Good Things Come in Threes: The Impact of Robot Responsiveness on Workload and Trust in Multi-User Human-Robot Collaboration

Francesco Semeraro[1], Jon Carberry[2], James Leadbetter[2] and Angelo Cangelosi[1], *Senior Member, IEEE*

***Abstract*— Human-robot collaboration has the potential of unlocking new manufacturing paradigms thanks to the introduction of a robotic architecture in a production chain that involves human workers. A possible innovative declination of this is the use of collaborative robots to enable two workers to concurrently act on the same manufacturing target without causing mutual disturbances. By doing so, the efficiency of the process would be preserved while reducing the production times. This work designs a physical collaborative task that involves two users and one collaborative robot. The users act in the scenario in a concurrent way on the same target object, while the robot physically intervenes in the scene as a mediator by adjusting the position and orientation of the object to accommodate both users at the same time. Through this experimental setup, 78 apprentices and teachers of the BAE Systems Academy for Skills and Knowledge Centre were recruited to investigate the users' perception of the task workload and trust towards the robotic system. Specifically, they performed the same task under two experimental conditions, in which the robot responded to changes in the interaction in a reactive or timed way, respectively. The statistical analysis showed that a timed response of the robot was associated with lower perceived workload and higher predictability of the system.**

Fig. 1. Experimental setup of the multi-user human-robot collaboration explored. The users need to perform different operations at the same time on a shared target object. The robot adjusts position and orientation of the object according to the progression of the interaction. Behind the user on the right, there was another table with a box on it (see Subsection III-A).

I. INTRODUCTION

Advances in human-robot interaction led to a growing research interest in interaction scenarios that go beyond the classic dyad of one user and one robot [1]. There have been case studies of non-dyadic human-robot interaction reported in literature [2]. However, scenarios of this kind for physical human-robot collaboration (HRC) applications are still largely unexplored [3]. A potential application of these scenarios involves the manufacturing sector. In serial production chains, one worker at a time performs manufacturing operations on a target object. Enabling two workers to operate at the same time on it could allow us to parallelise sub-steps of manufacturing processes. This can consequently pave the way for innovative modalities of manufacturing with reduced takt time [4], [5], [6].

The presence of a second concurrent user in the manufacturing process heightens the complexity of the interaction [7]. Indeed, although every worker would always perform their own same set of operations on the object, the timings between different sub-steps of their sequences are susceptible to variations. Consequently, workers might often come across situations in which the target object is in configurations that they do not expect, due to the mutual disturbances they may cause to each other. This could cause delays in the completion of the manufacturing activity. These delays would be further worsened if the workers tried to time themselves to each other or if they had to continuously come to arrangements to mediate the target object between themselves [8]. In these cases, the cognitive workload of the task would increase, while it should be kept at a minimum in procedural steps like the ones of a manufacturing process.

The introduction of a collaborative robot in this scenario could mitigate these mutual disturbances (see Fig. 1) [4]. Acting as a mediator in the collaboration, the robot could adjust the target object to serve the requirements of both users according to their current state of the interaction [9]. By doing so, the workers can focus on the procedural part of the manufacturing and delegate any high-level reasoning to the robot. While other works focus on switching the robotic assistance between the users through task allocation processes

Corresponding author: Francesco Semeraro

Francesco Semeraro's work was supported by the UKRI DTP CASE-conversion "Human-Robot Collaboration for Flexible Manufacturing" (Ref. 2480772), sponsored by UKRI Engineering and Physical Sciences Research Council and BAE Systems plc.

Angelo Cangelosi's work was partially supported by the H2020 project TRAINCREASE (Ref. 210638710), the UKRI Trustworthy Autonomous Systems Node on Trust (Ref. EP/V026682/1) and the Horizon Projects TRAIL and eTALK.

[1]Francesco Semeraro and Angelo Cangelosi are with the Manchester Centre for Robotics and AI, The University of Manchester, Manchester, UK. E-mails: francesco.semeraro@manchester.ac.uk angelo.cangelosi@manchester.ac.uk

[2]Jon Carberry and James Leadbetter are with BAE Systems (Operations) Ltd., Warton, UK. E-mails: jon.carberry@baesystems.com james.leadbetter@baesystems.com

[10], in this scenario the robot considers the requirements of both workers at the same time, thus increasing the scalability of the HRC and the throughput of the process [11].

