scientific reports

Check for updates

OPEN

Dynamic velocity scaling for industrial collaborative robots: a gaze-driven approach

Matteo Manzardo[1], Federico Fraboni[2], Sofia Morandini[2], Luca Gualtieri[1]✉, Renato Vidoni[1,3]✉ & Luca Pietrantoni[2]

Implementing an effective human-robot seamless interaction (HRSI) is a critical challenge for the industrial robotics community, which requires capturing all human mental processes and turn them into suitable robot's actions. This work presents a method to adjust the robot behavior in real-time considering both human's low-frequency, i.e. cognitive workload, and high-frequency cognitive processes, i.e. visual attention, which are generally overlooked in literature, to optimize safety, productivity and ergonomics. Regarding high-frequency processes, the system monitors the operator's gaze - the closer they look to the robot, the more attentive they are, while attention drops when the robot leaves their field of view, posing potential risks. The speed of the manipulator is then dynamically modulated based on operator's visual attention. Regarding low-frequency processes, the robot's trajectory is adjusted to optimize operator's cognitive workload. An experimental validation involving 26 participants led to three key findings: the developed algorithm improved productivity (18% improvement), cognitive workload (5%), fluency (10%), usability (5%), reliability (5%), and acceptance of the system (9%); the exploitation of high-frequency cognitive processes leads to a significant improvement in all mentioned metrics, suggesting its relevance for future research in the field of HRSI; increasing the level of adaptability of the robotic system positively affects collaboration's effectiveness.

Keywords Human-robot seamless interaction, Collaborative robots, Robot authonomy, Robot adaptation

Industrial Human-Robot Seamless Interaction (HRSI) is considered one of the most significant research challenges for the robotics community[1,2] and a core objective of Industry 5.0[3,4]. This is motivated both by the economic and social impact associated with its effective implementation[5]. Indeed, a successful human-machine collaboration has demonstrated substantial potential for improving productivity[6], worker's safety[7] and both cognitive[8] and physical ergonomics[9].

Similarly to human teams cooperating with each other, an effective HRSI can be achieved if (i) the robot can understand the human co-worker's psycho-physical status and actions and (ii) adapt its behavior accordingly, i.e. autonomously modifying its trajectory or tasks, (iii) without interfering, degrading, or limiting the human operations[4,10].

Thus, it is essential to comprehensively capture the operator's cognitive processes to implement human-centric adaptability. Cognitive processes can occur at high- or low-frequency, based on how the human's responses vary over time. Specifically, high-frequency processes involve near real-time processing (e.g., the visual attention towards and object or situation). In contrast, low-frequency processes are characterized by slower response mechanisms (e.g., impacting cognitive workload). Considering the interplay between high-frequency and low-frequency cognitive processes, it is possible to do a parallelism to the well-established *Dual Process Theory* in cognitive psychology by[11]. Dual Process Theory posits two systems of cognition: System 1, which is fast, automatic, and reliant on immediate sensory input, and System 2, which is slower, deliberate, and effort-intensive. Similarly, high-frequency processes, such as gaze-driven visual attention, operate in real-time, aligning closely with the rapid, instinctual nature of System 1. On the other hand, low-frequency processes are very similar to System 2's thoughtful and reflective engagement. Notably, both theories emphasize the dynamic interaction between these layers of cognition, where rapid, high-frequency inputs feed into and inform slower, low-frequency adjustments, mirroring the interplay between Systems 1 and 2. By integrating these concepts,

[1]Faculty of Engineering, Free University of Bozen/Bolzano, 39100 Bolzano, Italy. [2]Department of Psychology, University of Bologna, 40126 Bologna, Italy. [3]Polytechnic Department of Architecture and Engineering, University of Udine, 33100 Udine, Italy. ✉email: luca.gualtieri@unibz.it; renato.vidoni@uniud.it

the present work extends the dual-process framework into the domain of human-robot interaction, offering a unified perspective on optimizing collaborative tasks. Overlooking either of these aspects would result in an incomplete understanding of the human cognitive state, preventing the effective implementation of HRSI. By considering both high- and low-frequency cognitive processes, the robotic behavior can be adapted to enhance overall operational effectiveness and productivity while also ensuring the safety and well-being of the operator. The effectiveness of an adaptive interaction, however, might vanish if the system introduces new potential risk factors in terms of safety or ergonomics, or reveals impractical to be deployed[12].

In the quest to solve this challenge, several different solutions have been studied and implemented to enable minimally invasive human-to-machine as well as machine-to-human communication, ranging from haptic devices[13–15] to physiological signals[16,17] and behavioral measurements[18–20]. While an effective human-machine communication is a critical aspect in HRSI, if it is not integrated in a wider framework considering also robot's behavior adaptation, it could not fully address the challenge.

However, among the researches considering robot's real-time adaptation, to the best of the authors' knowledge, there is no work that simultaneously improves safety, ergonomics as well as productivity considering both high and low-frequency cognitive processes. Due to the difficulties at online evaluating the human's awareness, rather than adjusting the robot's behavior according to the human's state, a widely adopted strategy to improve safety the enhancement of human awareness through haptic devices[21,22], extended reality[23], verbal communication[24,25] as well as visual signals[26]. However, those approaches target the reduction of the mechanical risk without acting on the robot's behavior. Considerable effort has also been expended in the development of solutions to online evaluate[27–29] and, subsequently, improve both the physical as well as the cognitive ergonomics of a collaborative operation, for instance informing the user of possible unhealthy positions is a strategy to reduce physical fatigue[30,31]. Another approach is to adapt the robot's behavior in such a way to suggest the operator to work under the proper conditions[32–34]. To online reduce mental fatigue, instead, it has been proposed to adapt accordingly and in real-time the robot behavior[35], the degree of autonomy[36], or the assistance level to human operators[37]. Moreover, to improve a trade-off between cognitive workload and productivity, the robot's trajectory has been adapted online, exploiting ideas based on Game Theory[38] and Pareto Front[39]. However, all those approaches consider only low-frequency cognitive processes. In[40], human real-time cognitive processes have been considered with the aim of enhancing safety: a multimodal approach has been developed to determine whether a collision is intentional or not to activate safety measures in possible unsafe conditions, thus optimizing the behavior of the robot with respect to mechanical risk. While this work considered high-frequency cognitive process, it mainly focused on safety aspects.

Aims and justification

The aim of this work is to develop and assess a methodology for robot's behavior adaptation taking into account both high- and low-frequency cognitive processes to implement HRSI considering productivity, safety, and well-being requirements at the same time. Such a method integrates both the adaptation of the robot's behavior based on the operator's visual attention (high-frequency cognitive process) and cognitive workload (influenced by low-frequency cognitive process). In that regard, three specific aspects of the human visual systems are considered: human's field of view (FOV), gaze behavior (where and how humans are looking at objects) and pupil dilation. The former two are exploited to estimate the operator's *visual attention*, while the latter is taken into account to evaluate the *cognitive workload*.

Adapting robot behavior according to visual attention.

The human FOV plays a significant role in determining how individuals perceive and interact with their environment[41]. It defines the part of space that can be sensed through vision. This, combined with the information provided by gaze behavior (i.e., the way by which eyes are moving and looking around), can be exploited to evaluate the operator's visual attention towards the robot. Visual attention can be understood as a collection of cognitive and physiological mechanisms that regulate visual information. These processes are essential for integrating visual data into coherent objects, which can then be identified, recognized, and remembered. Humans rely on visual attention to facilitate object recognition. Since object recognition mechanisms can't simultaneously process every object in the field, attention is crucial for selecting and delivering manageable subsets of input for recognition[42]. Visual attention refers to the cognitive processes that enable humans to efficiently manage information overload by selecting what is relevant and filtering out what is not[43]. Inadequate selective attention, mental errors, and susceptibility to distractions are recognized as key human factors that contribute to accidents when hazards go unnoticed[44,45]. In industrial HRI, preventing (mechanical) hazards from unexpected contacts is crucial for ensuring the operator's safety. The likelihood and severity of such contacts are closely related to the robot's speed within the collaborative workspace. Consequently, many designers opt to reduce the robot's speed, which, while enhancing safety, often compromises robotic performance and overall productivity. In this work, we aim to enhance both safety, productivity and operator's wellbeing by exploiting worker's *visual attention* toward the robot. *Visual attention* is considered a high-frequency cognitive process since it involves rapid, moment-to-moment adjustments in focus and processing, often requiring quick shifts between different stimuli in the environment. Indeed, this rapid processing aligns with the concept of a high-frequency cognitive process, as attention needs to be continuously regulated to manage visual information effectively.

Starting from the key idea that the operator's *visual attention* is directly related to what he/she is looking at (i.e., whether the robot is or is not in his/her FOV), as well as how the focused object is moving (i.e. *visual attention* decreases as the robot moves away from the gaze point), the following main concept has been implemented in the proposed robot's adaptive behavior algorithm: the speed of the manipulator is dynamically scaled according to the operator's *visual attention* towards the robot, thus influencing safety, productivity and operator's well-being (Fig. 1, right and Video SV1). Considering the principles of *visual attention*, it is expected that when the operator is observing the robot, he/she has a higher capability of recognizing potential contact and

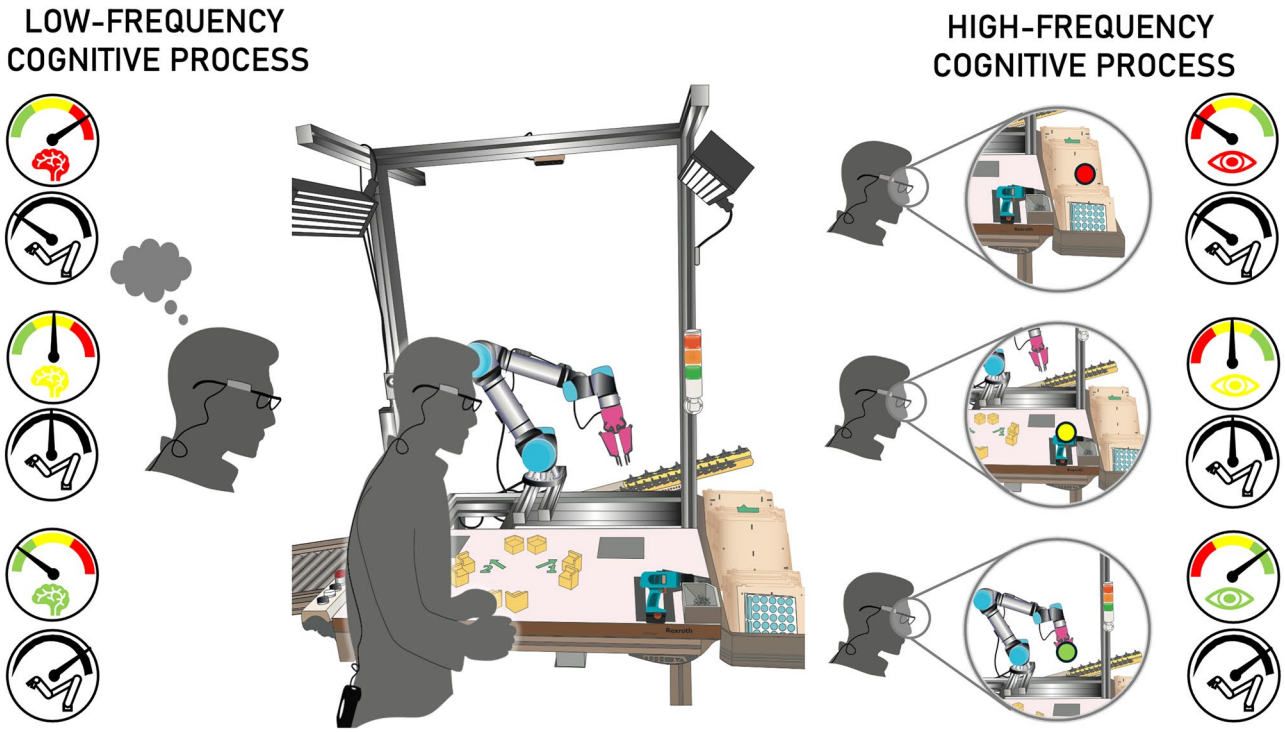

Fig. 1. representation of the experimental setup (center) and of the adaptation with respect to low- and high-frequency cognitive processes, respectively, on the left and on the right of the image. On the right, the gaze point is represented by the red, yellow and green circles.

thus preventing mechanical hazards. In this case, we suppose that the manipulator can operate at a higher speed while maintaining proper safety conditions given that the probability of a dangerous contact is supposed to be lower. At the same time, productivity is higher since it is correlated with robot speed in a collaborative cycle. *Visual attention*, then, decreases as the robot moves away from the gaze point (where the human is looking) until it reaches its minimum value as the manipulator exits the operator's FOV. In that case, the robot's speed is reduced accordingly to a safe value.

