Neuroergonomics, 2021

Mid-frontal theta oscillations as an index of cognitive conflict in a human-robot classification task

Abdulaziz Abubshait, Jairo Perez-Osorio & Agnieszka Wykowska

Italian Institute of Technology

Introduction. As technological advances improve, humans find themselves in situations where they interact with different machines in a collaborative manner. For instance, autonomous systems are used in search and rescue missions, space exploration and as decision aids (Hancock et al., 2011). While much of the research on human-robot interaction examines the dynamics of human robot trust, little is known about the effects of conflicting robot signals on the performance of human-robot teams. Understanding how conflicting robot signals influence human-robot interaction is of utmost importance since prior work has shown that robot signals can be interpreted as intentional (i.e., the signal was motivated by an internal state such as thoughts and feelings; e.g., Wykowska et al., 2014; Abubshait & Wiese, 2017). The present study examines the effects of conflicting robot cue signals during a handover-classification task. Specifically, we ask if neural oscillations that are correlated with conflict processing are observed when participants view correct or incorrect robot signals. Here, participants were asked to classify different objects based on their color after a humanoid robot (iCub; Metta et al., 2008), presented on a screen, simulated handing over the object to them. The robot proceeded to cue participants to the correct or incorrect target location. Since prior work has shown that social cues (i.e., gaze-cues) can interfere with oculomotor planning and influence responses mapping (e.g., Dalmaso et al., 2020; Hietanen, 2018), which can induce conflict, we expected that incongruent robot social signals' would interfere with the execution of actions and as such induce neural correlates of cognitive conflict as measured by mid-brain theta oscillations (Cohen & Donner, 2013; Cavanagh & Frank, 2014). Specifically, we hypothesized that mid-brain theta oscillations would be observed when participants see incongruent head movements.

Methods and materials. 35 participants saw one of 4 objects (i.e., a 100% blue object, a 75% blue and 25% yellow object, a 100% yellow object, and a 75% yellow and 25% object). After seeing the object, participants decided if it belonged to one of two categories (i.e., the blue or yellow category). Each trial started with a fixation presented for 1000 ms. Next, bins appeared at the top corners of the screen and remained there until the end of the trial. 1000 ms later, iCub appeared in the middle of the screen, looking straight with one of the 4 objects in front of it. After looking straight for 400 ms, iCub looked down at the object for 1000 ms. Next, iCub reached for the object and grasped it. 1000ms later, iCub extended its arm to indicate simulated "handing over" the object to the participant. 500 ms later, iCub looked either at the correct bin (i.e., congruent) or at the incorrect bin (i.e., incongruent). Participants were instructed to wait until iCub completed the head movement to categorize the object by pressing one of two keys that corresponded to the bins. If participants did not respond within 3000 ms, the trial timed out. After categorizing the object based on color, participants received "Correct" or "Incorrect" feedback. The inter-trial interval was set to 1000 ms. The experiment had 420 trials overall and took 35 minutes to complete.

Results & Discussion. EEG data was preprocessed by down-sampling the signal to 250 Hz, re-referencing to the average of the two mastoids, applying a band filter of .01-30 Hz. Afterwards, data were epoched 1000 ms prior to the onset of the robot's head movements and 7500 ms after. We then applied ICAs and rejected components that were clearly related to artifacts (e.g., eye blinks and saccades). Finally, artifact we rejected artifacts that had a maximal voltage of 20 µV/ms, a 200 µV difference in value, or low activity of .5 µV over a period of 100 ms. Prior to applying the wavelets, we re-epoched the data 1000 prior to the head cues to 5000 ms after the head cues. Theta-oscillations were time locked to the onset of the head cue and decomposed using Morlet Complex waveforms between 1-30 Hz. Averaged Event-Related Spectral Perturbation (ERSPs) were exported by finding the maximal voltage of all conditions for the entire duration of the epoch in which the power spectrum peaked for theta (i.e., 4-7 Hz) and using a 100 ms window around where theta oscillations peaked (i.e., 248 ms; from: 198 to 298 ms). A repeated measures 2 x 2 ANOVA with *Congruency* (congruent vs. incongruent) and *Object type* (100% color vs. 75% color) as factors was conducted to examine differences in theta oscillations. Analyses showed a significant effect of Congruency ($F(1, 33) = 48.92$, $\eta 2G = .20$, $p < .001$), with higher ERSP amplitudes in the theta band for incongruent vs. congruent signals (1.36 µV vs. .73 µV; see **Figure 1**). The analysis revealed no significant effect of Object type ($F(1, 33) = 1.51$, $\eta 2G = .002$, $p = .22$) or an interaction ($F(1, 33) = .76$, $\eta 2G < .002$, $p = .39$). Time-frequency analysis showed that viewing incongruent head movements elicited higher conflict-related theta amplitudes than viewing congruent signals, which mimic prior findings of error related ERPs (Chavarigga et al., 2014). The analysis did not show an effect of object type on theta amplitudes. This could be due to the fact that theta oscillations index cognitive conflict in general, but are not sensitive to the degree of conflict. These findings suggest, that robot signals can induce cognitive conflict when interacting with humans, which is important for HRI. The findings of this study illustrates how robot signals are important considerations in human-robot collaboration settings as incongruent robot signals can influence how people perform on a specific task, even when the robot signals are unrelated to the task.

