

Does it Feel Safer? A Pilot Study on the Stress Levels of Humans for Varied Robot Control Strategies and Collaboration Scenarios

Heiko Renz¹, Khazar Dargahi Nobari¹, Mohammed Faizan²,
and Torsten Bertram¹

¹Institute of Control Theory and Systems Engineering, TU Dortmund University,
Dortmund, Germany

²Former Student of TU Dortmund University, Dortmund, Germany

ABSTRACT

Human-robot collaboration is an essential factor in current industry and social applications. A key aspect of meaningful and effective collaboration is the safety of the human worker. Therefore, different objective metrics allow researchers to assess safety based on measurements like distance between humans and robots, speed of the robot, or force exerted by the robot. However, for an effective collaboration, objective safety metrics are essential, as well as the subjective perception of safety by the user. To investigate the subjective stress level of users during human-robot collaboration, we conducted a pilot study with 20 participants with varied control strategies and collaboration scenarios. Furthermore, a stress prediction model is proposed based on the collected data. The results show that the collaboration scenario significantly influences the subjective stress levels of users, and trends in the data indicate that the robot's collision avoidance strategy also impacts stress levels. The proposed stress prediction model shows the potential to forecast the stress levels of users based on the collected data, enabling possible feedback options for different control solutions. However, further studies are required to investigate generalized stress prediction models for various collaboration scenarios and control strategies.

Keywords: Human-robot collaboration, Human factors, Human stress levels

INTRODUCTION

Improving productivity and acceptance of human-robot collaboration (HRC) is a crucial factor for the success of robots in various areas, and it is increasingly important in industry and social applications. For the successful integration of robots into human environments, the perceived safety of the human worker is a key factor (Weiss et al., 2021). The perception of safety is influenced by various factors, not limited to the distance between the human and the robot or the robot's speed. Also, it includes factors like the robot's appearance, the task the robot is performing, and its behavior. However, measuring the subjective perception of safety is challenging, as various factors influence it and can vary between individuals. Extensive studies are required

to understand the impact of different factors on the subjective perception of safety and to develop strategies to include these factors in robot design and control strategies in HRC scenarios.

Related Work

The safety of HRC has been studied extensively in the past. Since this research area is interdisciplinary, it is approached from different perspectives, including psychology and engineering. Technical evaluations often focus on objective safety metrics to improve HRC by developing mechanical structures that increase safety (Malzahn and Bertram, 2014), strategies for impact minimization (Haddadin et al., 2012), or various collision avoidance algorithms (Renz et al., 2023a).

The present related work section focuses on psychological evaluations of users' subjective perceptions regarding different aspects of HRC. Due to the wide range of research in this area, this section is not exhaustive but aims to provide an overview of relevant human stress studies in collaboration regarding different scenarios and control strategies. Arai et al. (2010) investigate the impact of robot proximity and speed in industry collaboration scenarios and the user's notification about robot motions. The authors conclude their work with recommendations for the design of industrial HRC regarding the distance between robots and humans (> 2 m) and the speed of the robot toward the human (< 500 mm/s). Furthermore, a notification of robot motions is recommended to reduce the user's mental strain. The subject study of Dragan et al. (2015) examines whether different robot trajectories influence objective performance measures and subjective ratings regarding multiple aspects like fluency of team interaction, trust, and safety. The robot trajectories are classified into functional, predictable, and legible during a collaborative task but do not consider required adaptations for collision avoidance. The authors conclude that users prefer legible trajectories, leading to higher performance. To directly optimize the user's mental load, it is possible to integrate physiological signals into the control strategy of the robot (Messerli et al., 2021). A learning automaton optimizes the robot's production pace in the collaborative production task based on the user's stress level and task performance. Nevertheless, the authors do not consider different scenarios and online collision avoidance strategies. Su et al. (2023) investigate the effects of interactions on users' stress levels. The authors use NASA Task Load Index (NASA-TLX) questionnaires and electrodermal activity (EDA) to measure users' stress levels during human-robot interaction. The results show that different interaction modes between humans and robots lead to other stress levels for users that generally favor straightforward interactions. Instead of using solely normalized EDA measurements, machine learning approaches are beneficial for stress detection on EDA data (Zhu et al., 2023) and reason the usage of prediction models, including further input features, in this work. The inclined reader is also referred to the review from Lu et al. (2022), with details about research on stress and safety awareness in HRC.

