Applied Ergonomics 116 (2024) 104224

Contents lists available at ScienceDirect

Applied Ergonomics

journal homepage: www.elsevier.com/locate/apergo

Exploring the impact of human-robot interaction on workers' mental stress in collaborative assembly tasks

Bingyi Su, SeHee Jung, Lu Lu, Hanwen Wang, Liwei Qing, Xu Xu *

Edward P. Fitts Department of Industrial and Systems Engineering, North Carolina State University, USA

A R T I C L E I N F O

Keywords:
Human-robot collaboration
Interaction modality
Mental stress
Physiological responses

A B S T R A C T

Advances in robotics have contributed to the prevalence of human-robot collaboration (HRC). However, working and interacting with collaborative robots in close proximity can be psychologically stressful. Therefore, understanding the impacts of human-robot interaction (HRI) on mental stress is crucial for enhancing workplace well-being. To this end, this study investigated how the HRI factors – presence, complexity, and modality – affect the psychological stress of workers. We employed both the NASA-Task Load Index for subjective assessment and physiological metrics including galvanic skin responses, electromyography, and heart rate for objective evaluation. An experimental setup was implemented in which human operators worked together with a collaborative robot on Lego assembly tasks, using different interaction paradigms including pressing buttons, showing hand gestures, and giving verbal commands. The results revealed that the introduction of interactions during HRC helped reduce mental stress and that complex interactions resulted in higher mental stress than simple interactions. Meanwhile, using hand gestures led to significantly higher mental stress than pressing buttons and verbal commands. The findings provided practical insights for mitigating mental stress in the workplace and promoting wellness in the era of HRC.

1. Introduction

In recent years, human-robot collaboration (HRC) has been growing rapidly in the context of smart manufacturing and Industry 4.0. Collaborative robots (co-robots) are employed to assist human operators in achieving unprecedented flexibility, where human cognitive skills and dexterity are mutually reinforced with the physical capabilities of the co-robots. Typically, co-robots are used for repetitive and physically demanding tasks, while human operators are responsible for advanced decision-making and fine-tuning tasks (Cherubini et al., 2016). In HRC, human operators and co-robots perform tasks concurrently or jointly within a collaborative workspace. For example, in an HRC assembly task, the co-robot delivers and holds a part with pre-drill screw holes and the human operator turns screws into it. As the level of collaboration continues to increase, workspaces are shared more intensively, leading to a symbiotic HRC (Wang et al., 2019).

Close collaboration with robots raises concerns related to workplace health and safety (Murashov et al., 2016). Physical safety issues such as collisions may occur as isolating workers from co-robots is not an option in HRC tasks. One way to ensure safe coexistence between humans and robots is through proactive retraction of the co-robot end effector depending on the location of workers. This can be accomplished using motion-tracking devices, such as inertial sensors (Meziane et al., 2014), depth sensors (Mohammed et al., 2017), or RGB cameras (Xie et al., 2022). In addition to physical collision, co-robots can also lead to mental stress for human operators, which can negatively affect interaction and collaboration performance (Gervasi et al., 2022). Improper handling of mental stress can lead to adverse consequences, spanning from errors and accidents in the short term to the development of various long-term health issues (Umer, 2022). Therefore, it is equally important to study and understand human mental stress in HRC to optimize the process and achieve workplace wellness.

Mental stress can be assessed through subjective and objective measures. Subjective ratings, such as self-report questionnaires, have been commonly used to estimate levels of mental stress in humans (Aigrain et al., 2018). Participants are asked to answer a variety of questions about their experiences in the experiment. The NASA-Task Load Index (NASA-TLX) has been utilized in numerous research studies to assess people's mental stress levels. For instance, Zheng et al. (2012) employed the NASA-TLX to investigate the mental workload

* Corresponding author.
E-mail address: xxu@ncsu.edu (X. Xu).

https://doi.org/10.1016/j.apergo.2024.104224
Received 13 September 2023; Received in revised form 4 December 2023; Accepted 2 January 2024
Available online 5 January 2024

experienced by surgeons during endoscopy training. In the context of smart factories, Zakeri et al. (2021) applied the NASA-TLX to examine various factors contributing to mental stress, such as task complexity, time constraints, and collaboration duration. However, it is important to acknowledge the limitations of self-reporting, as participants cannot report in real-time and may not express their true feelings (Bethel et al., 2007).

The use of objective measures, such as physiological signals, is an important complement to subjective measures and provides insight into the subconscious and psychobiological phenomena involved (Rubagotti et al., 2022). Galvanic skin response (GSR), also known as electrodermal activity (EDA), measures skin conductivity and effectively reflects one's emotional states. GSR readings significantly increase when the stress level increases (Shi et al., 2007). Many researchers have used GSR as an objective measure to assess mental stress. For example, Healey and Picard (2005) analyzed the GSR data collected during real-world driving tasks to determine the driver's relative stress levels. Giakoumis et al. (2012) implemented an automatic stress detection algorithm by using GSR data as an important basis. In the realm of human-robot collaboration, Lu et al. (2022) leveraged GSR signals to assess workers' mental stress in hand-over activities.

Electromyography (EMG) is also considered as an effective objective measure to detect mental stress. EMG signals refer to the collective electrical activity of muscle motor units (Chowdhury et al., 2013), which is controlled by the nervous system and generated during muscle contraction. The EMG signal is categorized into surface EMG and intramuscular EMG, which are measured by non-invasive electrodes and invasive electrodes, respectively. Wijsman et al. (2013) investigated the validity of EMG as a predictor of mental stress and showed that the amplitude of the EMG signal increases in stressful situations. In a driving simulator experiment, Zheng et al. (2015) utilized masseter EMG to assess drivers' mental stress. Another similar study conducted by Robles et al. (2022) examined the use of convolutional neural networks for stress detection via surface EMG of the trapezius muscle.

In addition, heart rate (HR) is a well-established measure of stress and mental workload, which is expressed as the count of heart beats per minute (bpm) (Waard, 1996). In a study conducted by Fauquet-Alekhine et al. (2016), the authors attempted to construct a mathematical relationship between the heart rate data and the intensity of short-term indicators of mental stress. Umer (2022) employed heart rate data for simultaneously monitoring workers' physical and mental stress during construction tasks. Linssen et al. (2022) conducted a study using accelerometry and heart rate data to monitor soldiers' stress in a virtual reality military scenario.

Previous studies have investigated relevant robot factors, such as robot attributes and motion characteristics, to explore their correlation with human mental status. The experimental study conducted by Rahimi and Karwowski (1990) indicated that robot sizes and initial speeds were the significant main effects on the perception of safe robot speed. This finding was later verified by Duffy et al. (2006) in a virtual reality environment. Arai et al. (2010) concluded that operators experience high mental stress when robots move toward them at high speed and recommended that additional notice should be provided before a robot moves. Dragan et al. (2015) analyzed the benefits of robot motion planning and found that legible motions planned to clearly express the robot's intent led to more fluid collaboration. Another study by Lu et al. (2022) examined human psychological stress during HRC handover tasks, noting that the end effector approaching within the worker's field of view at a low speed and with a restricted trajectory can cause significantly less mental stress.

In complex HRC scenarios, collaboration is facilitated by human-robot interactions (HRI). Several studies have investigated the HRI within HRC contexts, primarily focusing on their advantages in terms of performance and user experiences (Jokinen and Wilcock, 2013; Galin and Meshcheryakov, 2020; Ötting et al., 2022). However, there remain gaps in understanding how various HRI factors impact workers' mental stress. While the introduction of HRI helps human operators retain control over the HRC process, the complexity of interactions could impact operators' mental stress (Robelski and Wischniewski, 2018). Moreover, HRI can be applied through various sensory channels, including haptic, visual, and auditory channels (Bonarini, 2020), potentially resulting in varying levels of mental workload for workers. Therefore, it is essential to understand how (1) interaction presence, (2) interaction complexity, and (3) interaction modality and their interactions affect the mental stress of workers in HRC.

