Trust Assessment with EEG Signals in Social Human-Robot Interaction

Campagna, Giulio; Rehm, Matthias

Published in:
Proceedings 15th International Conference on Social Robotics (ICSR 2023)

DOI (link to publication from Publisher):
10.1007/978-981-99-8715-3_4

Publication date:
2024

Document Version
Accepted author manuscript, peer reviewed version

Citation for published version (APA):
Campagna, G., & Rehm, M. (2024). Trust Assessment with EEG Signals in Social Human-Robot Interaction. In A. A. Ali, J.-J. Cabibihan, N. Meskin, S. Rossi, W. Jiang, H. He, & S. S. Ge (Eds.), *Proceedings 15th International Conference on Social Robotics (ICSR 2023)* (pp. 33-42). Springer. https://doi.org/10.1007/978-981-99-8715-3_4

Trust Assessment with EEG Signals in Social Human-Robot Interaction ⋆

Giulio Campagna[1] [0000−0002−2422−1663] and Matthias Rehm[1] [0000−0001−6264−1688]

Aalborg University, Aalborg, Denmark
{gica,matthias}@create.aau.dk

Abstract. The role of trust in human-robot interaction (HRI) is becoming increasingly important for effective collaboration. Insufficient trust may result in disuse, regardless of the robot's capabilities, whereas excessive trust can lead to safety issues. While most studies of trust in HRI are based on questionnaires, in this work it is explored how participants' trust levels can be recognized based on electroencephalogram (EEG) signals. A social scenario was developed where the participants played a guessing game with a robot. Data collection was carried out with subsequent statistical analysis and selection of features as input for different machine learning models. Based on the highest achieved accuracy of 72.64%, the findings indicate the existence of a correlation between trust levels and the EEG data, thus offering a promising avenue for real-time trust assessment during interactions, reducing the reliance on retrospective questionnaires.

Keywords: Trust in Human-Robot Interaction · Social Robotics · Human-Robot Collaboration.

1 Introduction

In social scenarios, trust is an important aspect of human robot collaboration because the level of trust of the human towards the robot can seriously affect the performance during the interaction. In other words, it can cause unbalance workload, inefficient monitoring of the robot, or even disuse of the system [4]. For example, socially assistive robots provide assistance to elderly people for improving their quality of life and independence [22] or serving as robot companions for activating and stimulating the users [27]. In such scenarios, fostering trust in the robot is crucial not only for the user but also for care personnel and relatives, ensuring a successful interaction.

Numerous attempts have been investigated to define the characteristics of trust as an emergent phenomenon in human-human as well as human-robot interaction (HRI). A meta-review of trust-related studies of human robot interaction [15] showed that the two main variants of trust are performance-based and relation-based, which in turn are related to specific domains of applications. Performance-based trust is mostly explored in manipulative robot systems in industrial contexts. By contrast, relation-based trust is more relevant for social scenarios focusing on communication rather than on manipulation of objects.

⋆ The work described in this paper was funded by the Independent Research Fund Denmark, grant number 1032-00311B.

To be able to use trust as a control parameter in the context of human-robot collaboration, evaluating real-time trust levels throughout the interaction becomes essential. As result, the behavior of the robot can be adapted to match the trust level of the user, increasing efficiency or preventing potentially dangerous situations. The current trust measurements are based on subjective, post-hoc questionnaires (e.g. [25, 28]) that provide only summary information about the subjects' interpretation of the interaction but do not support a continuous trust assessment during the collaboration. For this reason, new methods must be developed for data-driven assessment of trust. Recent studies have shown some promising steps in this direction (e.g. [7, 9, 26, 30]). For instance, Hu et al. [9] successfully experimented with electroencephalogram (EEG) signals as a real-time measurement of trust for a scenario based on performance-based trust. While their results are promising, the trust rating that was used for labeling the data was an oversimplification relying only on one question.