Reducing the frequency of both hazardous and non-hazardous contacts also impacts interaction ergonomics. From a physical perspective, and in line with the requirements of ISO TS 15066[46], ergonomic limits may differ from biomechanical limits. For frequent contacts, even those compliant with safety standards, applicable threshold values can be further lowered to achieve an ergonomically acceptable level. From a cognitive perspective, factors such as trust in the robotic system and the fluency of collaboration can improve when interactions minimize nonfunctional and unintended contacts.

Adapting robot behavior according to cognitive workload.

Cognitive workload refers to the effort that an individual shows during a task to achieve a particular level of performance: the more effort a task requires, the higher the *cognitive workload*[47]. It affects safety, performance and well-being of the operator[48]. Pupil dilation is a widely used method to assess *cognitive workload*. This is evidenced by various studies that have observed correlations between task difficulty and pupil dilation[49]. In particular, increased cognitive workload is associated with greater pupil dilation[50]. Eye tracking studies have also demonstrated that higher task difficulty evokes higher peak pupil dilation and longer peak duration, suggesting that pupil diameter can be used as a physiological indicator of task workload, especially in visual-motor tasks[51]. In this work, it is assumed that a robotic system can be used to modulate the operator's *cognitive workload* in such a way as to always keep it at the optimal level by changing the speed and, therefore, the pace of collaborative tasks. When a machine and a human operator cooperate, they exhibit two different peculiarities: while the machine is generally characterized by a fixed and constant cycle time, the operator's cycle time, driven by both cognitive and physiological aspects, changes continuously. In this regard, saturation (how much of the available time is occupied by work tasks) strongly affects workers *cognitive workload* in manufacturing[52]. Moreover, it has been shown that the robot's speed can also influence the operator's *cognitive workload*, even considering velocities compliant with ISO TS 15066 requirements[53]. *Cognitive workload* is considered a low-frequency process since it reflects the overall mental effort over a longer period or across tasks. The buildup and management of *cognitive workload* is affected not only by moment-to-moment adjustments, but it mostly occurs over the course of a task.

According to the measured operator's pupil dilatation, the proposed algorithm adjusts the task execution speed to optimize the worker's *cognitive workload*. It modifies the robot's speed to maintain an appropriate pace while avoiding the introduction of new cognitive related risk factors (Fig. 1, left and Video SV1). Such adaptability is dynamic and human-centric, i.e., capable of considering the variations of the optimal *cognitive*

workload level considering different operator's individual features and working conditions, potentially varying from task to task.

Combining the key adaptation actions, both high- and low-frequency cognitive processes are taken into account to adapt online the robot's behavior to improve safety, productivity and ergonomics.

Experimental scenario

To implement and validate the proposed solution in collaborative assembly tasks, robot motion scaling methods are driven in real-time by the human's *visual attention* (measured through gaze behavior) and *cognitive workload* (retrieved as an indirect measure from pupil dilation).

The experimental activity, which will be detailed in Section "Materials and Methods", consisted of a within-subject experiment with 26 participants who collaborated with an industrial robot to assemble a wooden box. In particular, once they underwent a training procedure to learn how to assemble the product, they were asked to mount the box in three different scenarios provided in a randomized order. As it will be explained in the Methods section, scenario randomization has been introduced to prevent influences on results such as the learning effect, increased familiarity with the task, and fatigue. In the first scenario (S1), the robot moves along its standard reference trajectory without adapting its behavior according to human mental processes. In the second (S2), the behavior of the robot (i.e., its trajectory) is adjusted considering only human's visual attention (high-frequency processes), while, in the third scenario (S3), the adaptive behavior is implemented considering both visual attention as well as cognitive workload (both low- and high-frequency cognitive processes). The outcomes of the three scenarios are evaluated and compared employing qualitative and quantitative results in terms of production performance, safety, and ergonomics. In particular, the following indicators have been selected: productivity- and quality- related human performance (cycle time and errors), cognitive workload, robot's reliability, HRI fluency, system usability and acceptance. Such an experiment also had the goal of evaluating the effect of the optimization considering the joint action of high- and low-frequency cognitive processes.

Results summary

Results show how, as the level of adaptivity grows (from S1 to S3), productivity, fluency of the operation, usability, reliability, and acceptance of the robotic system improve, thus proving the effectiveness of the developed algorithm. In particular, a statistically significant increase in productivity has been found shifting from S1 (no adaptability) to S3 (higher adaptability), without compromising the quality of the assembly operation: the number of errors, even though they decreased from S1 and S2 to S3, showed no statistically significant dependence of the scenario and, consequently, on the level of adaptability. In other words, it has been proven that the developed algorithm significantly improves productivity without compromising assembly quality (1^{st} result) and that, as the level of adaptability grows, the productivity increases without any worsening the number of production errors (2^{nd} result). Cognitive workload showed a statistically significant decreasing trend as the level of adaptivity increased from S1 to S3. This result also confirmed the effectiveness of the control algorithm deployed in S3, which proved to be capable of driving the cognitive workload towards the desired level (3^{rd} result). Fluency and usability showed a trend similar to that of productivity: they significantly improved from S1 to S3, meaning that the collaboration quality perceived by the user is positively impacted both by the developed method (4^{th} result) as well as by the level of adaptability of the robot (5^{th} result). Also, the reliability of the robotics system perceived by the user and the acceptance of the robot showed a similar, respectively, significant and marginally significant trend: they improved from S1 to S3, i.e. as adaptivity grows. It confirmed that the developed control algorithms can foster the the operator's working conditions form a cognitive perspective of the user (6^{th} result) and that this can be achieved through a growing adaptability of the system (7^{th} result).

Such results can be summarized in three key findings: (1^{st} key finding) the developed algorithms led to improved productivity, operator's cognitive workload, fluency of the collaborative operation, usability, reliability, and acceptance of the system without affecting the quality of the production process; (2^{nd} key finding) the exploitation of high-frequency cognitive processes, which are almost overlooked in literature, leads to a significant and consistent improvement in all the above-mentioned metrics, thus suggesting its relevance for future research in the field of HRSI, (3^{rd} key finding) an increasing level of adaptability of the robotic system positively affects all the metrics considered, without compromising the quality of the manufacturing process.

Results

This section reports the metrics and the analysis of the effects of the developed control algorithms and, consequently, of the level of robot's adaptability, on productivity (cycle time and errors), cognitive workload, robot's reliability, HRI fluency, system's usability, and acceptance. These were evaluated by considering all the scenarios (S1, S2, and S3). Results have been summarized in Table 1 and Fig. 2.

Productivity

To measure productivity-related human performance, two indicators have been evaluated: the average cycle time and the number of assembly errors. These have been widely demonstrated to be relevant indicators to quantify productivity performance.

The lower the cycle time, the higher the amount of products assembled. Descriptive statics shows how the mean cycle time value generally decreases from S1 (mean M=142.66s, standard deviation SD=11.46s) to S2 (M=120.86, SD=16.77, -15.28% w.r.t. S1) and S3 (M=117.08, SD=13.23, -17.93% w.r.t. S1). To evaluate the statistical significance of the scenario and, consequently, of the control strategy, an Analysis of Variance (ANOVA) has been performed, leading to the following results: Sum of Squares SS=8507.90, F-value F=38.77, p=$2.42 \cdot 10^{-8} < 0.001$. Considering that productivity is inversely proportional to cycle time, this confirms a systematic increase in productivity due to the progressive scenarios change.

Measured Quantity	Measured Quantity	Optimal Results	S1		S2		S3		
			M	SD	M	SD	M	SD	*p*-value
Productivity	Cycle Time [s]	Lower is better	142.66	11.46	120.86	16.77	117.08	13.23	$\leqslant 0.001$
Productivity	Assembly Errors	Lower is better	0.846	1.262	0.885	1.154	0.500	0.635	0.248
Cognitive Workload	Stochastic Gaze Entropy (SGE)	Lower is better	1.065	0.056	1.017	0.063	1.008	0.048	$\leqslant 0.001$
Fluency	Paliga and Pollak[59]	Higher is better	5.63	1.27	6.19	0.69	6.22	0.67	0.02
Usability	Lewis and Sauro[60]	Higher is better	4.27	0.52	4.43	0.44	4.50	0.40	0.01
Reliability	Stochastic Gaze Entropy (SGE)	Lower is better	1.065	0.056	1.017	0.063	1.008	0.048	$\leqslant 0.001$
Acceptance	Vanderlaan[62]	Higher is better	3.85	0.99	4.18	0.64	4.20	0.64	0.06

Table 1. table summarizing the results of the experimental campaign.

