The Effect of First Impressions on Human Neural and Hormonal Responses during Longitudinal Human-Robot Teaming using fNIRS and Oxytocin

Yigit Topoglu[1,2], Kevin L. Ramirez-Chavez[1], Frank Krueger[3], Jonathan Gratch[4], Adrian A. Franke[5], Xingnan Li[5] ,Ewart J. de Visser[1,2], Hasan Ayaz[1,6-9]

[1]School of Biomedical Engineering, Science and Health Systems, Drexel University, Philadelphia, PA, USA.
[2]Warfighter Effectiveness Research Center, U.S. Air Force Academy, Colorado Springs, CO, USA.
[3]School of Systems Biology, George Mason University, Fairfax, VA, USA.
[4]Institute for Creative Technologies, University of Southern California, Los Angeles, CA, USA.
[5]University of Hawaii Cancer Center, Honolulu, HI, USA
[6]Department of Psychological and Brain Sciences, College of Arts and Sciences, Drexel University, Philadelphia, PA, USA.
[7]Drexel Solutions Institute, Drexel University, Philadelphia, PA, USA.
[8]A.J. Drexel Autism Institute, Drexel University, Philadelphia, PA, USA.
[9]Center for Injury Research and Prevention, Children's Hospital of Philadelphia, Philadelphia, PA, USA.

Synopsis

This study examines how human neural and hormonal activity responds to dynamic robot performance during longitudinal human-robot collaborations. Participants interacted with a humanoid robot in matching pairs and spelling bee tasks. During the sessions, participants' prefrontal cortex (PFC) activity and salivary oxytocin (OT) were recorded. Our results show that robots' first impression matters, where erroneous-first exposure diminished neural engagement of subsequent congruent performance, while congruent-first exposure mitigated negativity bias of subsequent erroneous performance on perceived rapport and prosociality. The study demonstrates that neural and hormonal biomarkers can provide valuable insights to improve longitudinal human-robot teaming.

Background

One of the biggest challenges in the human-robot interaction (HRI) field is to design autonomous adaptive agents that can form teams with humans in long-term [1; 2]. Similar to human-human teams, in order to have more positive human-robot teaming (HRT), longitudinal engagements that contain dynamic elements such as trust, rapport, and performance are essential [1]. However, there is still limited understanding on how human brain and behavior operates in HRT settings over time and how robot's first impression effects HRT, more so in naturalistic contexts. To assess human brain, functional near-infrared spectroscopy (fNIRS) can be utilized as a non-invasive neuroimaging technique which makes it suitable to be utilized in HRT settings [3]. Alongside neural measures, hormones such as OT can also provide insights on one's social attachment towards others [4; 5]. In this study, we examined how exposure order of robot performance influenced user's neural and hormonal responses over time during naturalistic human-robot teaming using fNIRS and OT.

Methods

Twenty-four adult participants (mean ± SD = 25.63 ± 4.11) participated in the study. Participants interacted with Pepper, a humanoid robot (Softbank Robotics) in two computerized collaborative tasks over the span of four back-to-back days (Fig. 1a). The participants were divided into two groups regarding exposure order, where one group interacted with the congruent robot in the first two days, and then the erroneous robot in the other two days, while the other group interacted with robot in the reverse order (Fig. 1b). In the matching the pair task, participant and robot took turns to match as many pairs as possible in a set of cards (Fig. 1c). In the spelling bee task, participant and the robot took turns spelling the given words letter by letter and spelled as many words as possible (Fig. 1d). During all sessions, participants' PFC activity was recorded via fNIRS, and salivary OT

with self-reported surveys regarding perceived rapport and prosociality were collected pre- and post-sessions. In all sessions, Pepper had two robot performances: In the congruent performance, Pepper made minimal errors, while in the erroneous performance, Pepper made multiple mistakes, which consists making wrong decisions.

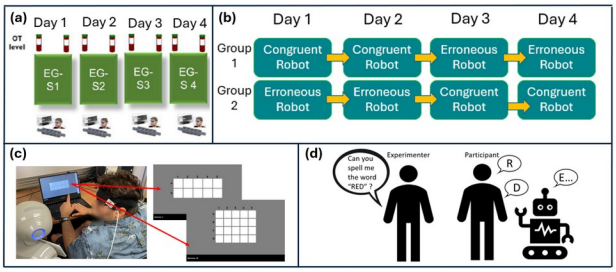

Fig. 1 (a) The experimental protocol. **(b)** The distribution for each session in two different exposure orders regarding robot performance. **(c)** Matching pairs task. **(d)** Spelling bee task.

Results

For self-reported measures (Fig. 2a), we found that in the erroneous-first group, perceived rapport ($F_{3,66}$ = 8.72, p < 0.001) and prosociality ($F_{3,66}$ = 5.53, p = 0.002) increases in the subsequent congruent performance, while there were no changes in the erroneous-first group. For OT (Fig. 2b), we found that in OT increased after the second congruent session in both groups (Congruent-first: $F_{3,250.6}$ = 7.76, p = 0.023; Erroneous-first: $F_{3,284.3}$ = 6.60, p = 0.043). For PFC activity in matching pairs task (Fig. 2c), we found significant interaction between exposure order and session in L dlPFC ($F_{3,354}$ = 3.21, p = 0.033), L mPFC ($F_{3,354}$ = 3.16, p = 0.033), and R dlPFC ($F_{3,354}$ = 3.65, p = 0.033). Post-hocs showed a significant decrease during the last session compared to the first one in the left dlPFC only for the congruent-first group ($F_{1,354}$ = 12.91, p = 0.004). For PFC activity in spelling bee task (Fig. 2d), we found significant interaction between exposure order and session for all PFC regions (L dlPFC: $F_{3,354}$ = 3.98, p = 0.011; L mPFC: $F_{3,354}$ = 5.45, p = 0.005; R mPFC: $F_{3,354}$ = 4.83, p = 0.005; R dlPFC: $F_{3,354}$ = 3.18, p = 0.024). Post-hocs showed that only for congruent-first group there was a significant decrease in the first erroneous (3rd) session compared to the first congruent (1st) session in L dlPFC ($F_{1,354}$ = 8.46, p = 0.046), L mPFC ($F_{1,354}$ = 9.33, p = 0.029and R mPFC ($F_{1,354}$ = 8.44, p = 0.047) and in the last session compared to the first session in mPFC (L mPFC: $F_{1,354}$ = 10.47, p = 0.016; R mPFC: $F_{1,354}$ = 12.24, p = 0.006).