This paper proposes trust assessment based on **EEG signals** and **Machine Learning** (ML) within the context of relation-based trust. For the purpose of having a solid basis for labeling the EEG data, the *Multi-dimensional Measure of Trust* (MDMT) questionnaire [28] is used as the baseline measurement to infer the trust levels. The decision to opt for MDMT hinges on the necessity for a more in-depth comprehension of trust. Trust, being multifaceted, is thoroughly examined by MDMT across a spectrum of dimensions, allowing to gain a more holistic insight into trust-related constructs. The aforementioned approach allows for the automatic labeling of the data, facilitating the subsequent real-time analysis of the correlation between trust and physiological responses, thereby overcoming the limitations associated with questionnaire responses.

2 Related Work

Trust can be defined as the operator's perception of the competence of the machine where it is essential that the operator is confident that the system appropriately accomplishes its tasks [18]. A correlation is present between the operator's trust in the machine and his willingness to use it, i.e. the more the human trusts the system, the more he is likely to use it. In [18], Muir and Moray reported that the level of trust of the human subject in a machine was heavily affected by the machine's performance. As a matter of fact, trust is a relevant factor that could affect the acceptance of a robot as assistant, co-worker or companion in social scenarios [12, 16]. Moreover, it can influence the human's perception of the capabilities of a robot [8, 23].

In HRI, the concept of trust is a timely and relevant topic for various reasons. First of all, there is a lack of a general understanding of the dynamic nature of trust and the methodologies to study it [14]. Trust is not a static phenomenon: it can be built, repaired, adjusted and it changes over time according to events that occur. Secondly, trust is an essential feature of human decision-making in collaboration tasks and it becomes crucial when robots are closely working together with human users [29]. Thirdly, mismatches in trust towards a robot can lead to severe consequences, ranging from unbalance workload, to loss of expensive equipment or even human life due to inefficient monitoring [4, 19]. Therefore, both under-trust and over-trust in robots have to be avoided.

Most trust assessments rely primarily on post-hoc questionnaires (e.g. [10, 25]). Only a limited number of methods have endeavored to gauge trust through performance, with a primary emphasis on task success and efficiency. Floyd et al. [6], for instance, estimated the robot's trustworthiness based on observable performance which was calculated as a comparison between the number of successful task completions and failed or interrupted attempts. Xu and Dudek [30] proposed a data-driven approach to infer trust levels, focusing specifically on a performance-centric definition of trust, which evaluated success or failure in performing the task. A HRI trust scale was developed by Yagoda et al. [31] that is based on a list of item dimensions including HRI attributes: team configuration, team process, context, task, and system. Salem et al. [24] explored the factors that affect human perception and trust towards an erroneous robot.

The post-hoc questionnaires are not sufficient to promptly determine the change of trust level over time. Therefore, it is essential to track the unfolding of trust (and distrust) in HRI to define reliable, real-time measures of trust [9] with the goal of adapting the robot's behavior to the user's trust level. According to [21], the use of psycho-physiological signals could be a solution to sense trust level. Among the several psycho-physiological measurements, EEG is identified as a non-invasive and convenient method for capturing brain signals and observing brain activity in response to a specific event-related potential (ERP). Researchers have conducted studies on trust by means of EEG signals. Hu et al. [9] introduced an initial attempt to establish a connection between real-time physiological signals and human-machine trust. Regrettably, their trust measurement remains superficial as it solely relies on asking participants about their trust in the system. In a coin toss experiment that simulated trust and distrust [3], Boudeau et al. found that ERP components had different peak amplitudes for the several participants involved. Akash et al. [2] analyzed approaches to develop a classification model to sense human trust using EEG and galvanic skin response measurements.

3 Methodology

3.1 The Human Subject Study

To assess the trust level of the human subjects based on EEG signals, a social experimental scenario was designed. A human and *EZ-robot JD Humanoid*[1] played a collaborative game inspired by the board game Activity. As shown in Fig. 1, the game consisted of five sections where, depending on the section, the humanoid robot either mimed or vocally described a word among four options presented to the participant. Throughout the interaction, the participant had to guess which word the robot was presenting. In case of correct answer, the participant received a candy as reward; otherwise one candy had to be returned. Each section was composed of two trials where, for each trial, a different set of four words were presented by the robot. Furthermore, each section implemented a different trust strategy that determined how the robot would behave. The first two sections were used for building up trust. Section 2 showed some situation awareness, e.g. by commenting on the participant's cloth with the purpose to establish