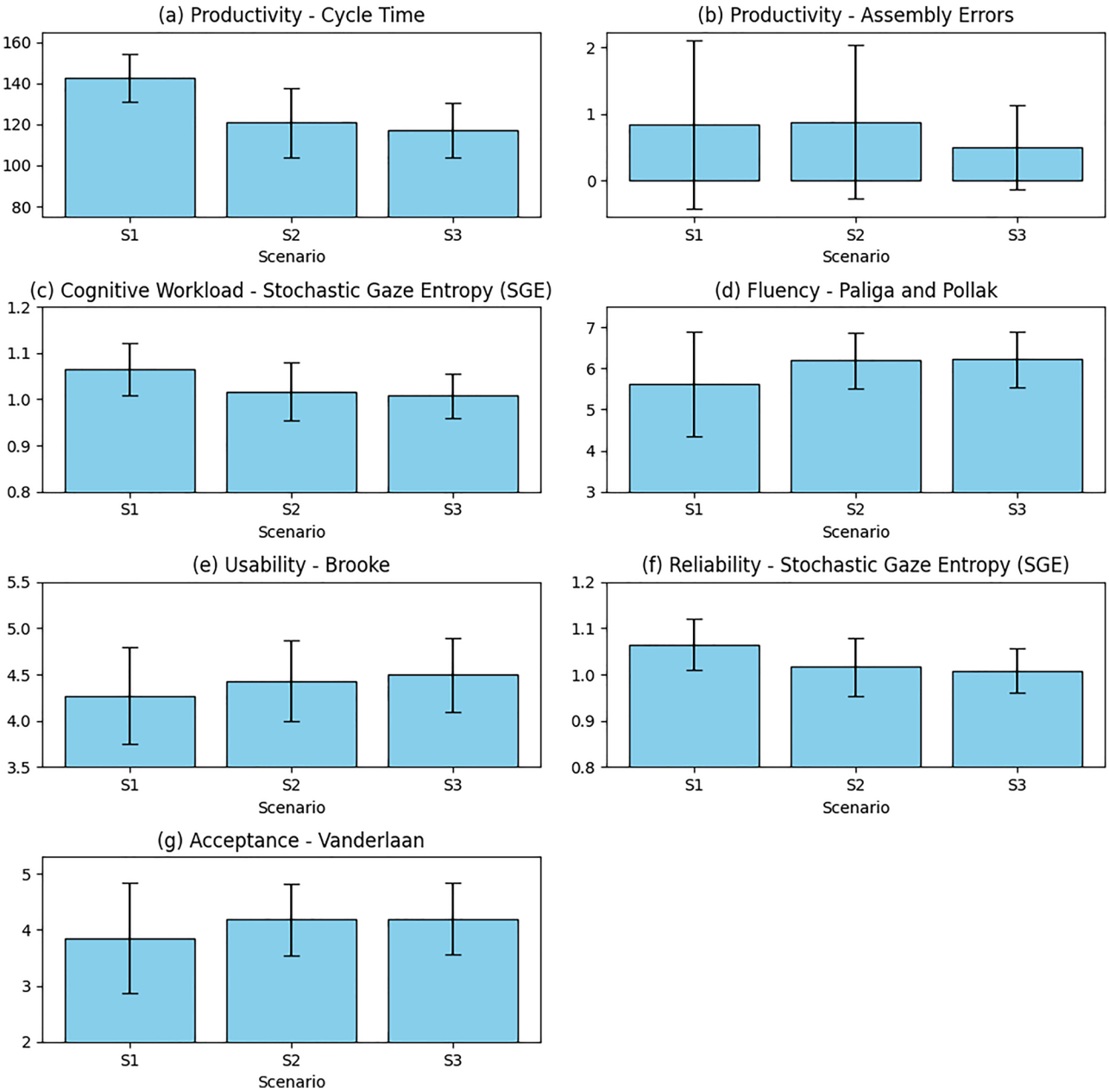

Fig. 2. Plots summarizing the results of the experiments. In particular, the error bar represents the average value and one standard deviation.

The number of human errors, i.e. the total number of wrongly executed assembly steps, is measured to evaluate the influence of the robot's adaptivity and, consequently, of the proposed control logic, on the effectiveness of the collaborative operation. The number of errors committed by an operator is heavily influenced by the operator's cognitive workload[50], which, in turn, varies depending on the method used to control the robot. Furthermore, the smaller the number of errors is, the better the performance of the human-robot dyad[54].

The descriptive statistics shows that the mean number of errors made by the participants for each scenario decreased from S1 (M =0.846, SD = 1.262) to S3 (M = 0.500, SD = 0.635, −40.9% w.r.t. S1), with S2 (M = 0.885, SD = 1.154, +4.61% w.r.t. S1) having a slightly higher mean than S1. An ANOVA analysis was performed to statistically evaluate the significance of such data, and the results were SS=1.558, F=1.356, p=0.248. Such results entail that there is no statistically significant trend.

Cognitive workload

This metric quantitatively describes gaze behavior and, in particular, gaze dispersion[55,56]. Multiple studies have found a correlation between Stochastic Gaze Entropy (SGE) and cognitive demand: as cognitive workload increases, SGE also increases[57,58]. SGE H_{SGE} has been computed using the Shannon equation[55]:

$$H_{SGE} = -\sum_{i=1}^{n} p_i \log_2(p_i)$$

where n is the number of areas of interest (AOI) within the FOV of the operator, p_i the probability that the gaze is observing the i^{th} AOI. The numeric values lie in the range $0 \leq H \leq H_{SGEmax} = \log_2(n)$, where 0 is associated with no gaze dispersion, i.e. the operator observes always the same AOI, and H_{SGEmax} indicates the maximum possible gaze dispersion, indicating randomness and unpredictability in the gaze behavior. In this study, SGE was normalized considering its baseline value (further details will be given in the "Material and Methods" Section). In this experiment 100 different possible AOIs have been defined to compute such a metric.

Regarding the experimental results, the descriptive statistics show a decreasing value of SGE, thus indicating a decreasing level of cognitive workload shifting from S1 (M=1.065, SD=0.056) to S2 (M=1.017, SD=0.063, −4.51% w.r.t. S1) and S3 (M=1.008 SD=0.048, −5.35% w.r.t. S1). The ANOVA test (SS=0.042, F=12.697, p=$6,37 \cdot 10^{-4}$) confirmed the statistical significance of such a trend.

Fluency in human-robot interaction

Fluency in HRI has been evaluated through the self-reported "Fluency in Human-Robot Interaction Scale," developed by Paliga and Pollak[59], which evaluates the smoothness and efficiency of interactions between humans and robots. Fluency in this context refers to how seamlessly human operators and robots work together as a team. The scale assesses this fluency considering three aspects: the human operator's perspective, the robot's performance, and the overall teamwork. It includes items that measure trust in the robot's decision-making, the robot's commitment to the team's success, the effectiveness of the robot as a team member, and the overall harmony of the human-robot team. This comprehensive approach provides valuable insights into the dynamics of human-robot collaboration.

The respondents are instructed: "The following statements concern your work with the robot. Please read them carefully and choose how much you agree with each statement. Use the scale provided below." The response format has a seven-point Likert scale (1 = I strongly disagree, 2 = I disagree, 3 = I somewhat disagree, 4 = I neither agree nor disagree, 5 = I somewhat agree, 6 = I agree, 7 = I strongly agree). Furthermore, the scale includes items such as "I trusted the robot to do the right thing at the right time" for assessing the human's trust in the robot, "The robot performed well as part of the team" for evaluating the robot's contribution, and "The human-robot team did well on the task" for gauging the overall effectiveness of the team. Each item is designed to reflect different aspects of fluency in the collaborative process.

To evaluate such data, a repeated measures analysis using a General Linear Model (GLM) has been performed to examine the differences in fluency scores across three scenarios. The mean scores increased from S1 (M = 5.63, SD = 1.27) to S2 (M = 6.19, SD = 0.69, +9.95% w.r.t. S1) and S3 (M = 6.22, SD = 0.77, +10.48% w.r.t. S1). The multivariate tests showed a marginally significant effect of the scenarios on fluency scores (Pillai's Trace = 0.22, F(2, 24) = 3.37, p = 0.05). However, Mauchly's test of sphericity was significant (p < 0.001), indicating a violation of the sphericity assumption. After applying the Greenhouse-Geisser and Huynh-Feldt corrections, the tests of within-subjects effects revealed a significant effect of scenarios on fluency scores (Greenhouse-Geisser: F(1.29, 32.27) = 5.99, p = 0.01; Huynh-Feldt: F(1.33, 33.28) = 5.99, p = 0.01).

The tests of within-subjects contrasts showed significant linear (F(1, 25) = 6.31, p = 0.02) and quadratic (F(1, 25) = 5.08, p = 0.03) trends across the scenarios. The estimated marginal means increased from S1 (M = 5.64, 95% CI [5.12, 6.15]) to S2 (M = 6.19, 95% CI [5.92, 6.47]) and S3 (M = 6.22, 95% CI [5.91, 6.53]), indicating a significant effect of scenarios on fluency scores, with scores increasing from S1 to S2 and S3, following both linear and quadratic trends.

Usability of the robotic system

In the evaluation of the robot system's usability, the research utilized the System Usability Scale (SUS), a widely recognized standardized questionnaire for assessing perceived usability. Developed by Brooke[60], the SUS is prevalent in industrial usability studies. The SUS consists of 10 items, each rated on a five-point Likert scale ranging from 1 (Strongly Disagree) to 5 (Strongly Agree), with statements alternating between positive and negative implications. Sample items from the SUS include assessments of system usage frequency, complexity, and ease of use ('I think that I would like to use this system frequently', 'I found the system unnecessarily

complex', 'I thought the system was easy to use'). Participants completed a questionnaire at the end of each scenario for SUS evaluation.

In S1, participants generally agreed or strongly agreed that they would like to use the robot frequently, found it easy to use, and felt confident using the system. They also disagreed that the robot was unnecessarily complex, inappropriate for the task, or required learning many things before use. In S2, SUS scores improved, with higher percentages strongly agreeing that they would like to use the robot frequently, found it easy to use, and felt confident using the system. They also strongly disagreed that the robot was unnecessarily complex, needed technical support, was inappropriate for the task, or required learning many things before use. S3 showed similar trends, with even higher percentages strongly agreeing across most dimensions, indicating that participants perceived the robot as user-friendly, easy to learn, and well-suited for the task, with a high level of confidence in using the system.

Regarding the quantitative analysis of the experimental results, a repeated measures analysis using a General Linear Model (GLM) has been performed to investigate the differences in SUS scores across three scenarios. The mean SUS scores increased from S1 (M = 4.27, SD = 0.52) to S2 (M = 4.43, SD = 0.44, +3.75% w.r.t. S1) and S3 (M = 4.50, SD = 0.40, +5.39% w.r.t. S1). The multivariate tests indicated a significant effect of the scenario on SUS scores (Pillai's Trace = 0.24, F(2, 24) = 3.75, $p = 0.04$). Mauchly's test of sphericity was not significant ($p = 0.15$), suggesting that the assumption of sphericity was not violated. The tests of within-subjects effects, assuming sphericity, revealed a significant effect of scenarios on SUS scores (F(2, 50) = 5.17, $p = 0.01$). The tests of within-subjects contrasts showed a significant linear trend (F(1, 25) = 7.80, $p = 0.01$) but no significant quadratic trend (F(1, 25) = 0.61, $p = 0.44$) across the scenarios. The estimated marginal means increased from S1 (M = 4.27, 95% CI [4.06, 4.48]) to S2 (M = 4.43, 95% CI [4.25, 4.61]) and S3 (M = 4.50, 95% CI [4.34, 4.66]), suggesting a significant effect of scenarios on SUS scores, with scores increasing linearly from S1 to S2 and S3.

Reliability of the robotic system

Another relevant peculiarity of SGE is its correlation with the reliability of the robot perceived by the human: as the robot's actions are more reliable, SGE decreases[61].

Analogously to what has been described for the analysis of the cognitive workload, the descriptive statistics show a decreasing value of SGE between the scenarios. This indicates that the reliability of the robot increases as the level of adaptability increases changing from S1 (M=1.065, SD=0.056) to S2 (M=1.017, SD=0.063, −4.51% w.r.t. S1) and S3 (M=1.008 SD=0.048, −5.35% w.r.t. S1). The ANOVA test (SS=0.042, F=12.697, p=$6,37 \cdot 10^{-4}$) revealed that this trend is characterized by a statistical significance.

Acceptance of the robotic system

The Acceptance Scores[62] were measured using a 5-point semantic differential scale, where participants rated their perceptions of the scenarios on various dimensions. Each dimension was presented as a pair of opposite adjectives, such as "Useful:Useless," "Pleasant:Unpleasant," "Bad:Good," "Pleasant:Annoying," "Effective:Superfluous," "Irritating:Likable," "Helpful:Worthless," "Undesirable:Desirable," and "Increases alertness:Causes drowsiness." For each pair of adjectives, participants selected a value from 1 to 5, with 1 representing the most negative perception (e.g., useless, unpleasant, bad) and 5 representing the most positive perception (e.g., useful, pleasant, good). In some cases, the scales were reverse to ensure consistency in the direction of the scores, with 1 representing the most positive perception and 5 representing the most negative perception. Using this semantic differential scale, it is possible to quantify participants' perceptions and attitudes toward the scenarios across multiple dimensions, providing a comprehensive measure of acceptance.

