User experience and physiological response in Human-Robot Collaboration: a preliminary investigation

Journal of Intelligent & Robotic Systems (2022) 106:36
https://doi.org/10.1007/s10846-022-01744-8

REGULAR PAPER

User Experience and Physiological Response in Human-Robot Collaboration: A Preliminary Investigation

Riccardo Gervasi[1] · Khurshid Aliev[1] · Luca Mastrogiacomo[1] · Fiorenzo Franceschini[1]

Received: 4 February 2022 / Accepted: 7 September 2022

Abstract
Within the context of Industry 4.0 and of the new emerging Industry 5.0, human factors are becoming increasingly important, especially in Human-Robot Collaboration (HRC). This paper provides a novel study focused on the human aspects involved in industrial HRC by exploring the effects of various HRC setting factors. In particular, this paper aims at investigating the impact of industrial HRC on user experience, affective state, and stress, assessed through both subjective measures (i.e., questionnaires) and objective ones (i.e., physiological signals). A collaborative assembly task was implemented with different configurations, in which the robot movement speed, the distance between the operator and the robot workspace, and the control of the task execution time were varied. Forty-two participants were involved in the study and provided feedbacks on interaction quality and their affective state. Participants' physiological responses (i.e., electrodermal activity and heart rate) were also collected non-invasively to monitor the amount of stress generated by the interaction. Analysis of both subjective and objective responses revealed how the configuration factors considered influence them. Robot movement speed and control of the task execution time resulted to be the most influential factors. The results also showed the need for customization of HRC to improve ergonomics, both psychological and physical, and the well-being of the operator.

Keywords Affective state · Human-robot collaboration · Industry 5.0 · Manufacturing · Physiological signals · User experience

1 Introduction

Human-Robot Collaboration (HRC) represents one of the cornerstones of Industry 4.0 and of the newly emerging Industry 5.0. HRC aims at combining the skills of the human with those of robots. On the one hand, humans have innate intelligence, flexibility, and problem-solving capabilities; on the other hand, robots provide repeatability, power, and precision [1]. Unlike classic industrial robots, collaborative robots (or cobots) are designed to work alongside humans, thus enabling the removal of confinement barriers in factories. As a result, production lines implementing HRC can be more flexible and easily reconfigurable, becoming capable of changing and adapting quickly to new products [2–4]. For an effective implementation of HRC, it is necessary to study and analyze its various aspects, focusing not only on the technical aspects of the robot, but also on the humans involved [5, 6]. The integration of new technologies is usually associated with a change in the way tasks are performed, with a trend toward increasingly passive and cognitive work [7–9]. Moreover, working closely with a robot may lead to the creation of stressful situations for operators, which can negatively affect the interaction and performance of a collaborative task. Industry 5.0 highlights the necessity of implementing a human-centric approach in future factories, where human needs and interests are placed at the center of the production process [10]. Therefore, human factors

This work has been partially supported by "Ministero dell'Istruzione, dell'Università e della Ricerca" Award "TESUN-83486178370409 finanziamento dipartimenti di eccellenza CAP. 1694 TIT. 232 ART. 6".

✉ Riccardo Gervasi
riccardo.gervasi@polito.it

Khurshid Aliev
khurshid.aliev@polito.it

Luca Mastrogiacomo
luca.mastrogiacomo@polito.it

Fiorenzo Franceschini
fiorenzo.franceschini@polito.it

[1] Department of Management and Production Engineering (DIGEP), Politecnico di Torino, Turin, Italy

Published online: 23 September 2022

need to be carefully analyzed when implementing HRC to preserve operators' well-being.

In order to best support the operator's well-being, it is first necessary to understand the operator's mental and physical state during HRC. To study these aspects, besides analyzing feedback obtained with classical self-assessment tools (e.g. questionnaires), the implementation of physiological measures represents an important resource. Through them it is possible to obtain information about the operator's state in real-time during the task execution, even revealing possible unconscious reactions. So far, a rather limited number of works have integrated physiological measures to understand the state of the operator in HRC [11–13]. Moreover, in such studies the interaction with the robot is often rather limited.

This paper aims to provide a novel study on how different configurations of a collaborative robotic system affect the user experience, determining the proficiency of HRC. The research questions that will be investigated in this paper are the following: (i) *In industrial HRC, what is the impact of different robot configuration factors (i.e., movement speed, distance from robot workspace, and control of execution time) on user experience?*; and (ii) *Which robot configuration factors affect operator stress?*

To address these research questions a collaborative assembly task of a mechanical component was implemented using the collaborative robot UR3e and involving 42 participants. Perceived interaction quality (i.e., perceived robot helpfulness, safety, naturalness, efficiency, fluency, comfort, and robot trustworthiness), participant affective state, and physiological response (i.e., Electrodermal Activity (EDA) and Heart Rate Variability (HRV)) were collected and analyzed to explore the joint effects of various HRC setting factors (i.e., robot movement speed, distance from robot workspace, and control of execution time). The inclusion of physiological response allowed to derive an objective measure for stress during HRC.

The major contributions of this paper can be summarized in the following elements:

(i) a novel study on the relationships among several HRC setting factors, perceived interaction quality, affective state, and stress,
(ii) preliminary results on the use of non-invasive biosensor to collect of participants' physiological response,
(iii) preliminary and qualitative insights on interaction effects of considered factors,
(iv) preliminary exploration of the effect of time (i.e., number of experimental trials performed) on response variables.

The results of this kind of study will have implications for collaborative task design, contributing to the identification of possible guidelines to improve cognitive ergonomics and make HRC more profitable from a human-centered perspective.

The paper is organized as follows. Section 2 provides an overview on HRC problem and challenges. In Section 3, experimental design and methodology are described. Section 4 contains preliminary results of the experiment, providing insights of the influence of the considered configuration factors on the perceived interaction quality, the participant's affective state and physiological responses. In addition, preliminary insights are provided on the interaction effects among factors and the influence of the habituation process on response variables. Section 5 focuses on discussion of the study results and possible implications for the improvement of HRC. Finally, Section 6 covers conclusions and future work.

2 Literature Review

HRC is a paradigm characterized by multiple aspects, both related to the robotic system and humans. The introduction of collaborative robots has allowed physical interaction with people, removing barriers between the human and robot workspace. However, the removal of these barriers also introduced new potential hazards to humans, requiring an evolution of safety standards. The introduction of ISO 10218-1 [14] and ISO 10218-2 [15] provided guidelines on workspace design and implementation of industrial robots, identifying a list of safety hazards. The subsequent ISO/TS 15066 [1] expanded the possibilities of HRC, allowing for the implementation of higher levels of robot autonomy in close proximity to humans.

HRC makes it possible to provide greater flexibility to manufacturing processes by combining human and robot skills. However, to exploit the full potential of HRC it is not enough just to comply with safety standards. Careful planning of the interaction is required so that it can best assist humans in the most demanding operations. Inkulu et al. [16] provided a review on HRC, highlighting some main challenges and opportunities. Natural modes of communication, such as voice and gestures, allow intuitive interaction with robots and potentially reduce idle time, but these recognition methods need to be made more robust to possible environmental disturbances. Power force limiting techniques are useful to efficiently collaborate with low-payload robot, however they may be not suitable for high-speed and high-payload robots which requires the implementation of additional flexible safety methods to allow collaboration with humans. Collaborative robots represent enabling technologies for reconfigurable production systems, however more research is needed on more advanced adaptive robot systems to reduce potential production downtime.

In recent years, increased attention has been focused on human factors involved in HRC. In addition to physical ergonomics, psychological and cognitive ergonomics play a key

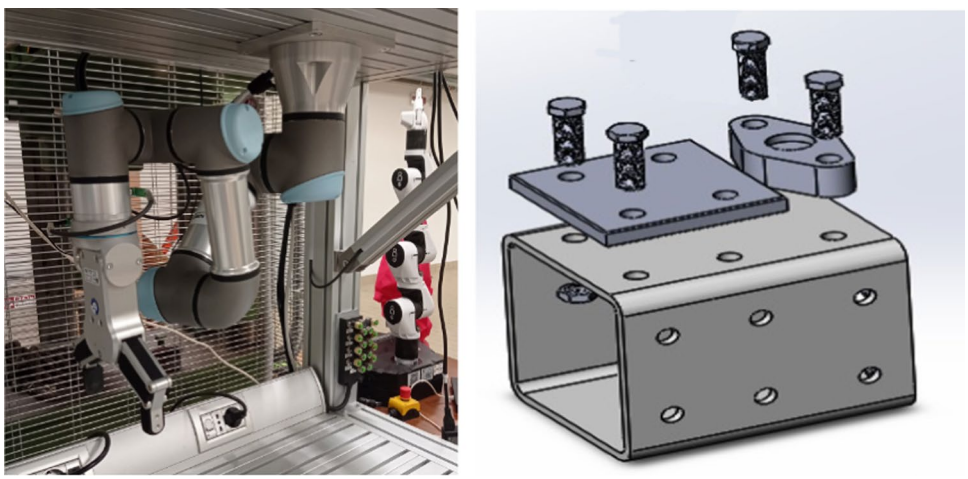
role in the effective implementation of HRC, analyzing the cognitive processes required to fulfill a given task as well as the psychological stressors involved [17, 18]. Galin and Meshcheryakov [19] focused on how to efficiently implement HRC, leading to the identification of the influencing factors for both cobots and humans. The perception of the robot by the human, as well as his/her emotional and cognitive aspects, were found influential for the effectiveness of HRC. Khalid et al. [20] presented an approach for the development of safe and cyber-secure HRC in the domain of heavy payload industrial robots. This led to the proposal of an integrated cyber-physical production system (CPPS) with also the identification of hazard sources. The main hazard categories for CPPS included hazards from the collaborating robot (e.g., robot characteristics, speed, and force), from the industrial process (e.g., ergonomic design and environmental conditions), and from the CPPS control layer malfunction (e.g., misuse of the system, cyber-attacks, and obstacles for active sensors).

Wang et al. [21] introduced the concept of symbiotic HRC. In traditional automation practice, humans must comply with rigid work procedures like the rest of the automated manufacturing environment. This rigidity is also present in some industrial HRC applications. To overcome this limitation a more responsive, intelligent and accessible collaborative system is required. In particular, symbiotic HRC aims to provide: (i) a multi-modal, intuitive programming environment that does not require in-depth knowledge of the system; (ii) natural communication with the robot through the use of inputs such as voice and gestures, potentially even allowing new tasks to be programmed; (iii) an immersive collaboration, allowing the operator to be involved in the tasks also through the use of wearable devices, such as Augmented Reality (AR) glasses or smart watches; (iv) an increased context dependency, where the system should be able to understand with the support of sensors the situation based on human and robotic conditions and actions, adapting accordingly.

In order to achieve an "advanced" HRC, several recent works have focused on proposing methods for dynamic adaptation of robot control during HRC. Mohammed et al. [22] presented a novel approach for effective online collision avoidance in an augmented environment, where virtual 3D models of robots and real images of human operators from depth cameras are used for monitoring and collision detection. Liu et al. [23] introduced a dynamic speed and separation monitoring (SSM) method for ensuring safe HRC while maintaining productivity as high as possible. The system included a dynamic risk-assessment and safe motion control based on the virtual model of the robot and human skeleton point data acquired from a vision sensor. In addition, through an augmented reality environment, the operator was able to visualize the risk-field around the robot, providing more transparency on the activation of the collision avoidance system. Roveda et al. [24] presented a model-based reinforcement learning variable impedance controller to assist human operators in collaborative tasks, such as a collaborative lifting. A set of neural networks was used to learn a human-robot interaction dynamic model, capturing also uncertainties. Subsequently, the learned model was kept updated through new information from collaborative tasks execution. Joseph et al. [25] proposed an aggregated "digital twin" solution for a collaborative workcell, which employed a collaborative robot and human workers. The architecture provides mechanisms to aggregate data and functionality in a manner that reflects reality; thereby enabling the intelligent, adaptive control of a collaborative robot.

2.1 Physiological Measures in HRC

In order to better understand and support the operator during HRC, cognitive and psycho-physical aspects should also be



(a) (b)

Fig. 1 (**a**) The collaborative robot UR3e with the gripper end-effector [34]. (**b**) Schema of the mechanical component with its components

Table 1 Operation list of the collaborative assembly task

Operation No.	Operation	Operation allocation	
		Control: *Human*	Control: *NoHuman*
OP1	Placement of the square flange on the base.	Robot	Robot
OP2	Fixing the square flange to the base with a pair of screws and nuts.	Human	Human
OP3	Commanding to continue the task.	Human	–
OP4	Placement of the oval flange on the base.	Robot	Robot
OP5	Fixing the oval flange to the base with a pair of screws and nuts.	Human	Human
OP6	Commanding to continue the task.	Human	–
OP7	Placement of the assembled component in another working area.	Robot	Robot

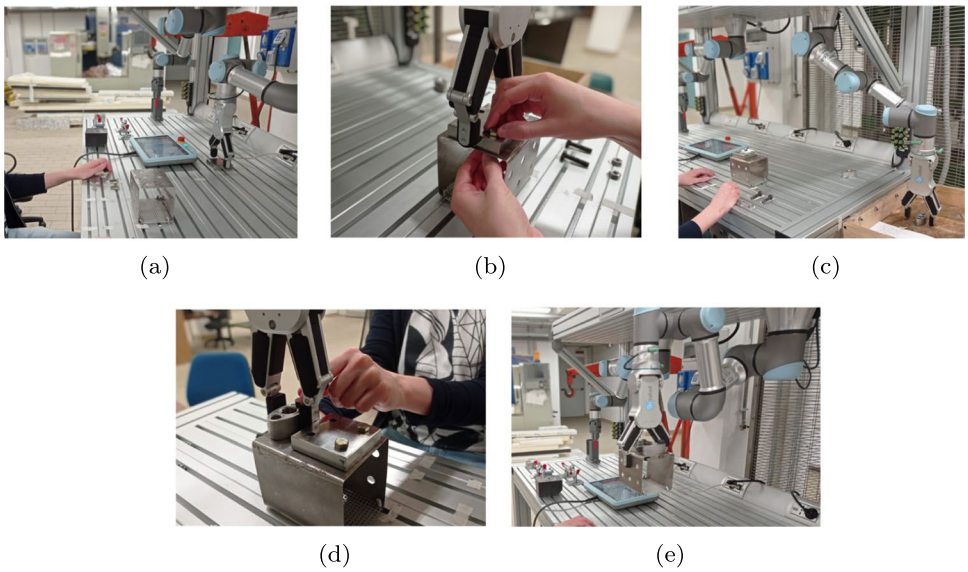


Fig. 2 Operations of the collaborative assembly task: (**a**) the UR3e picks up the square flange and places it on the base; (**b**) the operator takes the screws, inserts them into the holes and tightens them; (**c**) the robot takes the oval flange and places it on the base; (**d**) the operator takes the other two screws, inserts them into the holes and tightens them; (**e**) the robot takes the assembled component and places it in another area

taken into account. Concepts like mental workload, stress, demand, strain and fatigue have been widely discussed in literature, being particularly interesting for the manufacturing context [9, 26, 27]. Assessment of these constructs is often performed through self-reporting tools, such as the NASA-TLX [28] and the Subjective Workload Assessment Technique (SWAT) [29]. However, this kind of tools may suffer from post-task retrospective bias (i.e., a bias introduced by recalling an event) and are poorly suited to continuous monitoring in naturalistic settings, such as production lines [30, 31]. To overcome these limitations, in recent years there has been an increasing focus on physiological measures for the comprehension of the operator's state [32, 33]. So far, different works on this topic have been presented, however only few of them are focused on industrial HRC. Kulic and Croft [13] evaluated the impact of an industrial robot motion on subjective and physiological responses (i.e., Heart Rate Variability (HRV) and ElectroDermal Activity (EDA)) with various trajectory types presented to human participants. Results revealed an increased mental stress for fast and closely passing movements, but the scenario was static in terms of interaction with the robot. Arai et al. [11] conducted a similar study with an industrial manipulator,

Fig. 3 The Empatica E4 biosensor used for the experiment [35]

Table 2 Levels of prior experience with collaborative robots

Level	Statement
L0	I have never interacted with a cobot and I did not know them before now.
L1	I have never interacted with a cobot but I know what they are.
L2	I have interacted at least once with a cobot.
L3	I have already programmed and interacted with a cobot.

evaluating the impact of robot movement at different speeds and distances from the operator on EDA. However, also in this scenario, participants were not actively involved in the interaction with the robot and the number of participants was quite limited. Kühnlenz et al. [12] studied the impact of different trajectory profiles of a standard industrial robot on users' mental stress, assessed through HRV and EDA. Although the participant was actively involved in the task compared to other studies, there was still limited interaction with the robot.

3 Method

Building upon findings within the literature, we explored the differences in physiological response, self-reported affective state, and interaction quality to variations in HRC configuration throughout the implementation of a collaborative assembly task.

3.1 Experimental Design

In the present study, a collaborative assembly task was designed and implemented within the "Mind 4 Lab" (Manufacturing Industry 4.0 Laboratory) at the "Politecnico di Torino" to investigate user experience, operator affective state and stress in industrial HRC. The collaborative assembly task implemented was designed to recreate a workstation of a production cycle in an industrial context. The task consisted of assembling two mechanical flanges onto a base by tightening two pairs of screws (Fig. 1b) with the support of the collaborative robot UR3e [34] (Fig. 1a). The list of operations of the task is reported in Table 1. The collaborative robot began by picking up the square flange and placing it on the base in the correct position (Fig. 2a). Then, the operator took the screws, inserted them into the holes and tightened them (Fig. 2b). Next, the robot took the oval flange and placed it correctly on the base (Fig. 2c). The operator took the other two screws, inserted them into the holes and tightened them (Fig. 2d). Finally, the robot took the assembled mechanical component and placed it in another work area (Fig. 2e).

