



Understanding Rapport over Multiple Sessions with a Social, Teachable Robot

Xiaoyi Tian¹(✉), Nichola Lubold², Leah Friedman³, and Erin Walker¹

¹ University of Pittsburgh, Pittsburgh, PA, USA
{xiaoyi-tian,eawalker}@pitt.edu

² Arizona State University, Tempe, AZ, USA
nlubold@asu.edu

³ University of Pennsylvania, Philadelphia, PA, USA
leahf@pitt.edu

Abstract. Social robots have been shown to be effective educational tools. Rapport, or interpersonal closeness, can lead to better human-robot interactions and positive learning outcomes. Prior research has investigated the effects of social robots on student rapport and learning in a single session, but little is known about how individuals build rapport with a robot over multiple sessions. We reported on a case study in which 7 middle school students explained mathematics concepts to an intelligent teachable robot named Emma for five sessions. We modeled learners' rapport-building linguistic strategies to understand whether the ways middle school students build rapport with the robot over time follow the same trends as human conversation, and how individual differences might mediate the rapport between human and robot.

Keywords: Human-robot interaction · Multiple sessions interaction · Rapport · Case study · Long-term hri

1 Introduction

Intelligent social robots have been shown to have positive effects on learning and motivational outcomes [5, 12, 15] in part because of the socio-emotional support they provide [9, 11, 12]. One mechanism that may contribute to these positive effects is the rapport, or feeling of connection, that social robots engender with their human collaborators. However, over time, the nature of the relationship between the human and robot might shift (as human-human relationships do), and the importance of rapport may change [19]. Most research on human-robot rapport has been done in single-session studies [9, 13, 14], and has rarely investigated how learners develop and maintain rapport with a robot. Understanding how children build and maintain relationships during multiple encounters would help maintain engagement and personalize long-term learning experiences.

A widely-accepted human-human rapport framework comes from Tickle-Degnen and Rosenthal's three-factor theory [19], which includes mutual attention, positivity, and coordination. People start building rapport by expressing

mutual attentiveness and interests towards one another. High positivity plays a role in generating a feeling of mutual friendliness and warmth, but in the initial stage of an interaction, there may be less coordination (interlocutors are “in sync” with one another). Over the long term, positivity decreases, coordination increases, and mutual attentiveness remains stable. It’s not clear that the same phenomena can be observed in human-robot settings. Thus, our study aims to understand how students verbally build and maintain rapport with a robot in multiple sessions. We conducted an exploratory analysis of 7 middle school students interacting with a social teachable robot over 5 sessions. Our research question was: How do students differ from each other and differ from early to late interaction stages in the way that they build rapport with a teachable robot?

2 Multi-session Study

For this study, a Nao robot named Emma was taught by middle school students how to solve mathematics problems utilizing spoken language [10]. Students sat at a desk with a Surface Pro tablet in front of them. Emma stood on the desk to the right of the participant. Table 1 is an example of exchange between Emma and a learner on a ratio and proportions problem. More details on the system design can be found in [10]. Over multiple sessions, Emma mimicked the [19]’s model of rapport as follows. To implement coordination, we utilized an acoustic-prosodic entrainment module [9], which transforms Emma’s utterances to converge to the user’s pitch. Entrainment increased over the five sessions. Emma exhibited higher positivity in the initial sessions by exhibiting greater politeness and enthusiastic language (e.g., “Great! Thank you for teaching me”) than in later sessions. We operationalized attention as gaze behavior, and did not change Emma’s default gaze behavior throughout the sessions.

Table 1. Example dialogue between Emma and a learner, coded with conversational strategies.

Emma:	Interesting. Why do you think <i>we</i> [inclusive] do that?
Learner:	Because it’s going to be equivalent[responsive]...so that means two times two equals[ask question]?
Emma:	Sweet. I think I get it. <i>we</i> [inclusive] multiply two times two, so then is four the answer?
Learner:	Yes, <i>Emma</i> [name]. That’s correct.