A repeated measures analysis using a General Linear Model (GLM) has been performed to examine the differences in acceptance scores across three scenarios. The mean acceptance scores increased from S1 (M = 3.85, SD = 0.99) to S2 (M = 4.18, SD = 0.64, +8.57% w.r.t. S1) and S3 (M = 4.20, SD = 0.64, +9.09 % w.r.t. S1). The multivariate tests did not indicate a significant effect of the scenario on acceptance scores (Pillai's Trace = 0.14, F(2, 24) = 1.93, $p = 0.17$). However, Mauchly's test of sphericity was significant ($p < 0.001$), suggesting that the assumption of sphericity was violated. After applying the Greenhouse-Geisser and Huynh-Feldt corrections, the tests of within-subjects effects showed a marginally significant effect of scenario on acceptance scores (Greenhouse-Geisser: F(1.11, 27.79) = 3.77, $p = 0.06$; Huynh-Feldt: F(1.13, 28.15) = 3.77, $p = 0.06$).

The tests of within-subjects contrasts revealed a marginally significant linear trend (F(1, 25) = 4.01, $p = 0.06$) but no significant quadratic trend (F(1, 25) = 3.08, $p = 0.09$) across the scenarios. The estimated marginal means increased from S1 (M = 3.85, 95% CI [3.45, 4.25]) to S2 (M = 4.18, 95% CI [3.93, 4.44]) and S3 (M = 4.20, 95 % CI [3.94, 4.46]), suggesting a marginally significant effect of scenario on acceptance scores, with scores increasing linearly from S1 to S2 and S3.

Discussion

This work investigated the evolution of HRSI across the three scenarios, focusing on the following indicators: productivity-related human performance (cycle time and errors), cognitive workload, HRI fluency, system's usability, reliability and acceptance. The results provide valuable insights into the dynamics of human-robot collaboration and the potential for improvement over time. In the following, all the above-mentioned indicators are discussed considering the analyzed results. Finally, the future perspective of deploying this technological approach in industrial deployment is discussed.

Productivity

Results showed a statistically significant reduction of the cycle time, thus consistently enhancing the productivity of the collaborative assembly cycle. Such improvement can be explained by a concurrency of several factors. In particular, according to the scientific literature, both the improved cognitive workload of the operator, the

enhanced fluency of the collaborative operation, as well as the higher usability of the system are all elements that have been proven to positively influence the productivity of HRC[63,64].

One may argue that a systematic reduction in the cycle time, which is due to an increase of the robot's velocity, could be caused by an overly slow robot's nominal trajectory (i.e., the one used in the first scenario). However, the speed of the nominal trajectory has been chosen so as to be the maximum admissible safe speed according to the requirements of ISO/TS 16055[46]. Moreover, the collaborative cycle time is usually affected not only by the robot's velocity, but also by several other factors, such as the velocity of the operator in executing collaborative concurrent tasks, the number of potential human's errors to be fixed (that can also lead to the necessity of rework), idle times of both human and the machine and so on. Moreover, if the increase in productivity were due to an overly slow nominal trajectory rather than the control algorithm, the improvement observed from scenario S2 to S3 would not be justifiable.

Regarding errors, the analysis showed a general improvement in S3 compared to S1 and S2. However, ANOVA analysis showed that such a trend has no statistical significance. This result can also be interpreted as follows: the introduction of the new control algorithms does not compromise the quality of assembled products. In other words, the productivity increase is achieved without a corresponding rise in errors. In fact, such algorithms could shift the visual attention of the operator towards the machine, which, in turn, could compromise the correctness of the assembly operations[54]. However, this effect does not take place. This can be explained by the improvement in the cognitive workload level, which improves the operator's effectiveness[65], thus compensating the possible worsening effect due to the shift of the operator's visual attention.

Cognitive workload

Results showed a significant reduction in the cognitive workload from S1 to S3, i.e. the cognitive workload generally decreases as the level of adaptability increases. The improvement from S1 to S2, where the adaptation of the robot considering high-frequency cognitive processes is introduced, can be explained thanks to a better synchronization between the human and the robot, thus enhancing cognitive workload acting on saturation[52]. The additional improvement from S2 to S3, instead, can be explained by the further introduction of a control scheme that regulates the velocity of the robot to optimize the cognitive workload by acting on velocity[53]. In particular, this improvement confirms that the developed method is actually capable of effectively controlling and, consequently, optimizing the cognitive workload. The cognitive workload measured in S3 is the closest, among the cognitive workloads of the three scenarios, to the target cognitive workload, i.e. the baseline/nominal normalized one: 1 (see section "Materials and Methods" for further details).

One might expect the cognitive workload of S1 to be equal to the nominal one, since the robot's behavior was the same for both S1 and the baseline measurement. However, it is important to remember that the cognitive workload measured across the whole scenario S1 takes into account factors such as mental fatigue and potential errors, which might degrade the level of cognitive workload. This analysis also highlights the robustness of the developed method and how it proved to be capable of dynamically keeping the cognitive workload close to the desired value when other disturbance factors, such as fatigue or errors, happen.

Fluency in human-robot collaboration

The statistically significant improvement in HRI fluency scores across the scenarios is a key finding of this study. The increase from S1 to S3 indicates that participants perceived enhanced collaboration quality as the robot's adaptability level increased. This improvement manifested in multiple dimensions, including trust, perceived commitment, and coordination. The improvement from S1 to S2 can be explained by the fact that the method applied in S2 is designed to enhance the synchronization between human and machine, thus enhancing saturation[52]. Instead, the further improvement in fluency from S2 to S3, where also cognitive workload is taken into account by the adaptation algorithm, is probably due to the optimization of the cognitive workload of the operator, which allows the operator to work in a safe mental condition.

Usability of the robotic system

The System Usability Scale (SUS) scores showed a significant linear increase across the three scenarios, from S1 to S3. This improvement suggests that participants found the robotic system increasingly user-friendly and well-suited to the task as the level of adaptability increased. The high SUS scores, particularly in S3, indicate that the system design effectively supports user interaction and task completion. This result aligns with principles of human-centered design in robotics, which established how an intuitive interface and interaction improves the usability of a system[66]. The intuitiveness of the system, which has been perceived by the participants, could be motivated by the fact that the method is capable of tuning the robot's activities in an implicit way without requiring the operator to explicitly communicate his/her needs or wants.

Reliability of the robotic system

The results show a statistically significant increase in the reliability of the robotic system as the level of adaptability increases. According to the scientific literature, it has been shown how unpredictable[61] or excessively fast robot movements[67] compromise the reliability of the system. However, even if the developed methods lead to a constantly changing robot trajectory, its reliability has not been affected and, contrary to what could be expected, it improves. It can be explained by the predictability of the system: even if the robot's behavior changes continuously, the way it changes is predictable and intuitive for the user. Moreover, the system's adaptability according to the human state/gaze behavior allows the user to partially control the robot's behavior, thus making it more reliable. This fact have been reported by several participants during the non-structured interviews at the end of the experiment.

Acceptance of the robotic system

Although marginally significant, the trend towards improved acceptance scores (from M = 3.85 in S1 to M = 4.20 in S3 and M=4.18 in S2) suggests a positive shift in participants' attitudes towards the robotic system. In this context, acceptance can be ultimately interpreted as a key indicator of the operator's overall willingness to collaborate with the robot. This result could be related to improved operator trust in the robotic system. Participants might have developed a greater sense of trust in the system as the robot adapted to the human operator's cognitive state, particularly in scenario S3. Studies have shown that trust in automation is a critical factor in system acceptance[68]. Furthermore, the dynamic adaptivity in S2 and S3 may have fostered a stronger feeling of control over the system, contributing to a higher perception of usefulness, ease of use, and desirability—which are key factors influencing technology acceptance. Although the acceptance scores showed a positive trend, the fact that the results were only marginally significant suggests that there may be room for further improvements in the robot's adaptability or interaction design.

Robot's adaptation considering high-frequency cognitive processes

This subsection has been introduced to emphasize the effect of adapting the robot's trajectory considering high-frequency mental processes, which is often overlooked in the literature. In fact, to the best of the author's knowledge, none of the contributions proposed in literature adjust the robot behavior considering high-frequency mental processes. However, the results of this research highlight that this adaptation (implemented alone in S2) led to an improvement in productivity, cognitive workload, fluency of the operation, usability, reliability, and acceptance of the robotic system (without compromising the quality of the assembly operations) compared to a standard (i.e., not adaptive) collaborative applications (considered in S1). This is a relevant finding and leads to the conclusion that adaptation with respect to high-frequency mental processes can further enhance the effectiveness of HRC and should be further extended in future research.

Future prospects

When considering the potential future deployment of such technology in industrial scenarios, three topics must be addressed: (i) economic aspects, (ii) industrial safety as well as (iii) ethical and privacy considerations.

Regarding the first aspect, the use of such a technology led to a statistically significant increase in productivity, with a reduction in cycle time of up to 18%, and it did not affect the production quality. To this, a second consideration must be added concerning the purchase cost of an eye-tracking device. While a few years ago such devices were manufactured by only a small number of companies, in recent years we have witnessed an increase in the number of manufacturers, resulting in greater market competition that now allows head-mounted eye-trackers to be purchased at affordable prices. The observed increase in productivity, combined with the cost of such devices, makes this technological solution potentially suitable for future use in industrial settings.

Concerning industrial safety concerns, a safe and certifiable deployment of such a technological solution would need a certifiable data transmission protocol to interface the eye-tracking device with the robot controller. Moreover, it should have the capability to univocally identify the robot towards which the operator is looking. Such challenges can be overcome, respectively, by adopting a certifiable communication protocol, i.e. PROFISafe, and by covering the robot's skin with identifiable markers/patterns.

With respect to the last aspect, i.e. ethical and privacy concerns, it is worth mentioning that the developed technological solution, to be operated, can manage the data locally, storing them for a very limited amount of time. In particular, data regarding gaze movements are used in real-time and then discarded, while pupil diameter data are stored only for the duration of one assembly cycle, to update the parameter α_{max} for the following cycle. Therefore, such a technological solution would not lead to a mapping of the performance and/or physiological evolution of an operator. Processing data locally and deleting them after usage would reduce the risk associated with cyber-security attacks.

Methods

Robot's adaptive control

In this Subsection, the methods developed and deployed in the experiment to adjust the robot's behavior considering the human's low- and high- frequency cognitive processes are detailed. In particular, the first method dynamically scales the trajectory of the robot considering the operator's visual attention (scenario S2), while the second also integrates the cognitive workload within the control rules (S3).

Method 1 (applied in S2)

When considering a human interacting with a robot, which moves along a predefined reference trajectory $\mathrm{q}(t)$, the human observes a point (h_{gp}) within its FOV (Fig. 3).