[1] https://www.ez-robot.com/

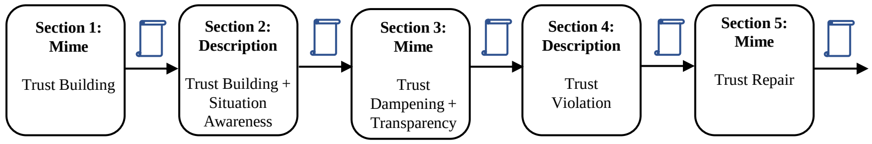

Fig. 1. Sequence of the sections with the different trust strategies associated.

a relation. In section 3 the robot introduced the participant to some technical limitations of the system (trust dampening + transparency). Section 4 violated trust by a deliberate malfunction of the robot while section 5 was used to regain trust. In the end, between each section and after the last section of the experiment, the MDMT questionnaire was used to assess the average trust score. The participants were asked to rate on a 8-point scale (0-7) how well some trust-based descriptors applied to the robot. Finally, an average trust score was calculated over the several dimensions measured by the MDMT. To conclude, it is essential to emphasize that the initial three sections were employed to establish a solid foundation of trust between the participants and the robot prior to any violation of trust taking place. Consequently, Section 3 serves as a comprehensive synthesis, bringing together all the trust-related insights derived from the earlier stages.

The experiment protocols were in accordance with the *Declaration of Helsinki*. Ethical approval was obtained from the institutional review board prior to the study. Twenty-one participants were recruited, 9 male and 12 female with an average age of 28.3 (*SD*=9.94). The sample size of 20 participants was chosen to initiate a preliminary assessment of trust levels, considering resource limitations and the exploratory nature of the study. Familiarity with robots was limited, three had previous practical experience with robots, while the remaining had only encountered them either in reality (12 participants) or through media (6 participants). The completion of the experiment required 30 minutes for each participant.

3.2 Experimental Setup

At the beginning of each session, a consent form and a description of the task were provided and, subsequently, the participant was equipped with an EMOTIV EPOC+ 14-Channel Wireless EEG Headset[2] connected to a computer to record brain waves with a rate of 128 Hz. To ensure a good contact quality, a saline liquid was applied to each electrode in order to have an efficient conductivity. A sensor map was used to check the location and contact quality of each sensor. The placement of the 14 electrodes is reported in Fig. 2 (left). Furthermore, participants were given instructions to avoid making pronounced movements to prevent any sensor shifts. The participant was seated on a chair facing a table with the robot (Fig. 2 right), which was controlled through the EZ-Builder software. In order to have a controlled data collection, a *Wizard of Oz protocol* was adopted, where the robot was manually controlled throughout the procedure. The participant remained unaware of this fact until the debriefing.

[2] https://www.emotiv.com/

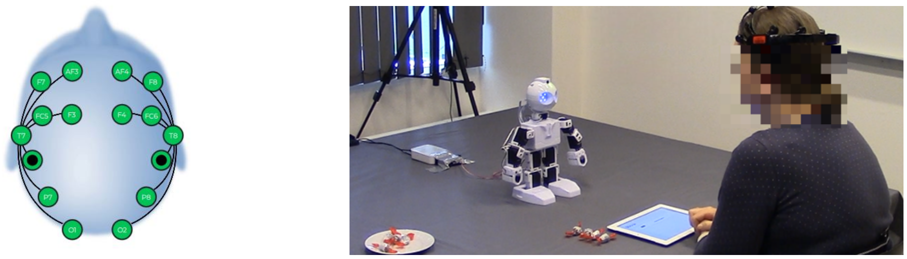

Fig. 2. Left: Placement of EEG electrodes. Right: Experimental scenario with the Humanoid Robot JD and the participant wearing the EEG headset.