Participants were 7 middle-school students (4 females, 3 males). The mean age was 12.7. Each participant interacted with Emma for five 30-min sessions over several weeks. We grouped session 1, 2 and 3 as early interaction stages, and sessions 4 and 5 as late stages. Participants solved 4–6 problems during each study session, resulting in 186 independent problems in the corpus.

Each problem contains 10.06 user utterances on average. Two coders manually coded conversational strategies indicating behavioral rapport in each utterance in the human-robot tutoring dialogue (Cohen’s kappa of all codes was higher than 0.8). The strategies consisted of *off-topic chat*, *inclusive pronouns* (e.g., use of “we” vs “I”), *use of Emma’s name*, *praise*, *apology*, *refer to past experience*, *ask a question*, *respond to Emma’s prompt*, and *adherence to social norms*, drawn from both human-human and human-robot rapport studies [1, 3, 7, 8, 17, 20]. Example of codes can be found in Table 1. To supplement our manual codes, we incorporated automatic linguistic feature detection using the 2015 LIWC [16] summary language variables (analytical thinking, clout, authenticity, and emotional tone).

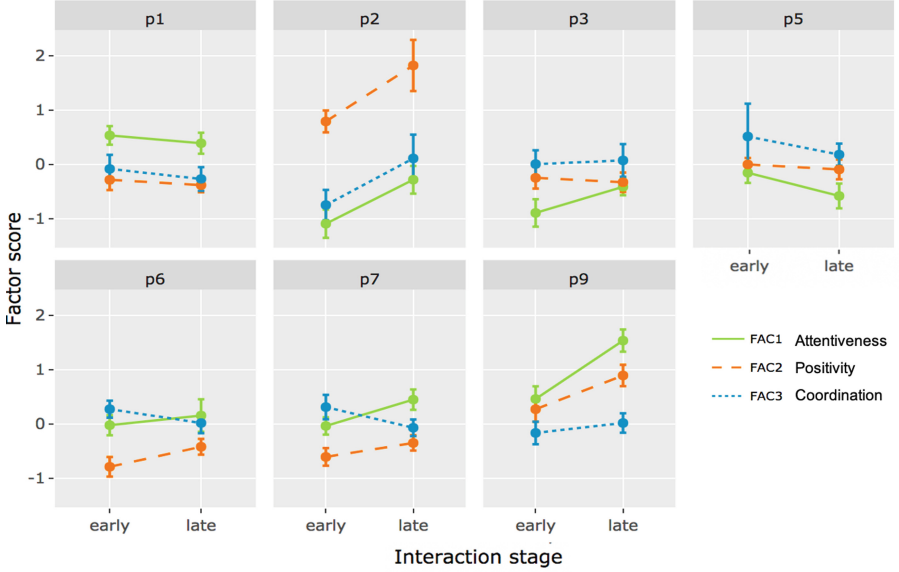


Fig. 1. Rapport factors (with standard errors) across 5 sessions separate by all seven participants. Each factor score across all participant problems has zero mean and unit standard deviation.

We used an Independent Component Analysis (ICA) to map strategies to three rapport factors, revealing how particular behaviors are used as ways of expressing and managing the underlying rapport-building constructs with the robot [2, 4]. Linguistic strategies that strongly loaded on factor 1 were *inclusive language*, *name usage*, *apology*, and *clout*, and we interpret the factor as **attentiveness**. *Name usage*, *praise*, *authenticity* and *emotional tone* were loaded in factor 2, and we interpret this as **positivity**. Markers of off-task coordination load strongly positively on factor 3 (*chat* and *adhere norm*), while markers of on-task coordination load strongly negatively (*responsiveness* and *ask question*). Nevertheless, we do interpret this factor to represent **coordination**, with

on-task and off-task coordination interestingly being negatively related. Finally, the *ask question* and *refer to past experience* loaded evenly amongst all factors.