In this study, we conducted an experimental study to investigate how different HRI paradigms affect the mental stress of human operators during HRC. A collaborative Lego assembly task was implemented and physiological signals, including GSR data, EMG data, and HR data, were applied to measure mental stress along with NASA-TLX self-report questionnaires. The aim is to evaluate how the presence of interactions during HRC affects human mental stress and the differences across interaction complexity and modalities, including pressing buttons (haptic channel), showing hand gestures (visual channel), and giving verbal commands (auditory channel). This exploration is essential in understanding the impacts of these HRI factors on mental stress, leading to more informed approaches in designing and implementing HRC systems that prioritize both efficiency and worker well-being.

2. Methods

2.1. Participants

A total of 24 healthy participants (10 females and 14 males) between 22 and 48 years old ($M = 26.3$ years, $SD = 5.2$ years) with no acute or chronic musculoskeletal disorders were recruited for this experimental study. All participants had no previous experience in HRC but were familiar with Lego assembly tasks. The experimental protocol was approved by the institutional review board at North Carolina State University (approval # 25,195).

2.2. Experiment setup

In this experiment, a Sawyer robot arm (Rethink Robotics) was used to complete the collaborative Lego assembly task, as shown in Fig. 1. The robot control was realized through the Robot Operating System (ROS noetic) framework in the Linux Ubuntu 20.04 environment.

The assembly task can be briefly described as follows: First, the co-robot picks up a large Lego block from Zone 1 and delivers it to the participant sitting in Zone 2. Then, the participant takes the large Lego block from the co-robot and secures small irregular-shaped Lego pieces on the large Lego block. The participant then needs to enter a serial number shown on the Lego block using an iPad. This procedure is repeated five times in each collaboration session to enhance statistical power.

To investigate the effects of HRI paradigms on human mental stress, different HRC scenarios were performed during the assembly tasks, as outlined in Fig. 2. Human operators collaborated with the co-robot either without interaction or with interaction in different modality. Depending on the scenario, the robot would execute repetitive actions independently or respond with specific actions to the commands issued by the human operators.

In detail, three relevant factors were involved in the HRC scenarios, namely (i) interaction presence, (ii) interaction complexity, and (iii) interaction modality. In terms of interaction presence, there were two levels: no interaction and with interaction. For no interaction, the co-robot repetitively delivered blocks at a specific time interval (16 s) and automatically opened the gripper in 2 s once it reached the operator's position. For with interaction, two levels of interaction complexity were considered: simple interaction and complex interaction. Specifically, in simple interaction, the human operator communicates with the robot only once when the operator needs the co-robot to open the

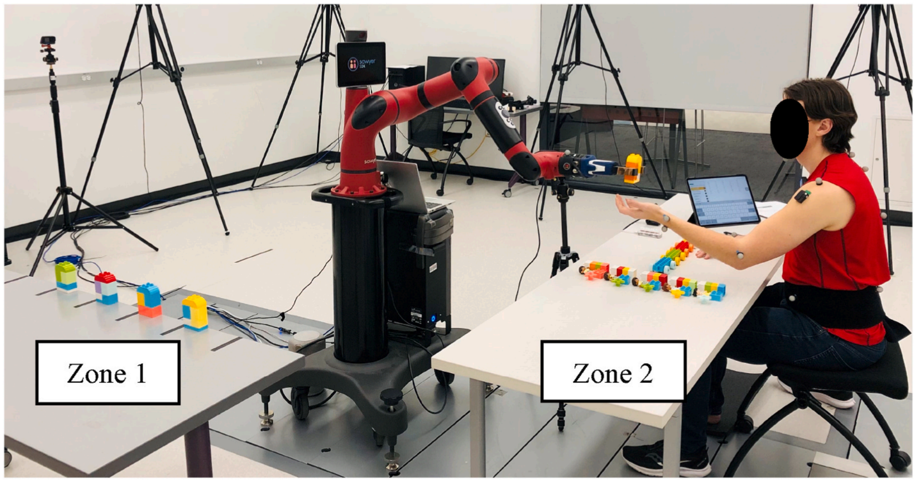

Fig. 1. The collaborative assembly task adopted in the experiment.

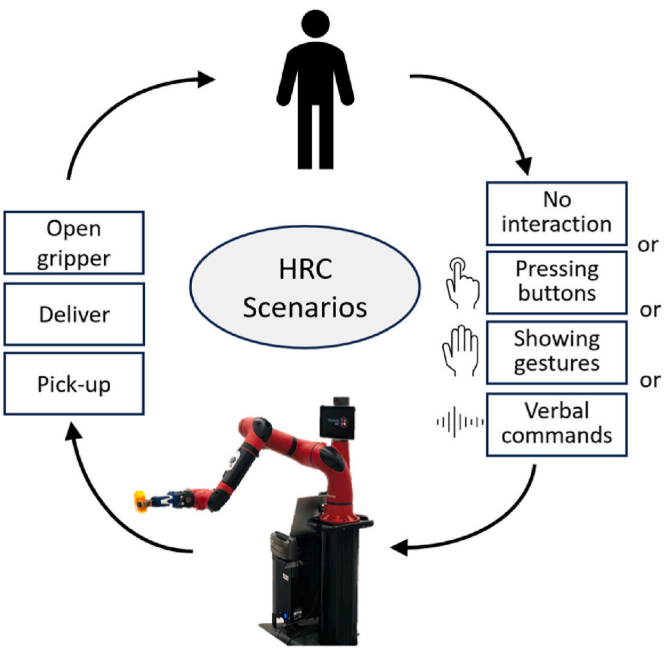

Fig. 2. HRC scenarios in the experiment.

gripper. In complex interactions, the human operator needs to communicate with the robot to initialize the delivery of the large Lego block in addition to requesting the robot to open the gripper. The robot would deliver the block immediately upon receiving the command, while open its gripper in 2 s after receiving the commands.

For both simple and complex interaction, three interaction modalities were adopted, including pressing buttons, showing hand gestures, and giving verbal commands. These three interaction modalities were achieved with off-the-shelf hardware, i.e., a numeric mini keyboard, a webcam, and a wireless microphone, as shown in Fig. 3. For pressing buttons, the numeric programmable keyboard was placed on the table and adjusted to the most comfortable position according to each participant's preference. It was connected to the same workstation that controls the co-robot, with key "1" programmed to deliver the next Lego block and key "4" programmed to open the gripper. In terms of gestures, an RGB webcam (Logitech BRIO) was also placed on the table to capture the hand movements and gestures. Twenty-one landmarks of the hand were first extracted using Google MediaPipe Hands and then fed into a self-trained gesture recognition model. The "OK" gesture was used to deliver the next Lego block, and the "Open" gesture was used to open the gripper, as is illustrated in Fig. 4. Furthermore, a wireless microphone (RODE) was clipped to the collar of the participant's clothes and connected to the workstation via a remote receiver for recording verbal commands. Google Speech API was adopted for speech recognition. The operator must utter a statement containing "open", such as "please open the gripper" to open the gripper, and a statement containing "next", such as "please deliver the next", to pass the next Lego block. All the algorithms for realizing different HRI paradigms were implemented in Python (ver. 3.8).

Fig. 3. Devices used to interact with the co-robot (from left to right): extended keyboard, webcam, and wireless microphone.