It is possible to compute the *visual distance*, i.e. the distance between the gaze point and center of the robot's end-effector (EE) (expressed in the space of the FOV) as:

$$d_{EEgp} = |\mathrm{r}_{EE} - \mathrm{h}_{gp}| \tag{1}$$

The *visual distance* d_{EEgp} quantifies the operator' *visual attention* by considering his/her capacity to prevent hazardous contacts concerning the actions that the robot is performing. The *visual attention* of the operator is higher when he/she observes the robot compared to the moments in which he/she isn't observing it or the robot is out of the field of view. To enhance productivity without affecting mechanical safety, the method foresees that the speed of the manipulator is scaled by a factor $\alpha(d_{EEgp}, t)$, which is a function of the distance d_{EEgp}, i.e. of the *visual attention*, and varies along time:

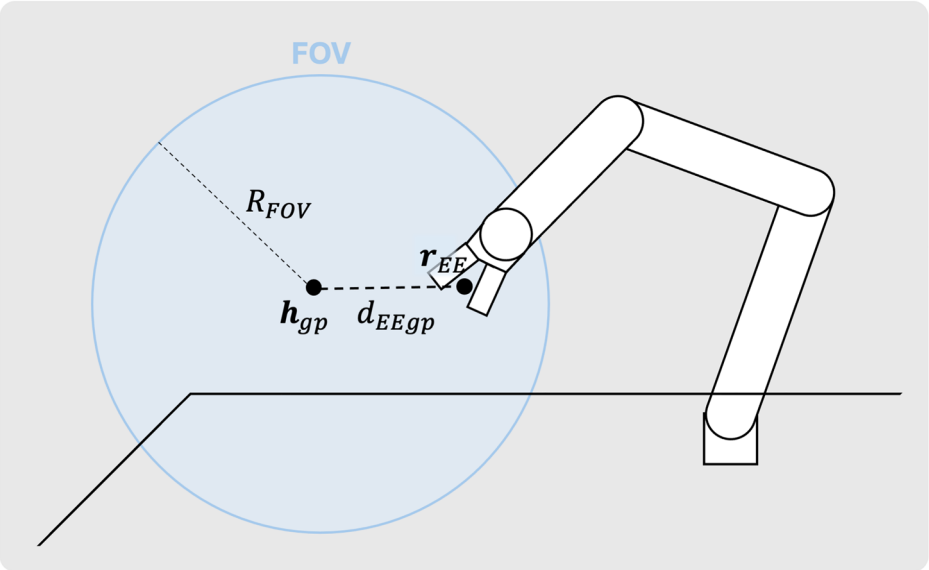

Fig. 3. representation of the field of view of the operator, its gaze point h_{gp}, the position of the center of the end-effector (EE) in the FOV of the operator r_{EE}, the distance between the operator's gaze point and the position of the center of the robot's EE d_{EEgp} and the border of the robot's EE (blue line).

$$\dot{\mathbf{q}}_s = \alpha(d_{EEgp}(t))\dot{\mathbf{q}} \tag{2}$$

where $\dot{\mathbf{q}}_s$ is the actual speed of the manipulator, i.e. the speed achieved by applying the velocity scaling. An effective velocity scaling $\alpha(d_{EEgp}(t))$ has to satisfy the following requirements:

- the velocity scaling has to be maximum ($\alpha_{max}, \alpha_{max} > 1$) when the human gaze point coincides with the EE of the robot:

$$\alpha(d_{EEgp} = 0) = \alpha_{max} \tag{3}$$

- the velocity scaling must decrease as the robot's EE goes further with respect to the gaze point of the operator:

$$\frac{d\alpha(d_{EEgp}, t)}{dd_{EEgp}(t)} \leq 0 \tag{4}$$

- the velocity scaling has to assume a value equal to one when the EE is out of the field of view of the operator:

$$\alpha(d_{EEgp} = R_{FOV}) = 1 \tag{5}$$

In other words, if the manipulator's EE is out of the FOV of the operator, the manipulator moves along the reference trajectory, thus operating at its nominal speed (defined according to safety standards requirements).

- The acceleration of the actual trajectory executed by the manipulator must be continuous. It has been shown that continuum acceleration trajectories, i.e. finite jerk trajectories, entail better results with respect to infinite jerk trajectories in terms of cognitive ergonomics[69,70]. Since this system is implemented to improve HRC, continuous acceleration trajectories are chosen to maximize interaction performances and system's ergonomics.

In the following paragraphs, the mathematical definition of a possible suitable scaling function is presented.

Let's consider the nominal trajectory $\mathrm{q}(t)$ and the trajectory obtained by applying the velocity scaling q_s, where, recalling eq. 2, the relation among them is

$$\dot{\mathbf{q}}_s = \alpha\dot{\mathbf{q}}$$

Consequently, the acceleration of the 'scaled' trajectory can be computed as:

$$\begin{aligned}\ddot{\mathbf{q}}_s =& \frac{d\dot{\mathbf{q}}_s}{dt} \\ =& \frac{d\alpha\dot{\mathbf{q}}}{dt} \\ =& \frac{d^2\dot{\mathbf{q}}}{dt^2}\alpha + \frac{d\dot{\mathbf{q}}}{dt}\dot{\alpha}\end{aligned} \tag{6}$$

Since it has been assumed that the reference trajectory $\mathrm{q}(t)$ has a finite jerk and, consequently, continuous accelerations, to obtain a finite jerk 'scaled' trajectory it is necessary that $\alpha(d_{EEgp}(t))$ and its first derivative are continuous and differentiable. Since α depends on the distance between the human gaze point and the EE's position, and that this distance varies along time, its derivative with respect to time can be expressed as:

$$\dot{\alpha} = \frac{d\alpha}{dd_{EEgp}}\frac{dd_{EEgp}}{dt} \tag{7}$$

Therefore, to achieve continuity of $\dot{\alpha}$, both $\frac{d\alpha}{dd_{EEgp}}$ and $\frac{dd_{EEgp}}{dt}$ must be continuous. Since the term $\frac{dd_{EEgp}}{dt}$ depends on the movement of three physical systems, it can be assumed to be continuous and differentiable. With respect to $\frac{d\alpha}{dd_{EEgp}}$, instead, it is necessary to mathematically ensure the fulfillment of continuity and differentiability. For this purpose, the following formulation of α can be exploited:

$$\begin{cases} \phi(d_{EEgp}) & \text{si } d_{EEgp} < R_{FOV} \\ 1 & \text{if } d_{EEgp} \geq R_{FOV} \end{cases} \tag{8}$$

where $\phi(d_{EEgp})$ is a third-order polynomial subject to the following boundary conditions, which are required to satisfy the constraints listed above:

$$\begin{aligned}\phi(d_{EEgp} = 0) =& \alpha_{max} \\ \dot{\phi}(d_{EEgp} = 0) =& 0 \\ \phi(d_{EEgp} = R_{FOV}) =& 1 \\ \dot{\phi}(d_{EEgp} = R_{FOV}) =& 0\end{aligned} \tag{9}$$

Considering the conditions above, it is possible to express $\phi(d_{EEgp})$ as:

$$\phi = \alpha_{max} - 3\frac{\alpha_{max} - 1}{R_{FOV}^2}d_{EEgp}^2 + 2\frac{\alpha_{max} - 1}{R_{FOV}^3}d_{EEgp}^3 \tag{10}$$

and its first derivative with respect to d_{EEgp} as:

$$\begin{aligned}\frac{d\phi}{dd_{EEgp}} =& \frac{6(\alpha_{max} - 1)}{R_{FOV}^2}d_{EEgp}^2\left(\frac{d_{EEgp}}{R_{FOV}} - 1\right) \\ \leq& 0 \quad \text{if } d_{EEgp} < R_{FOV}\end{aligned} \tag{11}$$

which entails that

$$\text{if} \quad d_{EEgp} > R_{FOV} \Rightarrow \frac{d\phi}{dd_{EEgp}} = 0$$

as it is required by the conditions stated above, is less or equal to zero in its entire domain of application.

Method 2 (applied in S3)

The goal of the second method is to extend what proposed in the first method to develop an interaction paradigm that can account for both the operator's visual attention and cognitive workload.

Method 1 leads to values of α in the range $1 \leq \alpha \leq \alpha_{max}$ and it determines the manipulator's increase of speed, where α_{max} defines the upper bound, i.e. the maximum possible speed increase. In other words, different values of α_{max} entail different velocities of the manipulator and, consequently, different levels of the operator's cognitive workload. The second method aims to online adjust the velocity scaling α_{max} and, consequently, α considering also the cognitive workload c_{wl} of the operator $\alpha(d_{EEgp}, c_{wl}, t)$, so that it automatically converges toward a *nominal* value which is assumed to be representative of a state-of-the-art collaborative application and does not represent a potential risk factor.

The developed method is structured as follows:

- a trained operator executes the task without applying the velocity scaling, i.e. in *standard/non adaptive conditions*. While executing the collaborative cycle, the cognitive workload of the operator is measured. The obtained value is considered as the baseline cognitive workload associated with the operation. This is the so-called *nominal* level of cognitive workload: $c_{wl,nominal}$, i.e. the level a worker has before starting the assembly activities implementing the adaptive behavior of the robot. The operator is trained to execute the assembly

operations so that the measurement of the *nominal* level of cognitive workload is not affected due to the effects of the learning curve. Such a baseline has been established for each operator working with the adaptive robot. This ensures a human-centered approach since the cognitive workload is a person-specific construct[71].

- the velocity scaling is then deployed. At the first cycle an initial arbitrary value of α_{max} is used: $\alpha_{max,0}$. At the end of each cycle *i*, the cognitive workload during that i^{th} cycle c_i is computed. The velocity scaling factor to be adopted in the subsequent cycle $\alpha_{max,i+1}$ is then computed as:

$$\alpha_{max,i+1} = \max(1 + (\alpha_{max,i} - 1)\frac{c_{wl,nominal}}{c_i}, \alpha_{limit}) \tag{12}$$

where α_{limit} is an upper bound introduced to avoid excessive velocities, so as to limit the mechanical risk for the operator.

From a practical standpoint, the method works as follows: if the level of cognitive workload measured in the previous cycle is larger than the *nominal* level, the speed increase is reduced. Instead, if the cognitive workload is lower than the *nominal* value, the speed increase rises. Consequently, this velocity scaling system tends to automatically adjust its parameters converging towards the *nominal* level of cognitive workload.

To numerically quantify the cognitive workload of the operator, the average diameter of the pupil ($\bar{d_p}$) is used:

$$c_s = \bar{d_p} \tag{13}$$

In fact, it has been shown that the pupil diameter is correlated with the cognitive workload associated with a specific task. In particular, as the cognitive workload increases, also the mean pupil diameter increases[49–51,72,73]

User study

Participants

The experiment involved 26 participants (13 male, 13 female; M_{age} = 24.5, SD = 3.2) with different educational levels and backgrounds. In fact, 30.8% of the participants achieved a high school diploma, 26.9% a bachelor's degree, 34.6% a master's degree, and 7.7% a doctorate or higher. Half of the participants were students, with others employed in academia, engineering, energy, and sustainability sectors. The majority (76.9%) were born in Italy, with the remainder from other EU (3.8%) and non-EU countries (19.2%).

Regarding participants' interest in technology, do-it-yourself (DIY) abilities, and familiarity with robots, 54% had 'some to high familiarity' with collaborative robots, 87% declared to be 'very interested' in technological innovations, 88% had a 'positive' or 'excellent' opinion about robots, 88% claimed to be 'skilled' or 'very skilled' in DIY and manual work skills.

Regarding participants' levels of self-efficacy in HRI, the Short version of the Self-Efficacy in Human-Robot Interaction Scale (SE-HRI), which includes various statements designed to gauge participants' current confidence in interacting with robots and emphasizes their present ability rather than potential future skills, showed that 42% believed they could make a robot perform a specific task, 54% agreed their technological familiarity would enable robot use, 85% felt confident in achieving goals with robot assistance, 69% believed they could teach a robot with sufficient effort, 62% felt capable of teaching a robot simple tasks.