Contribution

The contribution at hand presents a pilot study on the subjective stress levels of users during HRC. The following key contributions are made: First, the statistical evaluation of a pilot study with 20 participants is presented, investigating the impact of three stages of collaboration scenarios. Secondly, we extend the statistical evaluation to the effect of three different robot trajectory planners and controllers regarding their collision avoidance strategies on the subjective stress levels of users. Thirdly, we present the evaluation of neural network models that predict human stress levels based on the collected data, including two validation strategies to investigate the generalizability of stress prediction models.

STUDY DESIGN & DATA PROCESSING

This section describes the study design, data recording, and processing for the pilot study. Before the participants start the experiments, they are introduced to the study and sign a declaration of consent. The ethics committee of TU Dortmund University approved the study.

Study Design

The pilot study was conducted with 20 male participants aged 21 to 26 ($\mu = 23 \pm 1.38$) years. All participants are recruited from the university's electrical engineering department. Note that three students of the study are student assistants in the research group and have experience with the applied Universal Robot 10 (UR10). Nevertheless, they are included in the statistical analysis since they do not have any further experience with the specific experiments.

Each participant performs seven experiments with varied collaboration scenarios and collision avoidance strategies in the laboratory. Fig. 1 shows an exemplary image of the task setup in the robot lab. The participants are requested to accomplish assembly tasks sitting at a table. For each of the seven experiments, the participants constructed a different wooden toy based on a digital instruction manual. Two of them are visible in Fig. 2. After each experiment, the participants complete the NASA-TLX questionnaire to evaluate their stress levels subjectively.

The collaboration scenarios are divided based on the required collaboration between humans and robots. Each scenario is assigned to a level of human-robot interaction based on the taxonomy of Mukherjee et al. (2022). The first scenario belongs to level one (L1 – Coexistence), where the human and the robot work independently in a separate workspace without interaction. In the laboratory setup, L1 is realized by the participant working on the assembly task on one table and the robot working on an organization task on another table. The second scenario belongs to level two (L2 – Cooperation), where the human and the robot work together in the same workspace without any direct interaction but with a shared goal that requires subsequential actions of the human and the robot. In the laboratory setup, L2 is realized by the participant working on the task on one table and the robot delivering the required parts for the assembly task from another table and placing them in

front of the participant on the table. The third scenario belongs to level three (L3 – Collaboration), where the human and the robot work together in the same workspace with direct interaction and shared goals. In the laboratory setup, L3 is realized by the participant working on the assembly task on one table and the robot delivering the required parts for the assembly task from another table and placing them in the hand of the participant (see Fig. 1). The scenarios are abbreviated as L1, L2, and L3 in the remainder of this work.



Figure 1. Exemplary image of the task setup. The participant is working on scenario L3 while assembling a wooden toy.

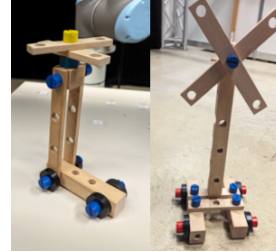


Figure 2. Two exemplary assembly tasks of the study.

Each participant executes L1 with the UR10 default controller (C1), that stops only at a collision. Since the workspace is separated, no collision avoidance strategy is required for L1, and the scenario is only executed once for each participant. L2 and L3 are executed with C1 and a different controller with two collision avoidance strategies (C2 and C3). The default controller C1 applies the teach pendant of the UR10 to move the robot. As a control strategy for C2 and C3, a Moving Horizon Planner (MHP) is applied (Krämer et al., 2020; Renz et al., 2024). The MHP plans a trajectory for the robot and optimizes the trajectory regarding a cost function, including terms to reach the goal, restrict robot joint velocities, and avoid collisions and proximity with the human. To consider dynamic obstacles like humans in the environment, the MHP plans a 3 s trajectory and replans it with 10 Hz. The static collision avoidance strategy C2 assumes the human is static for optimizing the 3 s trajectory for each planning cycle. Since this assumption is often invalid in real-world scenarios, the predictive collision avoidance strategy C3 predicts human motion (prediction horizon 0.4 s), including uncertainties for each planning cycle (Renz et al., 2023b).

The remainder of this work declares each experiment as a combination of the chosen scenario and controller type. For example, a cooperation scenario with the MHP controller and a static collision avoidance strategy are labeled as L2-C2.