The study implemented a three-factor within-subjects design approach to investigate the effects of varying robot movement speed, distance between robot workspace and operator, and control of the task execution time on the perceived interaction quality, self-reported affective state, and physiological response. The UR3e robot is an articulated arm whose movement is managed by the action of six rotating joints. Therefore, the movement speed of the robot's joints (*Speed*) was varied over three levels: *Low* (30°/s), *Medium* (90°/s), and *High* (270°/s). Speed levels were set by

Table 3 English version of NARS

Sub-scale	Questionnaire item
S1: *Negative Attitudes toward* Situations and Interactions with Robots	I would feel uneasy if I was given a job where I had to use robots.
	The word "robot" means nothing to me.
	I would feel nervous operating a robot in front of other people.
	I would hate the idea that robots or artificial intelligence were making judgements about things.
	I would feel very nervous just standing in front of a robot.
	I would feel paranoid talking with a robot.
S2: *Negative Attitudes* toward Social Influence of Robots	I would feel uneasy if robots really had emotions.
	Something bad might happen if robots developed into living beings.
	I feel that if I depend on robots too much, something bad might happen.
	I am concerned that robots would be a bad influence on children.
	I feel that in the future society will be dominated by robots.
S3: *Negative Attitudes* toward Emotions in Interaction with Robots	I would feel relaxed talking with robots.*
	If robots had emotions, I would be able to make friends with them.*
	I feel comforted being with robots that have emotions.*

Reverse items are indicated with "*"

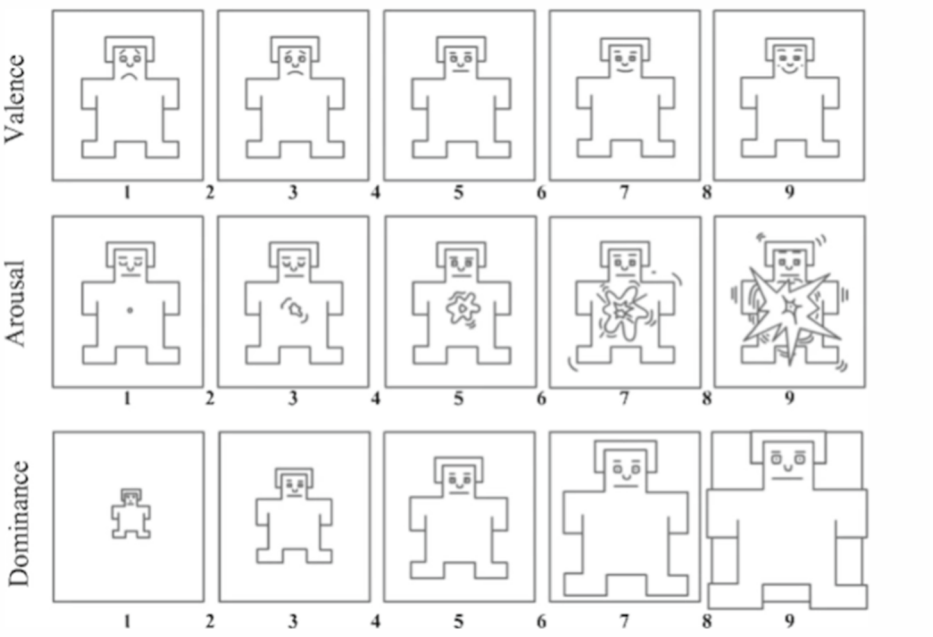


Fig. 4 Self-Assessment Manikin (SAM) with its three dimensions: valence, arousal, and dominance

Table 4 Questionnaire for the quality of interaction

Item No.	Questionnaire item	Dimension
Q1	The robot was helpful in accomplishing the task.	Robot helpfulness
Q2	I felt the interaction was not safe.	Interaction unsafety
Q3	The collaboration felt natural.	Interaction naturalness
Q4	The robot and I worked efficiently together.	Team efficiency
Q5	The robot and I worked fluently together.	Team fluency
Q6	I felt uncomfortable with the robot.	Discomfort
Q7	The robot was trustworthy.	Robot trustworthiness

accentuating the differences between them as much as possible. The distance between the robot workspace and the operator's chest (*Distance*) represents how close the robot can move in front of the operator during the task. It was varied on two levels so that the task was feasible: *Close* (30cm) and *Far* (40cm). Finally, the control of the task execution time (*Control*) was made to vary on two levels: *Human* (the operator has a button to communicate to the robot that he/

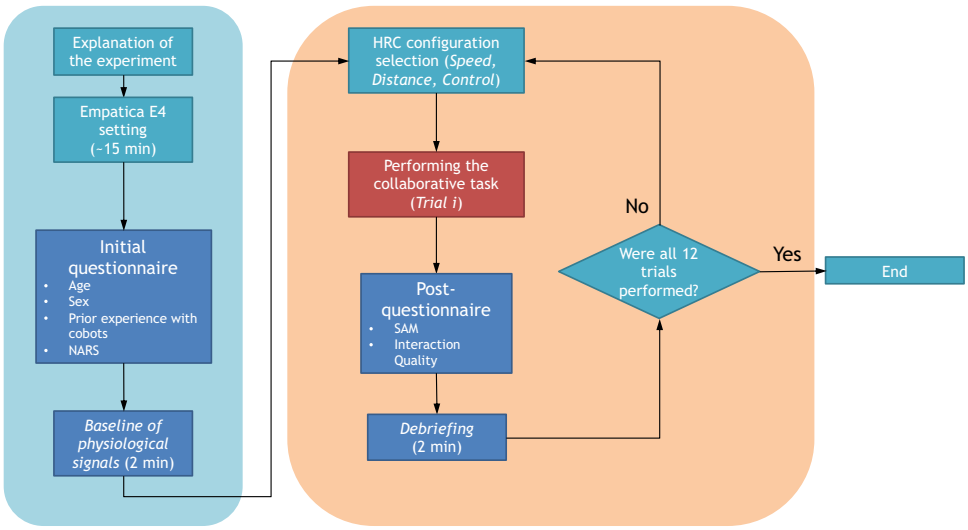


Fig. 5 Flow-chart of the experimental procedure

Table 5 Summary of the variables involved in the study

Type of variable	Variable category	Variable name	Description
Independent variables	Configuration factor	*Speed*	Robot movement speed (*Low*, *Medium*, *High*)
		Distance	Distance between the robot workspace and the operator (*Close*, *Far*)
		Control	Control of the task execution time (*Human*, *NoHuman*)
Response variables	Interaction quality	*Q1_Helpful*	Evaluation of perceived robot helpfulness
		Q2_NotSafe	Evaluation of perceived interaction unsafety
		Q3_Natural	Evaluation of perceived interaction naturalness
		Q4_Efficient	Evaluation of perceived team efficiency
		Q5_Fluid	Evaluation of perceived team fluency
		Q6_Uncomfortable	Evaluation of perceived discomfort
		Q7_Trustworthy	Evaluation of perceived robot trustworthiness
	Self-reported affective state	*Valence*	SAM dimension representing how much pleasant is an emotion
		Arousal	SAM dimension representing how much excited/agitated a person feels
		Dominance	SAM dimension representing how strong is the dominance feeling
	Physiological response	*Mean_SCR*	Average of Skin Conductance Response amplitudes (EDA indicator)
		RMSSD	Root Mean Square of Successive Differences between adjacent heart rate NN-intervals (HRV indicator)

she has finished his/her task and that it can continue) and *NoHuman* (the robot automatically continues with its operations after a predetermined waiting time of 25s). This factor distinguishes between two kinds of situations: one in which the operator completely manages the task time and another in which the task time is predetermined.

3.2 Equipment

The Empatica E4 biosensor wristband [35] (Fig. 3) was used to collect EDA data at 4Hz, heart data through Photopletismogram (PPG) at 64Hz, and 3-axis accelerometer data at 32Hz. The device also provided the heart rate NN-intervals. PPG and EDA data were used as arousal and stress indicators, evaluating HRV and average skin conductance response in each HRC configuration.

In addition to the physiological biosensor, the materials also included a questionnaire capturing previous experience with cobots and demographic data, the Negative Attitude toward Robots Scale (NARS) [36], the Self-Assessment Manikin (SAM) [37, 38], and a questionnaire for the interaction quality. Based on similar previous studies [39–41], the level of experience was evaluated on a four-level scale shown in Table 2, asking the participant to select which statement best represented his/her degree of experience with cobots.

The NARS [36] was used to further characterize the participant, evaluating his/her attitudes towards robot. The NARS was developed for measuring humans' attitudes toward robots, i.e., feelings or ways of thinking that affects a person's behavior toward robots. It is a 5-point Likert-scale (from "strongly disagree" to "strongly agree") composed by 14 items (Table 3). The NARS items are divided in the

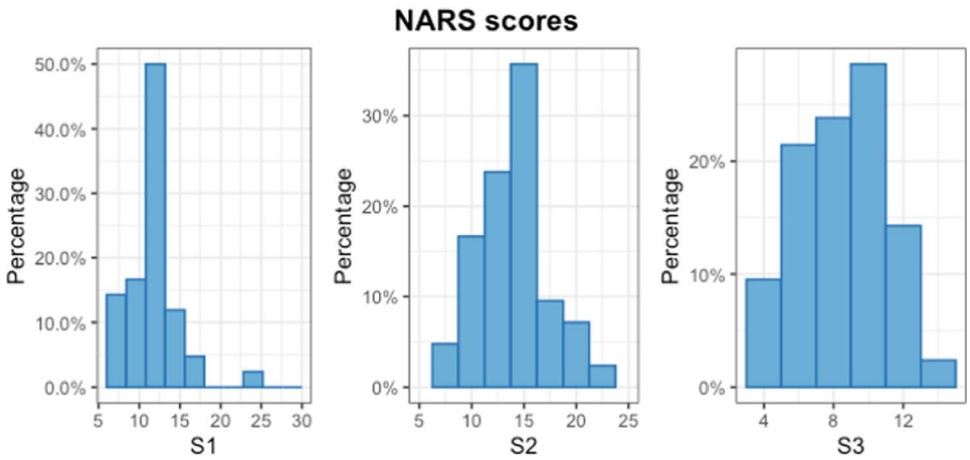


Fig. 6 Distribution of the three NARS scores: *Negative Attitudes toward Situations and Interactions with Robots* (S1), *Negative Attitudes toward Social Influence of Robots*(S2), and *Negative Attitudes toward Emotions in Interaction with Robots* (S3)

following three sub-scales: (i) *Negative Attitudes toward Situations and Interactions with Robots* (S1) (scoring from 6 to 30), (ii) *Negative Attitudes toward Social Influence of Robots* (S2) (scoring from 5 to 25), and (iii) *Negative Attitudes toward Emotions in Interaction with Robots* (S3) (scoring from 3 to 15) [36].

The SAM [37, 38] is a widely used image-based assessment tool to measure the affective reaction to a certain situation or event. It is based on the Pleasure-Arousal-Dominance (PAD) model, which represents affective states on three dimensions: valence, arousal, and dominance. Valence (or pleasure) describes the positivity or negativity of an elicited emotion (e.g., fear, anger, or boredom tend to be negative emotions, whereas relaxation or joy tend to be positive emotions). Arousal refers to how excited a person is, regardless of whether the excitement derives from a positive or negative emotion (e.g., boredom and relaxation are characterized by low arousal, whereas euphoria, fear, or anger tend to have a high arousal). Finally, dominance describes how much one feels in control of a situation, i.e., a feeling of control and influence over one's surroundings and others (e.g., fear or anxiety are usually characterized by low dominance, while relaxation or anger by a high dominance). Figure 4 shows the original 9-point scale SAM, which was used in the study to collect affective state of the participants during the different task configurations.

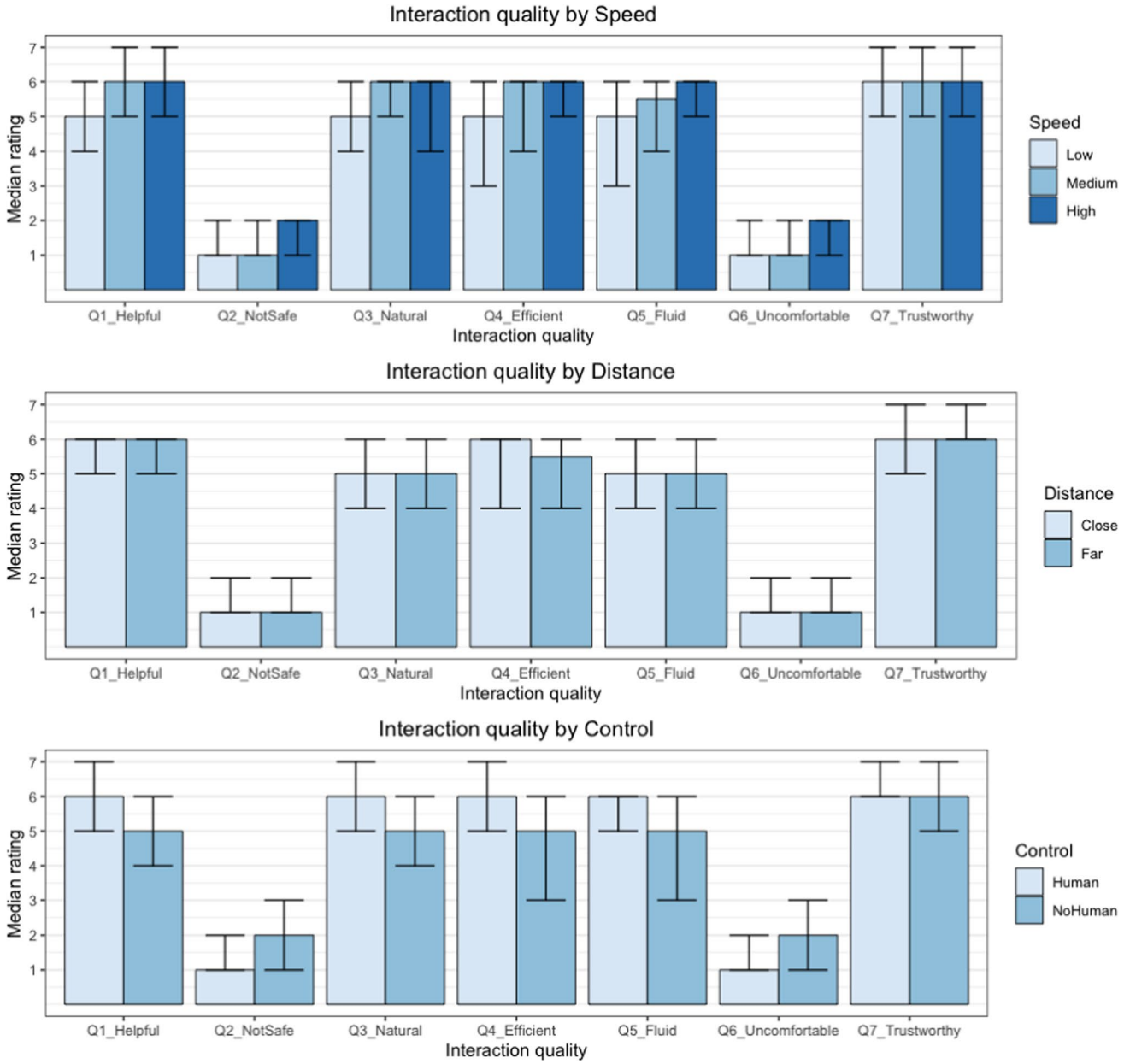


Fig. 7 Median ratings with interquartile range of interaction quality dimensions for the configuration factors *Speed*, *Distance*, and *Control*

A questionnaire on interaction quality (Table 4) based on Baraglia et al. [42] and Hoffman [43] was also administrated to the participant at the end of each trial. The questionnaire is composed of 7 items, in order to have a quick administration tool during the experiment that can collect different aspects of the interaction with the cobot. The 7 items collect participant's perception of robot helpfulness, interaction safety and naturalness, team efficiency and fluency, comfort, and robot trustworthiness. Robot helpfulness represents how helpful the robot is in accomplishing a certain task. Interaction safety refers to how safe the HRC is perceived. Interaction naturalness concerns the easiness of the interaction with the robot. Team efficiency represents how efficient the collaboration is. Team fluency refers to the level of coordination during the collaborative task. Comfort represents how at ease a person feels during HRC. Robot trustworthiness represents how reliable the robot is perceived to be during HRC. Each item was evaluated using a 7-point scale (from "strongly disagree" to "strongly agree").

3.3 Procedure

Figure 5 outlines the experiment procedure. Initially, the objectives of the study and its procedure were explained to the participant. The researcher subsequently seated the participant in the experiment location and explained the various steps of the collaborative assembly task, also illustrating the various configurations. After any questions were discussed, the Empatica E4 biosensor was firmly placed on the participant's left wrist and 15 minutes were waited for the electrodes to adhere well to the skin and to obtain reliable EDA data. The participant was asked to fill an initial questionnaire, which included demographics (gender and age), prior experience with cobots, and the NARS. Next, the participant was invited to relax and remain still to record 2 minutes of physiological signals at rest (i.e., the baseline of the physiological signals). After this phase was completed, the participant performed the collaborative assembly task

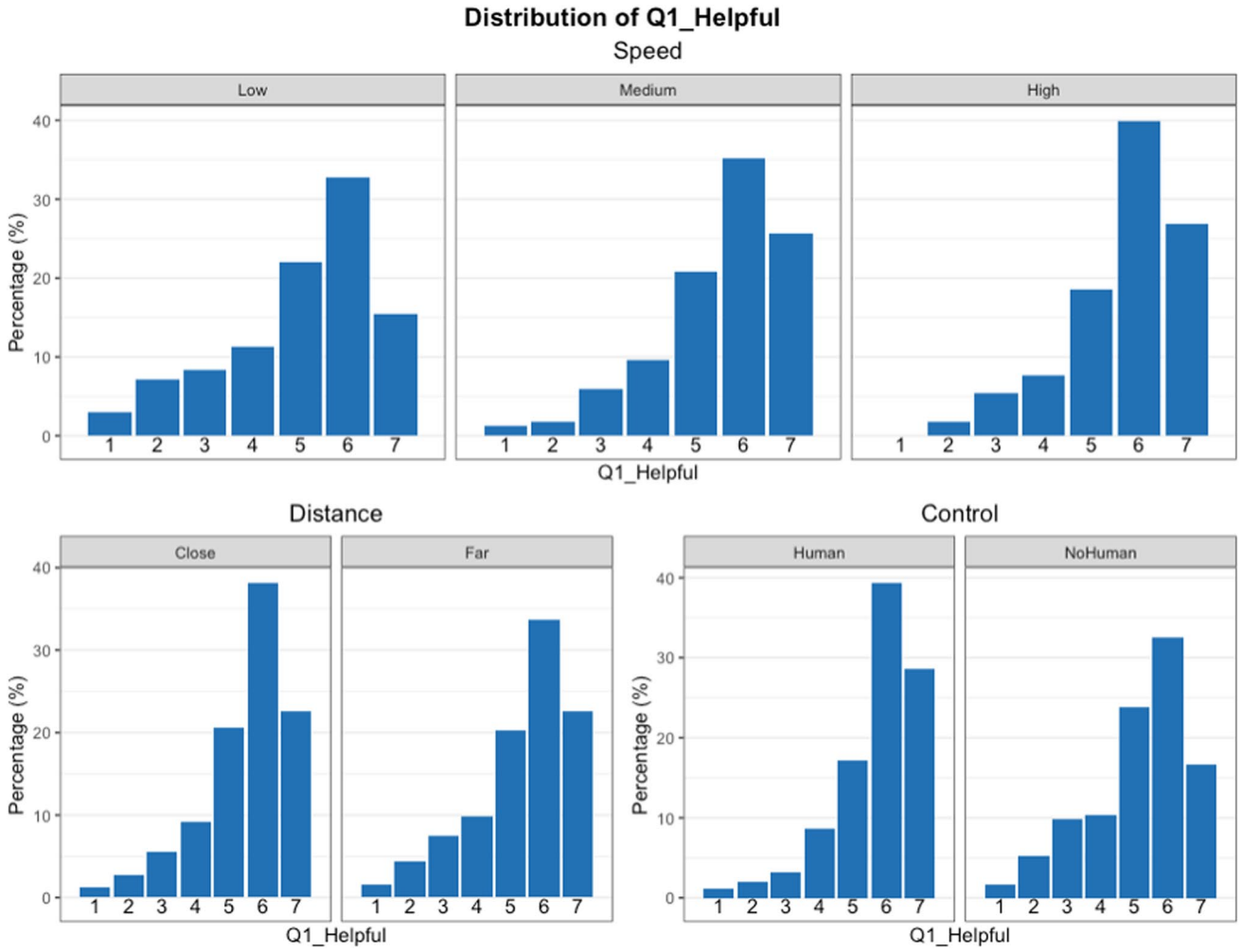


Fig. 8 Distribution of the perceived robot helpfulness (*Q1_Helpful*) for the configuration factors *Speed*, *Distance*, *Control*. Note that *Q1_Helpful* ranges from 1="Strongly disagree" to 7="Strongly agree"

with all the 12 possible configurations in a random order. The researcher supervised each trial and, between trials, set up the HRC configuration and the workpieces. At the end of each trial, the participant reported his/her affective state during the trial through the SAM, filled the questionnaire on the quality of interaction with the cobot, and was involved in a quick debriefing. At the conclusion of the experiment, the participant was asked for a general unstructured feedback about the overall experience. The entire experimental session lasted 90 minutes on average.

