. (2017). Listen to my body: Does making friends help influence people? In Proceedings of the Florida Artificial Intelligence Research Society Conference, 430–435.
17. Sherman, S. J., Ahlm, K., Berman, L., & Lynn, S. (1978). Contrast effects and their relationship to subsequent behavior. Journal of Experimental Social Psychology, 14(4), 340–350.