Experimental setup

The experimental area is divided in two parts: the first is the collaborative workstation (Fig. 4A) where the operator (Fig. 4B) assembled a wooden box (Fig. 4C), while the second is the place where the participants could complete the questionnaires (Fig. 4D).

The collaborative workstation consists of a worktable with a UR5e collaborative robot (produced by Universal Robots, 6 degrees of freedom and 5kg payload), equipped with a two finger 2F-85 gripper (produced by ROBOTIQ) (item 4 of Fig. 4A). The robotic system exploits a tower light (item 5 Fig. 4A) to communicate its status to the operator (working - green light, idle state - orange, error - red), an emergency stop for safety reasons as well as a push-button for interaction purposes (when required, the operator can command the robot to proceed to the following task by pressing the buttons) (item 9 Fig. 4A). Storage racks (items 6 Fig. 4A) are attached to the experiment's table. To facilitate the assembly process, a screwdriver (item 8 Fig. 4A) and some fixtures to hold the product in the correct position (item 7 Fig. 4A) have been used. In the left part of the workstation, a roller conveyor belt has been placed to deliver the finished products to storage (item 10 Fig. 4A). On the top of the workstation, a camera has been placed to record the experiment (item 1 Fig. 4A), and an artificial lightning system has been introduced to guarantee a constant brightness of the environment (item 2 Fig. 4A).

The operator is equipped with a Tobii Pro Glasses 2 device to track its gaze. The system is composed of a head unit (item 11 of Fig. 4B), which retrieves the gaze behavior, and a recording unit (item 12 of Fig. 4B), which streams the data to the control system exploiting the open-source library detailed in[74]. To calibrate the glasses, the marker provided by the manufacturer (Tobii) has been fixed to the collaborative workstation (item 3 of Fig. 4A).

Close to the assembly workspace, an isolated area where the participants filled the questionnaires have been prepared (Fig. 4D).

The wooden box developed for the experiment (Fig. 4C) has been designed with a structure and an assembly procedure similar to the one of a bedside table to test the control methods in realistic (almost) industrial scenarios. The procedure to assemble the wooden box consists of pick-and-place, screwing, labeling, and manual assembly operations. Some of them are executed in collaboration with the robot (e.g. screwing), and

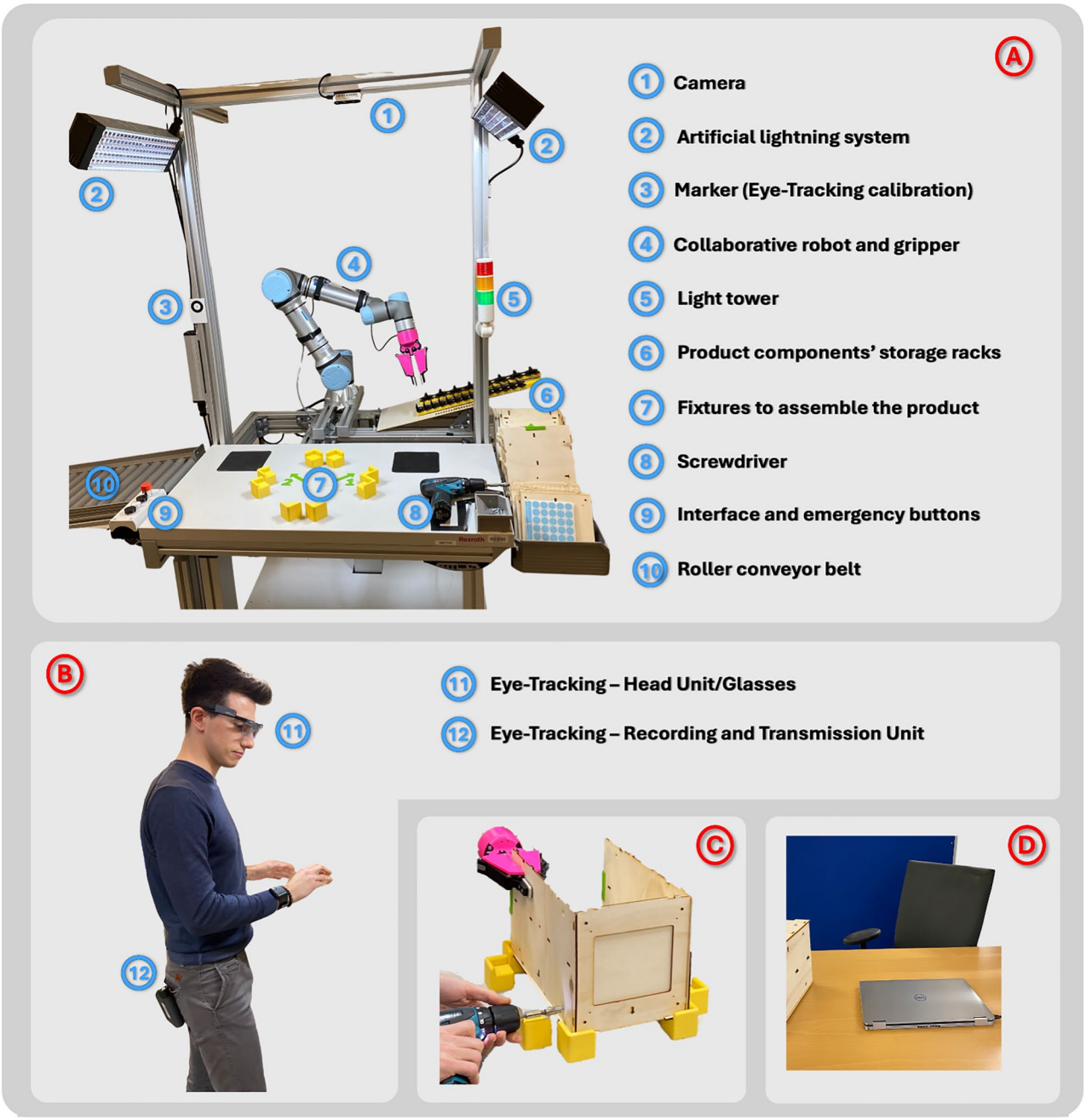

Fig. 4. in (**A**) representation of the collaborative workstation devoted to the assembly of the wooden box (which is shown in (**C**)). In (**B**), instead, it is possible to observe a picture of the operator and its equipment during the experiment, while, in (**D**) an image of the desk used by the participants to fill the surveys is reported.

some others in a cooperative fashion (e.g. pick-and-place). The assembly procedure has been designed with the goal of balancing those two different ways of interacting with the robot. The collaborative steps to assemble the products are described in Fig. S1 (Additional Material).

A risk assessment has been performed to prevent and mitigate potential mechanical risks. The workstation features and HRI modalities have been designed accordingly. For S1, the nominal speed of the robot has been defined by considering the requirements of ISO/TS 15066. Hands, wrists, and forearms are considered the main human body parts that can be involved in potential non-functional contact. The implementation of Power and Force Limiting criterion led to the definition of a maximum velocity associated with each different robot tasks, that have been imposed as upper bounds of the robot's velocity. In particular, general movements and pick and place operations are characterized by a maximum nominal speed equal to $0.73\frac{m}{s}$, while the collaborative operations, where the robot holds the wall of the box are executed by the robot at a maximum speed of $0.42\frac{m}{s}$.

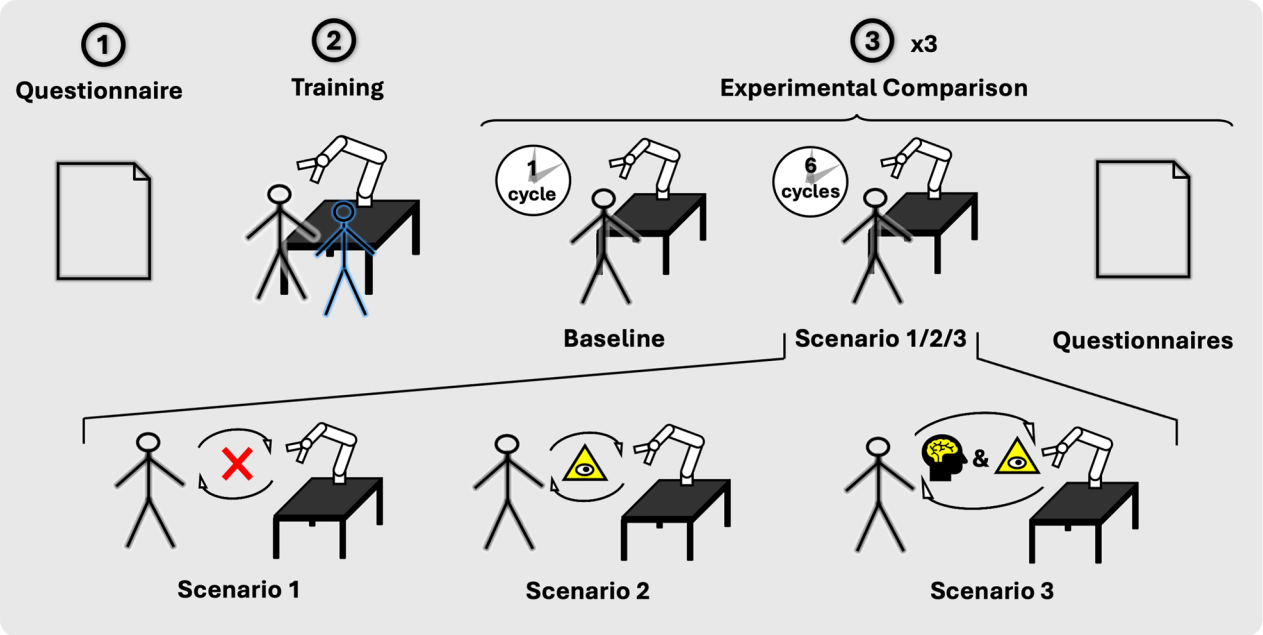

Fig. 5. schematic representation of the experimental procedure.

Considering S2, the maximum robot's speed increase has been set to $\alpha_{max} = 2$, while, in S3 the maximum acceptable velocity limit has been set to $\alpha_{limit} = 3$. Even if these values exceed those recommended by ISO TS 15066, the conducted risk assessment and applied safety measures confirms that risk levels remain tolerable and thus that the collaborative application is safe.

Experimental procedure

The study employed a within-subjects experimental design. The experiment and all methods were performed in accordance with the relevant guidelines and regulations. As participants arrived at the experimental site, they were provided with information about the study's objectives, procedures, safety protocols, and data management procedures (see Fig. 4D). Then, informed consent was obtained from all the study participants as well as consent to the retrieval of their personal data.

Once the bureaucratic steps have been completed, participants were asked to complete a questionnaire to assess their interest in innovations and technological developments, their self-assessed DIY and manual work skills, and their familiarity with collaborative robots. Additionally, perceived self-efficacy in HRI was measured at this stage (Fig. 5, point 1).

Subsequently, participants underwent a training session to familiarize themselves with the assembly task, the the workstation elements/equipment, and the robotic system (Fig. 5, point 2). At this stage, the functionalities of the systems were explained first. To ensure clear communication, operators were informed of the indicators of the robot's status, i.e. tower light, and of the push-button at their disposal to let the robot advance to its next tasks. Due to safety reasons, participants were informed that, in some scenarios, the behavior of the robot might change according to their gaze motions, but no explanations of the control laws were provided. Then, the assembly procedure needed to assemble the wooden box was explained. Finally, participants constructed the wooden box at least 6 times until they were familiarized with the operation. In this stage, the robot moved along its reference trajectory, i.e. at its nominal speed, and no adaptive algorithm has been deployed.