3.4 Data Processing

Table 5 summarizes the response and independent variables considered in the study. Physiological data were checked and cleaned of possible artifacts. EDA data were processed using the MATLAB-based software "Ledalab" [44]. Continuous Decomposition Analysis (CDA) [45] was performed to decompose the EDA signal into continuous signals of phasic and tonic activity. Tonic activity refers to long-term fluctuations in EDA that are not specifically elicited by external stimuli and is best characterized by changes in Skin Conductance Level (SCL). In contrast, phasic activity refers to short-term fluctuations in EDA which have been elicited by a usually identified and externally presented stimulus. Through the analysis of the phasic activity signal, Skin Conductance Responses (SCRs) (i.e., amplitude changes from the SCL to a peak of the response) can be identified. In this study, the average SCR was used as an arousal and stress indicator in each HRC configuration. From heart data, HRV measures can be derived and used as an arousal and stress indicator. In this study, the Root Mean Square of Successive Differences between adjacent NN-intervals (RMSSD) was considered as measure of HRV due to its common use [9, 46].

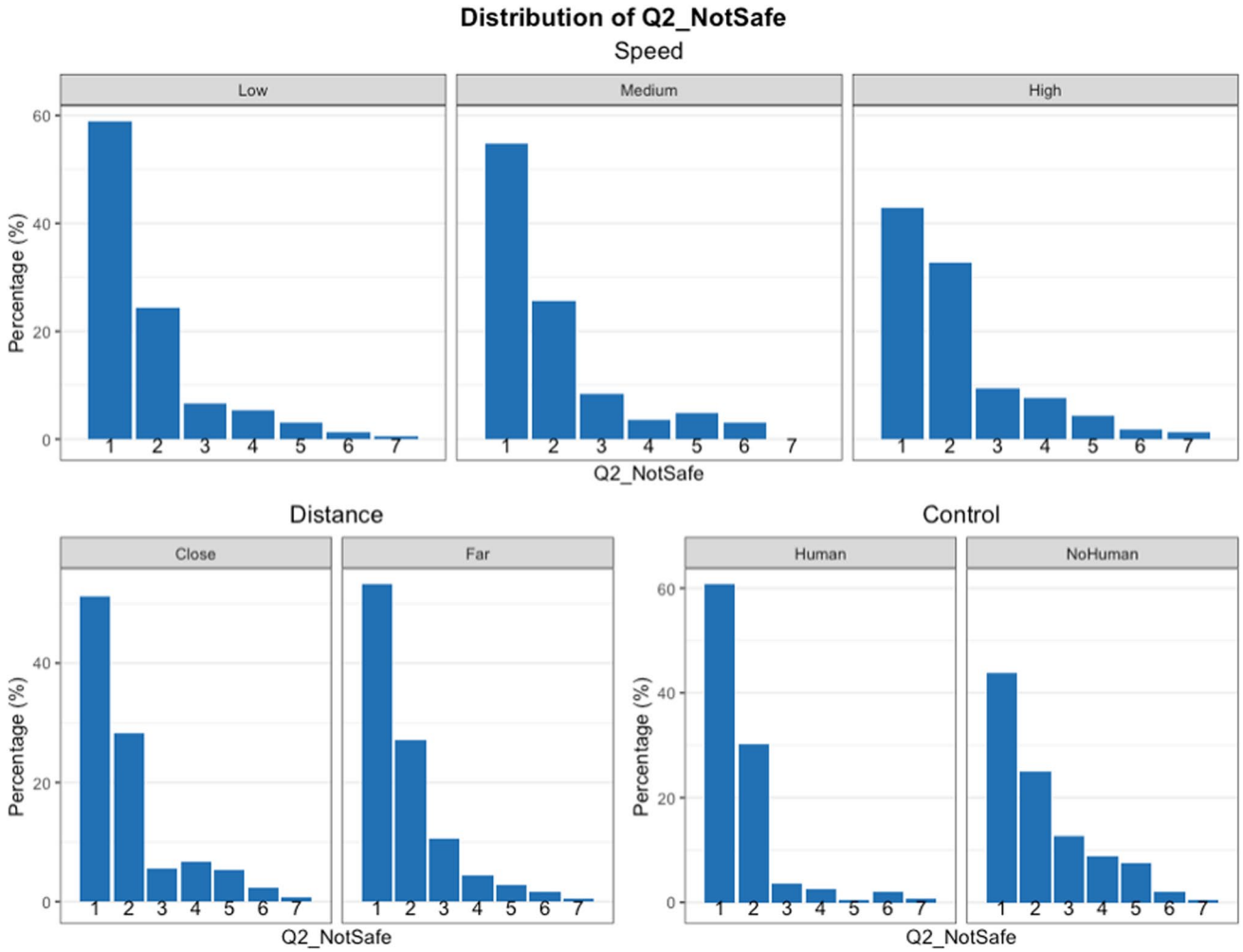


Fig. 9 Distribution of the perceived interaction unsafety (*Q2_NotSafe*) for the configuration factors *Speed*, *Distance*, *Control*. Note that *Q2_NotSafe* ranges from 1="Strongly disagree" to 7="Strongly agree"

4 Results and Analysis

In this section, the obtained results from the experiment are reported. After an overview of the participants involved in the study, analysis of the influence of the configuration factors on each response variable will be presented.

4.1 Participants

The study involved 42 participants (71.4% males and 28.6% females), with a mean average age of 28.24 years (sd = 8.1), who were recruited from the "Politecnico di Torino" and surrounding community. Regarding prior experience with cobots, 28.6% of participants had never interacted with a cobot and did not know them before the experiment; 45.2% of participants had never interacted with a cobot but knew them; 16.7% of participants had already interacted with a cobot; 9.5% of participants had already programmed and interacted with a cobot.

Figure 6 shows the distributions of the three NARS scores of the participants. The average score for *Negative Attitudes toward Situations and Interactions with Robots* (S1) was 11.64 (sd= 3.07), which is less than half of the maximum score (i.e., 30). Regarding *Negative Attitudes toward Social Influence of Robots* (S2), the average score was 13.95 (sd= 3.20), revealing some worries for the potential social influence of robots. Finally, the average score for *Negative Attitudes toward Emotions in Interaction with Robots* (S3) is 8.98 (sd= 2.54), showing some concern towards affective interaction with robots. These results show that the sample of participants is quite willing to interact with robots, however it seems to show some doubts about the social role of robots and the strong involvement of emotions during the interaction with them.

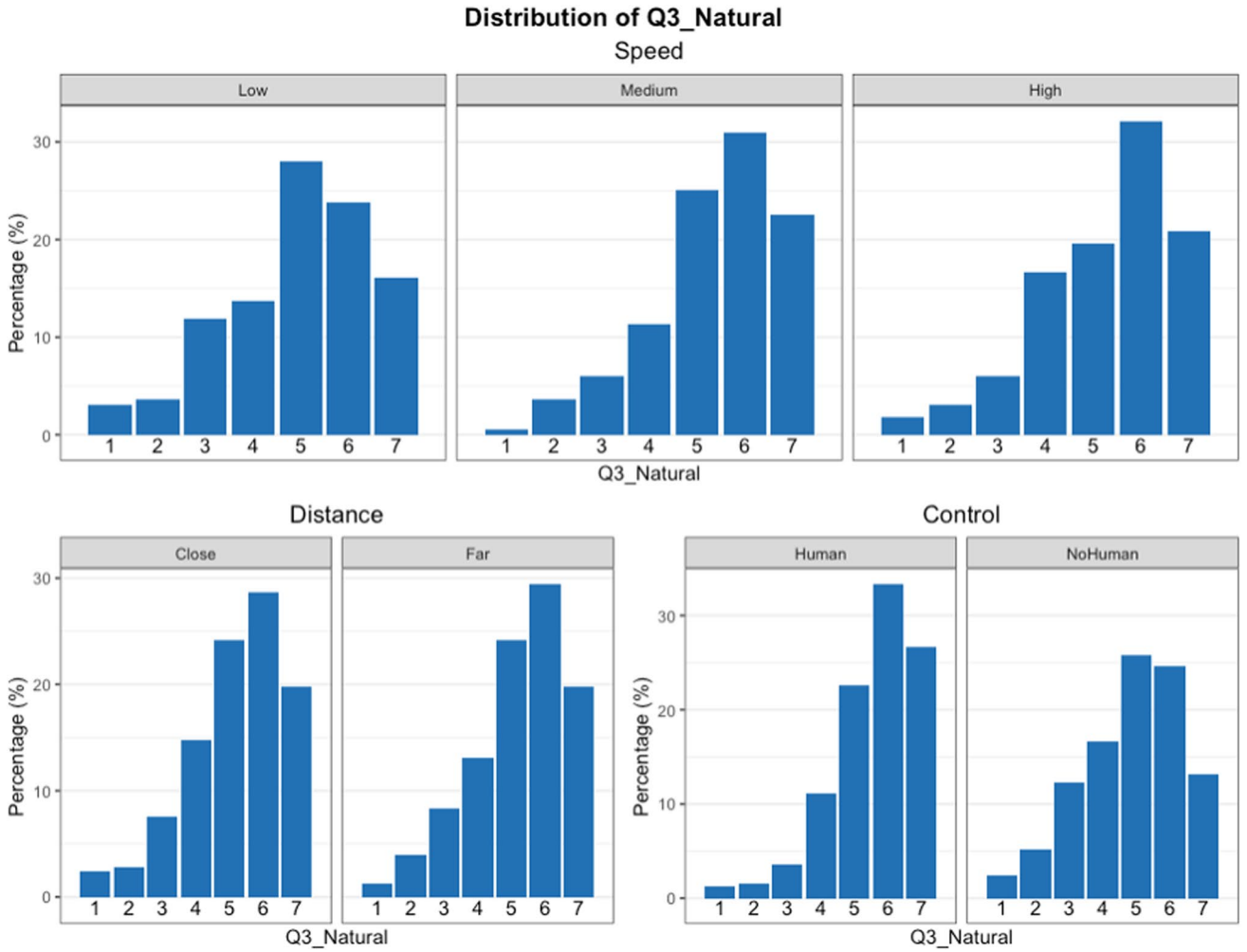


Fig. 10 Distribution of the perceived interaction naturalness (*Q3_Natural*) for the configuration factors *Speed*, *Distance*, *Control*. Note that *Q3_Natural* ranges from 1="Strongly disagree" to 7="Strongly agree"

4.2 Interaction Quality

In this section, a descriptive analysis of the perceived interaction quality is presented. Figure 7 provides an overall graphical comparison between the different levels of the configuration factors *Speed*, *Distance*, and *Control* for each interaction quality dimension. Additionally, significant differences in response variables by fixed factors is also highlighted through the Wilcoxon signed-rank test [47]. This test is suitable for analyzing paired ordinal data, hence allowing to consider the within-subject effect [48]. Since multiple pairwise comparisons were needed for the *Speed* factor to assess the differences between its three levels, Bonferroni's post-hoc correction has been applied [49, 50].

Figure 8 shows data distribution of perceived robot helpfulness (*Q1_Helpful*) as the configuration factors are varied. It can be observed that:

- As the level of *Speed* increased, the perceived robot helpfulness increased too. A great difference in the distributions can be seen when comparing the *Low* speed with the other two levels (see also Fig. 7). This result is confirmed by the Wilcoxon signed-rank test with Bonferroni correction, showing a highly significant difference between *Low* and *Medium* speeds ($p < 0.001$) and *Low* and *High* speeds ($p < 0.001$). The difference between *Medium* and *High*, instead, resulted not significant ($p = 0.471$).
- A slight difference can be seen between the two *Distance* levels, with a higher perceived robot helpfulness when the robot was closer to the operator. However, through the Wilcoxon signed-rank test, this difference was not found to be significant ($p = 0.144$) (Fig. 7).
- When the participant had no control of the task execution time (i.e., *Control=NoHuman*), robot helpfulness was perceived to be lower (Fig. 7). A highly signifi-

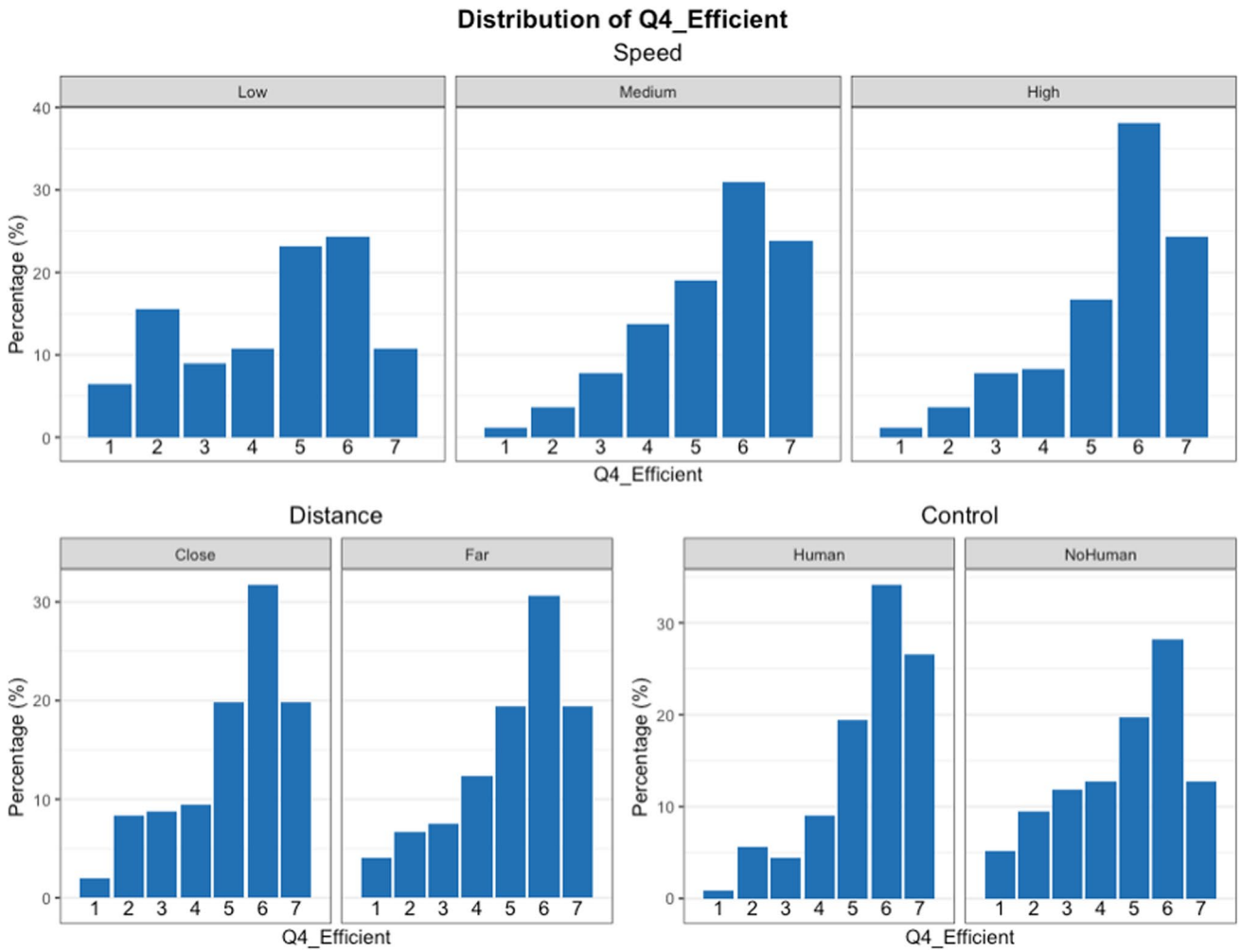


Fig. 11 Distribution of the perceived team efficiency (*Q4_Efficient*) for the configuration factors *Speed*, *Distance*, *Control*. Note that *Q4_Efficient* ranges from 1="Strongly disagree" to 7="Strongly agree"

cant difference was also highlighted by the Wilcoxon signed-rank test ($p < 0.001$).

Regarding the perceived interaction unsafety (*Q2_NotSafe*), Fig. 9 shows the response distributions for each configuration factor, and the following can be inferred:

- A high robot movement speed (*Speed*) resulted in a degradation in the perceived safety. The Wilcoxon signed-rank test with Bonferroni correction showed a significant difference between *Low* and *High* levels ($p = 0.007$), however the comparisons *Low-Medium* and *Medium-High* did not lead to a significant difference ($p = 0.645$ and $p = 0.159$, respectively).
- When the robot workspace was closer to the participant, the perceived unsafety was slightly higher. However, the response distributions for the two *Distance* levels are almost identical, and indeed the Wilcoxon signed-rank test confirmed a non-significant difference ($p = 0.171$).
- Not controlling the task execution time led to an increase of the perceived unsafety. Moreover, the difference resulted highly significant by the Wilcoxon signed-rank test ($p < 0.001$) (Fig. 7).