Data Collection and Processing

The participants' physiological parameters and upper body motions are recorded during the experiments. An Empatica E4 (Empatica Inc.,

MA, U.S.) wristband collects the physiological parameters of heart rate, skin temperature, and EDA. Furthermore, an Optitrack motion capture system (NaturalPoint Inc., OR, U.S.) records the upper body motions of the participants. Only the upper body motions are collected since the participants are sitting during the experiments, and the upper body motions are the most relevant for the assembly task. Since the measurements are available with different sampling rates, the data are synchronized and resampled in a primary-secondary scheme with the blood volume pulse (BVP) signal as the primary signal with a frequency of 64 Hz. Slower signals like the skin temperature and the EDA are set as fixed until a new value is available. The motion data exhibit a higher sampling rate and are only requested at the arrival of a new BVP value. The data for each experiment and participant is preprocessed before they are used to train and evaluate the stress prediction models. BIOBSS python package (Taşcı et al., 2024) provides functions for filtering, normalizing, segmenting, and extracting features from the physiological signals EDA and BVP. Furthermore, the average of BVP and EDA of experiment L1-C1 serve as a baseline for normalizing the physiological signals of all experiments. As a feature from the motion data, the mean velocity of each measured body part for each segment is extracted. All data are segmented to increase the data for the model training and to consider a time history required for meaningful stress prediction. A sliding window approach serves for segmentation with window lengths t_{win} from 10 s to 150 s and step sizes t_{step} between windows of 5 s and 10 s. Each window is labeled with the NASA-TLX score of the participant for the related experiment. The prediction models include six model types, namely Support Vector Machines (SVM), Multi-Layer Perceptrons (MLP), K-Nearest-Neighbours (KNN), Gaussian Naive Bayes (GNB), Random Forest (RF), and AdaBoost (AB), with various settings (kernel, layer sizes, activation functions, number of neighbors, etc.).

Each participant digitally fills out the NASA-TLX questionnaire (NASA-TLX questionnaire, n.d.) after each experiment. As a result, the questionnaire provides a weighted rating score to assess the task load of the participant. To normalize the scores, the z-score for the weighted rating is calculated for each participant (z_{part}) and for all participants (z_{all}).

EVALUATION & DISCUSSION

This section presents the evaluation of the pilot study and the results of the stress prediction models. The evaluation is divided into two parts. First, a two-way analysis of variances (ANOVA) (Girden, 1992) is applied to investigate the impact of the collaboration scenarios and the collision avoidance strategies on the subjective stress levels of the participants. The statistical evaluation is followed by discussing the results and interpreting Tukey's honestly significant difference (HSD) (Abdi and Williams, 2010) post-hoc test results. Secondly, the assessment of the stress prediction models is presented for different data segmentation and preparation, model architectures, and validation strategies.

Impact of Key Factors on Stress Levels

The questionnaire's statistical analysis includes all 20 participants. However, due to technical issues, three ratings are missing (one for each participant 6, 15, 20); therefore, three scenarios only include 19 samples instead of 20 (see Fig. 5).

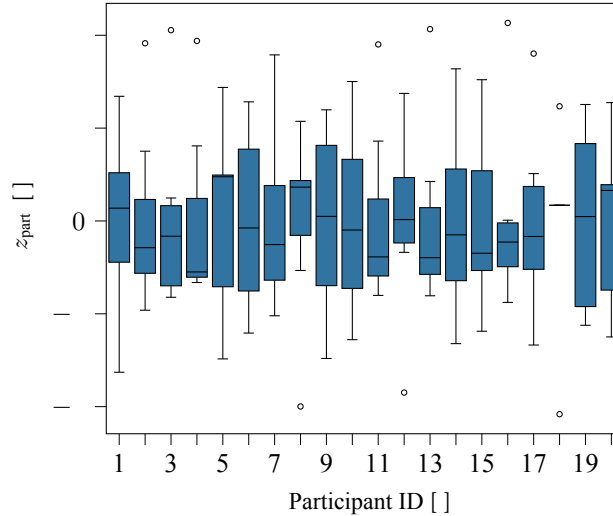


Figure 3: z_{part} of participants' NASA-TLX ratings for each participant.