Our next step was to understand how rapport varies from early stage to late stage of interaction. From our ICA model, we computed the source matrix for the 3 extracted rapport components for each 186 participant-problem pair, and aggregated the mean rapport score across all problems in each interaction stage. The results are represented in Fig. 1. We can observe that **attentiveness** went up across the majority of participants. **Positivity** appeared to vary from participant to participant, with some individuals who started from a high score seeing dramatic increases (p2 and p9). **Coordination** decreased in four out of seven learners with large variations (e.g., p5). It is worthy to note that, a decrease in coordination score meant more on-task behaviors, and thus coordination was the only factor to align with human-human rapport theory.

Based on learners’ rapport trends and their degree of variation, we clustered them into groups. The *flat* cluster including p1, p3, p6 and p7 tended to adopt one favorable strategy at the beginning of interaction and stick with it over the course of the sessions. These users did not sense their interaction mode changes over time (“At first I already feel how I was with Emma, I just kept that going with the routine.” -p6). P2 and p9 were grouped as an *increasing* cluster and p5 as a single *decreasing* case. These users had not interacted with a robot or AI system before but had different expectations and perceptions towards Emma. For example, p2 believed Emma is more friendly than robots he saw in movies. Over the sessions, p2 praised Emma more (frequency from 2% to 8%) and his apologizing behavior disappeared. Similarly, p9 stopped *asking questions* in later sessions. The disappearance of social strategies means the student started to adhere to a “personal norm” (non-apology, no questions) rather than socio-cultural norms [18]. This is a sign that the relationship between “increasing” participants and Emma had moved to more friend-like [20] dyads. On the other hand p5, the decreasing case, had a low expectation of Emma’s intelligence and socialness (“She’s a robot...I don’t think she would have background information or whatever.”).

It is important to note that participants’ behaviors from session to session are not only due to their rapport states, but also to contextual factors such as energy or mood. For example, in session 3, p6 seldom offered further elaboration except for saying “Yes”, seeming bored with the problems or upset on that study day. In session 5, he was very engaged in the task, and was more wordy and responsive (“Yes, but we also can convert it into a decimal, which is 0.125.”).

3 Discussion

Our goal was to investigate how middle school learners manage rapport with robots over multiple tutoring sessions. We demonstrated that rapport changes from early to late interaction stages in human-robot tutoring did not follow the same trends as human rapport theory [19]. The variation between individuals on positivity and the increase of attentiveness over time suggested that users may

be shifting how they express rapport as their expectations of Emma change. This corresponded to the contrast between users with *flat* rapport trends, who tended to stick with the same linguistic strategies, and users with either *increasing* or *decreasing* rapport trends, who articulated evolving perceptions towards Emma. Given the cross-session variability of individuals and that the majority of rapport studies' focus on an "instant" rapport [6], it is critical to conduct multiple session studies to understand more about human-robot rapport dynamics. This work is a first step towards personalizing rapport-based learning experiences over long-term human robot interactions.

Acknowledgements. This work is supported by the National Robotics Initiative and the National Science Foundation, grant #CISE-IIS-1637809. We would like to thank Mesut Erhan Unal for his guidance in the ICA modeling process.