The influence of the configuration factors on the perceived interaction naturalness with the cobot (*Q3_Natural*) can be observed in Fig. 10:

- A low robot movement speed led to a consistent degradation of perceived interaction naturalness compared to the other two Speed levels, even changing the mode of the distribution. The Wilcoxon signed-rank test confirmed a highly significant difference between *Low* and *Medium* speed ($p < 0.001$) and between *Low* and *High* speeds (p

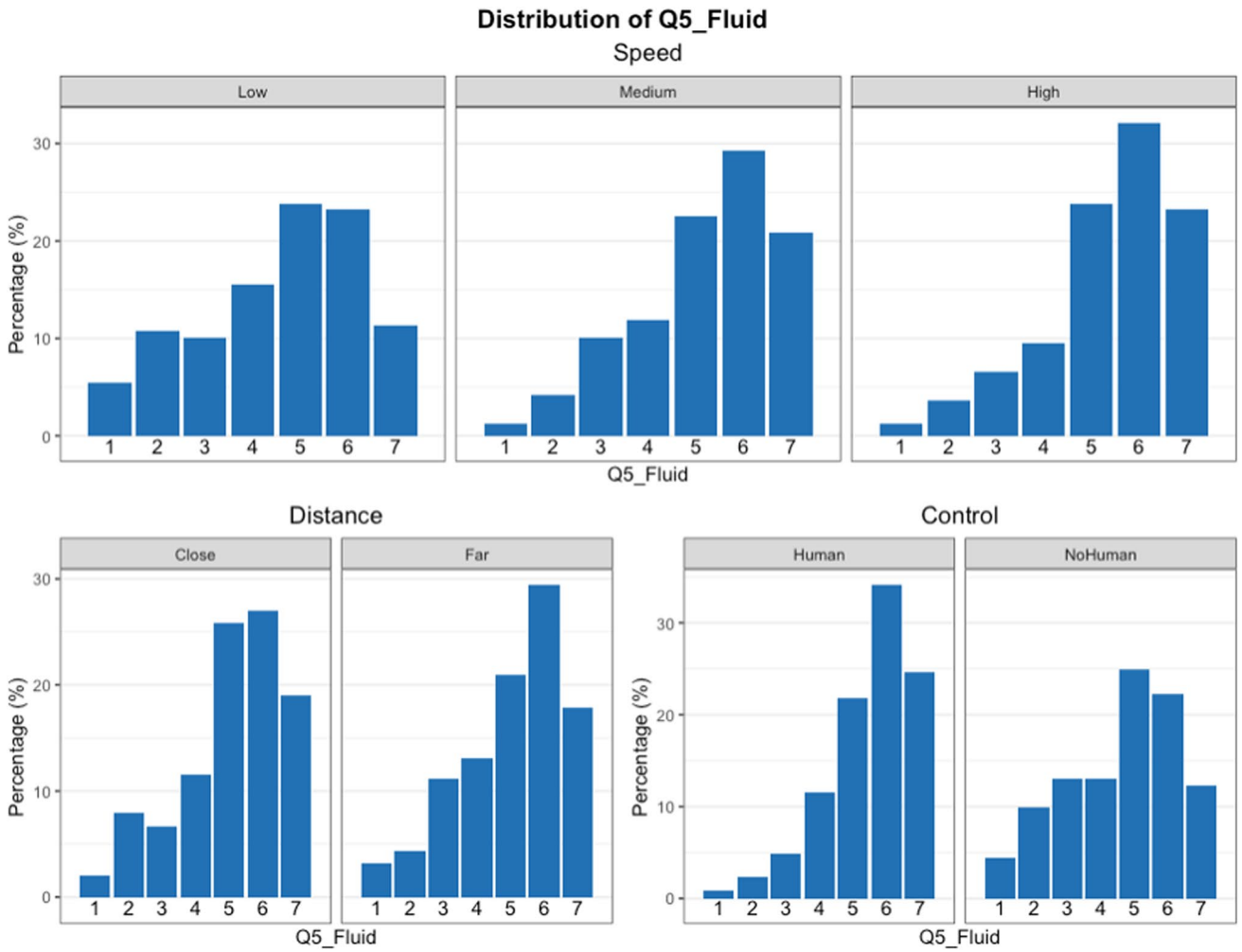


Fig. 12 Distribution of the perceived team fluency (*Q5_Fluid*) for the configuration factors *Speed*, *Distance*, *Control*. Note that *Q5_Fluid* ranges from 1="Strongly disagree" to 7="Strongly agree"

< 0.001). In contrast, the *Medium-High* comparison did not lead to a significant difference ($p = 1$).
- The *Distance* factor did not have a significant influence on the perceived interaction naturalness. This was also confirmed by the Wilcoxon signed-rank test ($p = 0.807$).
- Interaction in configurations without control of task execution time by the participant was perceived less natural. This can be observed in Fig. 10, with even a decrease in the mode of the response distribution. The Wilcoxon signed-rank test showed also a highly significant difference between the two *Control* levels ($p < 0.001$).

With respect to team efficiency (*Q4_Efficient*), Fig. 11 shows the response distributions as the configuration factors are varied:

- A consistent degradation of team efficiency can be observed for a low robot movement speed. Moreover, the Wilcoxon signed-rank test with Bonferroni correction confirmed a highly significant difference between *Low* and *Medium* speed ($p < 0.001$) and between *Low* and *High* speeds ($p < 0.001$). The difference between *Medium* and *High* resulted not significant ($p = 0.268$).
- The distance of the robot workspace from the operator did not have a significant influence on the perceived team efficiency, as can be seen by the almost identical distribution (Fig. 11). This was also confirmed by the Wilcoxon signed-rank test ($p = 0.922$).
- Not controlling the task execution time slightly decreased the perceived efficiency. This difference resulted highly significant by the Wilcoxon signed-rank test ($p < 0.001$).

Figure 12 shows data distribution of perceived team fluency (*Q5_Fluid*) for the configuration factors, and it can be noticed that:

- *Medium* or *High* robot speeds were associated with greater team fluency compared to the low one. The Wil-

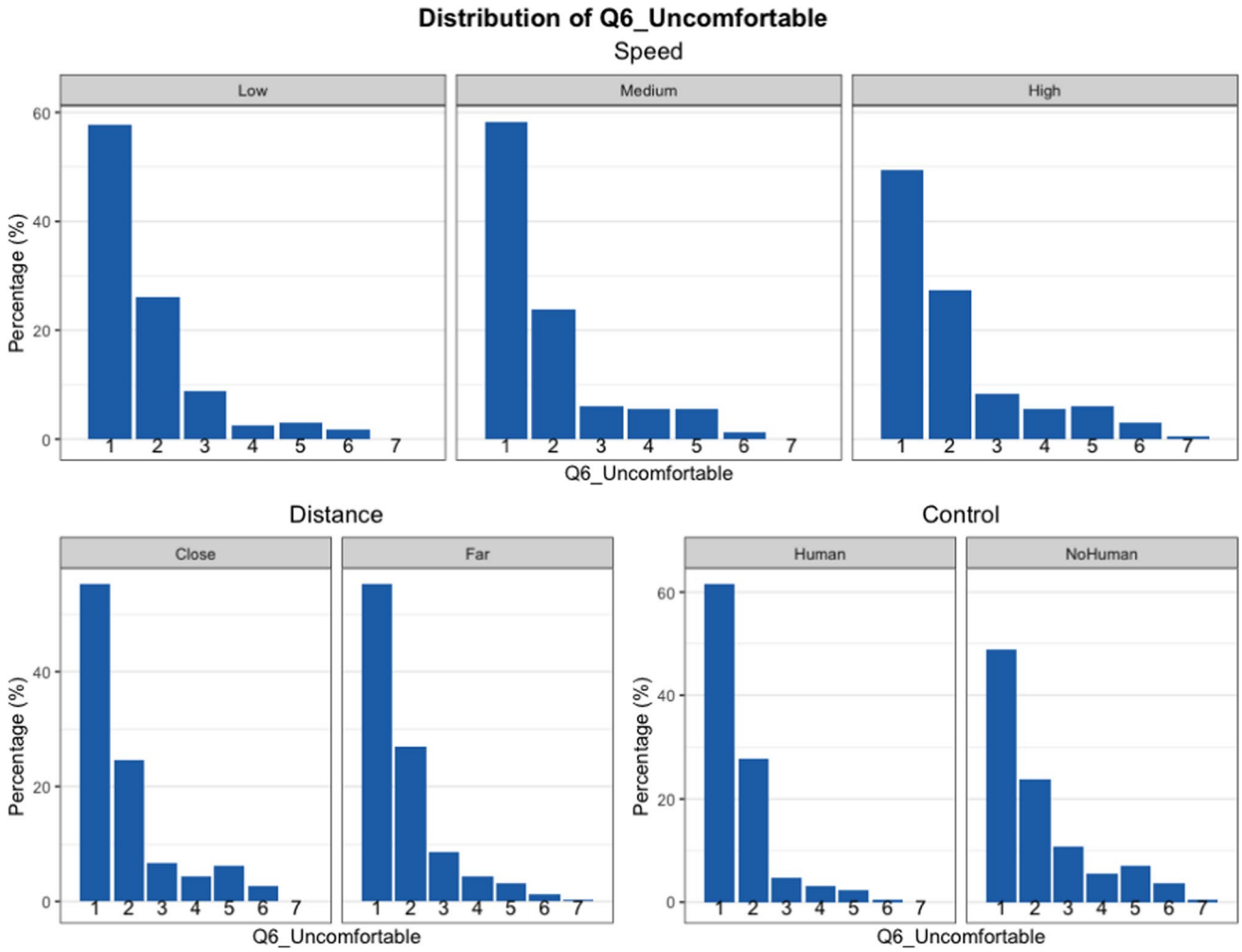


Fig. 13 Distribution of the perceived discomfort (*Q6_Uncomfortable*) for the configuration factors *Speed*, *Distance*, *Control*. Note that *Q6_Uncomfortable* ranges from 1="Strongly disagree" to 7="Strongly agree"

coxon signed-rank test with Bonferroni correction confirmed a highly statistical difference between *Low* and *Medium* speeds ($p < 0.001$) and between *Low* and *High* speeds ($p < 0.001$). The difference between *Medium* and *High* speeds was not found statistically significant ($p = 0.268$).

- A slight degradation of the perceived team fluency occurred when the robot workspace was closer to the participant. However, the Wilcoxon signed-rank test revealed a not significant difference between the two levels of *Distance* ($p = 0.542$).
- Team fluency in configurations without execution time control by the participant was negatively affected. Moreover, the Wilcoxon signed-rank test showed a significant difference between the two *Control* levels ($p < 0.001$).

The influence of the configuration factors on perceived discomfort (*Q6_Uncomfortable*) can be observed in Fig. 13:

- A high robot movement speed led to slightly more discomfort compared to lower ones. The Wilcoxon signed-rank test with Bonferroni correction confirmed a significant difference between *Low* and *High* speeds ($p = 0.009$), however the differences between *Low* and *Medium* speeds and between *Medium* and *High* speeds were not statistically significant ($p = 1$ and $p = 0.229$, respectively).
- Even if the *Close* level received slightly higher ratings, the distance of the robot workspace from the operator did not have a significant influence on the perceived discomfort, (Fig. 13). This was also confirmed by the Wilcoxon signed-rank test ($p = 0.204$).
- Perceived discomfort increased when the participant had no control of the task execution time. In addition, the Wilcoxon signed-rank test revealed a highly significant difference in perceived discomfort between the two *Control* levels ($p < 0.001$).

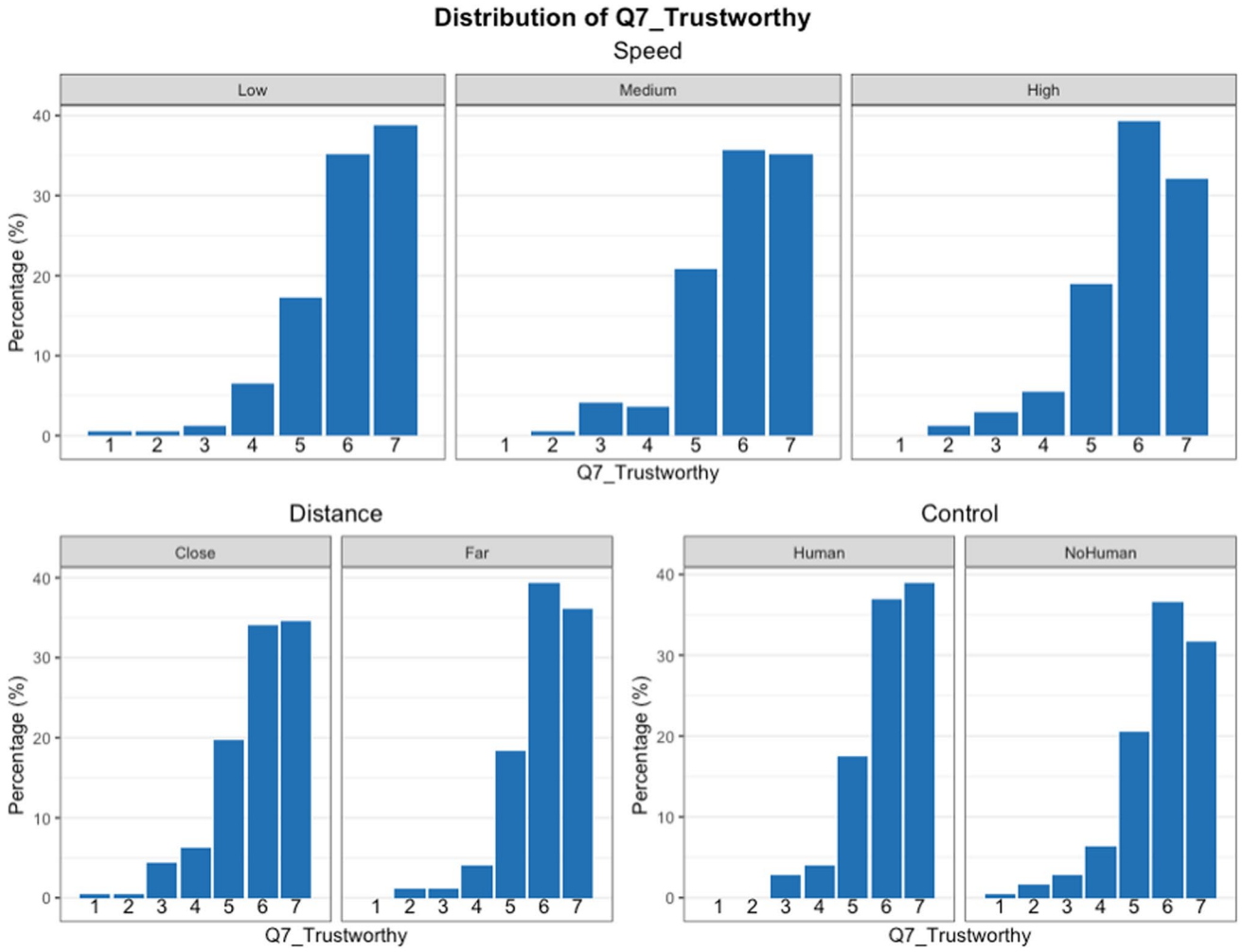


Fig. 14 Distribution of the perceived trustworthiness (*Q7_Trustworthy*) for the configuration factors *Speed*, *Distance*, *Control*. Note that *Q7_Trustworthy* ranges from 1="Strongly disagree" to 7="Strongly agree"

As regard robot trustworthiness (*Q7_Trustworthy*), Fig. 14 shows how the configuration factors influenced the response distribution:

- Higher robot movement speeds implied a slight decrease in robot trustworthiness. However, the Wilcoxon signed-rank test with Bonferroni correction revealed that this effect was not significant for any pairwise comparison (*Low-Medium*: $p = 1$; *Low-High*: $p = 0.564$; *Medium-High*: $p = 1$).
- A greater distance between the participant and robot workspace was associated with the cobot being more trustworthy. The Wilcoxon signed-rank test confirmed a significant difference between the two levels of *Distance* ($p = 0.022$).
- The lack of time execution control by the participant was associated with a loss of robot trustworthiness. This effect was confirmed to be highly significant by the Wilcoxon signed-rank test ($p = 0.001$).

4.3 Participant's Affective State

In this sub-section, the relationships between the experimental factors and the self-reported affective state through

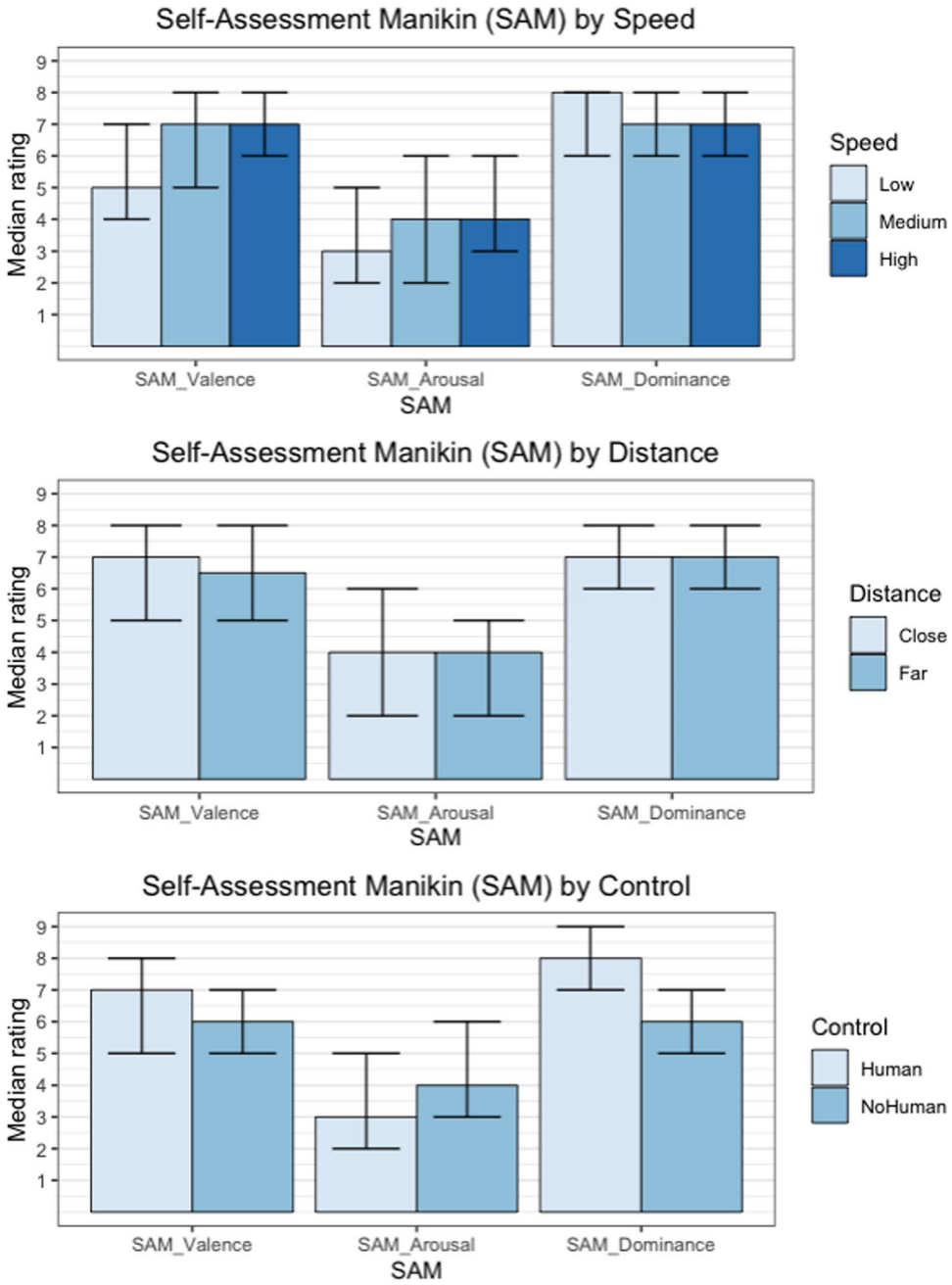


Fig. 15 Median ratings with interquartile range of SAM dimensions for the configuration factors *Speed*, *Distance*, *Control*

the SAM are presented. Figure 15 provides an overall graphical comparison between the different levels of the configuration factors *Speed*, *Distance*, and *Control* for each SAM dimension. Additionally, significant differences in response variables by fixed factors is also highlighted through the Wilcoxon signed-rank test. Also in this case, Bonferroni's post-hoc correction has been applied for the *Speed* factor.

