Applied Ergonomics 106 (2023) 103863

Contents lists available at ScienceDirect

Applied Ergonomics

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

Physiological and perceptual consequences of trust in collaborative robots: An empirical investigation of human and robot factors

Sarah K. Hopko [a], Ranjana K. Mehta [a,b,*], Prabhakar R. Pagilla [b]

[a] The Industrial and Systems Engineering Department, Texas A&M University, College Station, Tx, USA
[b] The Mechanical Engineering Department, Texas A&M University, College Station, Tx, USA

ARTICLE INFO

Keywords:
Cognitive fatigue
ECG
Sex
Human-robot collaboration
Reliability
Trust

ABSTRACT

Measuring trust is an important element of effective human-robot collaborations (HRCs). It has largely relied on subjective responses and thus cannot be readily used for adapting robots in shared operations, particularly in shared-space manufacturing applications. Additionally, whether trust in such HRCs differ under altered operator cognitive states or with sex remains unknown. This study examined the impacts of operator cognitive fatigue, robot reliability, and operator sex on trust symptoms in collaborative robots through both objective measures (i.e., performance, heart rate variability) and subjective measures (i.e., surveys). Male and female participants were recruited to perform a metal surface polishing task in partnership with a collaborative robot (UR10), in which they underwent reliability conditions (reliable, unreliable) and cognitive fatigue conditions (fatigued, not fatigued). As compared to the reliable conditions, unreliable robot manipulations resulted in perceived trust, an increase in both sympathetic and parasympathetic activity, and operator-induced reduction in task efficiency and accuracy but not precision. Cognitive fatigue was shown to correlate with higher fatigue scores and reduced task efficiency, more severely impacting females. The results highlight key interplays between operator states of fatigue, sex, and robot reliability on both subjective and objective responses of trust. These findings provide a strong foundation for future investigations on better understanding the relationship between human factors and trust in HRC as well as aid in developing more diagnostic and deployable measures of trust.

1. Introduction

The increased use of robotics and automation in the manufacturing industry is partially driven by the desire for more reliable and productive processes and/or products. Robots have a comparative advantage over their human counterparts for their ability to perform repetitive tasks consistently without fatiguing; however, automation can be expensive and inflexible (Akella et al., 1999), (Patel et al., 2012). Because humans are better at improvisation and creative decision making, more assembly lines are using both humans and robots in a shared space environment due to the need for increased flexibility (Krüger et al., 2009). Manufacturing operations such as surface finishing (e.g., grinding, chamfering, deburring, polishing) and material handling are common in many manufacturing sectors (Hu and Pagilla, 2021), (Jin and Pagilla, 2020), and these operations are predominantly performed manually by human operators as complete automation is difficult to achieve due to complex workpiece contour profiles and workpiece registration in the robot workcell. Robotic automation of such operations through collaboration between the human operator and automatic control are being envisioned to benefit from the high repeatability of the robots and the situation awareness of the human operator. With the use of HRC systems, emergent properties resulting from human-robot-interaction must be investigated and considered for improving collaborative performance.

1.1. Trust in collaborative robots

Collaborative robots (cobots) are intelligent assist devices defined by Akella et al. (1999) as “hybrid devices for direct, physical interaction with a human operator in a shared workspace.” The level of collaboration between a cobot and operator tends to increase as the proximity between the entities reduces (Vysocky and Novak, 2016); however, physical interaction can often lead to decreased system safety, performance, and efficiency resulting from inconsideration of emergent human factors, such as trust, anxiety, workload shifts, and corresponding utilization strategy of cobot assistance (Charalambous et al., 2016)–

* Corresponding author. Industrial and Systems Engineering, Texas A&M University, College Station, TX, USA.
E-mail address: rmehta@tamu.edu (R.K. Mehta).

https://doi.org/10.1016/j.apergo.2022.103863
Received 15 October 2021; Received in revised form 23 July 2022; Accepted 30 July 2022
Available online 30 August 2022
0003-6870/

(Lee and Seppelt, 2009). Trust is defined by (Lee and See, 2004) as the "attitude that an agent will help achieve an individual's goal in a situation characterized by uncertainty and vulnerable." Trust influences the operator's reliance on automation and the willingness to share or allocate tasks (Lee and See, 2004), thus proper trust measurement and calibration is important to optimize the efficiency and performance of the system. Miscalibrated trust can result in two undesirable behavior groups: overtrust and undertrust (Lee and See, 2004), (Hancock et al., 2011). Overtrust can result in the operator actively neglecting the cobot (i.e., becoming distracted during the task or reducing maintenance on the robot outside the task), misusing the automation beyond its intended purpose, or over utilizing the automation beyond its capability. In contrast, undertrust can cause operators to reject aid provided by the cobots and can cause significant discomfort to the operator. The initial adoption of cobots is not always successful due to low acceptance and trust of some manufacturing workers of cobot technology (Fletcher et al., 2017), (Meissner et al., 2020).

Trust is a topic discussed in many domains, most predominantly in human-human environments. While trust models have been thoroughly reviewed for this domain, it has been shown that human-human and human-automation trust environments have different fundamental components (Madhavan and Wiegmann, 2007). In the human-human domain there are three bases for trust: ability – the quality of skill or capability to do something, integrity – the quality of being honest and lawful, and benevolence – the quality of being good/kind (Colquitt and Salam, 2015). In contrast to human systems, robots lack intentionality as they are not capable of developing their own intents (Charalambous et al., 2016), (Madhavan and Wiegmann, 2007). Therefore, robots do not deceive people on purpose and cannot be inherently 'good' beings. As such, ability has been argued to be the main basis for trust in HRC (Chen et al., 2018a). However, humans have the tendency to personify collaborative agents and thus can perceive human-like characteristics in cobots such as deception or malevolence (Nass and Moon, 2000), so integrity and benevolence are still relevant bases especially in anthropomorphic, humanoid, or social robotics that intentionally parallel human characteristics. Additionally, trust in collaborative robotics has features directly related to the physical and psychological safety of the human-robot system due to its shared space nature. Perceptions of risk of an interaction can vary based on the perceived capability of the cobot to avoid collision. As such, the prominent ways to manipulate trust in HRC are through the decreased reliability of the automation (Chen et al., 2018a), (Wu et al., 2017; Rahman and Wang, 2018; Hamacher et al., 2016) and increased movement speed and approach behavior of the system (Fujita et al., 2010), (Hald et al., 2019), (Koppenborg et al., 2017), both rooted in perceptions of the cobot's ability.

1.2. Factors impacting trust in HRC

Trust is a subjective experience with three interdependent components: dispositional trust – the predisposition or tendency to trust in general, situational trust – trust based on a specific situation with internal human factors (i.e., mood, task engagement, fatigue) and external factors (i.e., distance from robot, robot behavior, work station design), and learned trust – which is dynamically impacted by past experiences with a specific situation or the technology (Lee and Seppelt, 2009), (Hoff and Bashir, 2015). Dispositional trust is impacted by demographics, personality, and societal characteristics that influence a general tendency to trust or distrust a system. Acceptance and perceptions of cobot behaviors have been shown to be impacted by the sex of the operator (Kuo et al., 2009), (Strait et al., Scheutz), and sex has been shown to differentially impact proxemic behaviors around robots (Syrdal et al., 2007). Despite these promising results, the influence of sex on trust perceptions and objective symptoms of trust, in addition to workload and fatigue, is understudied in HRC.

Within situational trust, specific task traits such as situation awareness, task engagement, workload, and fatigue can impact the operator's perceptions of the tasks difficulty with respect to their own ability (Mouloua and Hancock, 2019). Collaborative robotics are often designed to offload physical work and improve repeatability of a task (Baker and Yanco, 2004); however, a workload shift can demand larger cognitive load due to the increasing complexity of tasks or due to the need for increased situation awareness and decision making. Long-term sustaining of larger cognitive load increases susceptibility to cognitive fatigue, which can impact trust dynamics, resulting utilization strategy, complacency, and task performance (Mouloua and Hancock, 2019). Furthermore, sex is a known factor in cognitive processes and fatigue perceptions (Caplan et al., 1997), therefore the joint impact of sex and cognitive fatigue on trust is a compelling research topic.