3.3 Data Pre-Processing

The EEG data was pre-filtered to remove the DC offset with 0.16 Hz high-pass cutoff. Before computing the *Fast Fourier Transform* (FFT), the EMOTIV Cortex API was used to *minimize artifacts* in the EEG signal. EEG waves can be affected by intrinsic artifacts (e.g. movements of the eye, eye blinks, bio-electric potentials from muscle and heart) and by external artifacts (e.g. environmental noise). Moreover, a *Hanning window* function was applied to obtain good frequency resolution and leakage protection with fair amplitude's accuracy. In order to analyze the alpha frequency band power, the continuous time series EEG data were transformed through FFT to assess the involved frequencies. Afterwards, the *power spectral density* was computed to determine the power of each band. Subsequently, the extraction of the power, expressed in μV^2, was performed from the following frequency bands: Theta (4-8 Hz), **Alpha (8-12 Hz)**, Low Beta (12-16 Hz), High Beta (16-25 Hz) and Gamma (25-45 Hz). Considering the nature of the task, the **alpha brain waves** were analyzed since previous studies have shown a strong correlation between **attention** and this typology of waves. In [13], it is reported that alpha suppression reflects attentional processes. The **hypothesis** is that if the robot committed errors during the performance, the human would less trust the robot, thus paying more attention in order to correctly guess the answer.

3.4 Data Analysis

For the trust assessment, the focus lay on the break point between section 3 and 4, where a trust violation occurs. Firstly, it was necessary to validate whether the trust scores from the MDMT were impacted by the trust violation. If that was the case, the data collected in sections 3 and 4 could be utilized for training a trust assessment model, incorporating the labels derived from the MDMT. A *Shapiro-Wilk test* revealed that the difference in average trust scores between section 3 and section 4 did not follow a normal distribution (p=0.001). Therefore, instead of paired t-test, a *Wilcoxon Signed-Rank Test* was performed. The statistical test reported a significant difference in trust scores between section 3 (M=5.34, SD=1.31) and section 4 (M=4.64, SD=1.79). Specifically, the trust scores in section 3 (Mdn=5.45) were significantly higher than those in section 4 (Mdn=5.07), W=26, p=0.002. The results verified the usability of the trust scores for training the ML models.

Table 1. Features related to alpha frequency band power.

Features	Description
Mean Value	Average value of the power
Peak	Maximum value of the power
Standard Deviation	Dispersion of the power relative to its mean
Kurtosis	Sharpness of the peak

4 Trust Assessment

The categorization of the trust level was modeled as a **binary classification problem** (i.e. *high trust*, *low trust*). The ML models were based on features of the alpha frequency band power. For each participant, the contribution of the alpha frequency band power of each sensor was calculated and then averaged over the 14 channels. To **extract** the **features**, a *window size* of *1 second* (consisting of 128 data points) was defined. The initial features calculated on these windows were *mean*, *peak*, *median*, *standard deviation* and *kurtosis*. Based on the results of the MDMT questionnaire, the feature vectors were then labeled as one of the two classes, i.e. either as high or low trust. The labels were determined using the average of the results from the MDMT questionnaires as threshold for both section 3 (M=5.34) and section 4 (M=4.64). By applying *Univariate Feature Selection method*, the feature space was reduced by eliminating the median. The resulting features are summarized in Table 1. According to [17], classification algorithms are more suitable than regression models in brain computer interface applications. In this analysis, several supervised ML algorithms were selected to categorize the trust level of the participants. The adopted models were **Support Vector Machine** (SVM), **k-Nearest Neighbors** (kNN) and **Random Forest**. The data were normalized, shuffled and divided in 70% for the training set and 30% for the testing set. Tuning of hyperparameters was performed during the modeling phase. In the following, each algorithm is briefly presented along with the chosen evaluation metric (*classification accuracy*).

Support Vector Machine With reference to [20], SVM is a suitable model to classify physiological data. It is a discriminative classifier whose purpose is to provide a hyperplane in a N-dimensional space (N corresponds to the number of features) that distinctly classifies the data points of a binary classification problem. Many hyperplanes can be chosen and the selection depends on the maximum achieved margin, which is the maximum distance between the support vectors, i.e. the data points of the two classes closer to the hyperplane. Maximizing the margins aims to provide a wider confidence interval for classifying new data points into one of the two regions in the space, based on their respective class memberships. To determine the optimal hyperplane (i.e. decision boundary), the hyperparameters must be computed through (1):

$$\min_{w,b,\xi} \frac{1}{2} w^{\mathrm{t}} w + \frac{C}{n} \sum_{i=1}^{n} \xi_{\mathrm{i}} \quad \textit{for } i = 1,...,n \tag{1}$$