2.3. Experiment procedure

The operational procedure of the experiment can be summarized into three phases: pre-experiment, in-experiment, and post-experiment, as is shown in the flowchart in Fig. 5. Before the experiment, every

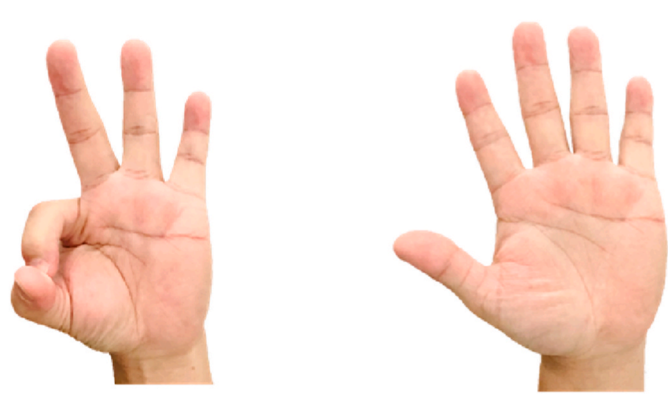

Fig. 4. Hand gestures: "OK" sign (left) and "Open" sign (right).

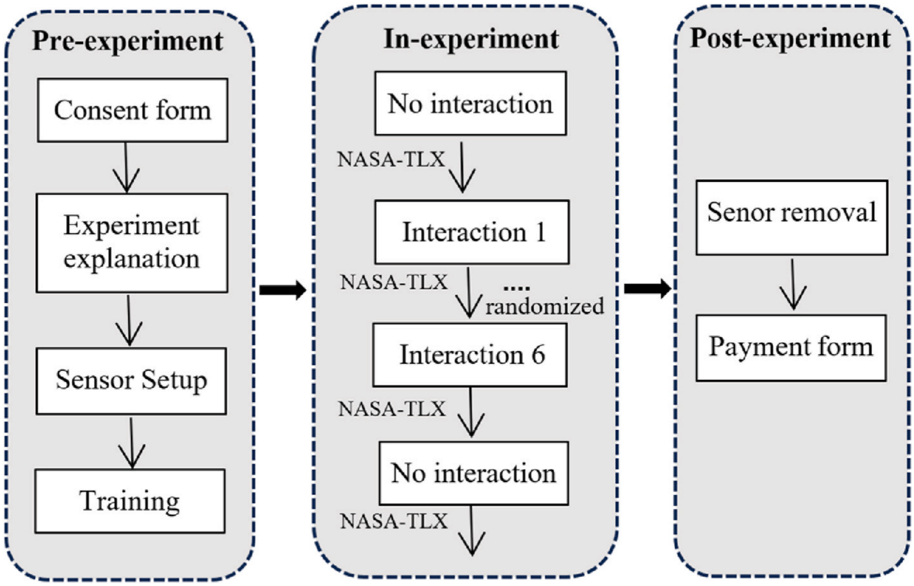

Fig. 5. Flow chart of the experimental procedure.

participant would sign a consent form. The researcher then introduced the experiment and explained how to use the devices to interact with the co-robot. Afterward, sensors were attached to the participants' bodies, and each participant received a training session. During this training, participants engaged in the assembly of Lego blocks to gain proficiency with the associated tasks.

Each experiment consisted of eight sessions, including two no-interaction sessions (one at the beginning and one at the end of the experiment) and six interaction combinations (2 interaction complexity × 3 interaction modality). The implementation of two sessions to assess the no-interaction mental stress level attempted to eliminate potential time order effects and aimed to enhance the statistical power for the no-interaction condition. Also, the sequence of the interaction combinations was randomized using a split-plot randomization approach, where interaction complexity served as the whole-plot factor and interaction modality as the subplot factor. Between each session, there was a 4-min period for participants to take a break and answer the self-report questionnaire.

2.4. Data acquisition and analysis

Galvanic Skin Response (GSR). In studies involving emotional arousal, skin conductance is a commonly used physiological measure that refers to the varying electrical properties of the skin in response to sweat secretion by sweat glands. In particular, eccrine sweat glands are mostly involved in emotional responses (Dawson et al., 2007). It is recommended that GSR should be recorded in areas with a high density of eccrine sweat glands, such as the palms and soles (Saga, 2002). Also, the findings of van Dooren et al. (2012) indicated that the feet, fingers, and shoulders were the most responsive recording locations. Therefore, we chose foot as the recording location since both hands were occupied in the collaborative assembly tasks. A non-invasive Shimmer 3 GSR + device was used in this experiment, and two electrodes were placed on the medial side of the foot sole, as shown in Fig. 6.

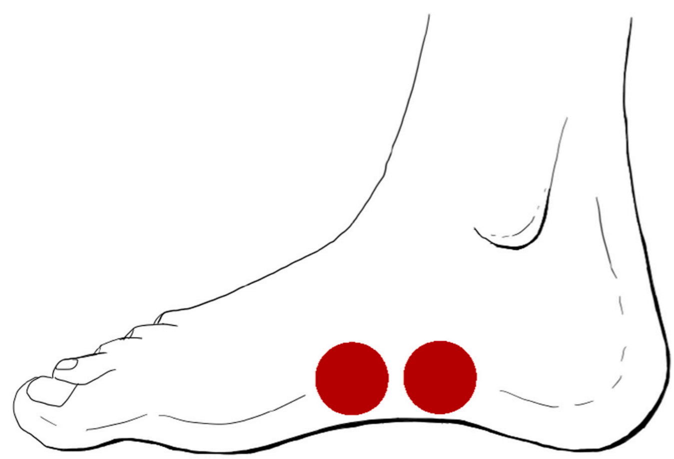

Fig. 6. Locations of GSR electrodes.

Skin conductance is composed of skin conductance level (SCL) and skin conductance response (SCR). SCL is the tonic level which refers to the absolute conductance level in the absence of a measurable stimulus. SCR is the phasic increases superimposed on SCL, reflecting the response to internal or external stimuli (Dawson et al., 2007). Therefore, the phasic component of the GSR data was extracted and used to evaluate the participant's mental stress level.

GSR data processing was performed by using Ledalab, a MATLAB-based software. The original signals measured in volts at a rate of 256 Hz were first down-sampled to 64 Hz and then decomposed into tonic and phasic components using continuous decomposition analysis (Benedek and Kaernbach, 2010). The mean values of SCR for each session were used as an indicator of mental stress levels. As SCR is susceptible to individual differences, the obtained data were normalized within each individual, i.e., divided by the maximum mean value of all sessions (Arai et al., 2010). Therefore, the normalized mean values were used for statistical analysis.

Electromyography (EMG). The validity of detecting mental stress through EMG data from the Trapezius muscles has been confirmed in previous studies (Lundberg et al., 1994; Robles et al., 2022). For the data acquisition, a wireless surface EMG sensor (Delsys Trigno) was employed and placed on the upper right trapezius muscle at the neck and shoulder junction (Perotto, 2005), as depicted in Fig. 7. The EMG signal was originally recorded in volts at a rate of 2000 Hz.

The preprocessing in MATLAB involved several steps, including band-pass filtering within the frequency range of 20–400 Hz, rectification, and the extraction of the EMG envelope through the calculation of the root mean square value using a sliding window of 200ms (Farfán et al., 2010). Subsequently, the mean values of post-processed EMG data for each session were used as an indicator of mental stress levels.

Fig. 7. The EMG sensor at upper right trapezius muscle.

Similarly, the mean data was normalized against the maximum value across all sessions for statistical analysis.

Heart rate (HR). In comparison to heart rate variability, HR measurements are less susceptible to motion artifacts (Królak et al., 2020). This characteristic makes heart rate measurements more suitable for application in dynamic assembly tasks. HR data was collected in bpm at 1 Hz using a Polar H10 heart rate monitor chest strap, which participants wore securely around their chest.