Lastly, learned trust is related to the expertise of the operator, and it is comprised of the amount of training and past experiences with operating cobots. As operators gain more familiarity with a system, they tend to have increased trust, unless the system consistently acts in a way to reduce their trust (Chen et al., 2018b); thus, it is important to consider expertise and training in determining trust levels. Therefore, the overall trust perception in the HRC environment is dependent on human factors, robot factors, and their interactions with each other.

1.3. Measurements sensitive to trust shifts

As trust calibration is a key for measuring trust, continuous non-invasive techniques are needed to capture and calibrate trust levels in real-time. Trust is a psychosocial construct, thus identification of measures sensitive to changes in trusting state require manipulation of trust and monitoring of corresponding responses. Historically, subjective questionnaires have been utilized to provide this quantification (Lewis et al., 2018); however, questionnaires are invasive, as they require the operator to stop production to complete them, and they only provide discrete measurement, which reduces the flexibility and adaptability of collaborative robots. Beyond their limitations for use in adaptive robots, surveys also suffer from validity and reliability concerns as they suffer from biases, recency effects, and cannot capture subconscious shifts in the operator's state (Jahedi and Méndez, 2014). As such, initial investigations based on physiological and neurological data have gained popularity in the trust domain due to the need for a direct continuous correlate as well as to reveal insights subjective responses are unable to capture (Jung et al., 2019), (de Visser et al., 2018), although their use in the context of HRC has not been studied. This study investigates the use of heart rate variability (HRV) as a potential measure sensitive to changes in trusting and cognitive fatigue state in a direct and continuous manner. HRV is linked to the autonomic nervous system consisting of the sympathetic system (fight-or-flight response; SNS) and parasympathetic system (rest-and-digest response; PNS). Both SNS and PNS features within HRV analysis have been shown to quantify cognitive fatigue state with moderate success (Melo et al., 2017), (Usui and Nishida, 2017), although few studies in any technology domain have used HRV analysis to quantify trust states; therefore, the sensitivity of HRV and the corresponding correlates of trust manipulation are widely unknown (Perelló-March et al., 2020). As such, this study aims to identify the effect of pertinent human factors (i.e., sex and cognitive fatigue) and pertinent robot factors (i.e., robot reliability) on perceptions, physiological response, and human performance metrics.

2. Methods

2.1. Participants

Sixteen participants, balanced by sex, were recruited from the engineering population at Texas A&M University with majors in industrial, biomedical, safety, chemical, aerospace, or mechanical engineering. All participants were right-handed with an age distribution of 25.12 (3.31) years, and three participants reported having prior experience with industrial robotics. On average, participants were 'moderately familiar' (i.

e., 2.71/5) with using any form of joystick controls.

Approval for IRB (IRB2020-0097DCR) and COVID-19 human subjects testing safety plan were received prior to starting the experiment. Each participant attended two sessions, on separate days, where each session focused on one level of the fatigue condition (no fatigue vs. fatigue). Within each session, participants underwent robot reliability conditions (reliable vs. unreliable). The study involved deception: the participants were not made aware of reliability conditions until the end of the second session. While the order of fatigue was counterbalanced, the reliable condition always preceded the unreliable condition due to the learned component of trust; similar to the effects of fatigue, a condition designed to induce poor trust can carry over effects into later conditions.

2.2. Collaborative task

Participants collaborated with a UR10 robot arm (Universal Robots, DK) to jointly perform a metal surface polishing task (Fig. 1). Participants controlled the robot through right-handed joystick inputs with control over six degrees of motion, but were only asked to control the X, Y, Z directional movements, which were then converted to joint motion commands. The speed of the robot was controlled as stepwise binary movements (on-1.5 cm/s, off-breaks to 0 cm/s); speed itself was not controllable by the participants. During each trial, participants navigated the end effector along a squared S-shape trajectory consisting of five events: three lateral movements and two U-turns that were traced in pink before each trial (Fig. 1, top left inset). Two assistance levels, low vs high, were performed; however the manual/low assistance level is discussed elsewhere (Hopko, Mehta, Pagilla). During high assistance, the participant was responsible for the lateral maneuvering (events 1, 3, 5), and automatic control was responsible for predicting and navigating U-turns (events 2, 4) in addition to preventing force over 15N into the surface at all times (multimedia video). This assistance was designed to keep the human operator in-the-loop while allowing automation to take over the more difficult aspects of the task: judging the distance before turning and the control of turning itself. To mitigate additional cognitive loading associated with automation takeover, a blue dialog box, in clear view of the participant, appeared when automatic control took over and disappeared when control was given back. To mitigate loading associated with poor physical ergonomics, participants were provided tubs (incremental by 2-inches) to place under the joystick to achieve approximately a 90-degree elbow angle. To account for initial task learning effects, participants were allowed to practice until they felt comfortable with the task and controls. Requested training practice runs ranged from one to three trials.

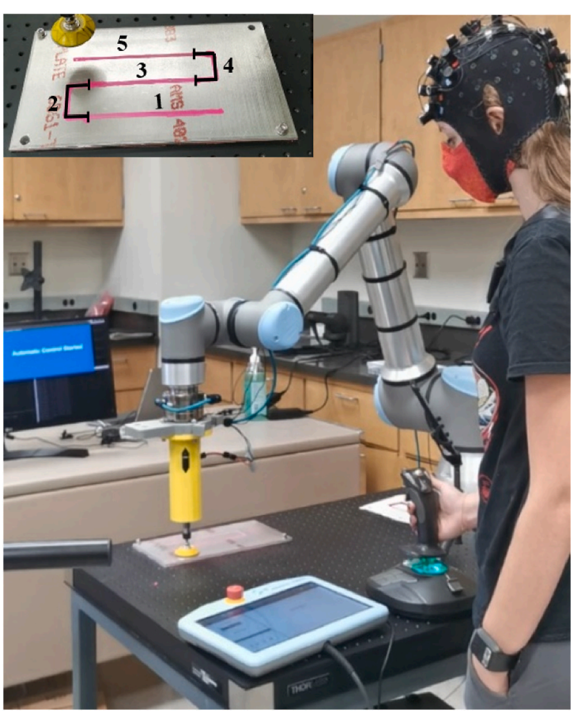

Fig. 1. Experimental setup of the HRC metal polishing task with the UR10 robot along a S-shaped trajectory consisting of three lateral events (1, 3, 5) and two U-turns events (2, 4; top left inset).

2.3. Study procedure

This study includes three major independent variables - operator sex (male, female), cognitive fatigue state (fatigued, not fatigued), and reliability (reliable, unreliable) – and three classes of dependent variables – human task performance, subjective responses, and physiological responses. Participants attended two sessions, split by the fatigue variable, and during each session underwent the two robot reliability conditions (Fig. 2).

2.3.1. Reliability manipulation

Reliability of the robotic assistance was manipulated to examine its effect on operator trust. This study included deception as neither the reliability manipulation nor the measures of trust were mentioned explicitly to the participants to prevent framing and to get a more naturalistic reaction to unreliability. Intentionally manipulated perturbations were not presented during the reliable trials, whereas the unreliable trials had six different predefined perturbations randomly distributed throughout the trials for an effective reliability rate of ~76% (Table 1). Each trial in the unreliable condition had an expectation of one perturbation, although actual perturbations per trial ranged from zero to three. Each participant experienced the same perturbations in the same location. As participants attended two sessions split by the fatigue variable, similar but different perturbations sets were reserved for each day to reduce failure familiarity. Perturbation sets were defined as combination of errors 1, 2, 3, 5 on one day and another combination of errors 1, 2, 3, 4, 6 on another (Table 1, multimedia video). When encountering perturbations and requesting guidance from the researchers, participants were instructed to continue with the trial as normal, maximizing their performance, and to tolerate/deal with the errors; no specific error tolerance strategy was provided. All perturbations were uncorrectable using joystick controls while the failure was occurring (based on distance or time as denoted in Table 1).