As a first step, the difference between the ratings of different participants is investigated. The hypothesis that the participants' ratings are significantly different is tested with a t-test. Fig. 3 shows z_{part} of the NASA-TLX ratings. The applied t-test shows that the participants' ratings are not significantly different regarding the participant-wise normalized NASA-TLX scores. Therefore, the normalized z_{part} is used for further statistical analysis, and the impact of the different scenarios and controllers is investigated. Subsequently, an investigation of z_{part} compares the seven experiments. Fig. 4 shows z_{part} of the NASA-TLX ratings for the different experiments, including the results of pairwise t-tests. The evaluation of this figure only treats some general trends and significant differences between the experiments. Fig. 4 shows that the participants' stress levels differ for the experiments. It is visible that L1-C1 has a lower median and upper quartile than the other experiments, followed by L2-C2 and L3-C2. The highest median and upper quartile are visible for L2-C1. The included t-test results show that the participants' stress levels statistically differ between six pairings. Since significant differences occur between scenarios and controllers, a detailed two-way ANOVA is applied to investigate this further. To classify different z_{part} into stress labels and enable a stress prediction, the scores are also divided into three categories: low stress ($z_{\text{part}} < 0$), medium stress ($0 \leq z_{\text{part}} \leq 1$), and high stress ($z_{\text{part}} > 1$). Fig. 5 shows the count of stress labels for each experiment and underlines the previous general statements about the experiments. Experiment L1-C1 has the highest count of low-stress labels, followed by L3-C2 and L2-C2. The

highest count of high and medium stress labels is visible for L3-C1, followed by L2-C1, indicating that C1 in the scenarios L2 and L3 leads to higher stress levels than C2 and C3.

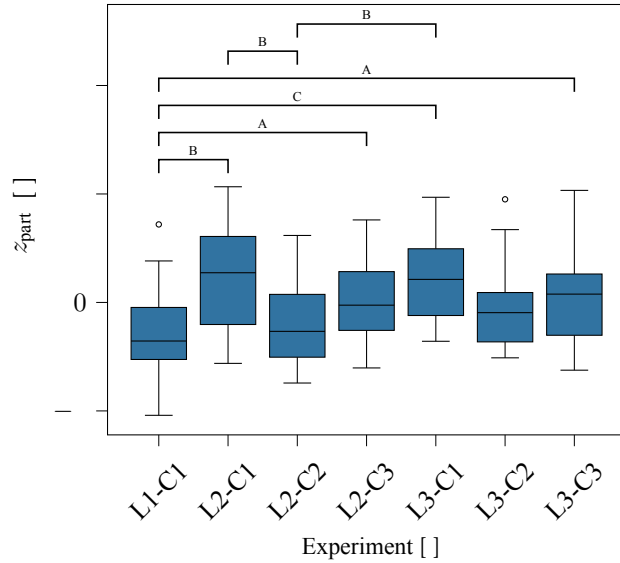


Figure 4: z_{part} of participants’ NASA-TLX ratings for each experiment, including results of pairwise t-test (A: $p < 0.05$, B: $p < 0.01$, C: $p < 0.001$).

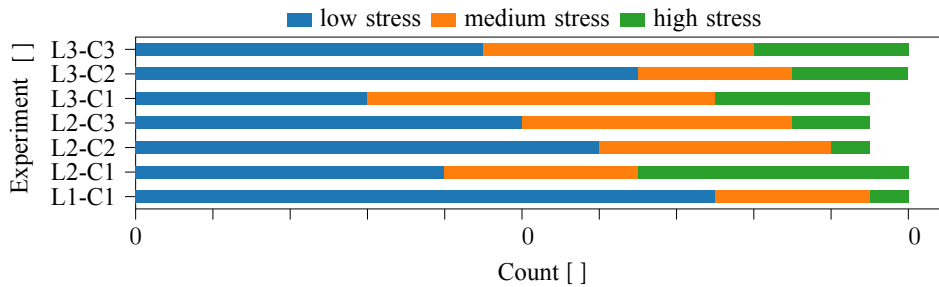


Figure 5: Count of stress labels for each experiment.