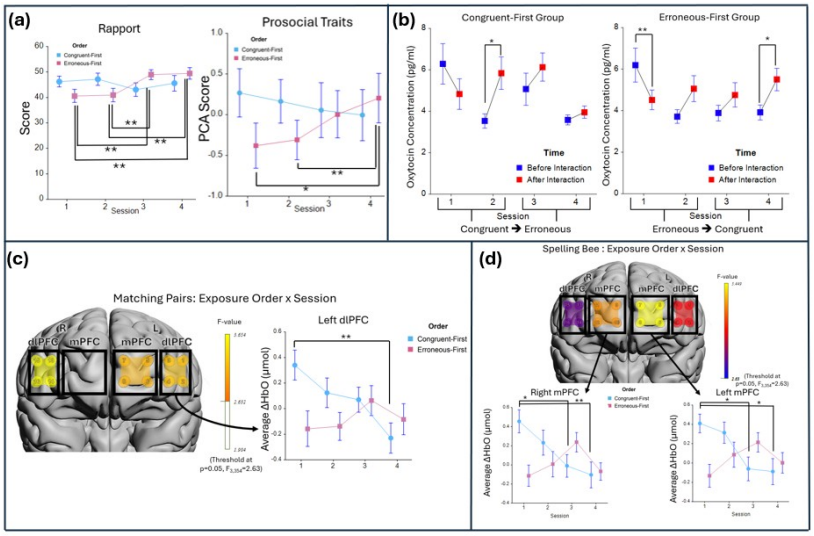

Fig. 2 (a) Perceived rapport and robot's prosociality. **(b)** Oxytocin levels. **(c)** PFC activity for the matching pairs task. **(d)** PFC activity for the spelling bee task.

Discussion

Our results show that the first impressions of robots affect on the user's neural, hormonal and psychologica response. In terms of self-reported measures, when the robot started congruent but made mistakes afterwards, changes on perceived rapport and prosociality towards the robot was less severe [6]. For OT, users tend to be guarded towards the erroneous robot in the beginning; however, as users spend more time with congruent robot, they also tend to build more social salience [7]. For neural activity, we showed that robot's initial competency matters, where for the congruent-first group, PFC activity increased at the start, possibly indicating task engagement [8]; however, PFC activity decreased as the robot shows errors in subsequent sessions. On the other hand, for erroneous-first group, there were no changes in the neural activity, possibly indicating a negativity bias [9]. This study demonstrated that neural and hormonal biomarkers can provide valuable insights on the human-robot teaming dynamics in longitudinal settings.

Disclosures

Funding from Air Force Office of Research. Funding #: FA9550-18-1-0455. The views expressed are those of the authors and do not reflect the official guidance or position of the U.S. Government, Department of Defense, U.S. Air Force, or U.S Space Force. The content or appearance of hyperlinks does not reflect the official position of the Department of Defense, U.S. Air Force, U.S. Space Force, or Air Force Research Laboratory, or endorsement of the external websites or the information, products, or services contained therein. HA has equity in fNIR Devices LLC, the company that licensed IP from Drexel University and manufactures the optical brain imaging sensor used in the study.

References

[1] E.J. de Visser, M.M.M. Peeters, M.F. Jung, S. Kohn, T.H. Shaw, R. Pak, and M.A. Neerincx, Towards a Theory of Longitudinal Trust Calibration in Human–Robot Teams. International Journal of Social Robotics 12 (2020) 459-478.

[2] G.-Z. Yang, J. Bellingham, P.E. Dupont, P. Fischer, L. Floridi, R. Full, N. Jacobstein, V. Kumar, M. McNutt, and R. Merrifield, The grand challenges of Science Robotics. Science robotics 3 (2018).

[3] H. Ayaz, and F. Dehais, Neuroergonomics: The brain at work and in everyday life, Academic Press, 2019.

[4] M. Kosfeld, M. Heinrichs, P.J. Zak, U. Fischbacher, and E. Fehr, Oxytocin increases trust in humans. Nature 435 (2005) 673-676.

[5] J.A. Bartz, J. Zaki, N. Bolger, and K.N. Ochsner, Social effects of oxytocin in humans: context and person matter. Trends in Cognitive Sciences 15 (2011) 301-309.

[6] J. Xu, and A. Howard, The Impact of First Impressions on Human- Robot Trust During Problem-Solving Scenarios, 2018 27th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN), 2018, pp. 435-441.

[7] S.G. Shamay-Tsoory, and A. Abu-Akel, The social salience hypothesis of oxytocin. Biological Psychiatry 79 (2016) 194-202.

[8] R. McKendrick, R. Mehta, H. Ayaz, M. Scheldrup, and R. Parasuraman, Prefrontal Hemodynamics of Physical Activity and Environmental Complexity During Cognitive Work. Human Factors 59 (2017) 147-162.

[9] J.B. Lyons, I.a. Hamdan, and T.Q. Vo, Explanations and trust: What happens to trust when a robot partner does something unexpected? Computers in Human Behavior 138 (2023) 107473.