2.3.2. Cognitive fatigue manipulation

Cognitive fatigue was manipulated through a 1-h sustained spatial 2-back test completed on a computer prior to starting the HRC task. During the test, circular stimuli are presented randomly within a 3 × 3 grid every 2 s and are visible for 1 s (Hopko, Mehta, Pagilla). The test requires participants to remember the location of the last two stimuli and press the space bar when the current stimulus matches one that appeared 2 back. The spatial n-back test was chosen for two reasons: the spatial n-back test requires sustained spatial and visual processing, which are necessary cognitive abilities to accurately navigate through the metal polishing task trajectory, and the sustained n-back has been shown to induce cognitive fatigue with associated performance reductions (Hopstaken et al., 2015).

2.4. Measurements

1) Task Performance: Task performance was measured using three metrics: 1) an efficiency metric, measured by the overall speed of the trial; overall speed was calculated by dividing the total traveled distance by the time to complete the trial, 2) an accuracy metric, measured by the deviation from the defined trajectory, and 3) a precision metric measured by the variance in deviation from the

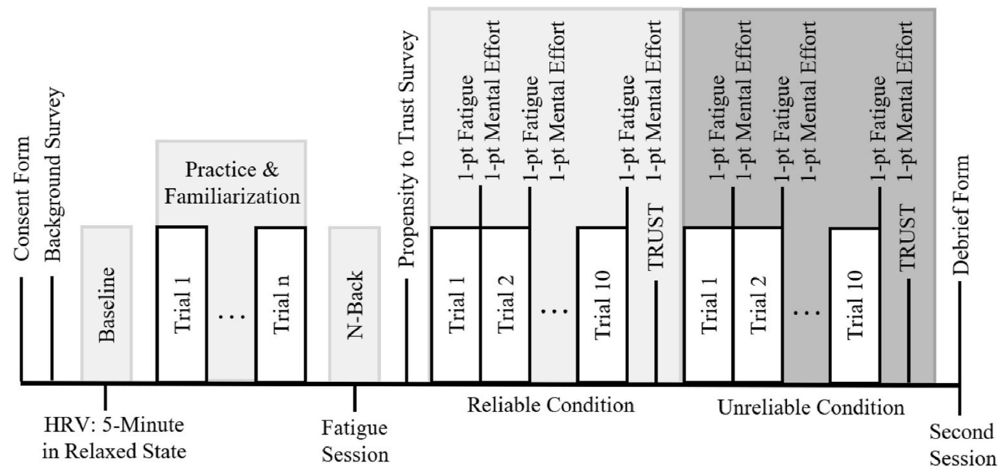

Fig. 2. Study Procedure Timeline. Participants attended two sessions, split by the fatigue variable, and at each session participants underwent reliability conditions (low/high) of 10 trials each, taking approximately 60–70 s each trial. Between trials participants answered one question about fatigue and one about mental effort load. After a condition, participants completed the trust in automation survey (TRUST).

Table 1
Robot perturbations.

S. No.	Perturbations
1	Tool/end-effector contact loss with surface for 2.5 cm
2	Robot not responding to joystick input for 2 s
3	Reduction in speed for 2.5 cm
4	Automatic control starts 2.5 cm too soon/late
5	Automatic control only completes half the turn
6	Automatic control does a straight turn (Not U)

trajectory. Because the unreliable condition had predefined perturbations that themselves reduce performance, task performance during unreliable conditions were normalized for the perturbations: full automation of the task was run without the perturbations to provide a baseline performance score and then again with the perturbations to provide the suboptimal performance score. The incremental difference between the two was subtracted from the actual participants performance during the unreliable trials. For the overall speed metric, performance was first normalized for travel distance and trial competition time.

2) Subjective Reponses: At the start of each session (after fatigue manipulation; Fig. 2), participants were asked to complete the propensity to trust automated agents questionnaire (Jessup et al., 2019) to capture their tendency to trust automation in general. The propensity to trust questionnaire consists of six questions rated on a 1 (low) to 7 (high) scale. The composite score is calculated as an average of all questions (after inverting negatively framed questions). Following each condition (i.e., after every 10 trials), the participants completed the trust in automation questionnaire (TRUST) by Jian et al. (2000) adapted to robots. The TRUST survey quantifies active trust perceptions based on the condition. The TRUST questionnaire consists of 12 questions rated on a scale from 1 (low) to 7 (high), with the composite score calculated as an average of all questions (after inverting negatively framed questions). Following each trial, participants were asked one question about fatigue, "What is your level of fatigue?" rated on an integer locked scale from 1 (low) to 7 (high) and one question about mental effort load from the subjective workload assessment technique (SWAT) rated on a continuous scale from 1 (low) to 3 (high) (Reid et al., 1988). The ratings for the two 1-pt questions were individually averaged across the ten trials for each condition before applying statistical analyses.

3) HRV Features and Analysis: Electrocardiogram (ECG) signals were recorded using a two-lead, chest affixed device, Actiheart (Actiheart 5, Camntech, UK). One male and two female participants were excluded from the HRV analysis due to missing data (i.e., equipment disconnection). Prior to analysis, the ECG signals were corrected for ectopics and missing beats, motion related artifacts, and interpolated using the recommended settings from the software packages by Marked, Strasser et al. and Li et al. (1995); Marked, 1995; Strasser et al., 2012). Data was segmented into two 5-min phases for each condition, the first starting at the beginning of trial one and the second starting at trial six. Frequency domain heart rate variability features were extracted including high frequency (HF; 0.15–0.40 Hz), low frequency (LF; 0.04–0.015 Hz), very low frequency (VLF; 0.003–0.04 Hz), and LF/HF Ratio. Frequency domain features were selected as they strongly correlate with the sympathetic nervous system (SNS) and the parasympathetic nervous system (PNS). Measurement of cognitive fatigue effects through an increase in PNS activity or decrease in SNS has been shown in other studies (Melo et al., 2017), (Delliaux et al., 2019); however few studies had looked into the effect of reliability on the automatic nervous system as measured through HRV (Gupta et al., 2020) as trust is traditionally measured though subjective surveys (Lewis et al., 2018). In the autonomous vehicle domain, lower reliability was shown to increase the LF band (Wetzel et al., 2016).

2.5. Statistical analysis

Separate repeated measures analyses of variance (RM ANOVAs) were run to determine the effects of the independent variables (fatigue, reliability, sex) on the task performance metrics (accuracy, efficiency, and precision) and subjective responses (namely, TRUST questionnaire, fatigue, and mental effort load). A RM ANOVA was conducted to assess the impact of fatigue and sex on propensity to trust. Additionally, separate RM ANOVAs were conducted on the HRV metrics to test the effects of fatigue, reliability, sex, and phase (early/late trials). Significance was determined at $\alpha = 0.05$. Where necessary, post hoc analysis and t-tests were conducted to determine significant interaction effects using Bonferroni corrections.

3. Results

3.1. Task performance

1) *Efficiency Metric – Overall Speed:* Fatigue had a significant impact on speed ($F(1, 14) = 22.00$, $p = 0.034$, $\eta^2 = 0.347$), where no fatigue (NF) had a larger overall speed at 1.444 ± 0.009 cm/s than fatigue (F) at 1.412 ± 0.012 cm/s. Reliability, normalized to remove the errors due to automation, was also significant ($F(1, 14) = 96.79$, $p < 0.001$, $\eta^2 = 0.724$), where reliable trials (R) had higher overall speed at 1.469 ± 0.008 cm/s than unreliable trials (UR) at 1.387 ± 0.014 cm/s. Sex and all variable interactions were not significant (all $p > 0.243$).
2) *Accuracy Metric – Deviation from Trajectory:* Reliability, normalized to remove the errors due to automation, was significant ($F(1, 14) = 88.232$, $p < 0.001$, $\eta^2 = 0.883$), where there was less deviation in R (0.7 ± 0.1 cm) than UR (2.0 ± 0.2 cm) trials. Fatigue, sex, and all other interactions were not significant (all $p > 0.495$).
3) *Precision Metric – Variation in Deviation from Trajectory:* While not significant, the interaction between fatigue and sex ($F(1, 14) = 18.29$, $p = 0.076$, $\eta^2 = 0.160$; Fig. 3) showed a trend with a small effect size, where males exhibited greater variance in deviation from trajectory in UR trials compared to R. All other main and interactive effects were not significant (all $p > 0.164$).