In this work, a collaborative assembly [3] that involves two users in the same scenario interacting with the same target object was set up. In this, the robot addresses the queries of the users at hand through direct physical intervention in the scene, modifying the configuration of the target object according to their states. In such a setup, users find themselves interacting with a collaborative robot in an unprecedented scenario. Such novel situation poses challenges regarding the design of the robotic system to improve the working conditions of the users. Different settings of the robotic system could induce a higher or lower workload of the task from the users [12]. They could also affect the trust the users would put in a robotic system of this kind. Trust plays a key role in the success of the collaboration: if users trust the robotic agent, they rely on the robot fulfilling its task, and focus solely on their own part of the manufacturing, leading to an improvement of the task performance [13].

This raises the question of how the collaborative robot should handle the concurrency of the sub-steps of different sets of operations. In the depicted scenario, there could be cases in which one user queries the robot to change the configuration of the object while the other user is already working on it [4], [7]. In this case, the best mediation strategy the robot should adopt could be different from the perspective of the two users. The user querying the robot would prefer an immediate response of the robot, but this could cause stress to the other worker, who would not have time to adjust to the sudden change in the object configuration, and would prefer to be given enough time to prepare themselves to the change. However, this could cause frustration to the user querying the robot, as this would cause delays in their part of the assembly process. Utilising the aforementioned setup, this work investigates through a user study how the responsiveness of the robot in these types of situations affects the users. More precisely, we surveyed how the timing of robot responses during conflictual situations, accompanied by a verbal explanation of the action about to be performed by the robot, could influence the users' perceived workload of the collaborative task and their trust towards the robotic system. For this purpose, we collaborated with the BAE Systems Academy for Skills and Knowledge (ASK) Centre, where apprentices are trained in manufacturing skills. Therefore, exploring these scenarios with such a pool of participants allowed us to get evaluations from workers who could potentially benefit from such a setup in future applications.

The contributions of this paper are as follows:

- *collaborative task setup* involving two users and one robotic manipulator; here, the users performed concurrent tasks on the same target object, which the robot had to physically move to accommodate the needs of both users at the same time, according to the ongoing state of the interaction;
- *user study* investigating the users' perception of workload and trust towards the robotic system in the depicted collaborative scenario under different levels of robot responsiveness to conflictual events in the collaboration.

This paper is structured as follows. Section II contextualises the work in the current literature. Section III describes the experimental design and the related user study. Section IV carries out a statistical analysis on the self-reported measurements probed during the study, discussed in Section V. Section VI summarises the main findings of this work.

II. RELATED WORK

There have been previous attempts in literature of setting up non-dyadic human-robot interactions [2]. However, there are very few cases in which a physical collaborative task was undertaken. Tsamis et al. [14] explored the benefits of using augmented reality in multi-user HRC. In their setup, a mobile manipulator collaborated with one of the users towards the same goal. In doing so, the second user passed by and was showed information regarding the behaviour of the robot in real time through an augmented reality device. In this case, there was coexistence of the users, but not actual collaboration [15].

Fu et al. [16] assessed the influence of affective robot behaviour on a non-dyadic collaborative task. In this, two users worked together to achieve a common goal. The robot produced an assessment of their performance during the task by means of facial expressions and gaze patterns. However, the robot did not intervene directly in the collaboration by physically affecting the workspace. Interestingly, there was a hierarchical relationship between the users, with one leading the other in the task by giving instructions. Therefore, one of the user was not directly intervening in the setup.

Faria et al. [17] explored a different multi-user collaborative task. They investigated the users' perception of the robot movements when three users had to collaborate with it in a mutually concurrent way. In this case, the users did not actively convey any request to the robot, which chose whom to focus on at a time without explicitly informing the users regarding its actions. In other terms, the robot was not sensitive to any inputs from the users that would enable it to change the outcome according to the situation, nor did it communicate its behaviour to them.

Schneiders et al. [9], [18] have started exploring the potential use of non-dyadic HRC for industrial tasks [19]. However, they have performed studies of interactions in collaboration of triads of participants performing pick-and-place tasks. This study drew conclusions regarding the interaction design that could be transferred to a scenario that replaces one of the users with a collaborative robot.