The preprocessing of HR data was also performed using MATLAB. Specifically, samples where heartbeats were recorded as 0 were excluded, which could happen if the chest strap failed to make reliable contact with the skin during the assembly tasks. Likewise, the mean values of HR data for each session were used as an indicator of mental stress levels, and the normalized mean data was used for statistical analysis.

NASA-Task Load Index (NASA-TLX). Participants were asked to fill out the NASA-TLX questionnaires after completing each HRC session. NASA-TLX performs a multi-dimensional assessment of the overall mental workload based on six subscales, including mental demand, physical demand, temporal demand, performance, effort, and frustration. For each dimension, the response scale is essentially a bipolar description (e.g., Low/High), with a line of 21 marks. Values were rounded up if a participant marked between two tick marks. The average of the six subscales was used for statistical analysis (Zakeri et al., 2021).

Statistical Analysis. The analyses were conducted in JMP Pro and included Analysis of variance (ANOVA) tests and Tukey HSD post hoc tests. Besides the factors we intended to study, participants were considered blocking factors, and time order was considered a covariate. The statistical significance level was set at 0.05.

3. Results

3.1. Interaction presence

Subjective measures (NASA-TLX) and objective measures (GSR, EMG, and HR data) were employed to analyze the impact of the interaction presence factor. Statistical analysis involved conducting one-way repeated measures ANOVA to analyze the effects of HRI presence, utilizing the mean values from various sessions for each measure. The results are presented in Table 1, and all assessments showed a significant effect of the introduction of HRI on human mental stress. In detail, for the subjective assessment using NASA-TLX, $F(1, 143) = 43.3695$, $p < 0.0001$; for GSR, $F(1, 167) = 39.0856$, $p < 0.0001$; for EMG, $F(1, 167) = 16.1766$, $p < 0.0001$; for HR, $F(1, 167) = 83.0726$, $p < 0.0001$.

Table 1
ANOVA results for interaction presence.

	NASA-TLX	GSR	EMG	HR
F ratio	43.3695	39.0856	16.1766	83.0726
p-value	<.0001*	<.0001*	<.0001*	<.0001*

The results demonstrated that interactions between human operators and the co-robot played a significant role in alleviating mental stress during collaborative assembly tasks, as illustrated in Fig. 8. Notably, the introduction of HRI significantly contributed to lowering mental stress levels among participants. Specifically, for the subjective measure, the NASA-TLX index decreased from 8.80 to 5.58. Similarly, for objective measures, GSR levels decreased from 0.89 to 0.58, EMG levels decreased from 0.14 to 0.12, and HR levels decreased from 0.99 to 0.94, correspondingly.

3.2. Interaction complexity

Furthermore, two-way repeated measures ANOVA was performed to analyze the effects of interaction complexity and interaction modality. The ANOVA results in terms of interaction complexity were detailed in Table 2. The subjective assessment using NASA-TLX revealed that the effects on mental stress were statistically significant, with $F(1,137) = 32.3288$, $p < 0.0001$. This finding was further supported by the objective GSR assessment, with $F(1,137) = 6.8893$, $p = 0.0099$. However, the other two objective assessments involving EMG and HR did not present significant differences, showing $F(1,137) = 0.1017$, $p = 0.7440$ for EMG and $F(1,137) = 0.0663$, $p = 0.7972$ for HR.

The outcomes from NASA-TLX and GSR indicated that as interactions became more complex, they negatively impacted and elevated mental stress levels. As is shown in Fig. 9, complex interactions resulted in a mental stress increase in the NASA-TLX index from 4.82 to 6.36, and an increase in GSR levels from 0.52 to 0.63, compared to simple interactions. In contrast, the mental stress levels assessed from EMG and HR appeared to be consistent across both simple and complex interactions.

3.3. Interaction modality

Likewise, regarding interaction modality, the results from the two-way repeated measures ANOVA are presented in Table 3. The assessments utilizing NASA-TLX, GSR, and EMG indicated significant effects of interaction modality on mental stress, whereas HR assessment showed non-significant effects (for NASA-TLX, $F(2,137) = 5.9333$, $p = 0.00344$; for GSR, $F(2,137) = 5.0339$, $p = 0.0080$; for EMG, $F(2,137) = 3.5576$, $p = 0.0317$; for HR, $F(2,137) = 0.0344$, $p = 0.9662$).

Post-hoc Tukey HSD tests were performed on measures that yielded significant results, which are detailed in Table 4. Objective and subjective assessments presented some disparities, as depicted in Fig. 10. For the subjective NASA-TLX evaluation, the tests confirmed that using hand gestures led to the highest mental stress levels among the three interaction modalities. It resulted in significantly higher mental stress than both button presses at $p = 0.0042$ and verbal commands at $p = 0.0262$. While verbal commands produced slightly more stress than button pressing, the difference was not significant. The GSR physiological data supported these findings, with hand gestures causing significantly more mental stress than pressing buttons at $p = 0.0061$. However, no significant difference was observed between hand gestures and verbal commands, nor between button pressing and verbal commands. On the other hand, the EMG data showed that hand gestures induced significantly higher stress than verbal commands at $p = 0.0437$, but there were no notable differences when comparing gestures to button pressing or verbal commands to button pressing. For HR assessments, all interaction methods resulted in non-significant stress effects. Collectively, the outcomes consistently showed non-significant effects between button pressing and verbal commands.