3.2. Subjective responses

1) *1-pt Fatigue:* Perception of fatigue was significantly impacted by fatigue ($F(1, 14) = 28.57$, $p = 0.001$, $\eta^2 = 0.557$) and reliability ($F(1, 14) = 29.71$, $p = 0.001$, $\eta^2 = 0.581$). Participants reported greater perceived fatigue in F trials (4.575 ± 0.580) compared to NF trials (2.625 ± 0.349) and greater fatigue during UR (4.219 ± 0.494) than R (2.981 ± 0.382). While not significant, there was a trend for the interaction between fatigue and sex ($F(1, 14) = 19.43$, $p = 0.059$, $\eta^2 = 0.232$); males and females perceived similar levels of fatigue in NF session (2.909 ± 1.941), but females reported significantly higher perceived fatigue (5.594 ± 0.820) than males in the F session (3.556 ± 0.820).
2) *Propensity to Trust Automation Questionnaire:* There were no statistical differences in the propensity to trust questionnaire for fatigue, sex, or their interaction (grand mean = 5.301 ± 0.718; all $p > 0.739$).
3) *Trust in Automation (Robots) Questionnaire:* Within the TRUST survey, reliability was a significant factor ($F(1, 14) = 34.29$, $p = 0.004$, $\eta^2 = 0.467$), where R had higher perceived trust. The main effects of fatigue and sex were not significant (all $p > 0.182$). While also not significant, the interactions between reliability and sex ($F(1, 14) = 17.93$, $p = 0.083$, $\eta^2 = 0.199$) and between reliability, fatigue, and sex ($F(1, 14) = 18.25$, $p = 0.077$, $\eta^2 = 0.207$; Fig. 4) were observed. Likely caused by females in NF reporting lower trust in UR than R (as compared to males in NF), although post hoc analysis did not reveal any significant pairwise differences after Bonferroni correction ($t = 32.252$, $p = 0.048$, $a_{Bonferroni} = 0.006$).
4) *1-pt Mental Effort Load:* Fatigue significantly impacted mental effort load ($F(1, 14) = 29.71$, $p = 0.008$, $\eta^2 = 0.401$), where higher mental effort load was reported in F trials (1.483 ± 0.083) than NF trials (1.312 ± 0.046). Similarly, reliability was a significant factor ($F(1, 14) = 23.43$, $p = 0.027$, $\eta^2 = 0.304$), where UR trials had higher mental effort load (1.441 ± 0.072) than R trials (1.354 ± 0.054). An interaction between fatigue and sex was found ($F(1, 14) = 21.14$, $p = 0.041$, $\eta^2 = 0.265$), where females rated higher mental effort load in F trials (1.619 ± 0.390) than males (1.386 ± 0.239), but similar mental effort load was reported in NF trials (1.357 ± 0.259). Sex and all other interaction effects were not significant (all $p > 0.137$).

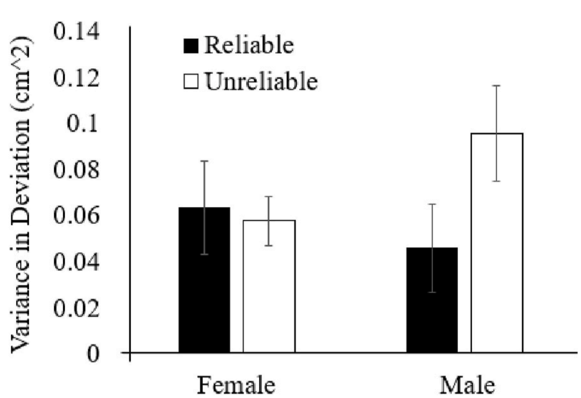

Fig. 3. The effects of reliability and sex in variance in deviation from trajectory. The error bars denote standard error. No significant pair-wise comparisons after Bonferroni.

3.3. HRV responses

1) *HF – Parasympathetic Activity:* While not significant, reliability showed a trend on parasympathetic activity ($F(1, 12) = 21.25$, $p = 0.06$, $\eta^2 = 0.309$) with greater parasympathetic activity during UR trials (792 ± 268) than R trials (484 ± 155). All other effects were statistically comparable (all $p > 0.199$).
2) *LF & VLF – Sympathetic Activity:* Reliability significantly impacted VLF ($F(1, 12) = 22.67$, $p = 0.047$, $\eta^2 = 0.340$) with higher sympathetic activity during UR (1613 ± 358) than R trials (1307 ± 239). All other effects were statistically comparable for VLF (all $p > 0.108$) and LF variables (grand mean = 1996.548 ± 265.442, all $p > 0.150$).
3) *LF/HF Ratio:* The LF/HF ratio was impacted by reliability ($F(1, 12) = 22.00$, $p = 0.05$, $\eta^2 = 0.325$) with lower ratios for UR trials (6.48 ± 1.56) than R trials (7.51 ± 1.48). The change in ratio was due to an increase in parasympathetic activity (i.e. HF) in UR trials as LF was statistically comparable. While not significant, the interaction between reliability and fatigue ($F(1, 12) = 20.00$, $p = 0.076$, $\eta^2 = 0.281$) showed an interesting trend. While not significant after Bonferroni ($t = 31.599$, $p = 0.045$, $a_{Bonferroni} = 0.0125$), the LF/HF ratio was similar for both UR and R trials when fatigued, but in NF the R trials had a higher ratio.

A significant three-way interaction between fatigue, reliability, and sex was observed ($F(1, 12) = 30.67$, $p = 0.012$, $\eta^2 = 0.486$; Fig. 5). Female participants exhibited significantly lower LF/HF ratio in UR trials compared to the R trials during the NF condition. Males did not exhibit such fatigue and reliability influences in LF/HF ratio.

While not significant, an interaction trend between fatigue, reliability, and phase was also observed ($F(1, 12) = 20.17$, $p = 0.076$, $\eta^2 =$

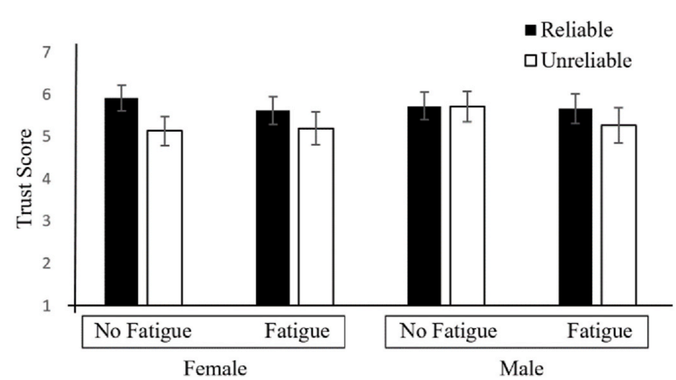

Fig. 4. The effects of fatigue, reliability, and sex on trust scores. The error bars represent standard error.

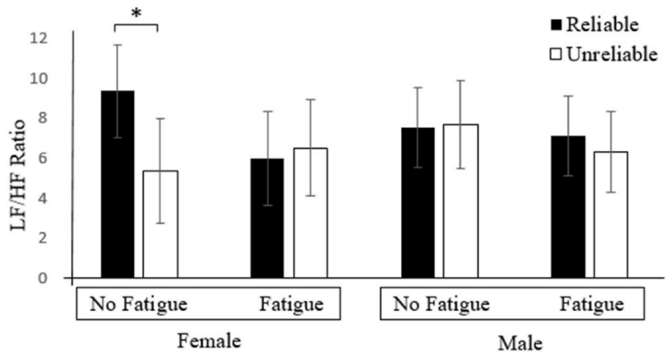

Fig. 5. The effects of fatigue, reliability, and sex on LF/HF ratio. The error bars denote standard error. * depicts significant difference.

0.282; Fig. 6). In the NF trials, late R trials had a higher LF/HF ratio than late UR trials (t = 30.126, p = 0.055). Additionally, late R trials were different across F and NF trials, although neither significant after Bonferroni (t = 31.920, p = 0.049, $a_{Bonferroni} = 0.006$).