After the training phase, each participant performed the assembly task in three different randomized scenarios (Fig. 5, point 3). Every scenario had the following structure:

- Baseline assembly: the wooden box is assembled one time in 'standard' conditions, i.e. with the robot moving along its nominal trajectory. This step is required to measure the baseline values required to compute the metrics as well as the parameters required to implement *Method 2*.
- Experimental Scenario: at this stage, the wooden box is assembled by the operator under each different randomized condition 6 times. In particular, the three possible different conditions are: (S1) the robot moves along its reference trajectory, i.e. no adaptive control algorithm is applied, (S2) the robot is controlled *Method 1*, i.e. the speed is adjusted according to the visual attention of the operator, while, (S3) the robot's behavior was adjusted according *Method 2*, i.e. the trajectory of the robot was changed based on both the cognitive workload as well as on the visual attention of the participant.
- Questionnaires: the participant was asked to fill out the questionnaires to evaluate the effectiveness of the control method.

It is important to remark that the order of the three different scenarios has been randomized for every participant so as to avoid possible dependencies of the results on the order in which the control methods were deployed. In particular, if the scenarios weren't randomized, the results could have been affected by the learning effect, increased familiarity with the system behavior, and fatigue of the participant. Such an experimental procedure has been adopted since it is among the standard and state-of-the-art procedures to compare different experimental scenarios in human-robot interaction studies[17,75–77]. Moreover, the learning and fatigue effect has been further mitigated by the selection of a suitable number of training cycles. In fact, during the training stage, the participant repeated the operation 6 times and, at the end, the user could execute additional training cycles if he was not already familiar with the operation. No participant requested to perform such additional training. The number of training cycles has been defined during a pilot study, where it has been evaluated that three cycles were sufficient for all users participating in the pilot study to complete the learning procedure. Following a more conservative approach, during the experiment, it has been decided to set the standard number of training cycles to six.

Statistical analysis

The choice of statistical analysis-either ANOVA or GLM-was guided by the specific research questions and data characteristics. ANOVA was employed for straightforward between-group comparisons where fixed effects were the primary focus (i.e. Productivity, Cognitive Workload, Reliability), ensuring clarity and interpretability. Conversely, GLM was used for analysis involving repeated measures, covariates, or more complex data structures that required additional flexibility (i.e. Fluency, Usability, Acceptance). For example, GLM facilitated the inclusion of subject-specific variability and corrections for sphericity violations in repeated-measures data, which would not have been possible using a simpler ANOVA framework. This combination of approaches allowed us to robustly address the varied demands of our experimental design while adhering to best practices in statistical modeling.

Pupil dilation measurements

Changes in pupil dilation are used to assess the operator's cognitive workload and to adjust the robot's speed accordingly. Pupil dilation measured with an eye tracker is a valid and widely used method for assessing cognitive workload. This is evidenced by various studies that have observed correlations between task difficulty and pupil dilation[49,51,78]. In particular, increased cognitive workload is associated with greater dilation of the pupils[50]. It is important to note that factors unrelated to workload, such as luminance, emotional state, and the specific biological peculiarities of each individual, might affect this measure[79]. Two countermeasures have been taken to avoid the influence of the two aforementioned aspects and ensure the correct estimation of the cognitive workload. Regarding the influence of brightness, the workstation has been illuminated employing an artificial lighting system to ensure a constant luminance of the environment, thus making the pupil diameter as much as possible independent of the ambient brightness. For the operator's emotional state and its biological characteristics, before every trial, a baseline measurement of the operator's pupil diameter is taken[80,81]. Such a baseline consists of a reference value that accounts for the operator's state and the anatomical dimension of its pupil. To compensate for potential interference due to such factors, instead of considering the absolute pupil diameter, its ratio is considered with respect to the baseline value, i.e., the nominal working condition.

When considering the robot's velocity adjustment, pupil dilation acts as an independent variable. Changes in pupil size, as a physiological response, indicate variations in the operator's cognitive workload. The robot's velocity is then adjusted in response to these changes, making pupil dilation a driving factor (independent variable) for the robot's behavior. On the other hand, pupil dilation can also be viewed as a dependent variable, especially when analyzing the effects of external factors on the operator. In our case, we are assessing how the operator is performing an assembly task in collaboration with a robot. Factors like task complexity, or the robot's behavior itself, might influence the operator's cognitive workload, which in turn could be reflected in their pupil size. In this case, pupil dilation is the outcome (dependent variable) influenced by the task, environment, and interaction dynamics. For this reason, being pupil dilation both a dependent and an independent variable, it has not been used in the statistical analysis, which would require it to be an independent variable.

Received: 7 January 2025; Accepted: 24 December 2025
Published online: 09 March 2026

References

1. Ajoudani, A. et al. Progress and prospects of the human–robot collaboration. *Auton Robots***42**(5), 957–975 (2018).
2. Wang, L. et al. Symbiotic human-robot collaborative assembly. *CIRP Ann***68**(2), 701–726 (2019).
3. Zafar, M. H., Langås, E. F. & Sanfilippo, F. Exploring the synergies between collaborative robotics, digital twins, augmentation, and industry 5.0 for smart manufacturing: A state-of-the-art review. *Robot Comput-Integr Manuf***89**, 102769 (2024).
4. Gualtieri, L. et al. Updating design guidelines for cognitive ergonomics in human-centred collaborative robotics applications: An expert survey. *Appl Ergon***117**, 104246 (2024).
5. Ajoudani, A. et al. Smart collaborative systems for enabling flexible and ergonomic work practices. *IEEE Robot Autom Mag***27**(2), 169–176 (2020).
6. Lasota, P. & Shah, J. A. Analyzing the effects of human-aware motion planning on close-proximity human-robot collaboration. *Human Factors***57**(1), 21–33 (2015).
7. Villani, V., Pini, F., Leali, F. & Secchi, C. Survey on human–robot collaboration in industrial settings: Safety, intuitive interfaces and applications. *Mechatronics***55**, 248–266 (2018).
8. Hopko, S. K., Khurana, R., Mehta, R. K. & Pagilla, P. R. Effect of cognitive fatigue, operator sex, and robot assistance on task performance metrics, workload, and situation awareness in human-robot collaboration. *IEEE Robot Autom Lett***6**(2), 3049–3056 (2021).

9. Peternel, L., Tsagarakis, N., Caldwell, D. & Ajoudani, A. Robot adaptation to human physical fatigue in human–robot co-manipulation. *Auton Robots***42**(5), 1011–1021 (2018).
10. Campagna, G., Chrysostomou, D. & Rehm, M. Investigating electrodermal activity for trust assessment in industrial human-robot collaboration, In: *21st International Conference on Ubiquitous Robots (UR 2024), IEEE* (2024).
11. Kahneman, D. *Thinking, fast and slow* (Straus & Giroux Inc, Published by Farrar, 2013).
12. Witte, T.E., Hasbach, J., Schwarz, J. & Nitsch, V. Towards iteration by design: An interaction design concept for safety critical systems, In: *Adaptive Instructional Systems: Second International Conference, AIS 2020*, Held as Part of the 22nd HCI International Conference, HCII 2020, Copenhagen, Denmark, July 19–24, 2020, Proceedings 22, pp. 228–241, Springer International Publishing (2020).
13. Liu, F. et al. Printed synaptic transistor–based electronic skin for robots to feel and learn. *Sci Robot***7**(67), abl7286 (2022).
14. Huang, Y. et al. A skin-integrated multimodal haptic interface for immersive tactile feedback. *Nat Electron***6**(12), 1020–1031 (2023).
15. Luo, Y. et al. Adaptive tactile interaction transfer via digitally embroidered smart gloves. *Nat Commun***15**(1), 868 (2024).
16. Dehais, F., Sisbot, E. A., Alami, R. & Causse, M. Physiological and subjective evaluation of a human-robot object hand-over task. *Appl Ergon***42**(6), 785–791 (2011).
17. Capponi, M., Gervasi, R., Mastrogiacomo, L. & Franceschini, F. Assembly complexity and physiological response in human-robot collaboration: Insights from a preliminary experimental analysis. *Robot Comput-Integr Manuf***89**, 102789 (2024).
18. Kulíc, D. & Croft, E. Affective state estimation for human-robot interaction. *IEEE Trans Robot***23**(5), 991–1000 (2007).
19. Rudovic, O., Lee, J., Dai, M., Schuller, B. & Picard, R. W. Personalized machine learning for robot perception of affect and engagement in autism therapy. *Sci Robot***3**(19), eaao6760 (2018).
20. Toisoul, A., Kossaifi, J., Bulat, A., Tzimiropoulos, G. & Pantic, M. Estimation of continuous valence and arousal levels from faces in naturalistic conditions. *Nat Mach Intell***3**(1), 42–50 (2021).
21. Sirintuna, D., Kastritsi, T., Ozdamar, I., Gandarias, J. M. & Ajoudani, A. Enhancing human–robot collaborative transportation through obstacle-aware vibrotactile warning and virtual fixtures. *Robot Auton Syst***178**, 104725 (2024).
22. Casalino, A. et al. Operator awareness in human-robot collaboration through wearable vibrotactile feedback. *IEEE Robot Autom Lett***3**(4), 4289–4296 (2018).
23. Hietanen, A., Pieters, R., Lanz, M., Latokartano, J. & Kämäräinen, J. K. AR-based interaction for human-robot collaborative manufacturing. *Robot Comput-Integr Manuf***63**, 101891 (2020).
24. St. Clair, A., Mataric, M. How robot verbal feedback can improve team performance in human-robot task collaborations, In: *ACM/IEEE International Conference on Human-Robot Interaction*, 2015-March, pp. 213–220 (2015).
25. Scalise, R., Li, S., Admoni, H., Rosenthal, S. & Srinivasa, S. S. Natural language instructions for human–robot collaborative manipulation. *Int J Robot Res***37**(6), 558–565 (2018).
26. Tang, G., Webb, P. & Thrower, J. The development and evaluation of robot light skin: A novel robot signalling system to improve communication in industrial human–robot collaboration. *Robot Comput-Integr Manuf***56**, 85–94 (2019).
27. Lorenzini, M., Kim, W. & Ajoudani, A. An online multi-index approach to human ergonomics assessment in the workplace. *IEEE Trans Human-Mach Syst***52**(5), 812–823 (2022).
28. Lagomarsino, M., Lorenzini, M., De Momi, E. & Ajoudani, A. An online framework for cognitive load assessment in industrial tasks. *Robot Comput-Integr Manuf***78**, 102380 (2022).
29. Lagomarsino, M., Lorenzini, M., Balatti, P., De Momi, E. & Ajoudani, A. Pick the right co-worker: Online assessment of cognitive ergonomics in human-robot collaborative assembly. *IEEE Trans Cognit Dev Syst***15**(4), 1928–1937 (2023).
30. Vignais, N. et al. Innovative system for real-time ergonomic feedback in industrial manufacturing. *Appl Ergon***44**(4), 566–574 (2013).
31. Kim, W., Ruiz, Garate V., Gandarias, J. M., Lorenzini, M. & Ajoudani, A. A directional vibrotactile feedback interface for ergonomic postural adjustment. *IEEE Trans Haptics***15**(1), 200–211 (2022).
32. Peternel, L., Fang, C., Tsagarakis, N. & Ajoudani, A. A selective muscle fatigue management approach to ergonomic human-robot co-manipulation. *Robot Comput-Integr Manuf***58**, 69–79 (2019).
33. Tassi, F., De Momi, E. & Ajoudani, A. An adaptive compliance hierarchical quadratic programming controller for ergonomic human–robot collaboration. *Robot Comput-Integr Manuf***78**, 102381 (2022).
34. Messeri, C., Bicchi, A., Zanchettin, A. M. & Rocco, P. A dynamic task allocation strategy to mitigate the human physical fatigue in collaborative robotics. *IEEE Robot Autom Lett***7**(2), 2178–2185 (2022).
35. Mitsunaga, N., Smith, C., Kanda, T., Ishiguro, H. & Hagita, N. Adapting robot behavior for human-robot interaction. *IEEE Trans Robot***24**(4), 911–916 (2008).
36. Villani, V., Sabattini, L., Secchi, C. & Fantuzzi, C. A Framework for affect-based natural human-robot interaction, In: *RO-MAN 2018 - 27th IEEE International Symposium on Robot and Human Interactive Communication*, pp. 1038–1044, 8525658 (2018).
37. Shah, J. K., Yadav, A., Hopko, S. K., Mehta, R. K. & Pagilla, P. R. Robot adaptation under operator cognitive fatigue using reinforcement learning. 1467–1474 (IEEE International workshop on robot and human communication, RO-MAN, 2023).
38. Messeri, C., Masotti, G., Zanchettin, A. M. & Rocco, P. Human-robot collaboration: Optimizing stress and productivity based on game theory. *IEEE Robot Autom Lett***6**(4), 8061–8068 (2021).
39. Lagomarsino, M., Lorenzini, M., De Momi, E. & Ajoudani, A. Robot trajectory adaptation to optimise the trade-off between human cognitive ergonomics and workplace productivity in collaborative tasks, In: *IEEE International Conference on Intelligent Robots and Systems*, 2022-October, pp. 663–669 (2022).
40. Wong, C. Y., Vergez, L. & Suleiman, W. Vision-and tactile-based continuous multimodal intention and attention recognition for safer physical human-robot interaction. *IEEE Trans Autom Sci Eng.***21**(3), 3205–3215 (2023).
41. Duncan, J. Selective attention and the organization of visual information. *J Exp Psychol: Gen***113**(4), 501–517 (1984).
42. Evans, K. K. et al. Visual attention. *Wiley interdiscip Rev: Cognit Sci***2**(5), 503–514 (2011).
43. McMains, S. & Kastner, S. Visual attention. In *Encyclopedia of neuroscience* (eds Binder, M. D. et al.) (Springer, Berlin, 2009).
44. Hasanzadeh, S., Esmaeili, B. & Dodd, M. D. Impact of construction workers' hazard identification skills on their visual attention. *J Construct Eng Manag***143**(10), 04017070 (2017).
45. Aroke, O., Hasanzadeh, S., Esmaeili, B., Dodd, M. D. & Brock, R. Using worker characteristics, personality, and attentional distribution to predict hazard identification performance: A moderated mediation analysis. *J Construct Eng Manag***148**(6), 04022033 (2022).
46. International Organization for Standardization, *ISO/TS 15066:2016. Robots and robotic devices — Collaborative robots.*
47. Hart, S. G. & Staveland, L. E. Development of NASA-TLX (task load index): Results of empirical and theoretical research. *Adv Psychol***52**(C), 139–183 (1988).
48. Van Acker, B. Mental workload monitoring in the manufacturing industry: conceptualisation, operationalisation and implementation, Doctoral dissertation, Ghent University (2020).
49. Klingner, J., Kumar, R. & Hanrahan, P. Measuring the task-evoked pupillary response with a remote eye tracker, In: Proceedings of the 2008 Symposium on Eye tracking research & Applications, (2008).
50. Knust, S.R., Marshall, S.P. & Ishizaka, K. Using eye activity to study cognitive processes underlying individual differences, In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, **44(17)**, pp. 129–132, Sage CA: Los Angeles, CA: SAGE Publications (2000).