Regarding *Valence* dimension, Fig. 16 shows how the configuration factors influenced the response distribution:

- A low robot movement speed led to more unpleasant emotions compared to the other two *Speed* levels. The Wilcoxon signed-rank test with Bonferroni correction confirmed a highly significant difference between *Low* and *Medium* speeds ($p < 0.001$) as well as between *Low* and *High* speeds ($p < 0.001$). The difference between *Medium* and *High* speeds was not significant ($p = 1$).
- When the robot workspace was farer from the participant, slightly less positive emotions were perceived. However, the Wilcoxon signed-rank test highlighted that the difference was not significant ($p = 0.879$).
- Configurations lacking time execution control by participant were associated with less pleasant emotions (see also Fig. 15). Moreover, the Wilcoxon signed-rank test highlighted a significant difference between the two control configurations ($p < 0.001$).

The influence of the configuration factors on *Arousal* can be observed in Fig. 17:

- Increasing robot speed had a general positive effect, meaning that participants got more aroused as the robot speed increased. This result is confirmed by the Wilcoxon signed-rank test with Bonferroni correction, where all the pairwise comparisons led to significant differences (*Low-Medium*: $p < 0.001$; *Low-High*: $p < 0.001$; *Medium-High*: $p = 0.004$).
- The distance of the robot workspace from the operator seemed to not have a significant influence on par-

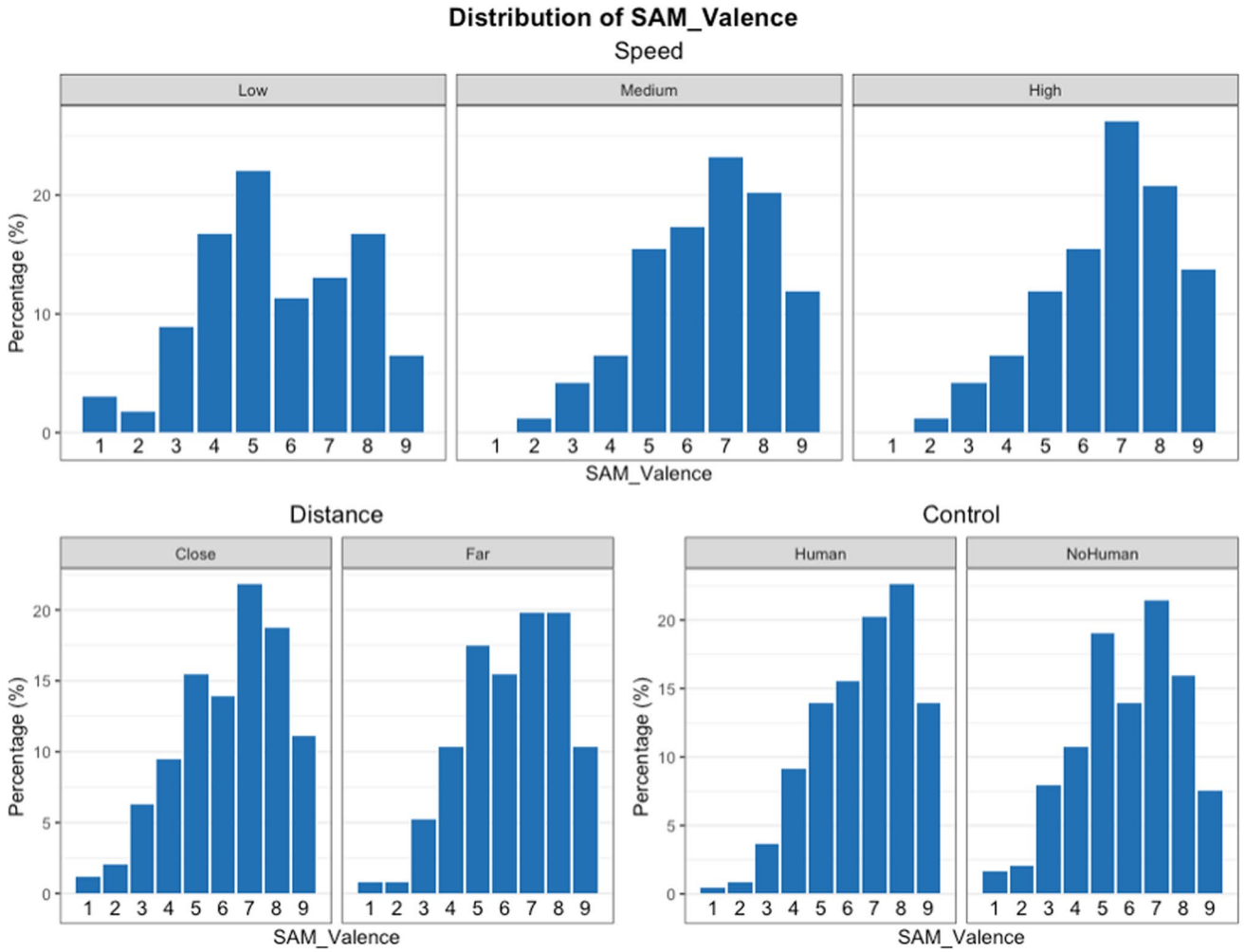


Fig. 16 Distribution of *Valence* for the configuration factors *Speed*, *Distance*, *Control*.

ticipants arousal. The Wilcoxon signed-rank test highlighted indeed a not significant difference between the two *Distance* levels ($p = 0.711$).
– The absence of time execution control by the participant was associated with an increase in participants arousal. This effect resulted significant according to the Wilcoxon signed-rank test ($p < 0.001$).

Figure 18 shows data distribution of *Dominance* for the configuration factors, and it can be observed that:

– Increasing the speed of the robot resulted in a slight lowering of the participant's sense of dominance. The Wilcoxon signed-rank test with Bonferroni correction highlighted a significant difference between *Low* and *Medium* speeds ($p = 0.039$), however not between *Low* and *High* speeds ($p = 0.180$) or between *Medium* and *High* speeds ($p = 1$).
– Even a slight degradation in the sense of dominance can be observed for the *Close* setting, the factor *Distance* seemed to not be particularly influential, and this was confirmed by the Wilcoxon signed-rank test ($p = 0.449$).
– In configurations without time execution control by participant, the sense of dominance was considerably lower (see also Fig. 15). The Wilcoxon signed-rank test confirmed a highly significant difference between the two *Control* levels ($p < 0.001$).

4.4 Physiological Responses

In this sub-section, the relationships between the experimental factors and physiological responses are presented. Figure 19 shows the distribution of the average SCR data for the configuration factors, while Fig. 20 allows to better compare with boxplots the response differences across the levels of each configuration factor. As can be observed, the data distribution has a high skewness and the hypothesis of the

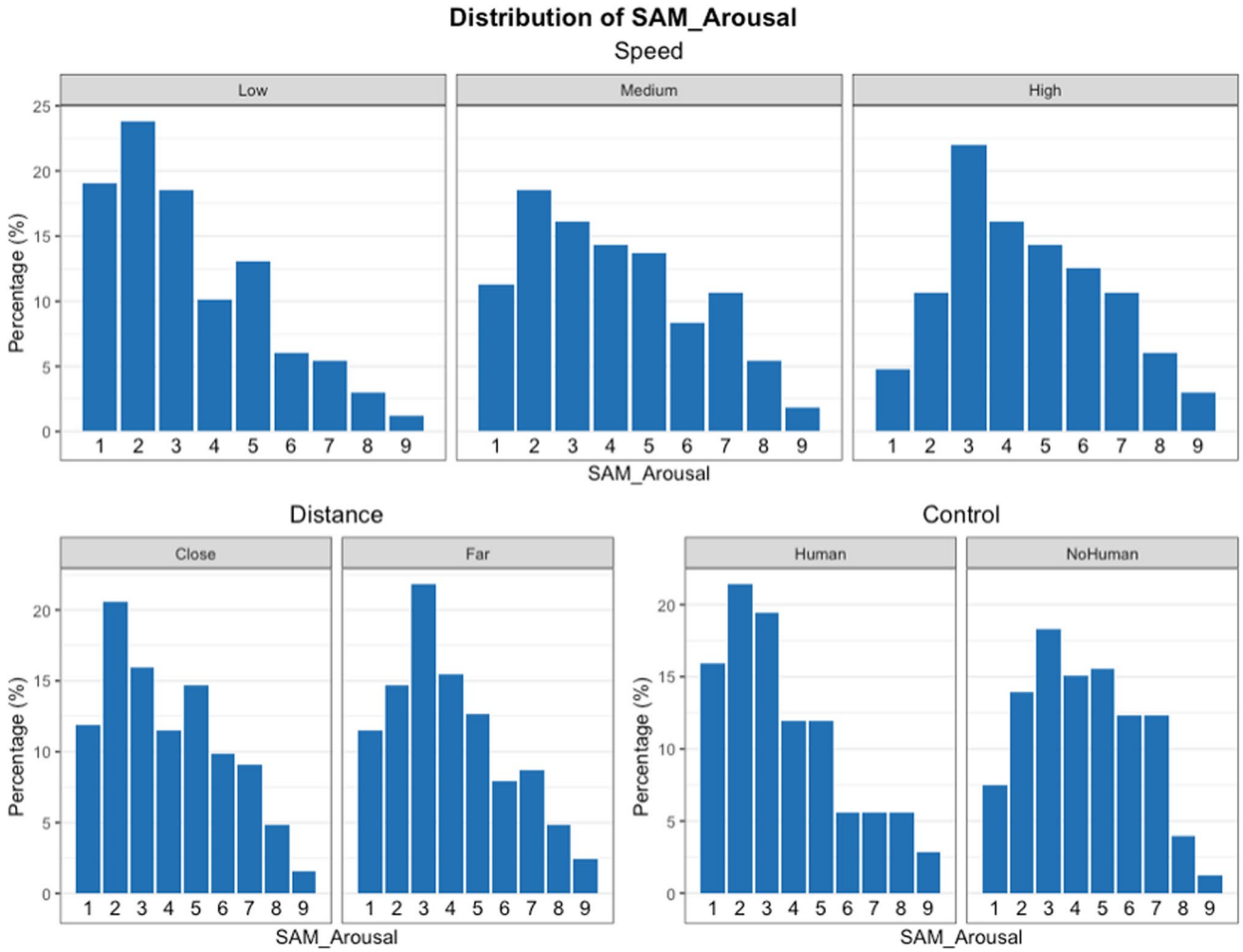


Fig. 17 Distribution of *Arousal* for the configuration factors *Speed*, *Distance*, *Control*

average SCR following a normal distribution was rejected by a Shapiro-Wilk test [51] ($p < 0.001$). It can be observed that:

- *Low* robot movement speed led to slightly lower average SCR, leading to slightly less mental strain. According to the Wilcoxon signed-rank test with Bonferroni correction, this difference resulted significant with respect to the *Medium* speed ($p = 0.039$), but not when compared with the *High* one ($p = 0.180$). The difference between *Medium* and *High* was also not significant ($p = 1$).
- The distance of the robot workspace from the operator seemed to not have a significant influence on the average SCR. This was also confirmed by the Wilcoxon signed-rank test ($p = 0.816$).
- Configurations without task execution control tended to generate slightly more stressful situations for the participant, leading to a slight increase in the average SCR. However, the Wilcoxon signed-rank test revealed an almost significant difference between the two settings ($p = 0.071$).

With respect to HRV, Fig. 21 provides the distribution of the RMSSD data for the configuration factors, while Fig. 22 shows boxplot comparisons of the response across the levels of each configuration factor. As can be noted, the data distribution has a high skewness and the hypothesis of the RMSSD following a normal distribution was rejected by a Shapiro-Wilk test ($p < 0.001$). It can be noted that:

- As the robot movement speed increased, a slight decrease in RMSSD was observed. This means that slightly more stressful situations happened with higher robot speed. However, the Wilcoxon signed-rank test with Bonferroni correction revealed that this effect was not significant for any pairwise comparison (*Low-*

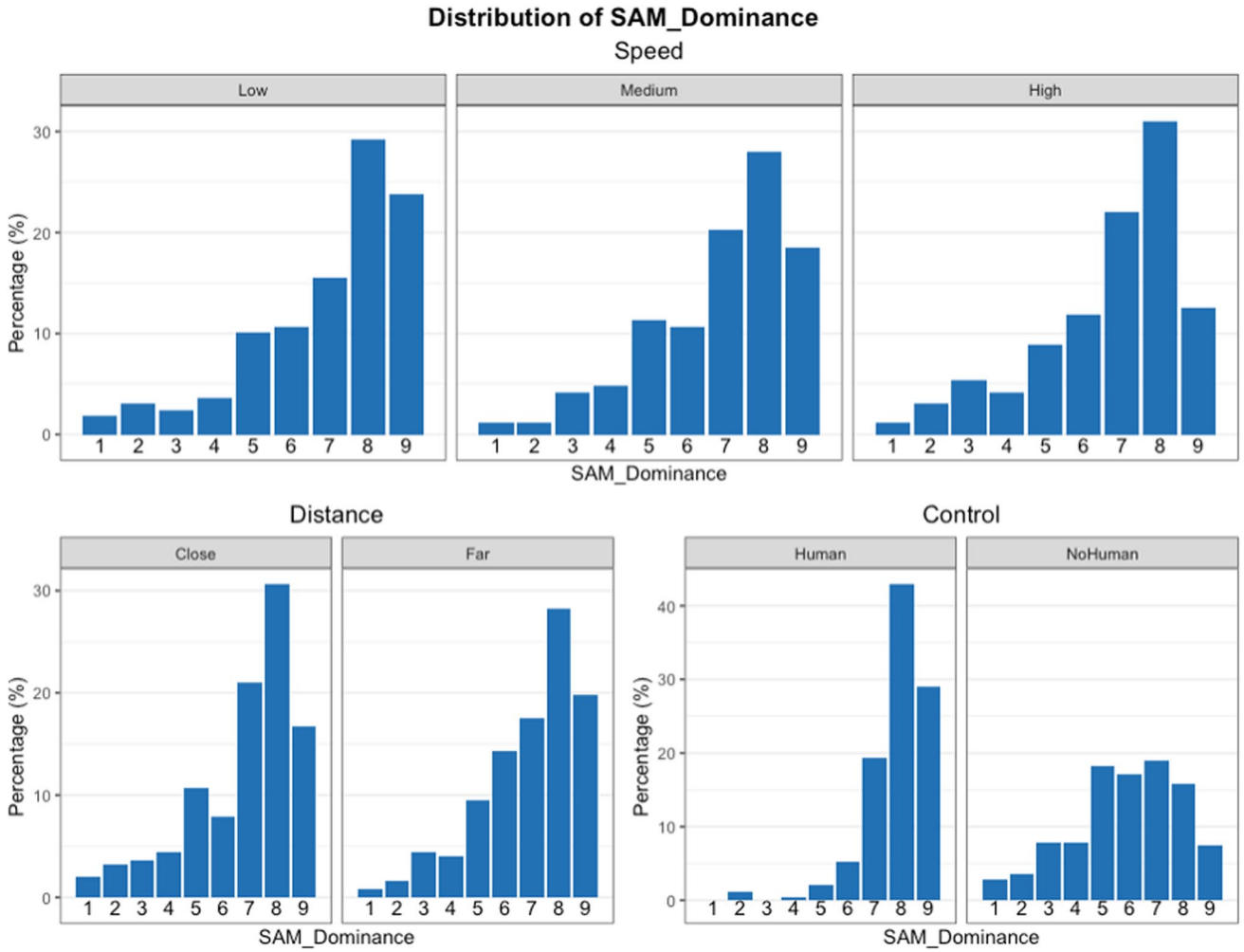


Fig. 18 Distribution of *Dominance* for the configuration factors *Speed*, *Distance*, *Control*

Medium: $p = 1$; *Low-High*: $p = 1$; *Medium-High*: $p = 1$).

- Increasing the distance between the robot workspace and the participant increased the *RMSSD*, leading to potentially slightly less stressful situations. However, the Wilcoxon signed-rank test showed that this difference was not statistically significant ($p = 0.897$).
- The absence of time execution control by the participant tended to generate more stressful situations, leading to a slight decrease of *RMSSD*. However, the Wilcoxon signed-rank test revealed that this effect was not statistically significant ($p = 0.769$).

4.5 Exploration of Two-Way Interactions

This section presents a preliminary and exploratory analysis of potential factor interactions on response variables.