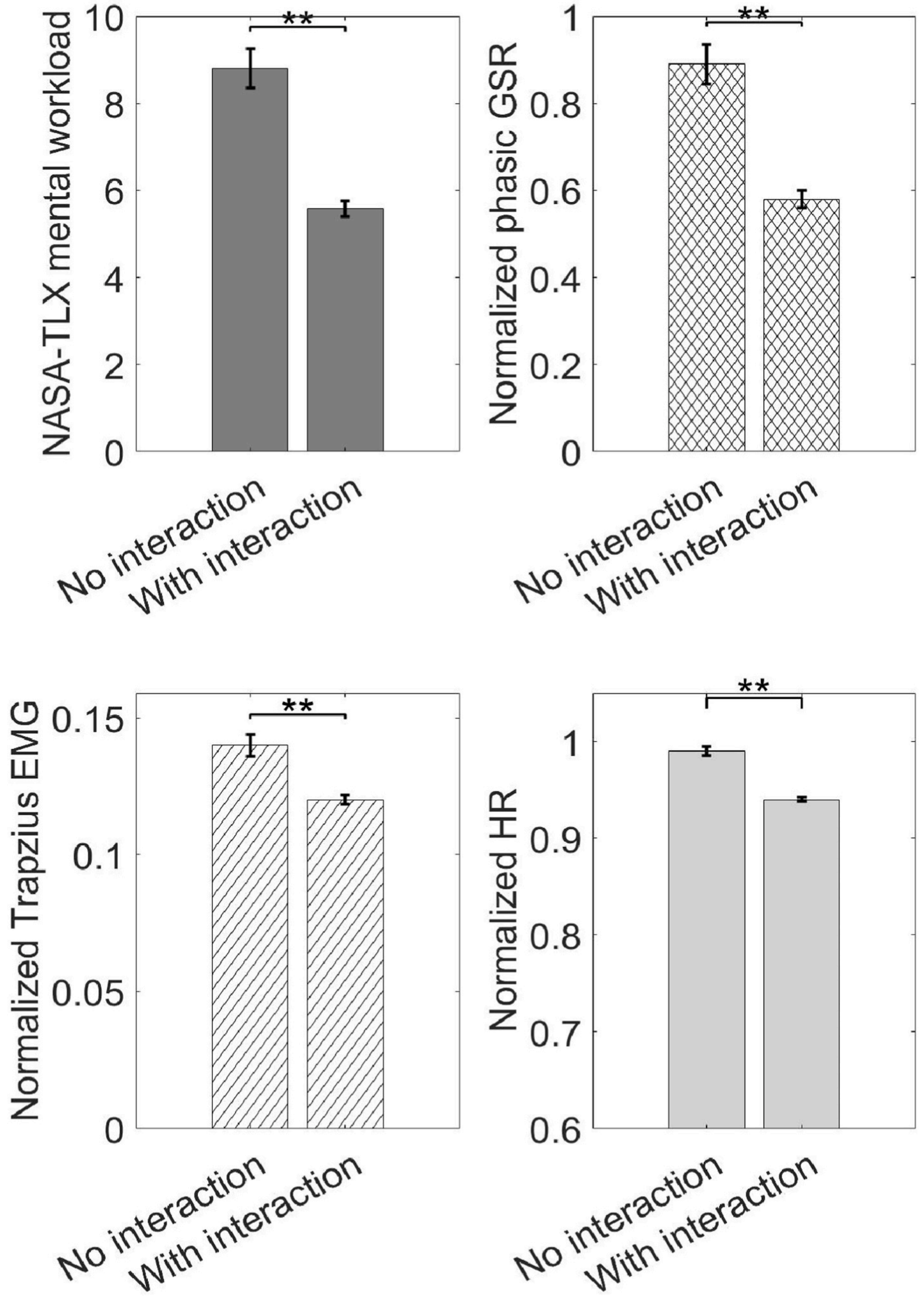

Fig. 8. Effects of the presence of interactions
(**: statistically significant with p < 0.01).

Table 2
ANOVA results for interaction complexity.

	NASA-TLX	GSR	EMG	HR
F ratio	32.3288	6.8893	0.1071	0.0663
p-value	<.0001*	.0099*	.7440	.7972

3.4. Interaction effects

The interaction effects of complexity and modality were also tested, while no significant effects were found, as shown in Table 5 (for NASA-TLX, $F(2,137) = 1.9001$, $p = 0.1537$; for GSR, $F(2,137) = 1.0548$, $p = 0.3516$; for EMG, $F(2,137) = 0.0269$, $p = 0.9734$; for HR, $F(2,137) = 0.0689$, $p = 0.9335$).

4. Discussion

Stress is commonly defined as a state of imbalance between environmental demands and an individual's capabilities, leading to a shift from a calm state to an excited state in order to preserve the organism's integrity (Alsuraykh et al., 2019). Some researchers distinguish between "eustress", a positive form of stress such as joy or excitement, and "distress", a negative form. In this study, mental stress is referred to as distress, emphasizing its negative nature on workers' health.

The subjective assessments using NASA-TLX indicated that participants experienced significantly higher mental stress when working with co-robots without interaction. This finding was also confirmed by all objective assessments including GSR, EMG, and HR. In the absence of interaction, human operators must adapt to the pace of the co-robot and take over the Lego blocks in time to prevent them from falling, which

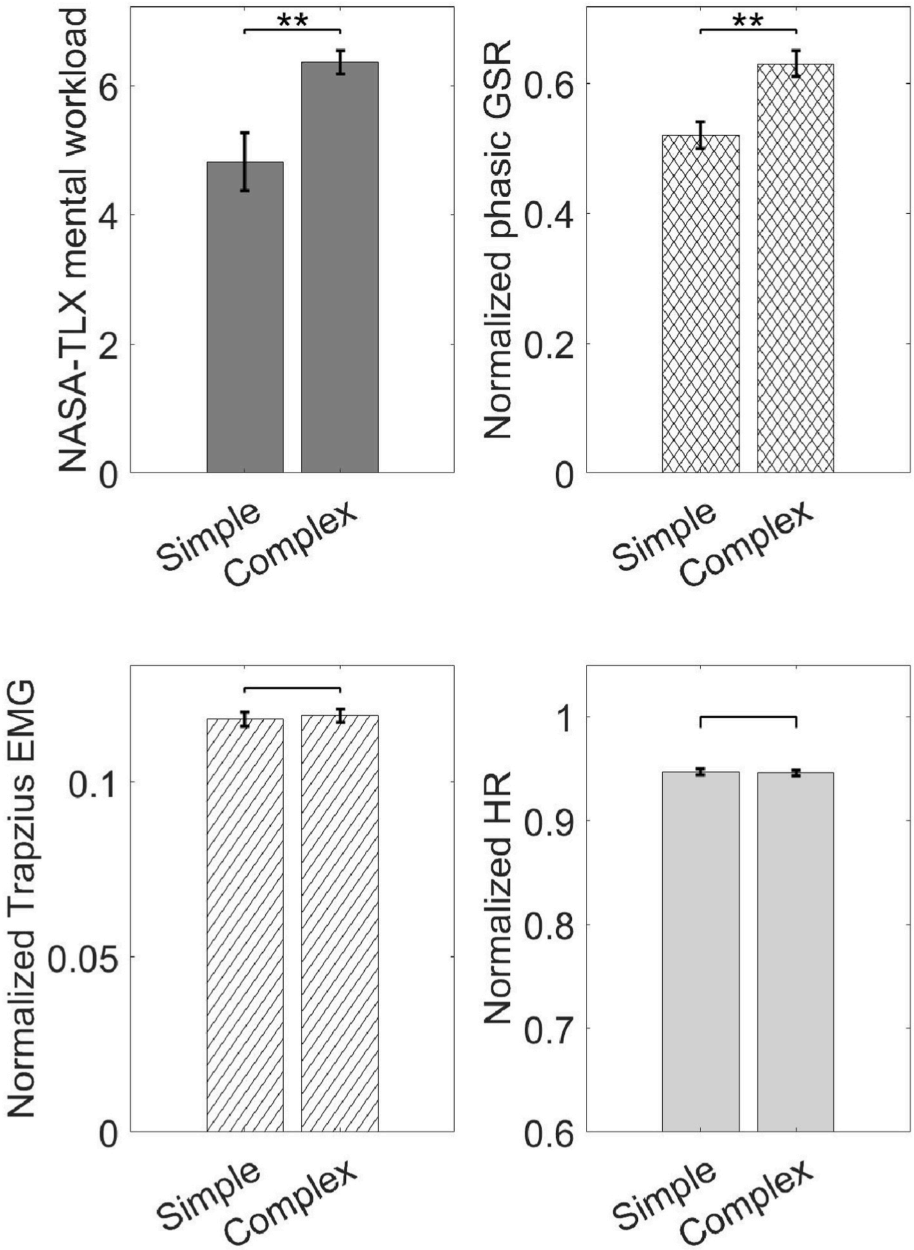

Fig. 9. ANOVA results of interaction complexity (*: statistically significant with p < 0.05; **: statistically significant with p < 0.01).

Table 3
ANOVA results for interaction modality.

	NASA-TLX	GSR	EMG	HR
F ratio	5.9333	5.0339	3.5576	0.0344
p-value	.0034*	.0080*	.0317*	0.9662

Table 4
Post-hoc Tukey HSD tests for interaction modality.

Group	p-values		
	NASA-TLX	GSR	EMG
Button - Gesture	.0042*	.0061*	.0437*
Gesture - Voice	.0262*	.4713	.9614
Voice - Button	.8074	.1239	.0816

imposes more mental demands on operators. In contrast, in the presence of interaction, the human operator can communicate with the co-robot when they need the co-robot to deliver the next Lego block or open its gripper. Thereby, human operators can adjust the pace of the HRC process according to their needs, which helps reduce mental stress during the collaboration. The current finding was aligned with the study conducted by Gervasi et al. (2022), in which it was found that the introduction of HRI in collaborative tasks reduced perceived mental stress, as the control of task execution time with robots is a significant influential factor for mental stress during HRC.