where w and b are hyperparameters, C is a tune parameter, n is the number of training samples and ξ_{i} is a variable that measures the extent of violation of constraint (2):

$$y_\mathrm{i}(\langle w, x_\mathrm{i}\rangle + b) \geq 1 - \xi_\mathrm{i} \tag{2}$$

where x_i is the training set while y_i are the categories. To conclude, SVM robustly mitigates overfitting due to the margin maximization, support vector utilization and regulation parameter control. In this analysis, the two regions of the space represented respectively the high trust and the low trust categories of the participants. The *regulation term C* was assigned a value of *0.5* and the *radial basis function* was selected as the *kernel*. The algorithm performs with 63.20% accuracy.

k-Nearest Neighbors kNN is a memory based classifier that exploits the similarity of the features between classes in order to predict the class of a new feature vector. kNN estimates the likelihood that a new data point belongs to a specific class (high or low trust) based on which class most of the data points closest to the new feature vector belong to. A distance function is used to determine the similarity between a new data point and its nearest neighbors [11], which is frequently the standard Euclidean distance (3):

$$d(x_\mathrm{i}, y_\mathrm{i}) = \sqrt{\sum_{i=1}^{d}(x_\mathrm{i} - y_\mathrm{i})^2} \tag{3}$$

where x_i are the unclassified samples, y_i are the labeled data and d is the dimension of the feature space. kNN is considered robust to overfitting due to its local nature and lack of assumptions about data distribution. Additionally, the choice of the hyperparameter k (*number of neighbors*) helps prevent overfitting. A number of *seven* neighbors was determined through experimentation, which resulted in an accuracy of 68.86%.

Random Forest Random Forest utilizes ensemble learning, a technique that combines many classifiers to make predictions. It consists of a large number of decision trees on various subsets of the given dataset. Each individual tree derives a class prediction and the class with the most votes is the output of the algorithm [5]. The robustness of Random Forest to overfitting is attributed to the random selection of features and sample subsets for each tree. This randomness ensures that no single tree have excessive influence over the ensemble, promoting generalization and reducing the risk of overfitting to unseen data. This model achieved the highest accuracy of 72.64% by utilizing *100 decision trees* and employing *Gini criterion* as the quality measure for evaluating splits in each tree.

5 Discussion

The data analysis indicated a statistically significant difference concerning the trust scores between the two analysed sections of the experiment (p=0.002). Therefore, considering that each of the two sections had a distinct trust strategy (i.e. the robot's performance varied), it can be concluded that the robot successfully elicited a noticeable shift in trust among the participants. Thus, the hypothesis about perceptible variations

Table 2. Supervised learning models and performance indicators.

Supervised Learning Model	Accuracy	Precision	Recall	F1-score
Support Vector Machine	63.20%	0.63	0.63	0.63
k-Nearest Neighbors	68.86%	0.70	0.69	0.69
Random Forest	72.64%	0.75	0.73	0.72

in trust scores around the break point when trust violation occurred has been validated. Based on this, three different algorithms for trust categorization were examined. The analysis involved various performance indicators, namely *accuracy*, *precision*, *recall*, *F1-score*. The corresponding outcomes are presented in Table 2. The *Random Forest* model achieved a higher accuracy (72.64%), likely attributed to its ensemble nature, indicating a greater number of correct predictions compared to the other models. Its elevated precision suggested a lower incidence of false positives, which is essential for minimizing misclassifications. Moreover, the model demonstrated a higher recall value, signifying its ability to identify a larger proportion of actual positive instances, i.e. high sensitivity. The second-best performing model was *kNN*. Its competitive performance might be attributed to its simplicity and the absence of strong assumptions about the underlying data distribution. Lastly, *SVM* exhibited low performance possibly due to the complexity of the dataset and limited feature space. SVM's performance might improve with more diverse and higher-dimensional data. For instance, other features concerning alpha band could be peak-to-peak amplitude and alpha band reactivity.