For the two-way ANOVA, we assume that the results from experiment L1-C1 are independent of the controller since the workspace is separated and no collision avoidance strategy is required. Based on this assumption, the results of L1-C1 are repeated for L1-C2 and L1-C3 to achieve a balanced number of samples for L1. Table 1 shows the results of the two-way ANOVA for z_{part} . The result of the two-way ANOVA shows that the interaction of the scenario and the controller is insignificant ($p = 0.18664$), and an analysis of the main effects is sufficient. The scenario’s impact is significant ($p = 0.00002$), while the controllers’ impact is not substantial for a significance with $p = 0.05042$. Since the controller impact is still close to being significant, a post-hoc test examines the controller effect in detail.

Table 1. Results of the two-way ANOVA for z_{part} (ΣSq : sum of squares, df: degrees of freedom, bold indicates statistical significance with $\alpha = 0.05$).

	ΣSq	df	F-stat.	p-val.
Scenario	17.935	2	11.333	0.000
Controller	4.809	2	3.039	0.050
Scenario:Controller	4.942	4	1.562	0.187
Residual	138.471	175		

The post-hoc test is based on Tukey's HSD test and investigates the impact of the controller and scenario regarding significant differences inside each group. Tukey's HSD test is common to investigate the significance of differences between group members after ANOVA proved a significant difference for one group. The results of the post-hoc test with a threshold $\alpha = 0.05$ are shown in Table 2. The first three rows of Table 2 show the post-hoc test results for the different scenarios, and the last three rows for the different controllers. It can be seen that L1 statistically differs from L2 and L3 regarding the participants' stress levels, while the pair L2 and L3 do not differ significantly. One reason for this is that the separation of the workspace influences the participants' stress levels more than the difference between cooperation and collaboration. Regarding the controllers, C1 statistically differs from C2, while the other pairs do not show significant differences. This shows that the static collision avoidance strategy C2 impacts the participants' stress levels more than the predictive collision avoidance strategy C3.

Table 2. Results of the post-hoc test for z_{part} for scenario and controller comparison (Hyp. Reject: hypothesis rejection (bold indicates a rejection)).

Group 1	Group 2	Group Mean Diff.	Adj. p-val.	95% Confidence Interval		Hyp. Reject
				Lower	Upper	
L1	L2	0.609	0.001	0.225	0.992	True
L1	L3	0.716	0.000	0.334	1.098	True
L2	L3	0.107	0.798	-0.288	0.503	False
C1	C2	-0.436	0.029	-0.836	-0.036	True
C1	C3	-0.231	0.382	-0.642	0.180	False
C2	C3	0.205	0.447	-0.195	0.605	False

In the following, we discuss the results of the statistical evaluation, the impact of the collaboration scenarios, and the collision avoidance strategies on the subjective stress levels of the participants, connecting the results of the different evaluation steps. The significant difference between the separated workspace and the shared workspace with the post-hoc test underlines the first statements regarding the higher stress levels of L1, also discussed for Fig. 4 and Fig. 5. The stress level in shared workspaces is higher than in separated workspaces, while the difference between cooperation and collaboration is not significant. L2 and L3 are not significantly different because of the experiment design. In both scenarios, the proximity between

the human and the robot is similar, and the separation by a table and the shared goal are, except for the assembly task, the same. This indicates that the task load is not significantly influenced by whether the robot hands over the required parts or places them close to the human. This results in the conclusion that industry settings with shared goals can be designed cooperative or collaborative while achieving other optimization goals like production efficiency. When discussing the control and collision avoidance strategies, the results of the two-way ANOVA indicate that the impact of the control and collision avoidance strategy is briefly below the significance level. Nevertheless, a trend is visible, and we executed a post-hoc test to investigate the impact of the controllers in more detail and show that only the static collision avoidance strategy C2 significantly differs from the default controller C1. Connecting this difference to the stress results of each experiment (Fig. 4 and Fig. 5), it is visible that the static collision avoidance strategy C2 significantly reduces participants' stress. The predictive collision avoidance strategy C3 results in a higher performance regarding objective metrics (Renz et al., 2023b) but does not significantly reduce the participants' stress levels compared to C1. A reason for this is robot motions to avoid collision with possible future human poses that are not intuitive for the participants and cause a higher stress level. Furthermore, various parameters in C2 and C3 require further investigation to optimize the participants' stress levels. In particular, parameters for C3 regarding prediction horizon and uncertainty consideration are currently set to default values and lack individual optimization considering participants' preferences. Possible solutions to this problem are also hybrid approaches of C2 and C3 to combine the higher performance on objective metrics of C3 and the better individual interpretability of C2 during different periods of experiments. For industry settings utilizing HRC, the results show that interpretability and predictability of the robot's motion are essential for the participants' stress levels and should be considered during the design phase.