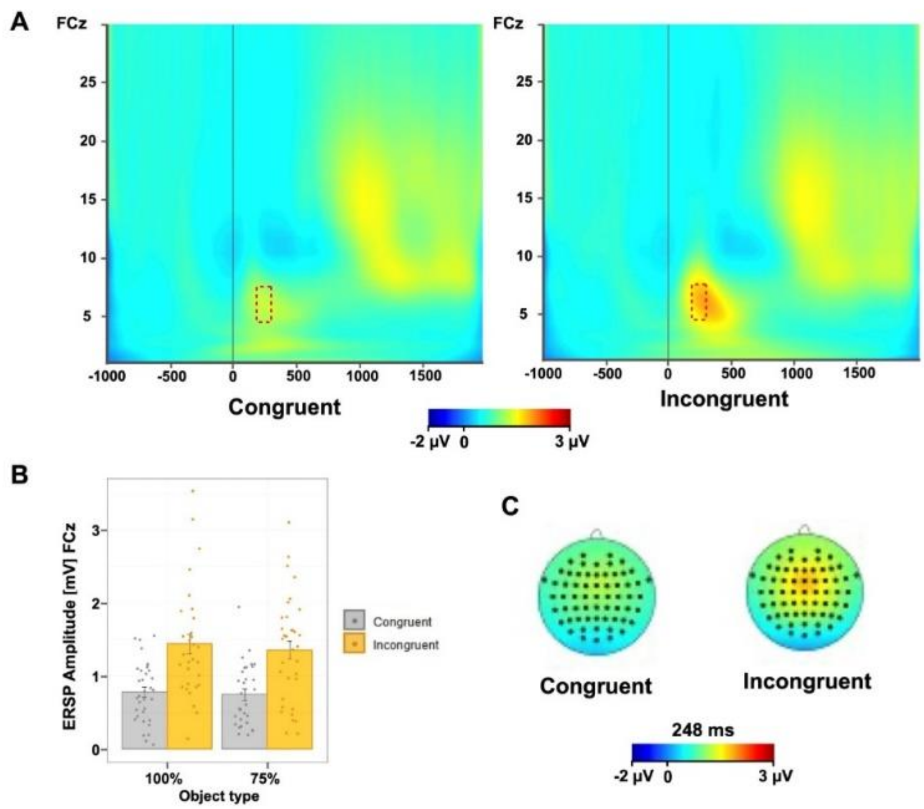

Figure 1. ERSP amplitudes time-locked to the onset of the head cue: The graph illustrates ERSP amplitudes at FCz that were time-locked to the robot's head cue (A). Analysis of the ERSP amplitudes showed larger frontal theta amplitudes for incongruent vs. congruent signals (B). The graph also illustrates the topographies for each condition, which suggests the mid-frontal origin of theta frequencies (C).

References.

Abubshait, A., & Wiese, E. (2017). You look human, but act like a machine: Agent appearance and behavior modulate different aspects of human–robot interaction. Frontiers in Psychology, 8.

Cavanagh, J. F., & Frank, M. J. (2014). Frontal theta as a mechanism for cognitive control. *Trends in cognitive sciences*, *18*(8), 414-421.

Chavarriaga, R., Sobolewski, A., & Millán, J. D. R. (2014). Errare machinale est: the use of error-related potentials in brain-machine interfaces. *Frontiers in neuroscience*, *8*, 208.

Cohen, M. X., & Donner, T. H. (2013). Midfrontal conflict-related theta-band power reflects neural oscillations that predict behavior. *Journal of neurophysiology*, *110*(12), 2752-2763.

Dalmaso, M., Castelli, L., & Galfano, G. (2020). Social modulators of gaze-mediated orienting of attention: A review. Psychonomic Bulletin & Review.

Hancock, P. A., Billings, D. R., Schaefer, K. E., Chen, J. Y., de Visser, E. J., & Parasuraman, R. (2011). A meta-analysis of factors affecting trust in human-robot interaction. *Human factors*, *53*(5), 517-527.

Hietanen, J. K. (2018). Affective Eye Contact: An Integrative Review. Frontiers in Psychology, 9, 1587.

Wykowska, A., Wiese, E., Prosser, A., & Müller, H. J. (2014). Beliefs about the minds of others influence how we process sensory information. PLoS ONE, 9(4), e94339.