51. Jiang, X., Zheng, B., Bednarik, R. & Atkins, M. S. Pupil responses to continuous aiming movements. *Int J Human-Comput Stud***83**, 1–11 (2015).
52. Thorvald, P., Lindblom, J. & Andreasson, R. On the development of a method for cognitive load assessment in manufacturing. *Robot Comput-Integr Manuf***59**, 252–266 (2019).
53. Story, M. et al. Do speed and proximity affect human-robot collaboration with an industrial robot arm?. *Int J Soc Robot***14**(4), 1087–1102 (2022).
54. Huang, C. M., Cakmak, M. & Mutlu, B. Adaptive coordination strategies for human-robot handovers. *Robot: Sci Syst***11**, 1–10 (2015).
55. Shiferaw, B., Downey, L. & Crewther, D. A review of gaze entropy as a measure of visual scanning efficiency. *Neurosci Biobehav Rev***96**, 353–366 (2019).
56. Holmqvist, K. et al. *Eye tracking : A comprehensive guide to methods and measures* (Oxford University Press, Oxford, 2011).
57. Di Stasi, L. L. et al. Catena, gaze entropy reflects surgical task load. *Surg Endosc***30**(11), 5034–5043 (2016).
58. Di Stasi, L. L. et al. Quantifying the cognitive cost of laparo-endoscopic single-site surgeries: Gaze-based indices. *Appl Ergon***65**, 168–174 (2017).
59. Paliga, M. & Pollak, A. Development and validation of the fluency in human-robot interaction scale. *A two-Wave Study on Three perspect Fluency, Int J Human-Comput Stud***155**, 102698 (2021).
60. Brooke, J. SUS: A "quick and dirty" usability scale. In *Usability evaluation in industry* (eds Jordan, P. et al.) 189–194 (Taylor Francis, Milton Park, 1996).
61. Hopko, S. K., Zhang, Y., Yadav, A., Pagilla, P. R. & Mehta, R. K. Brain-behavior relationships of trust in shared space human-robot collaboration. *ACM Trans Human-Robot Interact***13**(1), 5 (2024).
62. Van Der Laan, J. D., Heino, A. & De Waard, D. A simple procedure for the assessment of acceptance of advanced transport telematics. *Transp Res Part C: Emerg Technol***5**(1), 1–10 (1997).
63. Paliga, M. The relationships of human-cobot interaction fluency with job performance and job satisfaction among Cobot operators—the moderating role of workload. *Int J Environ Res Public Health***20**(6), 5111 (2023).
64. Lorenzini, M., Lagomarsino, M., Fortini, L., Gholami, S. & Ajoudani, A. Ergonomic human-robot collaboration in industry: A review. *Front Robot AI***9**, 813907 (2023).
65. Landi, C. T. et al. Relieving operators' workload: Towards affective robotics in industrial scenarios. *Mechatronics***54**, 144–154 (2018).
66. Khan, A. Human-robot interaction: Designing intuitive interfaces for automation. *Front Robot Autom***1**(02), 89–110 (2024).
67. Halme, R. J. et al. Review of vision-based safety systems for human-robot collaboration. *Proced CIRP***72**, 111–116 (2018).
68. Ghazizadeh, M., Lee, J. D. & Boyle, L. N. Extending the technology acceptance model to assess automation. *Cognit, Technol Work***14**, 39–49 (2012).
69. Masayoshi, K., Satoru, S. & Tomonori, Y. Modeling of handling motion reflecting emotional state and its application to robots, In: *SICE Annual Conference* (2008).
70. Rojas, R. A., Garcia, R., Wehrle, M. A. & Vidoni, R. E. A variational approach to minimum-jerk trajectories for psychological safety in collaborative assembly stations. *IEEE Robot Autom Lett***4**(2), 823–829 (2019).
71. Longo, L., Wickens, C. D., Hancock, P. A. & Hancock, G. M. Human mental workload: A survey and a novel inclusive definition. *Front Psychol***13**, 883321 (2022).
72. Ahlstrom, U. & Friedman-Berg, F. J. Using eye movement activity as a correlate of cognitive workload. *Int J Ind Ergon***36**(7), 623–636 (2006).
73. Van Orden, K. F., Limbert, W., Makeig, S. & Jung, T. P. Eye activity correlates of workload during a visuospatial memory task. *Human Factors***43**(1), 111–121 (2001).
74. De Tommaso, D. & Wykowska A. TobiiglassespySuite: An open-source suite for using the Tobii Pro Glasses 2 in eye-tracking studies, In: *Eye Tracking Research and Applications Symposium (ETRA)*, a46 (2019).
75. Mitsuoka, R., Kubota, T., Sato, S. & Ogawa, K. Semi-autonomous touch method merging robot's autonomous touch and user-operated touch for improving user experience in robot touch. *Sci Rep***14**, 17667 (2024).
76. Belkaid, M., Kompatsiari, K., De Tommaso, D., Zablith, I. & Wykowska, A. Mutual gaze with a robot affects human neural activity and delays decision-making processes. *Sci Robot***6**, eabc5044 (2021).
77. Rosen, E. et al. Communicating and controlling robot arm motion intent through mixed-reality head-mounted displays. *Int J Robot Res***38**(12–13), 1513–1526 (2019).
78. Samara, A., Galway, L., Bond, R. & Wang, H. Tracking and evaluation of pupil dilation via facial point marker analysis, In: IEEE *International Conference on Bioinformatics and Biomedicine (BIBM)*, pp. 2037–2043 (2017).
79. Xu, J., Wang, Y., Chen, F. & Choi E. Pupillary response based cognitive workload measurement under luminance changes, In: *IFIP Conference on Human-Computer Interaction*, pp. 178–185, Springer Berlin Heidelberg (2011).
80. Mathôt, S., Fabius, J., Van Heusden, E. & Van der Stigchel, S. Safe and sensible preprocessing and baseline correction of pupil-size data. *Behav Res Methods***50**(1), 94–106 (2018).
81. Steinhauer, S. R., Bradley, M. M., Siegle, G. J., Roecklein, K. A. & Dix, A. Publication guidelines and recommendations for pupillary measurement in psychophysiological studies. *Psychophysiology***59**(4), e14035 (2022).

Acknowledgements

This research was funded by the European Commission, Horizon Europe project SESTOSENSO - Physical Intelligence for Smart and Safe Human-Robot Interaction (Grant Agreement n. 101070310) and by the Italian Doctorate in Robotics and Intelligent Machines.

Author contributions

MM, LG and RV developed the method; MM, FF, SM and LG conceived the experimental design; MM, FF, SM and LG conducted the experiment, MM and FF analyzed the results, LP and RV supervised the research activities. All authors contributed to the writing process and revision of the manuscript.

Declarations

Competing interests

The authors declare no competing interests.

Ethical approval

The study strictly adhered to ethical principles and legislation, encompassing the EU Charter of Fundamental Rights and GDPR. The data retrieval and processing procedure has been reviewed and approved by the Data

Protection Officer of the Free University of Bolzano, while the experimental procedure has been evaluated and approved by the Research Ethics Committee of the Free University of Bolzano (prot. "HRAC_Code_2023_31"). Participants have been informed on both the structure and risks associated with the experiment as well as with the data treatment. The participant's consent to take part in the experiment and to allow the collection of their personal data is an essential and binding requirement. Without this consent, participation in the study and the gathering of any personal information was not considered.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-34142-9.

Correspondence and requests for materials should be addressed to L.G. or R.V.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.