Stress Level Prediction

For stress level prediction models the data of three participants (1, 3, 4) are excluded due to data collection issues. Due to missing stress labels and physiological signals the data from four more participants (2, 6, 15, 20) are impaired for single experiments. Therefore, the training and evaluation process is repeated with 13 participants D_{13} and 17 participants D_{17} , including the unaffected data of the four participants. This work applies the six mentioned models to prove the feasibility of predicting participants' stress levels based on the data. The results are only shown for the varying best-performing parameter settings for each model type. Furthermore, the models are evaluated using different data segmentations and preparation strategies (see Section Data Collection and Processing), and only the best results are shown to keep the evaluation concise. Since extracting features from the physiological signals and the motion considers different time, frequency, and statistical features of EDA and BVP and the mean velocity of the motion data for each segment, the number of possible features is 95 in the current

setting. However, to reduce the number of features and to avoid overfitting, a feature selection is applied based on the correlation of the features with the stress labels. The threshold for the feature selection is set to 0.1 for the stress labels using z_{all} . For predicting stress labels z_{all} serves as the score since it calculates the z-score over all participants and not only for each participant, and a preliminary test shows an increased prediction performance. All models are trained and evaluated with a 10-fold cross-validation (Cross) and a Leave-One-Participant-Out (LoPo) validation and test strategy. The classification accuracy and the F1-score for Cross are visible in Table 3 and for LoPo in Table 4, including the data preprocessing parameters. Bold values indicate the best performance in the column.

Table 3. Accuracy and F1-score for prediction models with cross.

Type	D_{13}				D_{17}			
	t_{Win}	t_{Step}	Accuracy	F1-Score	t_{Win}	t_{Step}	Accuracy	F1-Score
SVM	100	5	0.997	0.997	100	5	0.927	0.927
MLP	150	5	0.998	0.998	100	5	0.913	0.913
KNN	100	5	0.999	0.999	100	5	0.926	0.926
GNB	100	10	0.618	0.611	100	5	0.565	0.552
RF	150	5	0.998	0.998	100	5	0.902	0.902
AB	150	5	0.900	0.899	150	5	0.806	0.806

Table 4. Accuracy and F1-score for prediction models with LoPo.

Type	D_{13}				D_{17}			
	t_{Win}	t_{Step}	Accuracy	F1-Score	t_{Win}	t_{Step}	Accuracy	F1-Score
SVM	100	10	0.476	0.519	80	10	0.472	0.710
MLP	150	5	0.527	0.548	60	10	0.451	0.676
KNN	100	10	0.501	0.539	100	10	0.460	0.678
GNB	100	5	0.569	0.559	40	5	0.513	0.670
RF	150	10	0.526	0.543	100	10	0.491	0.688
AB	150	10	0.542	0.586	150	10	0.529	0.722

The results indicate different opportunities and limitations. Performance is higher using D_{13} and Cross than using D_{17} . A reason could be that the models overfit on D_{13} and are capable of predicting with high accuracy. A possible solution is a larger dataset that includes more participants. Furthermore, different models reach high results, only differing in the second or third decimal place, showing that different model types are capable of the prediction tasks. Models like MLP, KNN, and RF perform solidly on both strategies compared to the others and could be a starting point for further investigations. For more extended window sizes t_{Win} , the prediction mostly performs better, and it is reasoned that models predict better considering an extended history. When using stress prediction models in generalized settings, e.g., for humans not included in the training data, a significant decrease

in performance is visible (compare Tables 3 and 4). Different reasons for this limitation are possible and include besides others the parameter settings of the models, the model types themselves, or the data preprocessing and normalization since all of them potentially influence the generalizability.

CONCLUSION & OUTLOOK

The work at hand presents the results of a pilot study with 20 participants to investigate the impact of robot control and collision avoidance strategies during different levels of HRC on human stress. The study shows a significant effect of collaboration, and trends in the data justify further investigations of stress-optimized control strategies and recommendations regarding the design of collaborative tasks. Furthermore, the data collected during the study is utilized for training and evaluating simple prediction models for stress-label forecasting. The evaluation reveals limitations regarding the generalization of tested models.