4. Discussion

This study investigated the impacts of cognitive fatigue, robot reliability, and operator sex on subjective experience, physiological response (HRV), and resulting performance and human behaviors. Important contributions from this investigation include:

1) Unreliability in robot behavior causes additional performance decrements beyond automation failure (attributable to humans), and it is accompanied by increased perceived workload and fatigue.
2) Physiological HRV responses are sensitive to trust influencers (i.e., reliability), with large effect sizes.
3) Human perceptual and physiological responses provide joint insights into the influence of robot unreliability on trust under different cognitive states.
4) Propensity to trust is not shown to be impacted by operator fatigue or their sex.

4.1. Robot unreliability causes operator-induced performance decrements

The reliability of a robot has long been shown to manipulate trust perceptions as further validated in this investigation (Chen et al., 2018a), (Wu et al., 2017; Rahman and Wang, 2018; Hamacher et al., 2016); however, the resulting impact of reliability on operator performance has not been thoroughly investigated. The performance of a collaborative system is influenced by the individual performances of the cobot and operator agents, in addition to the interactive/collaborative

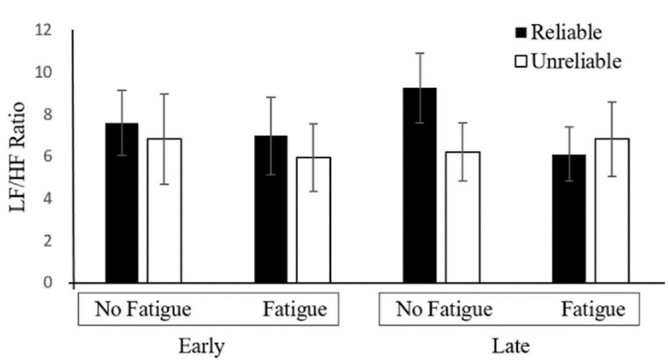

Fig. 6. The effects of fatigue, reliability, and phase on LF/HF Ratio. The error bars denote between participant standard error.

performance between the two. Novel to this investigation, reduced performance of a cobot (through unreliability) was shown to influence operator performance with large effect sizes around 0.8. During robot unreliability, operators reduced their task efficiency and accuracy, but not precision. Because the participants did not have experience with the error modes and were not provided with a tolerance strategy, their compensatory reactions caused additional performance decrements. These actions included: trying to directly correct the issue, stopping and re-providing input controls, waiting for the error to pass, randomly providing joystick inputs to confirm control, or other strategies. Previous investigations into false alarm rates have shown that operators are less willing to comply with alarms when the system is false-alarm prone (Wickens et al., 2005), (Dixon et al., 2007). A similar mechanism may influence operator's decisions during cobot unreliable errors by reducing the operator's willingness to comply with imperfect automation. This failure to comply is then presented as attempts to override or 'check' if controls are functioning, waiting on the cobot to function as expected, etc. These findings implicate the need to augment operator performance, either by providing formalized training on error mitigation strategies or through robot design considerations and adaptation. Trust adaptive robotics designed to calibrate overtrust must consider the impact of unreliability on human behavior and performance when no tolerance strategy is provided.

Suboptimal operator performance during unreliable conditions, via behaviors observed here, was also accompanied by an attitude of lower trust, and a perception of increased workload and increased mental fatigue, suggesting robot unreliability caused the participants to perceive higher task demand. In a human-multiple-mobile-robot teaming simulation task to search and detect subjects, Chen et al. (2011), found that the reliability of the mobile robots' detection similarly impacted workload perceptions, with higher workload and lower reliance when the robot had lower reliability. As unreliability is a pertinent factor in all robotic systems, collaborative systems should consider the variability in human responses and how system performance will likely reduce beyond predicted error caused by the automation alone. These findings have implications on the design of trust adaptive robots as the manipulation of robot performance is connected to how the operator will perform and their resulting trust, workload, and fatigue perceptions.

4.2. The autonomic nervous system is influenced by unreliability (and fatigue)

Trust has been shown to influence an operator's utilization strategies of robot assistance, their complacency, or other unsafe behaviors adopted around cobots. As such, continual measurement of trust has gained popularity to provide insight into when to implement strategies that adjust trust or for use to improve our understanding of why trust is manipulated. While trust is a highly subjective construct influenced by many factors (Lee and Seppelt, 2009), (Lee and See, 2004), some factors, such as the reliability and approach behavior of a cobot, have consistently shown to manipulate trust; however, very few studies evaluate the accompanying autonomic responses, such as HRV features. Utilization of new response metrics sensitive to trust manipulation (e.g., HRV, eye-tracking, brain activity) can reveal new perspectives subjective response is unable to capture, such as potential shifts in attention, changes in information processing strategy, or subconscious shift in cognitive state that may help explain why and how trust is manipulated. To reveal the sensitivity of HRV to trust, this study investigated the physiological heart measurements using frequency domain heart rate variability features to understand the impact of the trust influencer, reliability, on the SNS and PNS. Our findings indicate a significant effect of reliability on VLF, a measure that increases with sympathetic activity and stress (Shaffer et al., 2014). There was no significant effect seen on the LF metric, which is influenced by both SNS and PNS, in contrast to a previous investigation that reported an increase in LF with unreliability (Wetzel et al., 2016). While not significant, the same trend in LF was observed in our investigation (R = 1850 ± 226 ms^2, UR = 2142 ± 327 ms^2).

Unreliability also influenced the HF metric illustrating an increase in parasympathetic activity in UR, and while not significant at p = 0.06, had a large effect size of 0.309. The increase in HF drove a decrease in the LF/HF metric with a similarly large effect size. Previous investigations into cognitive fatigue and time-on-task effects have reported the opposite effect, i.e., lower HF with higher LF/HF, to increased cognitive fatigue, reduced effort/will, or task-disengagement (Melo et al., 2017), (Mizuno et al., 2011, 2014; Matuz et al., 2021). In line with these studies, our findings on these HRV features suggest participants required increased task engagement or effort in the presence of perturbations. This response is also supported by the perception of increased workload during the unreliable conditions. Participants also reported increased fatigue perceptions during unreliable trials. The use of subjective and physiological measurements can provide insight into how participants react to cognitive fatigue and unreliability, where both measurements showed an interaction between reliability, fatigue, and sex. Both the perceptual and HRV symptoms of trust found that females experienced a greater trust breach than males during the no fatigue condition; females reported lower trust in response to unreliable automation than males alongside experiencing lower LF/HF than males. Analysis based on human perception can provide implicit information where additional objective measures, such as HRV analysis, can be collaboratively implemented alongside subjective perceptions for explicit information.

4.3. Cognitive fatigue and its influence on trust perceptions

Within the manufacturing industry, the leading root causes of fatigue include task related stressors, scheduling shifts, and lack of sleep, where workers often attempt task disengagement fatigue mitigation strategies such as talking to coworkers (Lu et al., 2017). All leading root causes of fatigue directly relate to cognitive fatigue, as physically demanding tasks and task related physical stressors have cognitive implications (Aitken and MacMahon, 2019). Thus, the n-back test was used in this study to simulate industry type fatigue prior to the participants interacting with the cobot. In consensus with existing literature (Hopko, Mehta, Pagilla), (Hopstaken et al., 2015), the effects of cognitive fatigue were visible in all classes of response features: including subjective perceptions, physiological measurements (interacting with other independent variables), and task performance through a reduction of efficiency, but not accuracy or precision. The subjective questionnaires on fatigue showed a significant increase in ratings when the participants were cognitively fatigued confirming the successful manipulation of fatigue, and females reported higher levels of fatigue than males as expected (Bensing et al., 1999).

Based on our study pool, we did not find any impact of cognitive fatigue on the propensity to trust measure. Additionally, operator's sex did not seem to impact propensity to trust. While the existing empirical literature on factors impacting propensity to trust in automation is limited, in the interpersonal domain, the literature is divided on the impact of sex. Females have reported to be more trusting than males (Feingold, 1994), but contradictory findings show that males are more trusting than females (Buchan et al., 2008). The dearth of knowledge on the magnitude and uncertainties of factors (such as fatigue and sex) on propensity to trust warrants further investigation, especially in the automation domain. Propensity to trust is a strong predictor of the magnitude of trust attitudes during situational measurements (Jessup et al., 2019). As such, it is imperative to understand what dispositional, situational, or learned factors impact propensity to trust measures in addition to conditional trust measures (i.e., TRUST survey), as this influences how and what factors manipulate trust attitudes.