Furthermore, the consistency between the results of subjective assessments and objective measurements showcased the positive correlation between the employed subjective and objective measures. It was noted that participants tended to have increased physiological responses, as measured objectively by GSR, EMG, and HR, when they reported higher levels of mental stress through subjective measures like NASA-TLX. This finding aligns with the circumplex model of affect, which interprets emotions as a combination of two neurophysiological

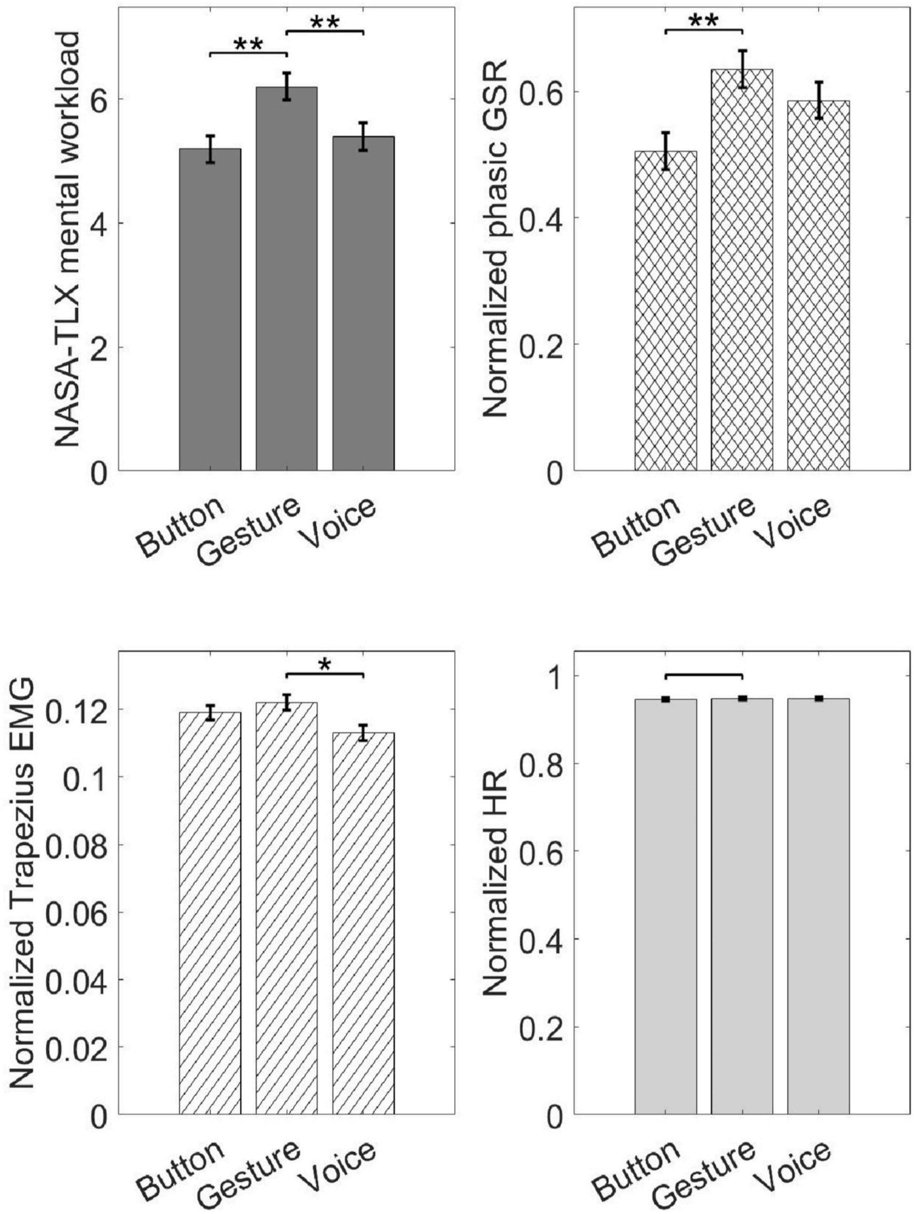

Fig. 10. ANOVA results of interaction modality: button, gesture, and voice (*: statistically significant with $p < 0.05$).

Table 5
ANOVA results for interaction effects of complexity and modality.

	NASA-TLX	GSR	EMG	HR
F ratio	1.9001	1.0548	0.0269	0.0689
p-value	.1537	.3516	.9734	.9335

dimensions: valence and arousal (Colibazzi et al., 2010; Russell, 1988). Valence refers to the subjectively experienced emotions, ranging from highly negative to extremely positive, while arousal reflects the state of responsiveness to sensory stimuli. In this context, mental stress is characterized by elevated arousal and negative valence. Moreover, studies have shown that there is a positive correlation between emotional arousal and physiological responses (Lang and Davis, 2006; Gjoreski et al., 2022). These findings are consistent with the trends observed in our study that higher perceived mental stress is associated with an elevation in physiological responses.

In terms of interaction complexity, subjective assessment and objective GSR assessment came to the same conclusion. While the introduction of interactions in HRC can help reduce mental stress to some extent, complex interactions can lead to higher levels of mental stress compared to simple interactions. Nomura et al. (2008) also found similar results that the anxiety increased after repeated interaction with a robot. It is noted that interactions with collaborative robots can introduce their own sources of stress for human operators. In addition, empirical findings highlight a significant positive correlation between mental stress and mental workload (Hou et al., 2015), which is described as the relationship between primary task performance and the resources required for the task (Wilson and Sharples, 2015). Consequently, as individuals experience a higher mental workload as interaction complexity increases, they are more likely to encounter elevated levels of mental stress. In HRC context, it becomes important to thoughtfully design the level and scope of interactions to ensure the resultant additional mental workload does exceed workers' mental capacity.

The results also revealed statistically significant differences among different interaction modalities. Specifically, the utilization of hand gestures resulted in significantly higher levels of mental stress than both button pressing and verbal commands, as indicated by subjective NASA-

TLX assessments. Meanwhile, objective GSR and EMG assessments also highlighted that showing hand gestures induced significantly higher mental stress during collaborative assembly tasks. These observations were in line with the findings of a study conducted by Xuan et al. (2019), it was noted that using gestures to control a TV in a home environment setting caused more mental stress than using a remote button control panel. It is reasonable to expect that pressing buttons would result in lower mental stress compared to using hand gestures. Pressing buttons is a more reliable method, while showing hand gestures relies on artificial intelligence (AI) which has recognition accuracy issues, leading to confusions.

Nonetheless, using verbal commands, another AI-based interaction method, also induced lower mental stress compared to using hand gestures. This outcome can be explained by the dual-coding theory (Paivio, 1971) and the multiple memory system concept (Constantinou, 2019), both of which relate to mental workload. These theories suggest that humans process verbal and visual/spatial information through distinct channels of the brain. Overloading one channel can lead to decreased performance and increased mental workload. In the context of assembly tasks, which are inherently spatial, incorporating the processing of hand gestures (another spatial task) places additional demand on the brain's spatial processing channel. This dual demand on the same cognitive pathway leads to a more significant mental workload and consequently, higher mental stress. In contrast, verbal commands are processed through a different cognitive channel. This separation allows more efficient parallel processing and reduces the overall mental workload, thus lowering mental stress. As a result, during assembly tasks, the use of hand gestures leads to higher mental stress compared to using verbal commands.

Regarding the effectiveness of different physiological measures, GSR emerges as a more sensitive indicator for detecting mental stress, showing most consistent results compared to subjective stress assessments using the NASA-TLX. Conversely, heart rate appears to be less sensitive in capturing subtle variations in mental stress levels. In addition, there are certain limitations associated with the use of EMG in this study. Ideally, the placement of EMG sensors should avoid muscles involved in the tasks being performed. In this work, the EMG sensor was positioned on the upper trapezius muscle, which is minimally involved during sedentary assembly tasks or when delivering commands such as verbal instructions, gestures, and button pressing. However, upon reviewing the results, it was noted that participants tended to lower their heads while pressing buttons, which inadvertently increased the EMG signal magnitude. This occurred because the mini keyboard was placed on the table, prompting participants to lower their heads. In contrast, participants maintained a forward-facing head posture when showing hand gestures or giving verbal commands. This limitation is also reflected in the Results section, where EMG measurements for button pressing yielded higher values than for verbal commands, even though NASA-TLX and GSR data indicated that button pressing resulted in lower perceived stress. Therefore, it is crucial to carefully consider the placement of EMG sensors to accurately measure mental stress.