In this paper, we designed a multi-user human-robot collaboration task in which two users perform tasks in a concurrent way in collaboration with a robotic manipulator in the same shared workspace. There is no hierarchy between the users and they both directly affect the scenario. The robot addresses the needs of both users to best accommodate them at the same time, to enable them to work efficiently despite the concurrency of their tasks.

III. MATERIALS AND METHODS

This section describes the technical contributions of the paper. First, the task specifics are outlined. To the best of our knowledge, there is no previous attempt in literature at designing a concurrent human-robot collaborative task that involves two users performing different physical operations at the same time on the same target object. Then, the performed user study, which made use of such a testbed, is described.

A. Collaborative task setup

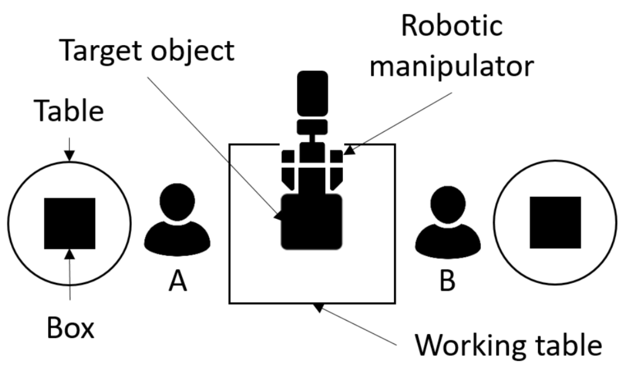

Fig. 2. Schematics of the experimental setup of the multi-user human-robot collaboration task explored. For its real-life deployment, see Fig. 1.

In the experimental setup, two different locations, namely A and B, at opposite sides of a working table, were associated to a different set of operations for the participants to perform on the same target object at the same time (see Fig. 2). The collaborative robot constantly held the target object throughout the task and hovered it above the working table. The sequence of operations associated with each location are summarised in Table I. The steps and their estimated duration were chosen so that the participants could experience all the possible states in an interaction of this kind, i.e. working on the object, preparing themselves to the next step of the sequence and requesting the robot to change the configuration of the object [4], [7]. The participants had to start together, but were free to move on to the next phase of the assigned sequence as soon as they completed the previous one, without worrying about the other participant. From this point on, the ensemble of the two sequences of operations will be referred to as the task of the experiment. The target object was built so that each one of the operations had to be performed on a different face of it (see Fig. 1). Because of this, the object was not always properly oriented towards the user to allow them to perform the operation at hand comfortably. When this happened, each participant was instructed to look at the object and raise a hand at shoulder height to request a change in the position and orientation of the object. This would result in the robot moving the object to a different pose, according to the case. Once the information regarding which step both participants were at in their respective sequence of operations reached the robot through the experimenter's command, the robot moved the

TABLE I
SUMMARISED SEQUENCES OF OPERATIONS TO PERFORM ON THE OBJECT ACCORDING TO THE ASSIGNED LOCATION.

LOCATION A	LOCATION B
Turn the top gear in front of you 5 times	*Move the beads at the top of the object from one end to the other of both wires one by one*
Collect the sphere and the stick from inside the box in location A and screw them together	*Fetch the shapes from inside the box in location B*
Use the previously assembled tool to play the tiles in front of you 25 times each	*Place the shapes inside the object through the related holes*

object to the associated pose. Several poses, according to every possible cases the interaction with the users during the task could unfold, were recorded beforehand. In terms of position, the object was placed above the middle of the table when shared between the users and closer to a user when the robot was focusing solely on that user. As for the orientation, the side of the object on which the user needed to perform the operation at hand was set to face the user. When the object had to be shared, the orientation was set in between the ones needed to align the specific sides of interest with the users.

Fig. 3. Sequences of operations performed by the two users. For the user in location A (top), from left to right: rotating gears, screwing stick with sphere, playing tiles. For the user in location B (bottom): moving beads, fetching shapes, placing shapes. In both cases, the second operation was performed by the user while facing its associated table.