Figure 23 shows interaction plots for interaction quality. Some interaction effects between robot movement speed (*Speed*) and distance from robot workspace (*Distance*) are present in perceived robot helpfulness (*Q1_Helpful*), team fluency (*Q5_Fluid*), and discomfort (*Q6_Uncomfortable*). For low movement speeds, the robot was perceived to be more helpful when working closer to the operator. For medium speeds, the interaction was perceived to be slightly smoother when the workspace was further from the operator. In sessions with high speeds greater proximity to the robot workspace was associated with a slight increase in discomfort. Some interaction effects between *Speed* and *Control* are present in perceived robot helpfulness (*Q1_Helpful*), interaction unsafety (*Q2_NotSafe*), naturalness (*Q3_Natural*), team efficiency (*Q4_Efficient*), fluency (*Q5_Fluid*), and discomfort (*Q6_Uncomfortable*). For low movement speeds the robot was perceived to be less helpful when the participant was not in control of execution time. However, when the execution time control was absent, perceived interaction

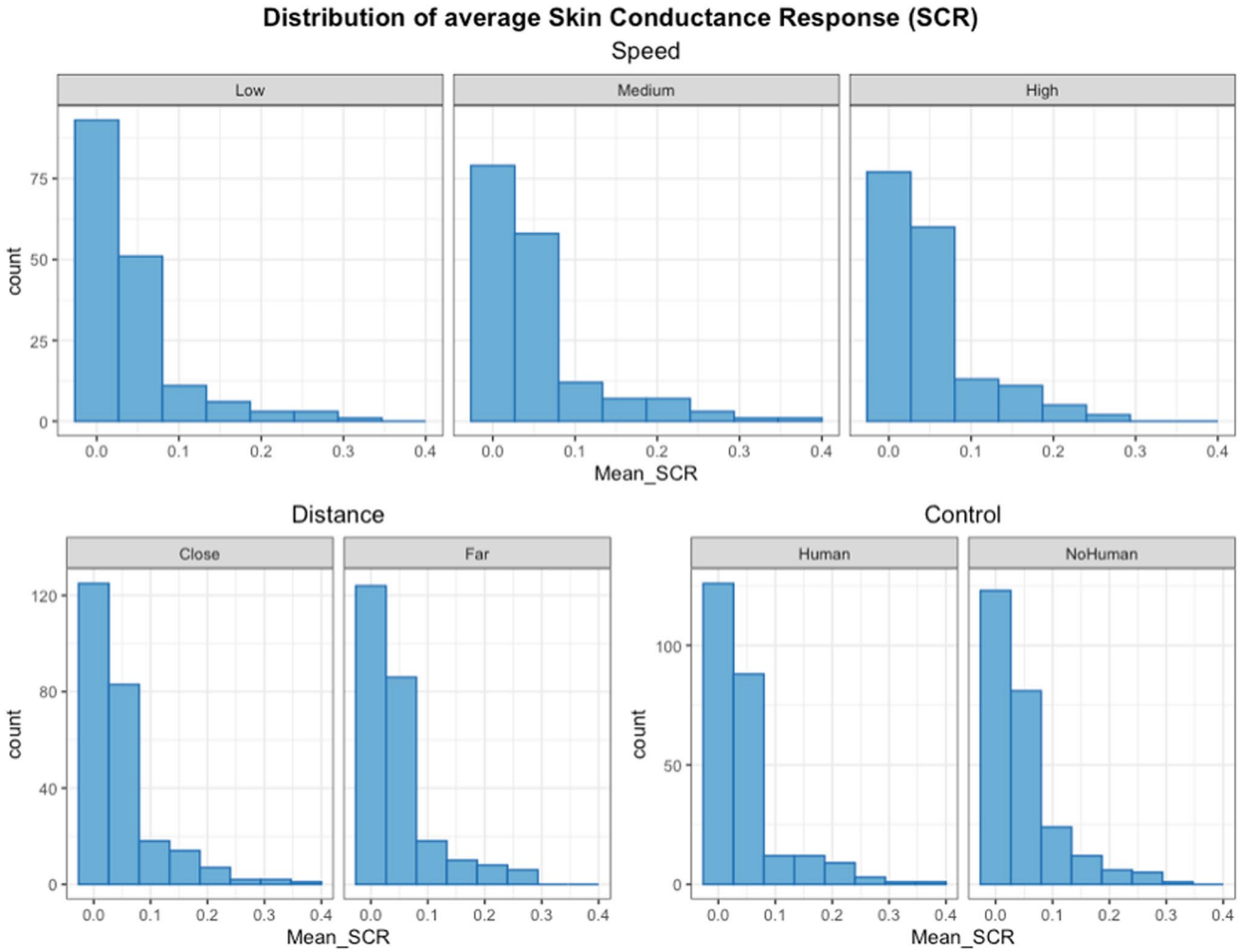


Fig. 19 Distribution of the average SCR (*Mean_SCR*) by the configuration factors *Speed*, *Distance*, *Control*

unsafety and discomfort were greater for medium and high speeds. Moreover, an improvement in perceived interaction naturalness, team efficiency and fluency were observed for medium and high speeds when the participant had execution time control. No interaction effects between *Control* and *Distance* are present in the dimensions of interaction quality, except in perceived robot helpfulness (*Q1_Helpful*). When the robot workspace was more distant from the participant, the absence of time execution control the robot implied a slight decrease in perceived helpfulness of the robot.

Interaction plots for SAM dimensions are reported in Fig. 24. Some interaction effects between *Speed* and *Distance* can be observed in *Valence* and *Arousal*. When the robot workspace was closer to the participant, low speeds led to slightly more positive emotions and lower arousal, attributable to a slightly greater sense of relaxation. Interaction effects between *Speed* and *Control* are present in *Valence* and *Arousal*. In general, as *Speed* increased, an increase in *Valence* and *Arousal* can be noticed. When the participant was in control of execution time, more positive emotions were perceived with medium robot speeds. The absence of execution time control caused a general increase of arousal, but for low speeds a smaller difference can be noticed. No interaction effects between *Control* and *Distance* can be observed in the SAM dimensions.

Figure 25 contains the interaction plots for physiological responses. Some interaction effects between *Speed* and *Distance* can be observed in both the HRV indicator *RMSSD* and the average SCRs (*Mean_SCR*). When the robot workspace was closer to the participant, a slight decrease in HRV was observed for medium and high speeds, potentially leading to slightly increased stress. This effect was also observed in the average SCRs for medium speeds, where a slight increase was present. When the speed was low, HRV was slightly higher than in configurations where the robot workspace was further from the participant, leading to potentially less stressful situations. However, this effect does not seem to be confirmed by the average SCRs, as it was slightly higher in these configurations.

Some interaction effects between *Speed* and *Control* are present in *RMSSD* and *Mean_SCR*. When the participant was not in control of execution time, a slight increase of average SCRs was observed for low robot speeds, leading to potentially more stressful situations. However, this effect is not confirmed by the HRV.

Some small interaction effects between *Distance* and *Control* can be observed in *RMSSD* and *Mean_SCR*). When the participant was closer to robot workspace, the absence of execution time control caused a slight decrease in HRV and an increase in mean SCRs, thus leading to slightly more stressful situations.

4.6 Exploration of Time (*Trial*) Effect

In this subsection, a preliminary analysis of the effect of time (i.e., the progression of experimental trials) on response variables is presented.

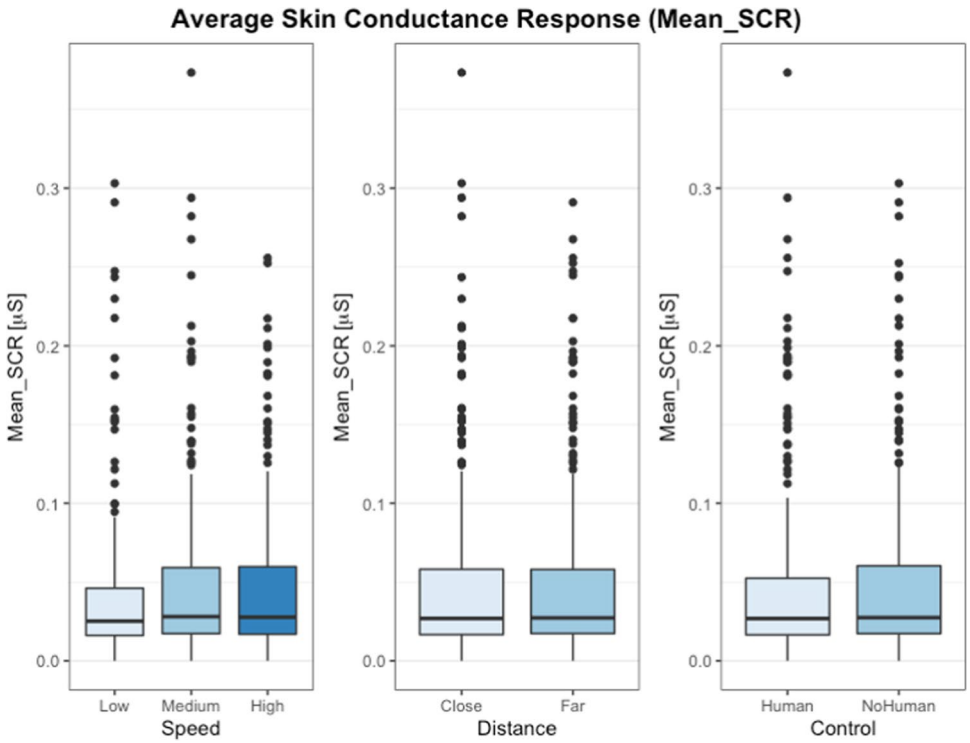


Fig. 20 Boxplot comparison of the average SCR (*Mean_SCR*) for each configuration factor.

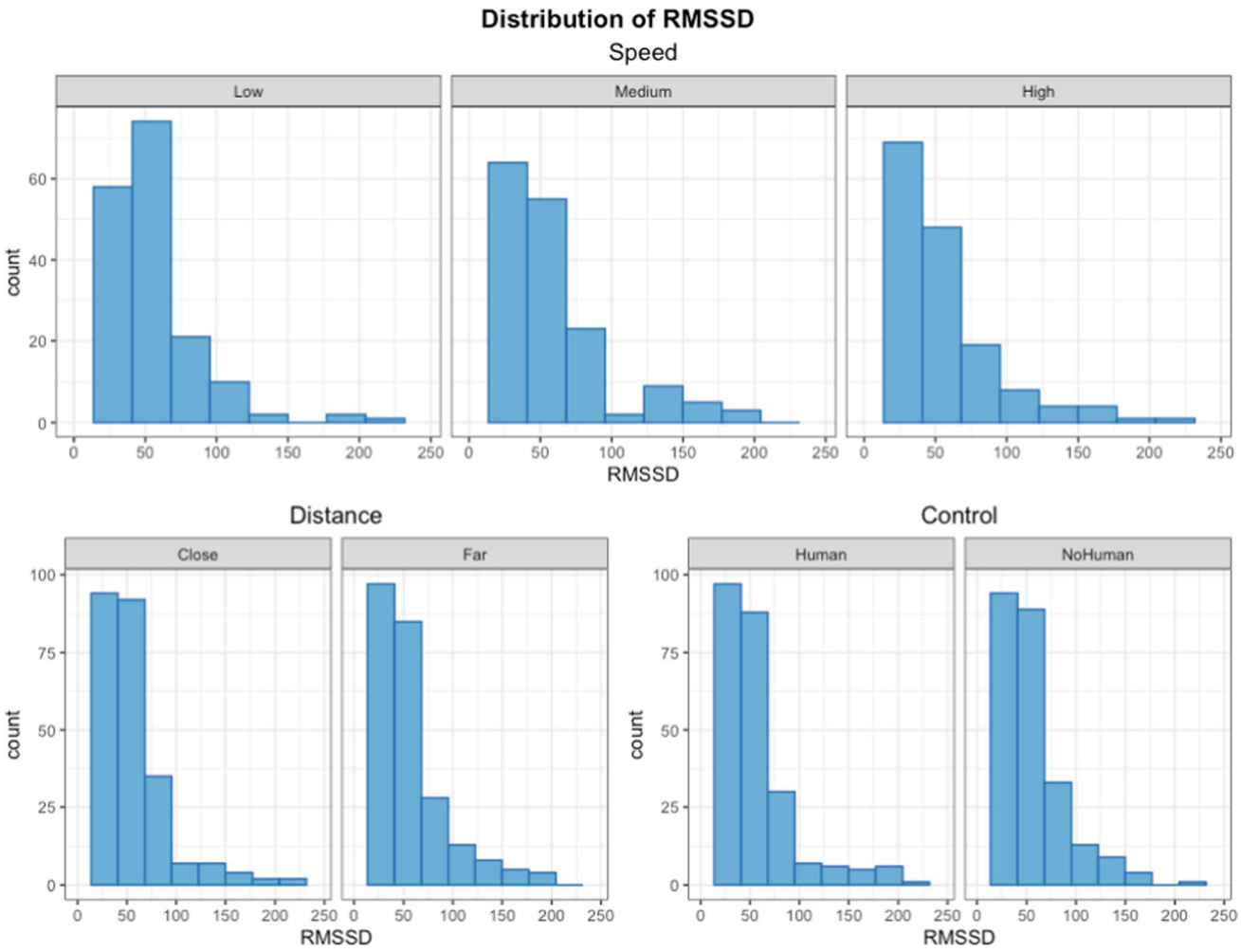


Fig. 21 Distribution of the HRV indicator *RMSSD* by the configuration factors *Speed*, *Distance*, *Control*

Figure 26 shows the evolution of the interaction quality dimensions with respect to the progression of the experiment (*Trial*). As the experiment progressed through the trials, the perception of various aspects related to interaction quality tended to improve. It can be noticed that there is a generally positive effect in perceived robot helpfulness (*Q1_Helpful*), interaction naturalness (*Q3_Natural*), team efficiency (*Q4_Efficient*), fluency (*Q5_Fluid*), and robot trustworthiness (*Q7_Trustworthy*). Moreover, a general decrease in interaction unsafety (*Q2_NotSafe*) and discomfort (*Q6_Uncomfortable*) can be observed.

In Fig. 27 the evolution of affective state collected through the SAM is presented. It can be seen that as the experiment progressed, the sense of control and relaxation generally increased. In fact, *Arousal* gradually decreased, while *Dominance* gradually increased. For *Valence*, the effect is not clear. However, a slight negative trend can be noticed, which can be attributed to a potential increasing sense of boredom.

Figure 28 shows the evolution of physiological response with respect to *Trial*. A progressive decrease in the average SCRs (*Mean_SCR*) and its dispersion can be observed. This phenomenon can be interpreted as a gradual decrease in stress level probably due to the task-related habituation effect. Regarding HRV, no particular trend can be noticed for the RMSSD.

5 Discussion

The analysis revealed some interesting relationships among the configuration factors and both subjective and objective responses. Among the configuration factors, *Speed* and *Control* were found to be most significant overall.

Several significant differences emerged between *Low* and higher robot movement speed levels (Fig. 7). *Low* speed was associated with lower perceived robot helpfulness,

team efficiency, interaction naturalness, and team fluidity. However, a significant increase in the perceived safety and comfort was also observed compared to *High* speed. Interestingly, no significant differences in the dimensions of interaction quality emerged between the *Medium* and *High* levels. The *Low* robot movement speed was also associated with lower valence, implying more negative emotions such as boredom (Fig. 15). In addition, as the speed of the robot increased, there was a significant increase in participants arousal. This effect was also partially detected physiologically, with an increase of average SCRs (*Mean_SCR*). This result is in agreement with findings form previous studies, such as Kulić and Croft [13, 52] and Arai et al. [11]. In addition, unstructured feedback from participants confirmed these findings, showing a preference for *Medium* and *High* robot movement speeds. *Low* speed led participants to become bored and distracted more easily. From a task design perspective, a low speed may be useful in the beginning to familiarize with robot movements. However, once the operator is familiar with the task, higher robot speeds can improve efficiency and operator involvement.

The lack of control of the task execution time had a significant negative effect on the interaction quality (Fig. 7) and participants' affective state (Fig. 15). In this kind of configurations, the robot was perceived by participants less helpful and trustworthy. Moreover, the interaction was felt significantly less natural, efficient, safe, fluid, and comfortable. Regarding the participants' affective state, control of the task execution time was associated with higher levels of valence and dominance and with a lower level of arousal. This means that participants were more relaxed during this kind of configuration, thus leading to less stressful situations for the participants. From a physiological point of view, an almost significant increase in average SCRs was observed in configurations without control of the task execution time. This result is in line with the findings of Arai et al. [11], in which it was pointed out that the absence of information regarding the robot motion speed generated more stress in the operator. Participants' unstructured feedback supported these findings, showing a preference for having control of the task execution time. When they did not have control of task timing, participants tended to experience more pressure and anxiety. However, some participants who finished their operation well ahead of schedule had to wait for the robot to continue, and this also negatively impacted the experience. From a task design perspective, it is important for the operator to have control of the task execution time to avoid possible process defects or idle time. This form of control, however, does not exclude the possibility for the robot to make autonomous decisions. For instance, the introduction of a robot adaptivity system based on the state of advancement of the operations of the human would allow to reach a more advanced form of HRC, while maintaining the operator in control of the task.