Yet, it is essential to acknowledge the disparity between subjective and objective assessments. Objective GSR assessments demonstrated significant differences exclusively between hand gestures and button pressing, while EMG assessments solely indicated significant distinctions between hand gestures and verbal commands. Additionally, no significant effects on mental stress were found through objective HR assessments. One potential explanation for this difference is that stress can actually occur even when physiological changes are not present because the body's physiological responses are more slowly recognized by the brain than its function to release emotional responses (Dalgleish, 2004). Therefore, it is possible that subjective assessments may yield significant effects, whereas measures reliant on physiological signals may not yield statistically significant results under certain circumstances. Consequently, it is recommended that a combination of multiple physiological measures should be employed in the detection of mental stress to enhance the reliability and accuracy of results.

Overall, the combination of both subjective assessment and various physiological objective assessments in this work has provided convincible insights into the effects of HRI on mental stress. These findings have practical implications for the design of HRC systems. Firstly, it is suggested to incorporate interactions as an integral part of the collaboration process, allowing human operators to exert control over the pace of task execution. Secondly, simple and straightforward interactions are preferred, while complex or redundant interactions should be avoided. Thirdly, the selection of interaction modalities should be task specific. For instance, in assembly tasks, the inclusion of verbal commands as a substitute for traditional buttons is advisable, while the use of hand gestures should be avoided. By considering these recommendations, more effective and user-friendly HRC systems that prioritize mental well-being and overall performance can be designed.

5. Conclusion

Collaborative robots are increasingly being deployed in a variety of industrial settings, sharing a common workplace with human operators. Research focusing on aspects of human psychological effects is critical to achieving workplace wellness for HRC. This study investigated the impact of human-robot interactions on mental stress during collaborative Lego assembly tasks, considering different levels of interaction presence, interaction complexity, and interaction modality. Subjective measures using the NASA-TLX index and objective measures using the physiological data (GSR, EMG, and HR) were applied to assess the level of mental stress. The findings provided practical insights for optimizing the HRC process with a focus on reducing mental stress on human teammates.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This manuscript is based upon work supported by the National Science Foundation under Grant #2024688.

References

Aigrain, J., Spodenkiewicz, M., Dubuisson, S., Detyniecki, M., Cohen, D., Chetouani, M., 2018. Multimodal stress detection from multiple assessments. IEEE Transactions on Affective Computing 9 (4), 491–506. https://doi.org/10.1109/TAFFC.2016.2631594.

Alsuraykh, N.H., Wilson, M.L., Tennent, P., Sharples, S., 2019. How stress and mental workload are connected. In: Proceedings of the 13th EAI International Conference on Pervasive Computing Technologies for Healthcare, pp. 371–376. https://doi.org/10.1145/3329189.3329235.

Arai, T., Kato, R., Fujita, M., 2010. Assessment of operator stress induced by robot collaboration in assembly. CIRP Annals 59 (1), 5–8. https://doi.org/10.1016/j.cirp.2010.03.043.

Benedek, M., Kaernbach, C., 2010. A continuous measure of phasic electrodermal activity. J. Neurosci. Methods 190 (1), 80–91. https://doi.org/10.1016/j.jneumeth.2010.04.028.

Bethel, C.L., Salomon, K., Murphy, R.R., Burke, J.L., 2007. Survey of Psychophysiology Measurements Applied to Human-Robot Interaction. *RO-MAN 2007 - the 16th IEEE International Symposium On Robot And Human Interactive Communication*, pp. 732–737. https://doi.org/10.1109/ROMAN.2007.4415182.

Bonarini, A., 2020. Communication in human-robot interaction. Current Robotics Reports 1 (4), 279–285. https://doi.org/10.1007/s43154-020-00026-1.

Cherubini, A., Passama, R., Crosnier, A., Lasnier, A., Fraisse, P., 2016. Collaborative manufacturing with physical human–robot interaction. Robot. Comput. Integrated Manuf. 40, 1–13. https://doi.org/10.1016/j.rcim.2015.12.007.

Chowdhury, R., Reaz, M., Ali, M., Bakar, A., Chellappan, K., Chang, T., 2013. Surface electromyography signal processing and classification techniques. Sensors 13 (9), 12431–12466. https://doi.org/10.3390/s130912431.

Dalgleish, T., 2004. The emotional brain. Nat. Rev. Neurosci. 5 (7), 583–589. https://doi.org/10.1038/nrn1432.

Dawson, M.E., Schell, A.M., Filion, D.L., 2007. The electrodermal system. In: Handbook of Psychophysiology, third ed. Cambridge University Press, pp. 159–181. https://doi.org/10.1017/CBO9780511546396.007.

Dragan, A.D., Bauman, S., Forlizzi, J., 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, pp. 51–58. https://doi.org/10.1145/2696454.2696473.

Duffy, V.G., Or, C.K.L., Lau, V.W.M., 2006. Perception of safe robot speed in virtual and real industrial environments. Hum. Factors Ergon. Manuf. 16 (4), 369–383. https://doi.org/10.1002/hfm.20058.

Farfán, F.D., Politti, J.C., Felice, C.J., 2010. Evaluation of EMG processing techniques using Information Theory. Biomed. Eng. Online 9 (1), 72. https://doi.org/10.1186/1475-925X-9-72.

Fauquet-Alekhine, P., Rouillac, L., Bertoni, J., Granry, J.-C., 2016. Heart rate vs stress indicator for short term mental stress. Br. J. Med. Med. Res. 17 (7). Article 7.

Galin, R.R., Meshcheryakov, R.V., 2020. Human-robot interaction efficiency and human-robot collaboration. In: Kravets, A.G. (Ed.), Robotics: Industry 4.0 Issues & New Intelligent Control Paradigms. Springer International Publishing, pp. 55–63. https://doi.org/10.1007/978-3-030-37841-7_5.

Gervasi, R., Aliev, K., Mastrogiacomo, L., Franceschini, F., 2022. User experience and physiological response in human-robot collaboration: a preliminary investigation. J. Intell. Rob. Syst. 106 (2), 36. https://doi.org/10.1007/s10846-022-01744-8.

Giakoumis, D., Drosou, A., Cipresso, P., Tzovaras, D., Hassapis, G., Gaggioli, A., Riva, G., 2012. Using activity-related behavioural features towards more effective automatic stress detection. PLoS One 7 (9), e43571. https://doi.org/10.1371/journal.pone.0043571.

Gjoreski, M., Kiprijanovska, I., Stankoski, S., Mavridou, I., Broulidakis, M.J., Gjoreski, H., Nduka, C., 2022. Facial EMG sensing for monitoring affect using a wearable device. Sci. Rep. 12 (1), 1. https://doi.org/10.1038/s41598-022-21456-1.

Healey, J.A., Picard, R.W., 2005. Detecting stress during real-world driving tasks using physiological sensors. IEEE Trans. Intell. Transport. Syst. 6 (2), 156–166. https://doi.org/10.1109/TITS.2005.848368.

Hou, X., Liu, Y., Sourina, O., Tan, Y.R.E., Wang, L., Mueller-Wittig, W., 2015. EEG based stress monitoring. In: 2015 IEEE International Conference on Systems, Man, and Cybernetics, pp. 3110–3115. https://doi.org/10.1109/SMC.2015.540.