A subject study with more and diverse participants is required to further examine the impacts and the option of stress label prediction in different scenarios and collision avoidance strategies. Furthermore, future work should include different model types and parameter settings for the controllers and the models.

ACKNOWLEDGMENT

The authors gratefully acknowledge the financial support of the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – 497071854.

REFERENCES

- Abdi, H. and Williams, L. J. (2010), ‘Tukey’s honestly significant difference (HSD) test’, *Encyclopedia of research design* 3(1).
- Arai, T., Kato, R. and Fujita, M. (2010), ‘Assessment of operator stress induced by robot collaboration in assembly’, *CIRP Annals* 59(1), pp. 5–8.
- Dragan, A. D., Bauman, S., Forlizzi, J. and Srinivasa, S. S. (2015), ‘Effects of robot motion on human-robot collaboration’, in *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction*, ACM, pp. 51–58.
- Girden, E. R. (1992), *ANOVA: Repeated measures*, number 84, Sage Publications.
- Haddadin, S., Haddadin, S., Houry, A., Rokahr, T., Parusel, S., Burgkart, R., Bicchi, A. and Albu-Schäffer, A. (2012), ‘On making robots understand safety: Embedding injury knowledge into control’, *The International Journal of Robotics Research* 31(13), pp. 1578–1602.
- Krämer, M., Rösmann, C., Hoffmann, F. and Bertram, T. (2020), ‘Model predictive control of a collaborative manipulator considering dynamic obstacles’, *Optimal Control Applications and Methods* 41(4), pp. 1211–1232.
- Lu, L., Xie, Z., Wang, H., Li, L. and Xu, X. (2022), ‘Mental stress and safety awareness during human-robot collaboration - review’, *Applied Ergonomics* 105, 103832.
- Malzahn, J. and Bertram, T. (2014), ‘Collision detection and reaction for a multi-elastic-link robot arm’, *IFAC Proceedings Volumes* 47(3), pp. 320–325.

- Messeri, C., Masotti, G., Zanchettin, A. M. and Rocco, P. (2021), ‘Human-robot collaboration: Optimizing stress and productivity based on game theory’, *IEEE Robotics and Automation Letters* 6(4), pp. 8061–8068.
- Mukherjee, D., Gupta, K., Chang, L. H. and Najjaran, H. (2022), ‘A survey of robot learning strategies for human-robot collaboration in industrial settings’, *Robotics and Computer-Integrated Manufacturing* 73, 102231.
- NASA-TLX questionnaire (n.d.). Last Access: 2024-05-14. URL: <https://nasa-tlx.web.app/>
- Renz, H., Krämer, M. and Bertram, T. (2023a), ‘Comparing human motion forecasts in moving horizon trajectory planning of collaborative robots’, in *2023 IEEE International Conference on Robotics and Biomimetics (ROBIO)*.
- Renz, H., Krämer, M. and Bertram, T. (2023b), ‘Uncertainty estimation for predictive collision avoidance in human-robot collaboration’, in *2023 IEEE International Conference on Robotics and Biomimetics (ROBIO)*.
- Renz, H., Krämer, M. and Bertram, T. (2024), ‘Moving horizon planning for human-robot interaction’, in *Proceedings of the 2024 ACM/IEEE International Conference on Human-Robot Interaction*, pp. 939–943.
- Su, B., Jung, S., Lu, L., Wang, H., Qing, L., Xie, Z. and Xu, X. (2023), ‘The effects of human-robot interaction modality on workers’ mental stress’, *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 67(1), pp. 446–452.
- Taşcı, Ç., Karakuş, İ., Çavuşoğlu, D. and Akyön, F. Ç. (2024), ‘Biobss’. Accessed: 2024-04-29. URL: <https://github.com/obss/BIOBSS>
- Weiss, A., Wortmeier, A.-K. and Kubicek, B. (2021), ‘Cobots in industry 4.0: A roadmap for future practice studies on human-robot collaboration’, *IEEE Transactions on Human-Machine Systems* 51(4), pp. 335–345.
- Zhu, L., Spachos, P., Ng, P. C., Yu, Y., Wang, Y., Plataniotis, K. and Hatzinakos, D. (2023), ‘Stress Detection Through Wrist-Based Electrodermal Activity Monitoring and Machine Learning’, *IEEE Journal of Biomedical and Health Informatics* 27(5), pp. 2155–2165. 11