In summary, the results highlighted the utility of EEG data in estimating trust levels in relation-based trust in human robot collaboration. Being able to classify trust based on sensor data has two main advantages over traditional trust evaluation through questionnaires. Firstly, questionnaire results do not always align with user behavior [1]. Trust assessment relying on sensor data becomes more objective, eliminating the need to rationalize the entire interaction with the robot. Secondly, while questionnaires offer summative evaluation after interaction, sensor-based trust assessment enables continuous evaluation, capturing trust dynamics as it develops during interaction. This facilitates the potential to react in real-time to over/undertrust by adapting the robot's interaction or communication behavior. The study's main limitation was the fluency of the robot's speech, which significantly influences human attention and trust in the robot's communication abilities. When a robot communicates smoothly, it appears more competent and reliable, leading to increased trust from humans. To further explore trust, the next step could involve increasing the risk for human participants during interactions to elicit stronger trust responses toward the robot.

6 Conclusion

This paper explored the correlation between robot performance and levels of human trust by leveraging EEG signals as input to trust assessment models. To this end, a scenario fostering social collaboration was designed, requiring the human to engage with a robot to successfully accomplish a game. Results show the successful manipulation of the participants' trust levels through the robot performance. Using the identified trust levels, a connection between the EEG data and trust was identified through ML models. These findings serve as the foundation for future research endeavors. As part of their ongoing work, the authors intend to delve into the effects of the trust repair section on the process of regaining trust from user participants following the robot's performance error. Additionally, sensor fusion techniques, incorporating EEG data, will be employed to enhance the robustness of trust level categorization.

References

1. Adamik, M., Dudzinska, K., Herskind, A.J., Rehm, M.: The difference between trust measurement and behavior: Investigating the effect of personalizing a robot's appearance on trust in hri. In: 2021 30th IEEE International Conference on Robot & Human Interactive Communication (RO-MAN). pp. 880–885. IEEE (2021)
2. Akash, K., Hu, W.L., Jain, N., Reid, T.: A classification model for sensing human trust in machines using eeg and gsr. ACM Transactions on Interactive Intelligent Systems (TiiS) **8**(4), 1–20 (2018)
3. Boudreau, C., McCubbins, M.D., Coulson, S.: Knowing when to trust others: An erp study of decision making after receiving information from unknown people. Social cognitive and affective neuroscience **4**(1), 23–34 (2009)
4. De Visser, E.J., Peeters, M.M., Jung, M.F., Kohn, S., Shaw, T.H., Pak, R., Neerincx, M.A.: Towards a theory of longitudinal trust calibration in human–robot teams. International journal of social robotics **12**(2), 459–478 (2020)
5. Edla, D.R., Mangalorekar, K., Dhavalikar, G., Dodia, S.: Classification of eeg data for human mental state analysis using random forest classifier. Procedia computer science **132**, 1523–1532 (2018)
6. Floyd, M.W., Drinkwater, M., Aha, D.W.: Learning trustworthy behaviors using an inverse trust metric. In: Robust Intelligence and Trust in Autonomous Systems, pp. 33–53. Springer (2016)
7. Hald, K., Rehmn, M., Moeslund, T.B.: Human-robot trust assessment using motion tracking & galvanic skin response. In: 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). pp. 6282–6287. IEEE (2020)
8. Hancock, P.A., Billings, D.R., Schaefer, K.E., Chen, J.Y., De Visser, E.J., Parasuraman, R.: A meta-analysis of factors affecting trust in human-robot interaction. Human factors **53**(5), 517–527 (2011)
9. Hu, W.L., Akash, K., Jain, N., Reid, T.: Real-time sensing of trust in human-machine interactions. IFAC-PapersOnLine **49**(32), 48–53 (2016)
10. Jessup, S.A., Schneider, T.R., Alarcon, G.M., Ryan, T.J., Capiola, A.: The measurement of the propensity to trust automation. In: Virtual, Augmented and Mixed Reality. Applications and Case Studies: 11th International Conference, VAMR 2019, Held as Part of the 21st HCI International Conference, HCII 2019, Orlando, FL, USA, July 26–31, 2019, Proceedings, Part II 21. pp. 476–489. Springer (2019)
11. Jiang, L., Cai, Z., Wang, D., Jiang, S.: Survey of improving k-nearest-neighbor for classification. In: Fourth international conference on fuzzy systems and knowledge discovery (FSKD 2007). vol. 1, pp. 679–683. IEEE (2007)
12. Khavas, Z.R., Ahmadzadeh, S.R., Robinette, P.: Modeling trust in human-robot interaction: A survey. In: International conference on social robotics. pp. 529–541. Springer (2020)