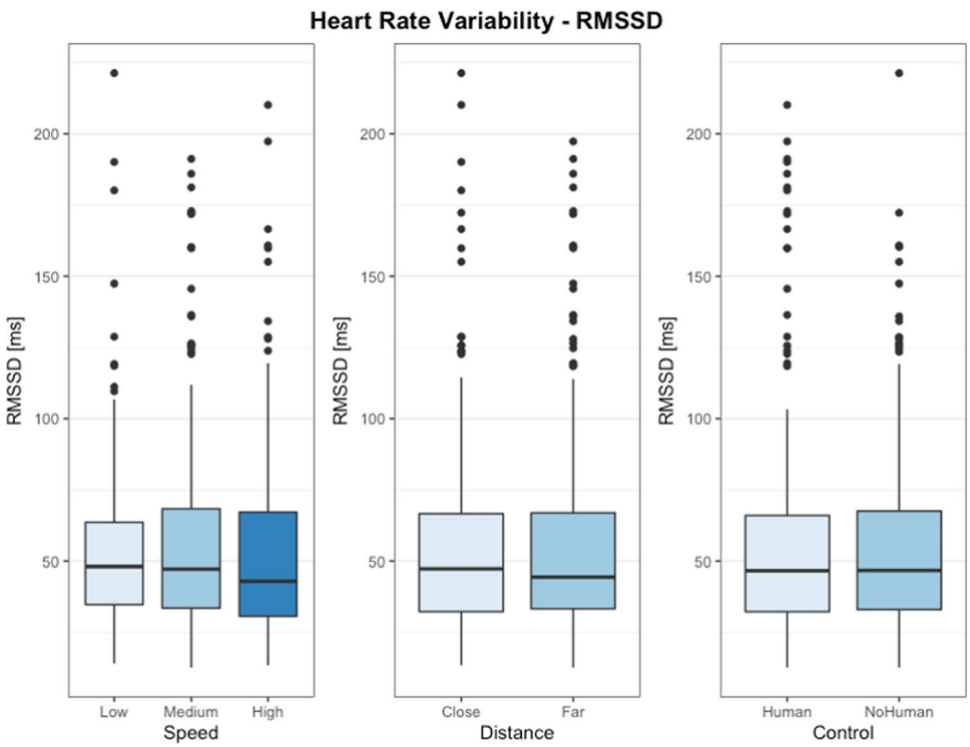


Fig. 22 Boxplot comparison of the HRV indicator *RMSSD* for each configuration factor

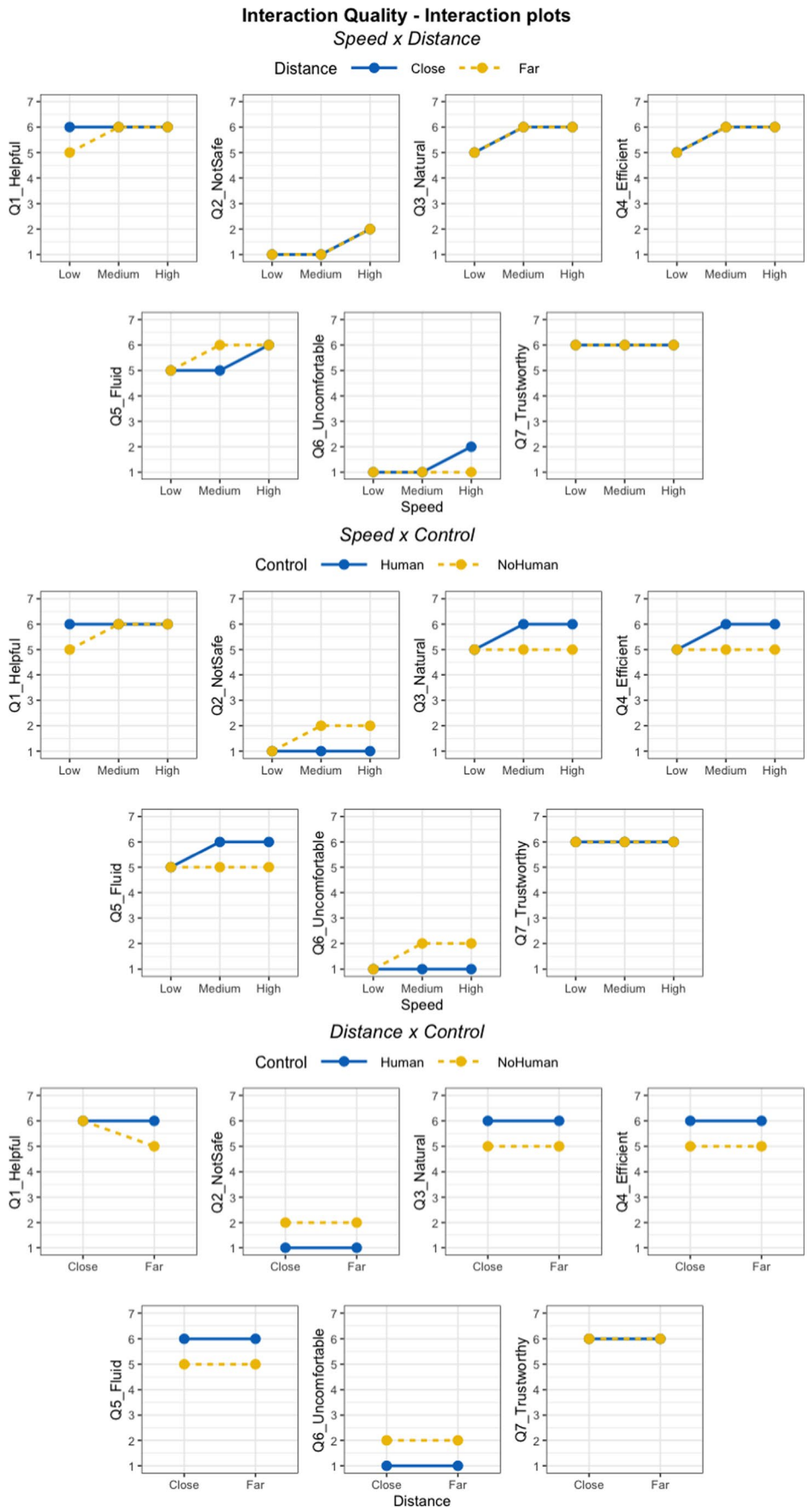


Fig. 23 Interaction median plots of interaction quality dimensions for the configuration factors *Speed*, *Distance*, *Control*.

The distance between the operator and the robot workspace was generally found to be not particularly influential (Figs. 7 and 15). However, a significant slight increase in robot trustworthiness was observed with a higher distance. Indeed, from the unstructured feedback, it was apparent that the distance factor was mostly a matter of preference and that, for the purposes of the collaborative task, it did not particularly affect the experience. This is partially in contrast to

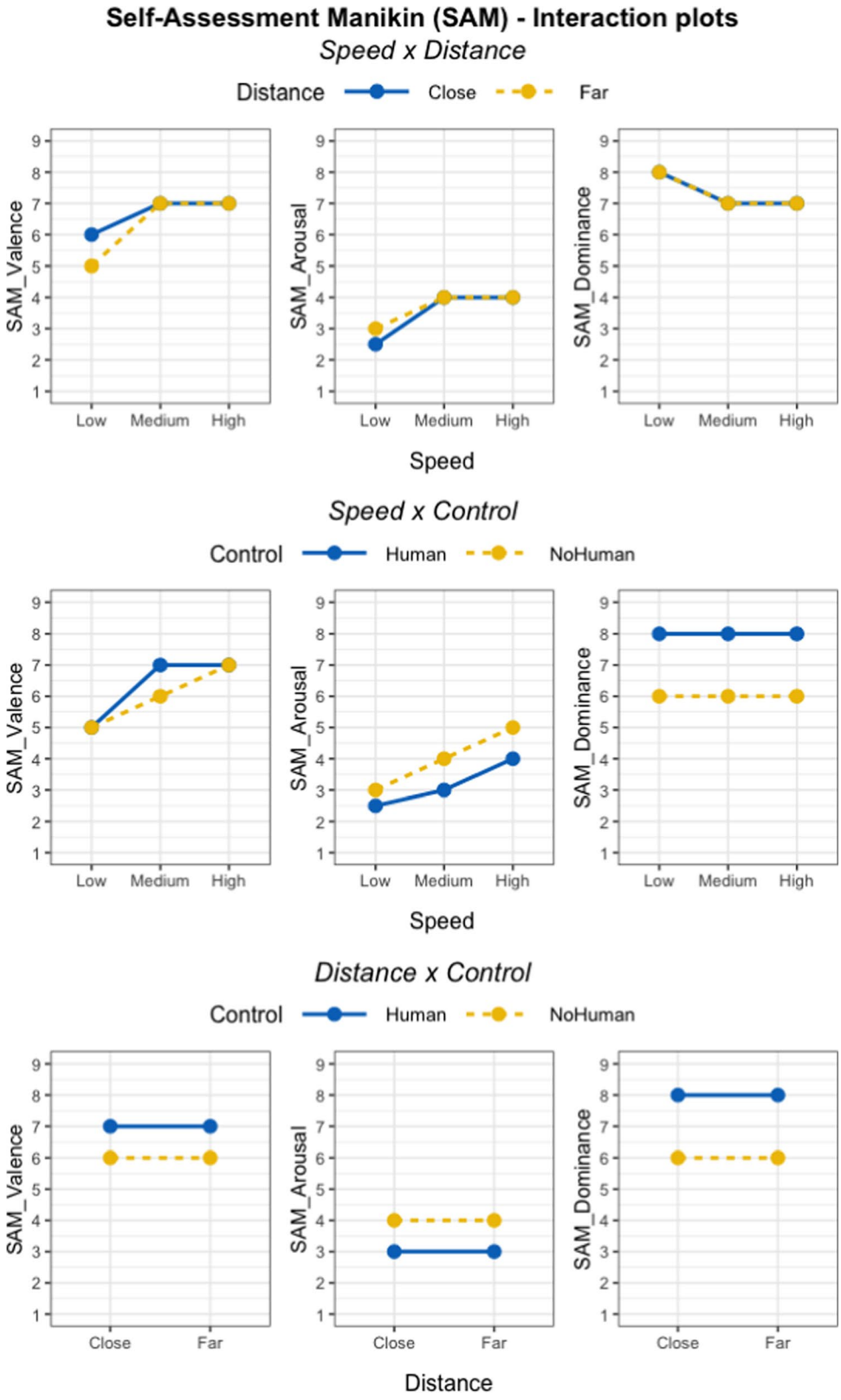


Fig. 24 Interaction median plots of SAM dimensions for the configuration factors *Speed*, *Distance*, *Control*

what was shown in Arai et al. [11], where greater proximity of the operator to the moving robot induced greater stress. It should be noted, however, that the size and type of robot implemented were different from those in the present study. In fact, the low significance of *Distance* on many response variables may also be due to the rather small size of the UR3e robot. Future studies will be needed to confirm this hypothesis.

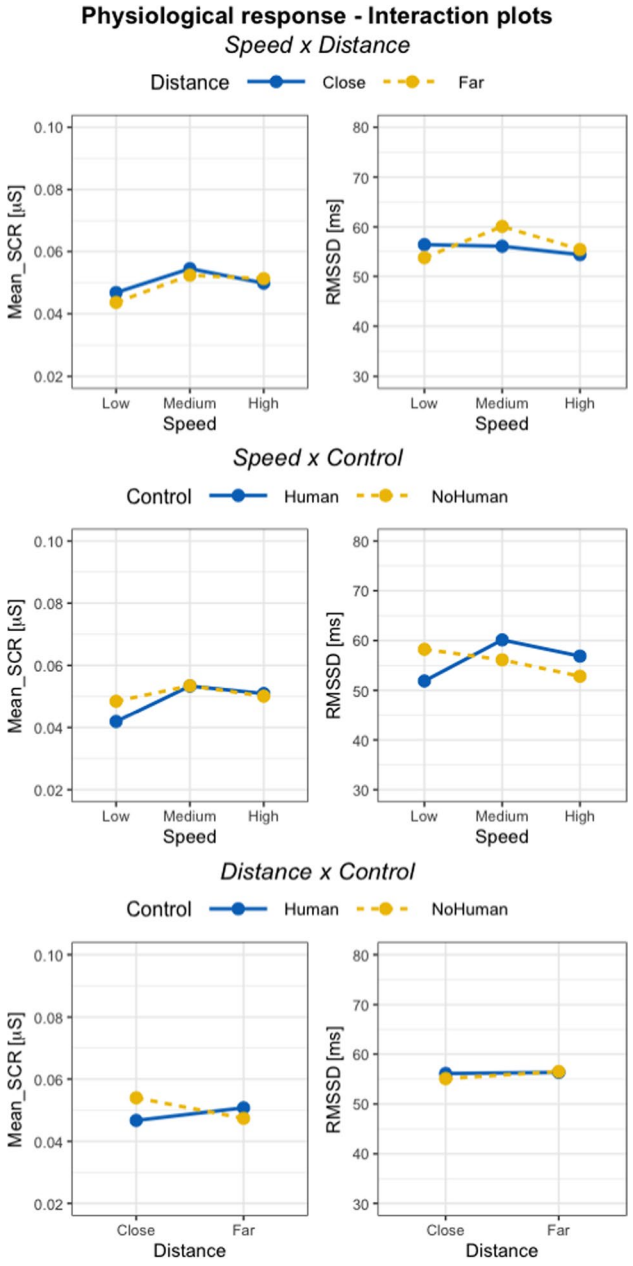


Fig. 25 Interaction mean plots of physiological response for the configuration factors *Speed*, *Distance*, *Control*

A qualitative analysis on factor interactions revealed some potential effects on response variables. Some slight interaction effects between *Speed* and *Distance* emerged in perceived robot helpfulness, interaction fluency, discomfort, valence, and arousal. For *Low* speed, a general preference for configurations where the operator is closer to the robot workspace emerged. However, for higher speeds, greater discomfort was felt with greater proximity. Physiologically, no clear effect on stress emerged. Interaction effects between *Speed* and *Control* revealed that with the absence of control of task execution time, increased speed had an overall negative effect on user experience, with a degradation of most dimensions of perceived interaction quality. For the interaction between *Distance* and *Control*, a slight effect on physiological response was noted. Specifically, when the robot workspace was closer to the operator, slightly more stress was observed in configurations where the participant was not in control of the execution time. These preliminary results provide a starting point for further investigation through in-depth quantitative analysis in future work.

Experimental outcomes revealed potential effects of time (*Trial*) on response variables. As the experimental trials progressed, a gradual improvement in user experience was observed. Various dimensions of interaction quality gradually improved, with also an increasing feeling of control and relaxation by participants. This phenomenon indicates an influence of the habituation process, which will be further investigated in future work through in-depth quantitative analysis.

Overall, the need for customization of the HRC was hinted from the participant comments. Technology that can adapt to the user's needs and preferences can provide a significant ergonomic benefit, both physically and psycho-cognitively. Further investigation on this front is needed, also exploring possible relationships with operator characteristics.

6 Conclusions

The aim of this research was to present a preliminary study focused on user experience in industrial HRC, by exploring the relationships among the perceived interaction quality, operator's affective state, physiological response and several HRC setting factors.

The study results revealed an important influence of the robot movement speed and the participant's control of the task execution time, but not of the distance between the operator and the robot workspace.

Medium and *High* speeds were found to be the most appreciated. Although most participants initially appreciated slower speeds, once they understood the trajectory of the robot's movements they wanted higher speeds to be more efficient and feel more engaged. This fact highlights that when the trajectories of the robot are not known to the operator, it is preferable to maintain slower speeds. However, higher robot speeds can be preferable to maintain engagement once the operator feels confident with the task.

Configurations without human control of execution time were less appreciated and tended to generate more anxiety. However, interestingly, some participants appreciated the absence of direct control as not having to press the button represented one less operation for them. This fact seems to

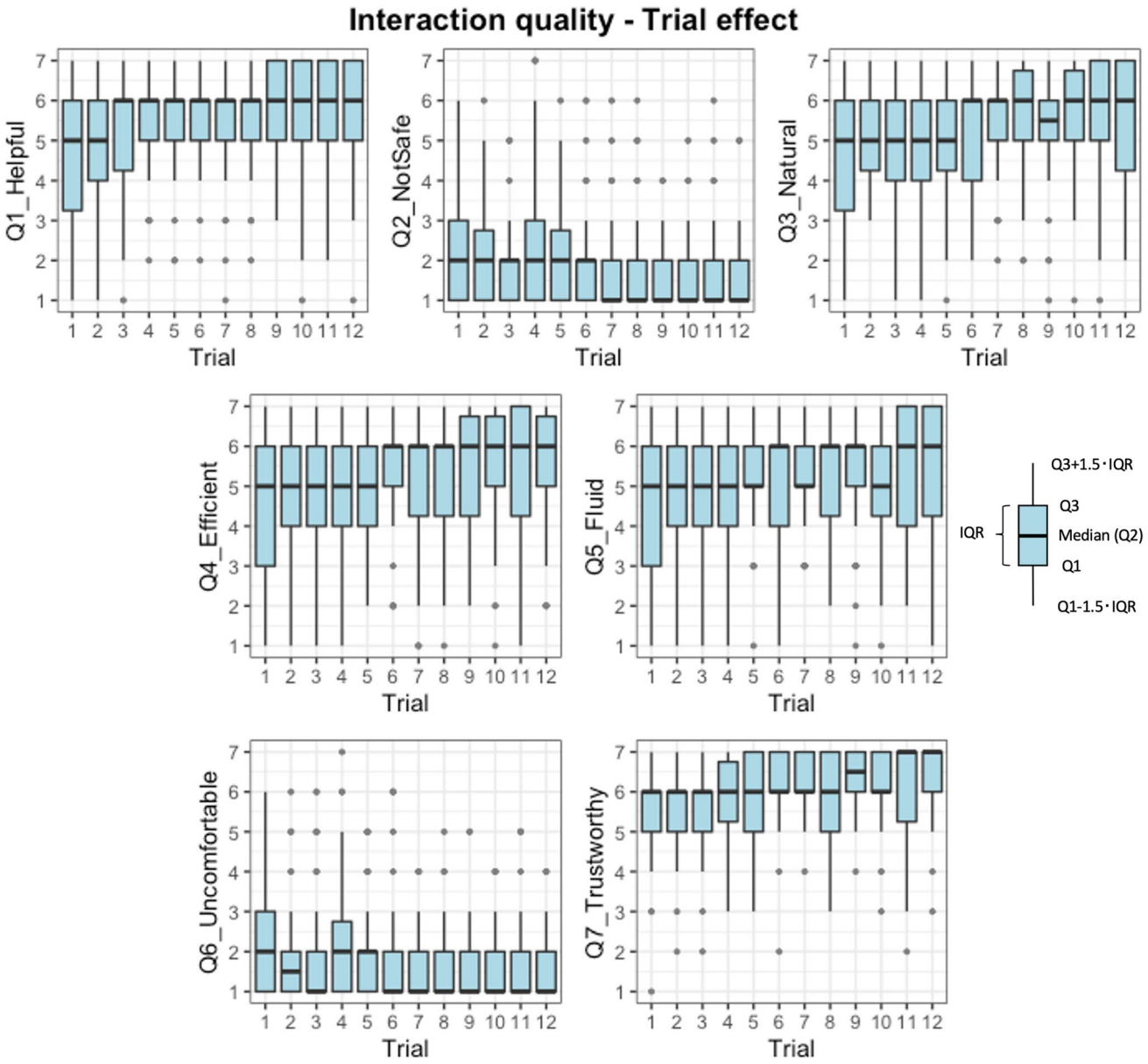


Fig. 26 Effect of time (*Trial*) on interaction quality dimensions

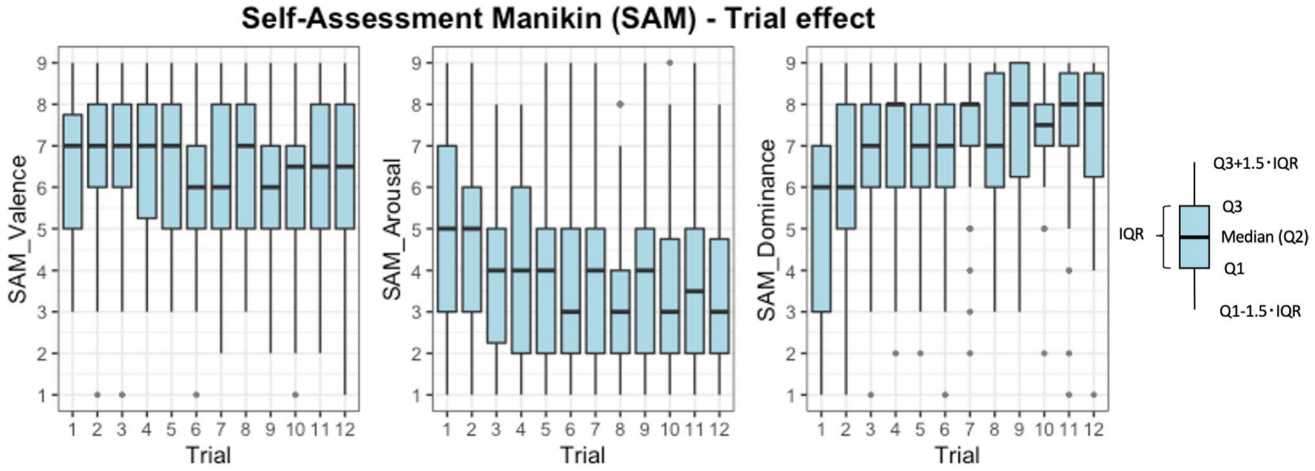
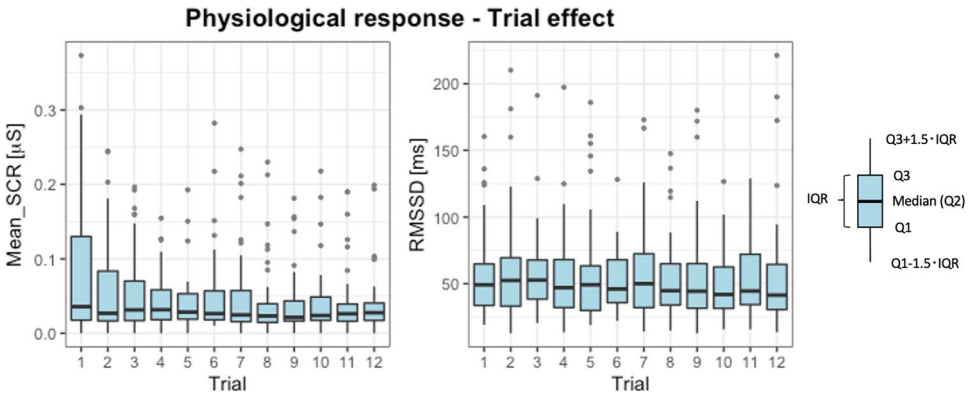