The lack of impact of sex or cognitive fatigue on the propensity to trust questionnaire suggest that fatigue or sex-related changes in subjective trust during the HRC tasks were likely not caused by a baseline change in propensity to trust. Within the TRUST questionnaire, neither fatigue, nor sex were main effects, although an interaction between fatigue, reliability, and sex was observed. Other studies have found females to be more trusting in dangerous robots than males (Gallimore et al., 2019), although this study did not find this as a main effect. Under the no fatigue conditions, females and males did react differently to unreliability; males showed no significant trust reduction from the unreliable actions of the robot, whereas females showed a reduction in subjective trust when the robot made errors, supporting sex differences in trusting behavior. When mentally fatigued, both males and females showed similar reduced trust behavior in the presence of robot error. While fatigue and sex have individually been identified as important factors impacting trust by three separate review papers in HRC (Hancock et al., 2011), (Schaefer et al., 2014), (Khavas et al., 2020), none of the reviews identified any papers that jointly evaluated fatigue and sex on trust. To the author's knowledge, this is the first evidence for an investigation on the interaction between sex and fatigue on situation-based trust perceptions in addition to a propensity to trust measure for HRC.

The implications of human factors, such as fatigue, on the collaborative task performance are important to consider in HRC systems. Regardless of the simplicity of the task, and the visually guided trials, we observed an effect of cognitive fatigue on task performance through a reduction of task efficiency. As the efficiency of the task is affected by cognitive fatigue, it is important to investigate the dynamics of operator states such as fatigue and robot reliability when designing trust calibration models to better optimize the performance of the HRC system. Further, such calibration models should offer equitable performance for both males and females.

4.4. Study limitations and future work

The majority of participants in this study were seeking advanced degrees in engineering, and all participants were college students, which need not be representative of the target population of industry workers. Further investigation is needed to compare applications of trust using data reported here from college students to applications using industrial workers or technical operators who may have more experience with these failure modes. Experience, specifically, with the failure modes and task will impact the strategies operator choose to respond with and how they interaction with the robot. Given the findings that trust-related responses induce additional performance decrements, the difference in strategy between experienced and inexperienced users can impact how human performance perturbs under unreliable robot actions or fatigued users. While experienced users may have compensation strategies, faulty automation is still a perturbation to the human's state and requires human adaptation to the robot. This shift in adaptation may cause the symptoms of reduced trust perceptions, changes in HRV, and changes in human performance. Thus, future research is needed to explore how established strategy and experience impacts the response metrics utilized here.

Additionally, the protocol did not offer participants to adopt compensatory strategies such as controlling the speed of the robot or to switching off the robot assistance, to document the initial impact of the perturbations on human experiences. Compensatory measures can provide insight into robot utilization and its impact on performance as an extension of trust. As such, our next step is to examine ways to quantify human or system performance within collaborative robotics when operators are provided the flexibility to adapt to robot perturbations. Furthermore, while one specific task, namely surface finishing, was employed in this study, the pertinence of the factors here studied are relevant to other HRC tasks or even other automation domains, although their magnitude and uncertainties warrant further investigation.

Lastly, trust is a construct, a label researchers use to explain human experience and behavior; it is not the subjective perception of trusting or distrusting behavior itself. As such, trust can be derived from or impact changes in the operator's attention, information and cognitive load, physical ergonomics, arousal state, and many other human factors. Capturing the physiological, neurological, perceptual, and behavioral shifts can reveal new insights about how trust connects to aspects of human-robot collaboration (Yadav et al., 2022). As such, future research is needed to evaluate the sensitivity of various response measures able to capture underlying shifts in the human state (e.g., attention shifts) with different perceptions of risks with automation (Stuck et al., 2021). Subjective responses only capture a subset of the story at play, and metrics sensitive to these other factors (e.g., shifts in attention or cognitive resources) can reveal the mechanisms that connect trust and human performance and behavior right when trust shifts occur. For example, it is possible that participants shifted their attention placed on the task as a response to trusting/distrusting experiences. The metrics we utilized do not directly capture attentional shifts, thus future research can leverage multimodal measures such as eye tracking or neuroimaging to reveal this part of the story (Hopko and Mehta, 2022). Beyond the need to real-time measure or adaptive robotics, these new perspectives can help researchers better explain why pertinent HRC factors, such as were studied here (i.e., operator sex, robot reliability, cognitive fatigue), impact trust perceptions and corresponding performance.

5. Conclusion

The joint consideration of operator sex, cognitive fatigue, and robot reliability on various human robot collaboration metrics (i.e., task performance outcomes, operator subjective perceptions, and physiological responses) is necessary in advanced manufacturing robotics for designing more robust, adaptive, and human-aware robots. This study illustrates operator performance degradations associated with lower robot reliability and cognitive fatigue state. Operators induce additional accuracy and efficiency performance reductions beyond reductions caused directly by robot unreliability, and operators further reduce efficiency under cognitive fatigue. Faulty robot automation perturbs the human's state by requiring compensatory strategy to identify, understand, and respond to the shift in robot behavior. This results in a change in human behavior that impacts human performance as well as directly impacts perceptual and physiological symptoms of robot unreliability manipulation. The impact of more experienced and familiarized populations with these failure modes may impact the specific compensatory strategy and effort needed to respond the shift in robot state; however faulty automation will still have a perturbance on human factors, thus these findings implicate novel and experienced workforce designs.

The impact of low cobot reliability was additionally captured via physiological markers, i.e., a significant increase in sympathetic activity and an increasing trend in parasympathetic activity, all with large effect sizes. These findings demonstrate the feasibility and promise of using HRV metrics as a continuous and explainable measure of trust, and the need for human factor designs in trust adaptive HRC systems. Utilization of new response metrics to study trust, such as HRV employed here, can reveal new insights that subjective measures alone may be unable to capture. This work shows the promise of utilizing physiological response to understand the interaction between operator sex, cognitive fatigue, and robot reliability. Additional investigation can utilize this same approach to deploy measures sensitive to other human-state shifts, such as eye-tracking – that can capture shifts in attention and information processing, brain imaging – that can reveal changes in load and recruitment shifts in functionally-varied neurons, and performance/observational data – that can reveal the types of errors operators make (e.g., mistakes, slips, lapses) when perturbed by robot malfunction. Such investigations will advance our understanding of trust and improve the metrics that can be deployed to improve human-robot collaboration systems.

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.

Acknowledgment

This study was supported by the National Science Foundation under grant number 1900704 (Augmenting Human Cognition with Collaborative Robots).

References

Aitken, B., MacMahon, C., 2019. Shared demands between cognitive and physical tasks may drive negative effects of fatigue: a focused review. Front. Sports Act. Living 1. https://doi.org/10.3389/fspor.2019.00045.

Akella, P., et al., 1999. Cobots for the automobile assembly line. In: Proceedings 1999 IEEE International Conference on Robotics and Automation (Cat. No.99CH36288C), Detroit, MI, USA, vol. 1, pp. 728–733. https://doi.org/10.1109/ROBOT.1999.770061.

Baker, M., Yanco, H.A., Oct. 2004. Autonomy mode suggestions for improving human-robot interaction. In: 2004 IEEE International Conference on Systems, Man and Cybernetics (IEEE Cat. No.04CH37583), vol. 3, pp. 2948–2953. https://doi.org/10.1109/ICSMC.2004.1400781.

Bensing, J.M., Hulsman, R.L., Schreurs, K.M.G., Oct. 1999. Gender differences in fatigue: biopsychosocial factors relating to fatigue in men and women. Med. Care 37 (10), 1078–1083.

Buchan, N.R., Croson, R.T.A., Solnick, S., Dec. 2008. Trust and gender: an examination of behavior and beliefs in the Investment Game. J. Econ. Behav. Organ. 68 (3–4), 466–476. https://doi.org/10.1016/j.jebo.2007.10.006.