13. Klimesch, W., Doppelmayr, M., Russegger, H., Pachinger, T., Schwaiger, J.: Induced alpha band power changes in the human eeg and attention. Neuroscience letters **244**(2), 73–76 (1998)
14. Kok, B.C., Soh, H.: Trust in robots: Challenges and opportunities. Current Robotics Reports **1**, 297–309 (2020)
15. Law, T., Scheutz, M.: Trust: Recent concepts and evaluations in human-robot interaction. Trust in human-robot interaction pp. 27–57 (2021)
16. Martelaro, N., Nneji, V.C., Ju, W., Hinds, P.: Tell me more designing hri to encourage more trust, disclosure, and companionship. In: 2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI). pp. 181–188. IEEE (2016)
17. McFarland, D.J., Anderson, C.W., Muller, K.R., Schlogl, A., Krusienski, D.J.: Bci meeting 2005-workshop on bci signal processing: feature extraction and translation. IEEE transactions on neural systems and rehabilitation engineering **14**(2), 135–138 (2006)
18. Muir, B.M., Moray, N.: Trust in automation. part ii. experimental studies of trust and human intervention in a process control simulation. Ergonomics **39**(3), 429–460 (1996)
19. Parasuraman, R., Riley, V.: Humans and automation: Use, misuse, disuse, abuse. Human factors **39**(2), 230–253 (1997)
20. Rani, P., Liu, C., Sarkar, N., Vanman, E.: An empirical study of machine learning techniques for affect recognition in human–robot interaction. Pattern Analysis and Applications **9**, 58–69 (2006)
21. Riedl, R., Javor, A.: The biology of trust: integrating evidence from genetics, endocrinology, and functional brain imaging. Journal of Neuroscience, Psychology, and Economics **5**(2), 63 (2012)
22. Robinson, H., MacDonald, B., Broadbent, E.: The role of healthcare robots for older people at home: A review. International Journal of Social Robotics **6**, 575–591 (2014)
23. Rossi, A., Dautenhahn, K., Koay, K.L., Walters, M.L.: How the timing and magnitude of robot errors influence peoples' trust of robots in an emergency scenario. In: Social Robotics: 9th International Conference, ICSR 2017, Tsukuba, Japan, November 22-24, 2017, Proceedings 9. pp. 42–52. Springer (2017)
24. Salem, M., Lakatos, G., Amirabdollahian, F., Dautenhahn, K.: 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 (2015)
25. Schaefer, K.E., Chen, J.Y., Szalma, J.L., Hancock, P.A.: A meta-analysis of factors influencing the development of trust in automation: Implications for understanding autonomy in future systems. Human factors **58**(3), 377–400 (2016)
26. Shayesteh, S., Ojha, A., Jebelli, H.: Workers' trust in collaborative construction robots: Eeg-based trust recognition in an immersive environment. Automation and robotics in the architecture, engineering, and construction industry pp. 201–215 (2022)
27. Shishehgar, M., Kerr, D., Blake, J.: A systematic review of research into how robotic technology can help older people. Smart Health **7**, 1–18 (2018)
28. Ullman, D., Malle, B.F.: What does it mean to trust a robot? steps toward a multidimensional measure of trust. In: Companion of the 2018 acm/ieee international conference on human-robot interaction. pp. 263–264 (2018)
29. Wu, J., Paeng, E., Linder, K., Valdesolo, P., Boerkoel, J.C.: Trust and cooperation in human-robot decision making. In: 2016 aaai fall symposium series (2016)
30. Xu, A., Dudek, G.: Optimo: Online probabilistic trust inference model for asymmetric human-robot collaborations. In: Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction. pp. 221–228 (2015)
31. Yagoda, R.E., Gillan, D.J.: You want me to trust a robot? the development of a human–robot interaction trust scale. International Journal of Social Robotics **4**, 235–248 (2012)