Fig. 27 Effect of time (*Trial*) on SAM dimensions

Fig. 28 Effect of time (*Trial*) on physiological response



presage that a system capable of automatically recognizing when the operator has finished their operations may be preferable to a traditional control system. In fact, such a modality would allow to approach the natural human-human collaboration.

The distance between the operator and the robot workspace turned out to be not particularly influential and from participant feedback it seemed to be simply a matter of preference. This fact may be due to the rather small size of the UR3e robot, however this hypothesis needs more investigation.

The introduction of physiological measures to analyze user experience in the HRC allowed for a better understanding of participant reactions, while also capturing unconscious ones. Through the analysis of HRV and EDA, physiological responses provided some initial insights, partially confirming the self-reported affective state and the unstructured feedback of participants. More studies on the implementation of non-invasive biosensors will be necessary and useful to achieve a profitable and human-centered HRC.

Future work will provide a deeper and quantitative analysis of the HRC configuration factors, also exploring their interactions. Moreover, further investigation on the relationship between affective state and physiological response in HRC will be necessary. Unstructured interviews with participants revealed different opinions about the preferred HRC configuration, which may hint to the need for a more personalized HRC. For this reason, the effects of the habituation process and participant's characteristics will be deeply explored in following studies.

Author Contributions All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by R. Gervasi and K. Aliev. The first draft of the manuscript was written by R. Gervasi with contribution of K. Aliev under the supervision of L. Mastrogiacomo and F. Franceschini. All authors read and approved the final manuscript.

Funding This work has been partially supported by "Ministero dell'Istruzione, dell'Università e della Ricerca" Award "TESUN-83486178370409 finanziamento dipartimenti di eccellenza CAP. 1694 TIT. 232 ART. 6".

Code Availability The datasets generated and analysed during this study are not currently publicly available.

Declarations

Ethics approval The authors respect the Ethical Guidelines of the Journal.

Consent for Publication Not applicable.

Conflict of Interests The authors declare that they have no conflict of interest.

Consent to participate Informed consent was obtained from all individual participants included in the study.

References

1. ISO/TS 15066:2016: Robots and robotic devices – Collaborative robots. Standard ISO/TS 15066:2016, International Organization for Standardization, Geneva, CH. https://www.iso.org/standard/62996.html (2016)
2. Mishra, R.: Confirmation of a measurement model for manufacturing flexibility development practices. Int. J. Qual. Reliab. Manag. **38**(1), 317–338 (2020). https://doi.org/10.1108/IJQRM-01-2019-0027
3. Rabbani, M., Behbahan, S.Z.B., Farrokhi-Asl, H.: The collaboration of human-robot in mixed-model four-sided assembly line balancing problem. J. Intell. Robot. Syst. https://doi.org/10.1007/s10846-020-01177-1 (2020)
4. Vicentini, F.: Collaborative robotics: A survey. J. Mech. Des. **143**(040802). https://doi.org/10.1115/1.4046238(2020)
5. Gervasi, R., Mastrogiacomo, L., Franceschini, F.: A conceptual framework to evaluate human-robot collaboration. Int. J. Adv. Manuf. Technol. **108**(3), 841–865 (2020). https://doi.org/10.1007/s00170-020-05363-1
6. Gervasi, R., Mastrogiacomo, L., Maisano, D.A., Antonelli, D., Franceschini, F.: A structured methodology to support human–robot collaboration configuration choice. Prod. Eng. **16**(4), 435–451 (2022). https://doi.org/10.1007/s11740-021-01088-6
7. McColl, D., Hong, A., Hatakeyama, N., Nejat, G., Benhabib, B.: A survey of autonomous human affect detection methods for social robots engaged in natural HRI. J. Intell. Robot. Syst. **82**(1), 101–133 (2016). https://doi.org/10.1007/s10846-015-0259-2
8. Xu, D., Wu, X., Chen, Y.L., Xu, Y.: Online dynamic gesture recognition for human robot interaction. J. Intell. Robot. Syst. **77**(3), 583–596 (2015). https://doi.org/10.1007/s10846-014-0039-4
9. Young, M.S., Brookhuis, K.A., Wickens, C.D., Hancock, P.A.: State of science: mental workload in ergonomics. Ergonomics **58**(1), 1–17 (2015). https://doi.org/10.1080/00140139.2014.956151
10. Wang, L., Gao, R., Váncza, J., Krüger, J., Wang, X.V., Makris, S., Chryssolouris, G.: Symbiotic human-robot collaborative assembly. CIRP Ann. **68**(2), 701–726 (2019). https://doi.org/10.1016/j.cirp.2019.05.002
11. Arai, T., Kato, R., Fujita, M.: Assessment of operator stress induced by robot collaboration in assembly. CIRP Ann. **59**(1), 5–8 (2010)
12. Kühnlenz, B., Erhart, M., Kainert, M., Wang, Z.Q., Wilm, J., Kühnlenz, K.: Impact of trajectory profiles on user stress in close human-robot interaction. at - Automatisierungstechnik **66**(6). https://doi.org/10.1515/auto-2018-0004 (2018)
13. Kulić, D., Croft, E.: Physiological and subjective responses to articulated robot motion. Robotica **25**(1), 13–27 (2007). https://doi.org/10.1017/S0263574706002955
14. ISO 10218–1:2011: Robots and robotic devices – Safety requirements for industrial robots – Part 1: Robots. Standard ISO 10218-1:2011, International Organization for Standardization, Geneva, CH. https://www.iso.org/standard/51330.html (2011)
15. ISO 10218–2:2011: Robots and robotic devices – Safety requirements for industrial robots – Part 2: Robot systems and integration. Standard ISO 10218-2:2011, International Organization for Standardization, Geneva, CH. https://www.iso.org/standard/41571.html (2011)
16. Inkulu, A.K., Bahubalendruni, M.R., Dara, A., K., S.: Challenges and opportunities in human robot collaboration context of Industry 4.0 - a state of the art review Industrial Robot: the international journal of robotics research and application. https://doi.org/10.1108/IR-04-2021-0077 (2021)
17. Hollnagel, E.: Cognitive ergonomics: it's all in the mind. Ergonomics **40**(10), 1170–1182 (1997). https://doi.org/10.1080/001401397187685
18. Vink, P.: Advances in Social and Organizational Factors. CRC Press, Boca Raton (2012). https://doi.org/10.1201/b12314
19. Galin, R.R., Meshcheryakov, R.V.: Human-robot interaction efficiency and human-robot collaboration. In: Kravets, A.G. (ed.) Robotics: Industry 4.0 Issues & New Intelligent Control Paradigms, Studies in Systems, Decision and Control, pp 55–63. Springer International Publishing, Cham (2020). https://doi.org/10.1007/978-3-030-37841-7_5
20. Khalid, A., Kirisci, P., Ghrairi, Z., Pannek, J., Thoben, K.D.: Towards implementing safety and security concepts for human-robot collaboration in the context of industry 4.0. In: 39th International MATADOR Conference on Advanced Manufacturing (2017)
21. Wang, L., Liu, S., Liu, H., Wang, X.V.: Overview of human-robot collaboration in manufacturing. In: Wang, L., Majstorovic, V.D., Mourtzis, D., Carpanzano, E., Moroni, G., Galantucci, L.M. (eds.) Proceedings of 5th International Conference on the Industry 4.0 Model for Advanced Manufacturing, Lecture Notes in Mechanical Engineering, pp 15–58. Springer International Publishing, Cham (2020). https://doi.org/10.1007/978-3-030-46212-3_2
22. Mohammed, A., Schmidt, B., Wang, L.: Active collision avoidance for human–robot collaboration driven by vision sensors. Int. J. Comput. Integr. Manuf. **30**(9), 970–980 (2017). https://doi.org/10.1080/0951192X.2016.1268269
23. Liu, Z., Wang, X., Cai, Y., Xu, W., Liu, Q., Zhou, Z., Pham, D.T.: Dynamic risk assessment and active response strategy for industrial human-robot collaboration. Comput. Ind. Eng. **141**(106302), 1–15 (2020). https://doi.org/10.1016/j.cie.2020.106302
24. Roveda, L., Maskani, J., Franceschi, P., Abdi, A., Braghin, F., Molinari Tosatti, L., Pedrocchi, N.: Model-based reinforcement learning variable impedance control for human-robot

collaboration. J. Intell. Robot. Syst. **100**(2), 417–433 (2020). https://doi.org/10.1007/s10846-020-01183-3
25. Joseph, A.J., Kruger, K., Basson, A.H.: An aggregated digital twin solution for human-robot collaboration in industry 4.0 environments. In: Borangiu, T., Trentesaux, D., Leitão, P., Cardin, O., Lamouri, S. (eds.) Service Oriented, Holonic and Multi-Agent Manufacturing Systems for Industry of the Future, Studies in Computational Intelligence, pp 135–147. Springer International Publishing, Cham (2021). https://doi.org/10.1007/978-3-030-69373-2_9
26. Gawron, V.J.: Human Performance, Workload, and Situational Awareness Measures Handbook, 2nd edn. CRC Press, Boca Raton (2008). https://doi.org/10.1201/9781420064506
27. Wickens, C.D.: Multiple resources and mental workload. Hum. Factors **50**(3), 449–455 (2008). https://doi.org/10.1518/001872008X288394
28. Hart, S.G., Staveland, L.E.: Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In: Hancock, P.A., Meshkati, N. (eds.) Advances in Psychology, *Human Mental Workload*, vol. 52, pp 139–183. North-Holland (1988). https://doi.org/10.1016/S0166-4115(08)62386-9
29. Reid, G.B., Nygren, T.E.: The subjective workload assessment technique: a scaling procedure for measuring mental workload. In: Hancock, P.A., Meshkati, N. (eds.) Advances in Psychology, *Human Mental Workload*, vol. 52, pp 185–218, North-Holland (1988). https://doi.org/10.1016/S0166-4115(08)62387-0
30. Franceschini, F., Galetto, M., Maisano, D.: Designing Performance Measurement Systems: Theory and Practice of Key Performance Indicators. Management for Professionals. Springer International Publishing, Cham Switzerland (2019). https://doi.org/10.1007/978-3-030-01192-5
31. Marinescu, A.C., Sharples, S., Ritchie, A.C., Sánchez López, T., McDowell, M., Morvan, H P.: Physiological parameter response to variation of mental workload. Hum. Factors **60**(1), 31–56 (2018). https://doi.org/10.1177/0018720817733101
32. Argyle, E.M., Marinescu, A., Wilson, M.L., Lawson, G., Sharples, S.: Physiological indicators of task demand, fatigue, and cognition in future digital manufacturing environments. Int. J. Hum.-Comput. Stud. **145**, 102522 (2021). https://doi.org/10.1016/j.ijhcs.2020.102522
33. Charles, R.L., Nixon, J.: Measuring mental workload using physiological measures: A systematic review. Appl. Ergon. **74**, 221–232 (2019). https://doi.org/10.1016/j.apergo.2018.08.028
34. Universal Robots: Collaborative robotic automation | Cobots from Universal Robots (2020). https://www.universal-robots.com/
35. Empatica: E4 wristband. https://www.empatica.com/research/e4 (2022)
36. Nomura, T., Kanda, T., Suzuki, T., Kato, K.: Psychology in human-robot communication: an attempt through investigation of negative attitudes and anxiety toward robots. In: RO-MAN 2004. 13th IEEE International Workshop on Robot and Human Interactive Communication (IEEE Catalog No.04TH8759), pp 35–40 (2004). https://doi.org/10.1109/ROMAN.2004.1374726
37. Bradley, M.M., Lang, P.J.: Measuring emotion: The self-assessment manikin and the semantic differential. J. Behav. Ther. Exp. Psychiatry **25**(1), 49–59 (1994). https://doi.org/10.1016/0005-7916(94)90063-9
38. Lang, P. J.: Behavioral treatment and bio-behavioral assessment: Computer applications. In: Sidowski, J.B., Johnson, J.H., Williams, T.A. (eds.) Technology in mental health care delivery systems, pp 119–137. Ablex, Norwood, NJ (1980)
39. Sanders, T.L., MacArthur, K., Volante, W., Hancock, G., MacGillivray, T., Shugars, W., Hancock, P.A.: Trust and prior experience in human-robot interaction. Proc. Hum. Factors Ergon. Soc. Ann. Meet. **61**(1), 1809–1813 (2017). Publisher: SAGE Publications Inc. https://doi.org/10.1177/1541931213601934
40. Schaefer, K.E.: Measuring trust in human robot interactions: development of the "Trust Perception Scale-HRI". In: Mittu, R., Sofge, D., Wagner, A., Lawless, W. (eds.) Robust Intelligence and Trust in Autonomous Systems, pp 191–218. Springer US, Boston, MA (2016). https://doi.org/10.1007/978-1-4899-7668-0_10
41. Witherspoon, E.B., Schunn, C.D., Higashi, R.M., Baehr, E C.: Gender, interest, and prior experience shape opportunities to learn programming in robotics competitions. Int. J. STEM Educ. **3**(1), 18 (2016). https://doi.org/10.1186/s40594-016-0052-1
42. Baraglia, J., Cakmak, M., Nagai, Y., Rao, R., Asada, M.: Initiative in robot assistance during collaborative task execution. In: 2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI), pp 67–74 (2016). https://doi.org/10.1109/HRI.2016.7451735
43. Hoffman, G.: Evaluating fluency in human–robot collaboration. IEEE Trans. Hum.-Mach. Syst. **49**(3), 209–218 (2019). https://doi.org/10.1109/THMS.2019.2904558
44. Ledalab: http://www.ledalab.de/ (2021)
45. Benedek, M., Kaernbach, C.: A continuous measure of phasic electrodermal activity. J. Neurosci. Methods **190**(1), 80–91 (2010). https://doi.org/10.1016/j.jneumeth.2010.04.028
46. Kim, H.G., Cheon, E.J., Bai, D.S., Lee, Y.H., Koo, B.H.: Stress and heart rate variability: A meta-analysis and review of the literature. Psychiatry Inv. **15**(3), 235–245 (2018). https://doi.org/10.30773/pi.2017.08.17
47. Wilcoxon, F.: Individual comparisons by ranking methods. Biom. Bull. **1**(6), 80–83 (1945). https://doi.org/10.2307/3001968
48. Conover, W.J.: Practical Nonparametric Statistics. Wiley, New York (1999)
49. Lee, S., Lee, D.K.: What is the proper way to apply the multiple comparison test? Korean J. Anesthesiol. **71**(5), 353–360 (2018). https://doi.org/10.4097/kja.d.18.00242
50. Schrum, M.L., Johnson, M., Ghuy, M., Gombolay, M.C.: Four years in review: Statistical practices of likert scales in human-robot interaction studies. In: Companion of the 2020 ACM/IEEE International Conference on Human-Robot Interaction, pp 43–52 (2020)
51. Shapiro, S.S., Wilk, M.B.: An analysis of variance test for normality (complete samples)†. Biometrika **52**(3-4), 591–611 (1965). https://doi.org/10.1093/biomet/52.3-4.591
52. Kulić, D., Croft, E. A.: Affective state estimation for human–robot interaction. IEEE Trans. Robot. **23**(5), 991–1000 (2007). https://doi.org/10.1109/TRO.2007.904899

Riccardo Gervasi is a Postdoctoral researcher at the Department of Management and Production Engineering at Politecnico di Torino. His main scientific interests currently concern the areas of Human-Robot Collaboration and Quality Engineering.

Khurshid Aliev is a Postdoctoral researcher at the Department of Management and Production Engineering at Politecnico di Torino. His main scientific interests currently concern the areas of Human-Robot Collaboration and Robotics.

Luca Mastrogiacomo is an Associate Professor at the Department of Management and Production Engineering at Politecnico di Torino. His main scientific interests currently concern the areas of Service Quality and Quality Engineering.

Fiorenzo Franceschini is a Full Professor of Quality Engineering at Politecnico di Torino (Italy) – Department of Management and Production Engineering. He is author or coauthor of 9 books and more than 300 published papers in prestigious scientific journals, and international conference proceedings. His current research interests are in the areas of Quality Engineering, Performance Measurements Systems and Service Quality.