Caplan, P.J., Crawford, M., Hyde, J.S., Richardson, J.T.E., 1997. Gender Differences in Human Cognition. Counterpoints: Cognition, Memory, and Language Series. Oxford University Press, 2001 Evans Road, Cary, NC 27513. -0-19-511291-1, $19.

Charalambous, G., Fletcher, S., Webb, P., Apr. 2016. The development of a scale to evaluate trust in industrial human-robot collaboration. Int. J. Social Robot. 8 (2), 193–209.

Chen, J.Y.C., Barnes, M.J., Kenny, C., Mar. 2011. Effects of unreliable automation and individual differences on supervisory control of multiple ground robots. In: 2011 6th ACM/IEEE International Conference on Human-Robot Interaction (HRI), pp. 371–378. https://doi.org/10.1145/1957656.1957793.

Chen, M., Nikolaidis, S., Soh, H., Hsu, D., Srinivasa, S., Feb. 2018. Planning with trust for human-robot collaboration. In: Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction, Chicago IL USA, pp. 307–315. https://doi.org/10.1145/3171221.3171264.

Chen, M., Nikolaidis, S., Soh, H., Hsu, D., Srinivasa, S., Nov. 2018. Trust-Aware Decision Making for Human-Robot Collaboration: Model Learning and Planning [Online]. Available:. *arXiv:1801.04099 [cs]*. http://arxiv.org/abs/1801.04099. (Accessed 27 October 2020).

Colquitt, J.A., Salam, S.C., 2015. Foster trust through ability, benevolence, and integrity. In: Handbook of Principles of Organizational Behavior. John Wiley & Sons, Ltd, pp. 389–404. https://doi.org/10.1002/9781119206422.ch21.

de Visser, E.J., et al., Aug. 2018. Learning from the slips of others: neural correlates of trust in automated agents. Front. Hum. Neurosci. 12 https://doi.org/10.3389/fnhum.2018.00309.

Delliaux, S., Delaforge, A., Deharo, J.-C., Chaumet, G., May 2019. Mental workload alters heart rate variability, lowering non-linear dynamics. Front. Physiol. 10 https://doi.org/10.3389/fphys.2019.00565.

Dixon, S.R., Wickens, C.D., McCarley, J.S., Aug. 2007. On the independence of compliance and reliance: are automation false alarms worse than misses? Hum. Factors 49 (4), 564–572. https://doi.org/10.1518/001872007X215656.

Feingold, A., 1994. Gender differences in personality: a meta-analysis. Psychol. Bull. 116 (3), 429–456. https://doi.org/10.1037/0033-2909.116.3.429.

Fletcher, S.R., Webb, P., 2017. Industrial robot ethics: the challenges of closer human collaboration in future manufacturing systems. In: Silva Sequeira, J., Tokhi, M.O., Kadar, E.E., Virk, G.S. (Eds.), A World with Robots: International Conference on Robot Ethics: ICRE 2015, M. I. Aldinhas Ferreira. Springer International Publishing, Cham, pp. 159–169. https://doi.org/10.1007/978-3-319-46667-5_12.

Fujita, M., Kato, R., Tamio, A., Sep. 2010. Assessment of operators' mental strain induced by hand-over motion of industrial robot manipulator. In: 19th International Symposium in Robot and Human Interactive Communication, RO-MAN, 2010. IEEE, pp. 361–366. https://doi.org/10.1109/ROMAN.2010.5598689.

Gallimore, D., Lyons, J.B., Vo, T., Mahoney, S., Wynne, K.T., 2019. Trusting robocop: gender-based effects on trust of an autonomous robot. Front. Psychol. 10 https://doi.org/10.3389/fpsyg.2019.00482.

Gupta, K., Hajika, R., Pai, Y.S., Duenser, A., Lochner, M., Billinghurst, M., Mar. 2020. Measuring human trust in a virtual assistant using physiological sensing in virtual reality. In: 2020 IEEE Conference on Virtual Reality and 3D User Interfaces. VR), pp. 756–765. https://doi.org/10.1109/VR46266.2020.00099.

Hald, K., Rehm, M., Moeslund, T.B., Oct. 2019. Proposing human-robot trust assessment through tracking physical apprehension signals in close-proximity human-robot collaboration. In: 2019 28th IEEE International Conference on Robot and Human