References

1. Ädel, A.: Rapport building in student group work. *J. Pragmat.* **43**(12), 2932–2947 (2011)
2. Beckmann, C.F., Smith, S.M.: Probabilistic independent component analysis for functional magnetic resonance imaging. *IEEE Trans. Med. Imaging* **23**(2), 137–152 (2004)
3. Bracewell, D.B., Tomlinson, M.T., Wang, H.: Identification of social acts in dialogue. Technical report, Language Computer Corporation Richardson United States (2012)
4. Delorme, A., Makeig, S.: Eeglab: an open source toolbox for analysis of single-trial eeg dynamics including independent component analysis. *J. Neurosci. Methods* **134**(1), 9–21 (2004)
5. Gordon, G., Breazeal, C., Engel, S.: Can children catch curiosity from a social robot? In: 2015 10th ACM/IEEE International Conference on Human-Robot Interaction (HRI), pp. 91–98. IEEE (2015)
6. Gratch, J., et al.: Virtual rapport. In: Gratch, J., Young, M., Aylett, R., Ballin, D., Olivier, P. (eds.) IVA 2006. LNCS (LNAI), vol. 4133, pp. 14–27. Springer, Heidelberg (2006). https://doi.org/10.1007/11821830_2
7. Gremler, D.D., Gwinner, K.P.: Rapport-building behaviors used by retail employees. *J. Retail.* **84**(3), 308–324 (2008)
8. Lakin, J.L., Chartrand, T.L.: Using nonconscious behavioral mimicry to create affiliation and rapport. *Psychol. Sci.* **14**(4), 334–339 (2003)
9. Lubold, N., Walker, E., Pon-Barry, H., Ogan, A.: Automated pitch convergence improves learning in a social, teachable robot for middle school mathematics. In: Penstein Rosé, C., et al. (eds.) AIED 2018. LNCS (LNAI), vol. 10947, pp. 282–296. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-93843-1_21
10. Lubold, N., Walker, E., Pon-Barry, H., Ogan, A.: Comfort with robots influences rapport with a social, entraining teachable robot. In: Isotani, S., Millán, E., Ogan, A., Hastings, P., McLaren, B., Luckin, R. (eds.) AIED 2019. LNCS (LNAI), vol. 11625, pp. 231–243. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-23204-7_20
11. McDaniel, B., D'Mello, S., King, B., Chipman, P., Tapp, K., Graesser, A.: Facial features for affective state detection in learning environments. In: Proceedings of the Annual Meeting of the Cognitive Science Society, vol. 29 (2007)

12. Michaelis, J.E., Mutlu, B.: Supporting interest in science learning with a social robot. In: Proceedings of the 18th ACM International Conference on Interaction Design and Children, pp. 71–82. ACM (2019)
13. Ogan, A., Finkelstein, S., Mayfield, E., D’Adamo, C., Matsuda, N., Cassell, J.: “Oh dear stacy!” social interaction, elaboration, and learning with teachable agents. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 39–48 (2012)
14. Ogan, A., Finkelstein, S., Walker, E., Carlson, R., Cassell, J.: Rudeness and rapport: insults and learning gains in peer tutoring. In: Cerri, S.A., Clancey, W.J., Papadourakis, G., Panourgia, K. (eds.) ITS 2012. LNCS, vol. 7315, pp. 11–21. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-30950-2_2
15. Park, H.W., Rosenberg-Kima, R., Rosenberg, M., Gordon, G., Breazeal, C.: Growing growth mindset with a social robot peer. In: 2017 12th ACM/IEEE International Conference on Human-Robot Interaction (HRI), pp. 137–145. IEEE (2017)
16. Pennebaker, J.W., Boyd, R.L., Jordan, K., Blackburn, K.: The development and psychometric properties of liwc2015. Technical report (2015)
17. Seo, S.H., Griffin, K., Young, J.E., Bunt, A., Prentice, S., Loureiro-Rodríguez, V.: Investigating people’s rapport building and hindering behaviors when working with a collaborative robot. *Int. J. Social Robot.* **10**(1), 147–161 (2018)
18. Spencer-Oatey, H.: (Im) politeness, face and perceptions of rapport: unpackaging their bases and interrelationships (2005)
19. Tickle-Degnen, L., Rosenthal, R.: The nature of rapport and its nonverbal correlates. *Psychol. Inq.* **1**(4), 285–293 (1990)
20. Zhao, R., Papangelis, A., Cassell, J.: Towards a dyadic computational model of rapport management for human-virtual agent interaction. In: Bickmore, T., Marsella, S., Sidner, C. (eds.) IVA 2014. LNCS (LNAI), vol. 8637, pp. 514–527. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-09767-1_62