Jokinen, K., Wilcock, G., 2013. User Experience in Human-Robot Interactions. 4th International Workshop on Perceptual Quality of Systems. https://doi.org/10.21437/PQS.2013–24. *PQS 2013)*, 125–130.

Królak, A., Wiktorski, T., Bjørkavoll-Bergseth, M.F., Ørn, S., 2020. Artifact correction in short-term HRV during strenuous physical exercise. Sensors 20 (21). https://doi.org/10.3390/s20216372. Article 21.

Lang, P.J., Davis, M., 2006. In: Anders, S., Ende, G., Junghofer, M., Kissler, J., Wildgruber, D. (Eds.), Progress in Brain Research, Vol. 156,. Elsevier, pp. 3–29. https://doi.org/10.1016/S0079-6123(06)56001-7.

Linssen, L., Landman, A., van Baardewijk, J.U., Bottenheft, C., Binsch, O., 2022. Using accelerometry and heart rate data for real-time monitoring of soldiers' stress in a dynamic military virtual reality scenario. Multimed. Tool. Appl. 81 (17), 24739–24756. https://doi.org/10.1007/s11042-022-12705-6.

Lu, L., Xie, Z., Wang, H., Su, B., Edward, P., Fitts, X.X., 2022. Assessing workers' mental stress in hand-over activities during human-robot collaboration. Proc. Hum. Factors Ergon. Soc. Annu. Meet. 66 (1), 537–541. https://doi.org/10.1177/1071181322661194.

Lundberg, U., Kadefors, R., Melin, B., Palmerud, G., Hassmén, P., Engström, M., Elfsberg Dohns, I., 1994. Psychophysiological stress and emg activity of the trapezius muscle. Int. J. Behav. Med. 1 (4), 354–370. https://doi.org/10.1207/s15327558ijbm0104_5.

Meziane, R., Li, P., Otis, M.J.-D., Ezzaidi, H., Cardou, P., 2014. Safer hybrid workspace using human-robot interaction while sharing production activities. In: 2014 IEEE International Symposium on Robotic and Sensors Environments (ROSE) Proceedings, pp. 37–42. https://doi.org/10.1109/ROSE.2014.6952980.

Mohammed, A., Schmidt, B., Wang, L., 2017. Active collision avoidance for human–robot collaboration driven by vision sensors. Int. J. Comput. Integrated Manuf. 30 (9), 970–980. https://doi.org/10.1080/0951192X.2016.1268269.

Murashov, V., Hearl, F., Howard, J., 2016. Working safely with robot workers: recommendations for the new workplace. J. Occup. Environ. Hyg. 13 (3), D61–D71. https://doi.org/10.1080/15459624.2015.1116700.

Nomura, T., Kanda, T., Suzuki, T., Kato, K., 2008. Prediction of human behavior in human–robot interaction using psychological scales for anxiety and negative attitudes toward robots. IEEE Trans. Robot. 24 (2), 442–451. https://doi.org/10.1109/TRO.2007.914004.

Ötting, S.K., Masjutin, L., Steil, J.J., Maier, G.W., 2022. Let's work together: a meta-analysis on robot design features that enable successful human–robot interaction at work. Hum. Factors 64 (6), 1027–1050. https://doi.org/10.1177/0018720820966433.

Perotto, A., 2005. Anatomical Guide for the Electromyographer: the Limbs and Trunk. Charles C Thomas Publisher.

RAHIMI, M., KARWOWSKI, W., 1990. Human perception of robot safe speed and idle time. Behav. Inf. Technol. 9 (5), 381–389. https://doi.org/10.1080/01449299008924252.

Robelski, S., Wischniewski, S., 2018. Human-machine interaction and health at work: a scoping review. Int. J. Hum. Factors Ergon. 5 (2), 93–110. https://doi.org/10.1504/IJHFE.2018.092226.

Robles, D., Benchekroun, M., Zalc, V., Istrate, D., Taramasco, C., 2022. Stress detection from surface electromyography using convolutional neural networks. In: 2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), pp. 3235–3238. https://doi.org/10.1109/EMBC48229.2022.9871860.

Rubagotti, M., Tusseyeva, I., Baltabayeva, S., Summers, D., Sandygulova, A., 2022. Perceived safety in physical human–robot interaction—a survey. Robot. Autonom. Syst. 151, 104047 https://doi.org/10.1016/j.robot.2022.104047.

Saga, K., 2002. Structure and function of human sweat glands studied with histochemistry and cytochemistry. Prog. Histochem. Cytochem. 37 (4), 323–386. https://doi.org/10.1016/S0079-6336(02)80005-5.

Shi, Y., Ruiz, N., Taib, R., Choi, E., Chen, F., 2007. Galvanic skin response (GSR) as an index of cognitive load. CHI '07 Extended Abstracts on Human Factors in Computing Systems 2651–2656. https://doi.org/10.1145/1240866.1241057.

Umer, W., 2022. Simultaneous monitoring of physical and mental stress for construction tasks using physiological measures. J. Build. Eng. 46, 103777 https://doi.org/10.1016/j.jobe.2021.103777.

van Dooren, M., de Vries, J.J.G., Gert, -J., Janssen, J.H., 2012. Emotional sweating across the body: comparing 16 different skin conductance measurement locations. Physiol. Behav. 106 (2), 298–304. https://doi.org/10.1016/j.physbeh.2012.01.020.

Waard, D. de, 1996. The Measurement of Drivers' Mental Workload. Rijksuniv.

Wang, L., Gao, R., Váncza, J., Krüger, J., Wang, X.V., Makris, S., Chryssolouris, G., 2019. Symbiotic human-robot collaborative assembly. CIRP Annals 68 (2), 701–726. https://doi.org/10.1016/j.cirp.2019.05.002.

Wijsman, J., Grundlehner, B., Penders, J., Hermens, H., 2013. Trapezius muscle EMG as predictor of mental stress. ACM Trans. Embed. Comput. Syst. 12 (4), 1–20. https://doi.org/10.1145/2485984.2485987.

Wilson, J.R., Sharples, S. (Eds.), 2015. Evaluation of Human Work, 4th ed. CRC Press https://doi.org/10.1201/b18362.

Xie, Z., Lu, L., Wang, H., Li, L., Fitts, E.P., Xu, X., 2022. A human-robot collision avoidance method using a single camera. Proc. Hum. Factors Ergon. Soc. Annu. Meet. 66 (1), 2244–2248. https://doi.org/10.1177/1071181322661540.

Xuan, L., Daisong, G., Moli, Z., Jingya, Z., Xingtong, L., Siqi, L., 2019. Comparison on user experience of mid-air gesture interaction and traditional remotes control. In: Proceedings of the Seventh International Symposium of Chinese CHI, pp. 16–22. https://doi.org/10.1145/3332169.3333570.

Zakeri, Z., Omurtag, A., Breedon, P., Hilliard, G., Khalid, A., 2021. Studying mental stress factor in occupational safety in the context of the smart factory. In: Proceedings of the 31st European Safety and Reliability Conference (ESREL 2021), pp. 92–99. https://doi.org/10.3850/978-981-18-2016-8_024-cd.

Zheng, B., Jiang, X., Tien, G., Meneghetti, A., Panton, O.N.M., Atkins, M.S., 2012. Workload assessment of surgeons: correlation between NASA TLX and blinks. Surg. Endosc. 26 (10), 2746–2750. https://doi.org/10.1007/s00464-012-2268-6.

Zheng, R., Yamabe, S., Nakano, K., Suda, Y., 2015. Biosignal analysis to assess mental stress in automatic driving of trucks: palmar perspiration and masseter electromyography. Sensors 15 (3). https://doi.org/10.3390/s150305136. Article 3.