Interactive Communication (RO-MAN), Robot and Human Interactive Communication (RO-MAN), 2019 28th IEEE International Conference on, pp. 1–6. https://doi.org/10.1109/RO-MAN46459.2019.8956335.
Hamacher, A., Bianchi-Berthouze, N., Pipe, A.G., Eder, K., Aug. 2016. Believing in BERT: using expressive communication to enhance trust and counteract operational error in physical Human-robot interaction. In: 2016 25th IEEE International Symposium on Robot and Human Interactive Communication. RO-MAN, pp. 493–500. https://doi.org/10.1109/ROMAN.2016.7745163.
Hancock, P.A., Billings, D.R., Schaefer, K.E., Chen, J.Y.C., de Visser, E.J., Parasuraman, R., Oct. 2011. A meta-analysis of factors affecting trust in human-robot interaction. Hum. Factors 53 (5), 517–527. https://doi.org/10.1177/0018720811417254.
Hoff, K.A., Bashir, M., May 2015. Trust in automation: integrating empirical evidence on factors that influence trust. Hum. Factors 57 (3), 407–434. https://doi.org/10.1177/0018720814547570.
Hopko, S.K., Mehta, R.K., 2022 Jun 16. Trust in shared-space collaborative robots: shedding light on the human brain. Hum. Factors, 00187208221109039.
Hopko, R. Khurana, Mehta, R.K., Pagilla, P.R., Apr. 2021. Effect of cognitive fatigue, operator sex, and robot assistance on task performance metrics, workload, and situation awareness in human-robot collaboration. IEEE Rob. Autom. Lett. 6 (2), 3049–3056. https://doi.org/10.1109/LRA.2021.3062787.
Hopstaken, J.F., van der Linden, D., Bakker, A.B., Kompier, M.A.J., 2015. A multifaceted investigation of the link between mental fatigue and task disengagement. Psychophysiology 52 (3), 305–315. https://doi.org/10.1111/psyp.12339.
Hu, J., Pagilla, P.R., Jun. 2021. Dual-edge robotic gear chamfering with registration error compensation. Robot. Comput. Integrated Manuf. 69, 102082 https://doi.org/10.1016/j.rcim.2020.102082.
Jahedi, S., Méndez, F., Feb. 2014. On the advantages and disadvantages of subjective measures. J. Econ. Behav. Organ. 98, 97–114. https://doi.org/10.1016/j.jebo.2013.12.016.
Jessup, S., Schneider, T., Alarcon, G., Ryan, T., Capiola, A., Jun. 2019. The Measurement of the Propensity to Trust Automation," Virtual, Augmented and Mixed Reality. Applications and Case Studies [Online]. Available: https://www.researchgate.net/publication/334344580_The_Measurement_of_the_Propensity_to_Trust_Automation. (Accessed 29 July 2020).
Jian, J.-Y., Bisantz, A.M., Drury, C.G., Mar. 2000. Foundations for an empirically determined scale of trust in automated systems. Int. J. Cognit. Ergon. 4 (1), 53–71. https://doi.org/10.1207/S15327566IJCE0401_04.
Jin, Z., Pagilla, P., Sep. 2020. Collaborative operation of robotic manipulators with human intent prediction and shared control for object inspection and handling. In: Presented at the IEEE Intl. Conf. On Human Machine Systems.
Jung, E.-S., Dong, S.-Y., Lee, S.-Y., Jul. 2019. Neural correlates of variations in human trust in human-like machines during non-reciprocal interactions. Sci. Rep. 9 (1), 1–10. https://doi.org/10.1038/s41598-019-46098-8.
Khavas, Z.R., Ahmadzadeh, S.R., Robinette, P., 2020. In: Modeling Trust in Human-Robot Interaction: A Survey. Social Robotics, Cham, pp. 529–541. https://doi.org/10.1007/978-3-030-62056-1_44.
Koppenborg, M., Nickel, P., Naber, B., Lungfiel, A., Huelke, M., Jul. 2017. Effects of movement speed and predictability in human-robot collaboration. Human Factors & Ergonomics in Manufacturing & Service Industries 27 (4), 197–209.
Krüger, J., Lien, T.K., Verl, A., 2009. Cooperation of human and machines in assembly lines. CIRP Annals 58 (2), 628–646. https://doi.org/10.1016/j.cirp.2009.09.009.
Kuo, I.H., et al., Sep. 2009. Age and gender factors in user acceptance of healthcare robots. In: RO-MAN 2009 - the 18th IEEE International Symposium on Robot and Human Interactive Communication, pp. 214–219. https://doi.org/10.1109/ROMAN.2009.5326292.
Lee, J.D., See, K.A., Mar. 2004. Trust in automation: designing for appropriate reliance. Hum. Factors 46 (1). https://doi.org/10.1518/hfes.46.1.50_30392.
Lee, J., Seppelt, B., 2009. Human Factors in Automation Design, pp. 417–436. https://doi.org/10.1007/978-3-540-78831-7_25.
Lewis, M., Sycara, K., Walker, P., 2018. The role of trust in human-robot interaction. In: Abbass, H.A., Scholz, J., Reid, D.J. (Eds.), Foundations of Trusted Autonomy. Springer International Publishing, Cham, pp. 135–159. https://doi.org/10.1007/978-3-319-64816-3_8.
Li, Cuiwei, Zheng, Chongxun, Tai, Changfeng, Jan. 1995. Detection of ECG characteristic points using wavelet transforms. IEEE (Inst. Electr. Electron. Eng.) Trans. Biomed. Eng. 42 (1), 21–28. https://doi.org/10.1109/10.362922.
Lu, L., Megahed, F.M., Sesek, R.F., Cavuoto, L.A., Nov. 2017. A survey of the prevalence of fatigue, its precursors and individual coping mechanisms among U.S. manufacturing workers. Appl. Ergon. 65, 139–151. https://doi.org/10.1016/j.apergo.2017.06.004.
Madhavan, P., Wiegmann, D.A., Jul. 2007. Similarities and differences between human–human and human–automation trust: an integrative review. Theor. Issues Ergon. Sci. 8 (4), 277–301. https://doi.org/10.1080/14639220500337708.
Marked, V., 1995. Correction of the Heart Rate Variability Signal for Ectopics and Missing Beats. Heart rate variability.
Matuz, A., van der Linden, D., Kisander, Z., Hernádi, I., Kázmér, K., Csathó, Á., Mar. 2021. Enhanced cardiac vagal tone in mental fatigue: analysis of heart rate variability in Time-on-Task, recovery, and reactivity. PLoS One 16 (3). https://doi.org/10.1371/journal.pone.0238670.
Meissner, A., Trübswetter, A., Conti-Kufner, A.S., Schmidtler, J., Jul. 2020. Friend or foe? Understanding assembly workers' acceptance of human-robot collaboration. J. Hum.-Robot Interact. 10 (1) https://doi.org/10.1145/3399433, 3:1-3:30.
Melo, H.M., Nascimento, L.M., Takase, E., 2017. Mental fatigue and heart rate variability (HRV): the time-on-task effect. Psychology & Neuroscience 10 (4), 428–436. https://doi.org/10.1037/pne0000110.
Mizuno, K., Tanaka, M., Yamaguti, K., Kajimoto, O., Kuratsune, H., Watanabe, Y., May 2011. Mental fatigue caused by prolonged cognitive load associated with sympathetic hyperactivity. Behav. Brain Funct. 7, 17. https://doi.org/10.1186/1744-9081-7-17.
Mizuno, K., Tajima, K., Watanabe, Y., Kuratsune, H., Jul. 2014. Fatigue correlates with the decrease in parasympathetic sinus modulation induced by a cognitive challenge. Behav. Brain Funct. 10, 25. https://doi.org/10.1186/1744-9081-10-25.
Mouloua, M., Hancock, P.A., 2019. Human Performance in Automated and Autonomous Systems: Current Theory and Methods. CRC Press.
Nass, C., Moon, Y., 2000. Machines and mindlessness: social responses to computers. J. Soc. Issues 56, 81–103. https://doi.org/10.1111/0022-4537.00153.
R. Patel, M. Hedelind, and P. Lozan-Villegas, "Enabling robots in small-part assembly lines: the 'ROSETTA approach' - an industrial perspective," in ROBOTIK 2012; 7th German Conference on Robotics, May 2012, pp. 1–5.
Perelló-March, J., Burns, C., Elliott, M., Birrell, S., 2020. Integrating Trust in Automation into Driver State Monitoring Systems, pp. 344–349. https://doi.org/10.1007/978-3-030-25629-6_53.
Rahman, S.M.M., Wang, Y., Oct. 2018. Mutual trust-based subtask allocation for human–robot collaboration in flexible lightweight assembly in manufacturing. Mechatronics 54, 94–109.
Reid, G.B., Nygren, T.E., 1988. The subjective workload assessment technique: a scaling procedure for measuring mental workload. In: Hancock, P.A., Meshkati, N. (Eds.), Advances in Psychology, vol. 52. North-Holland, pp. 185–218. https://doi.org/10.1016/S0166-4115(08)62387-0.
Schaefer, K.E., et al., 2014. A Meta-Analysis of Factors Influencing the Development of Trust in Automation: Implications for Human-Robot Interaction. Defense Technical Information Center, Fort Belvoir, VA, Jul. https://doi.org/10.21236/ADA607926.
Shaffer, F., McCraty, R., Zerr, C.L., Sep. 2014. A healthy heart is not a metronome: an integrative review of the heart's anatomy and heart rate variability. Front. Psychol. 5 https://doi.org/10.3389/fpsyg.2014.01040.
M. Strait, P. Briggs, and M. Scheutz, "Gender, More So than Age, Modulates Positive Perceptions of Language-Based Human-Robot Interactions," p. 8.
Strasser, F., Muma, M., Zoubir, A.M., Aug. 2012. Motion artifact removal in ECG signals using multi-resolution thresholding. In: 2012 Proceedings of the 20th European Signal Processing Conference. EUSIPCO), pp. 899–903.
Stuck, R.E., Tomlinson, B.J., Walker, B.N., 2021 Aug 30. The importance of incorporating risk into human-automation trust. Theor. Issues Ergon. Sci. 1–7.
Syrdal, D.S., Lee Koay, K., Walters, M.L., Dautenhahn, K., Aug. 2007. A personalized robot companion? - the role of individual differences on spatial preferences in HRI scenarios. In: RO-MAN 2007 - the 16th IEEE International Symposium on Robot and Human Interactive Communication, pp. 1143–1148. https://doi.org/10.1109/ROMAN.2007.4415252.
Usui, H., Nishida, Y., Aug. 2017. The very low-frequency band of heart rate variability represents the slow recovery component after a mental stress task. PLoS One 12, e0182611. https://doi.org/10.1371/journal.pone.0182611.
Vysocky, A., Novak, P., Jun. 2016. HUMAN – ROBOT COLLABORATION IN INDUSTRY, vol. 2016. MM SJ, pp. 903–906. https://doi.org/10.17973/MMSJ.2016_06_201611, 2.
Wetzel, J.M., Sheffert, S.M., Backs, R.W., Nov. 2016. Driver trust, annoyance, and acceptance of an automated calendar system. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting. https://doi.org/10.1177/154193120404801928.
Wickens, C., Dixon, S., Goh, J., Hammer, B., Jan. 2005. Pilot Dependence on Imperfect Diagnostic Automation in Simulated UAV Flights: an Attentional Visual Scanning Analysis, vol. 7.
Wu, B., Hu, B., Lin, H., 2017. Toward efficient manufacturing systems: a trust based human robot collaboration. In: 2017 American Control Conference (ACC), American Control Conference. ACC), pp. 1536–1541. https://doi.org/10.23919/ACC.2017.7963171. May 2017.
Yadav, A., Hopko, S., Zhang, Y., Mehta, R., Mar. 2022. Multimodal Bio-Behavioral Approaches to Study Trust in Human-Robot Collaboration.