Although the sequences differed from one another, they shared a common structure (see Fig. 3). Each sequence started with a task that could be performed with bare hands. After this, each participant had to turn around and fetch objects from the box associated to their location, that were needed to perform the next operation. After having completed the second operation, the task finished for that participant. In terms of structural differences, instead, the first operation associated to location A was designed to last significantly longer than the first one to perform in location B, whereas the second operation related to B lasted longer than the second one in A. This design choice was done

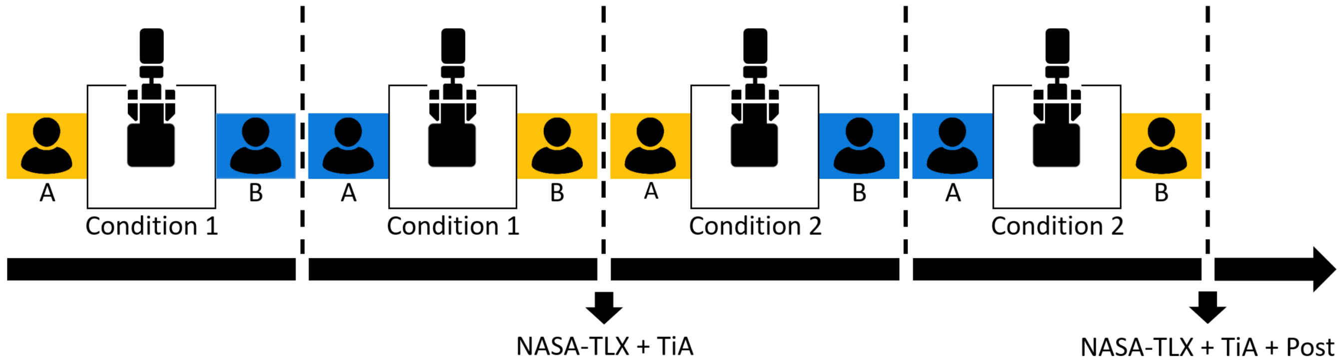

Fig. 4. Experiment timeline. In the image, letters A and B correspond to the sets of operations associated to the locations, while the colours to the users.

to allow participants to experience every possible case that could happen in any situation in which the target of different operations was shared between them [4], [7].

B. User study

The user study took place at the BAE Systems Academy for Skills and Knowledge Centre (Blackburn, UK), upon approval by the Department of Computer Science Research Ethics Committee of The University of Manchester (Ref. 2023-15441-30561). The Participant Information Sheet and the consent form were both GDPR compliant. In this study, 78 participants (59 males, 19 females), all of them at least 17 years old (mean: 22.26 years, standard deviation: 7.16 years), performed the collaborative task. They were all factory apprentices and teachers of the ASK Centre. Most of them were not particularly experienced with robotic technologies (see Table II), but were knowledgeable regarding manufacturing processes.

Materials and procedure are reported below.

1) Materials: A Universal Robots UR5e[1] equipped with a Robotiq Hand-E adaptive gripper[2] was used as collaborative robot. The whole system[3] was controlled through a ROS 2 Humble workspace [20] on Ubuntu 22.04 (Intel® Core™ i7-9750H CPU, 2.60 GHz) in a *Wizard of Oz* fashion. The target object was a prop which was modified to enable the robot to grasp it firmly (see Fig. 1). All the props and tools used during the task were mock ones to ensure safety of the participants during the procedure.

The users' perception of the experiment was probed through the NASA Task Load Index (NASA-TLX) [21] and the Trust in Automation (TiA) [22] questionnaires. The NASA-TLX is used to measure workload of tasks [21]. It asks the participant to rate the task in terms of physical, temporal and mental demand, performance, frustration and effort, on a 21-point Likert scale for each component. These scores can then be used to extract the overall workload score, by computing the average of the scores given to the components multiplied by 5. Alternatively, the questionnaire allows to weigh such scores based on the components deemed relevant by the single participant [21]. To get this information, the participant is shown cards with all the possible pairings of the components and asked to pick which of the two contributed more in terms of workload. The number of times the same component is chosen is used as weight for the component score of that specific participant (see Subsection IV-D).

The TiA questionnaire uses 19 5-point Likert scale items to assess the reliability, predictability and trust of the user towards automated systems, their perception of the developers' intention, and the user's familiarity and propensity to trust automated systems [22]. A post-experimental survey was used to collect demographic features of the participants [23] and their level of experience with robotic manipulators [12], and to verify whether they saw through the *Wizard of Oz* methodology. The consent form, questionnaires and post-experimental survey were administered through Qualtrics[4].

2) Experimental procedure: This user study applied a within-subjects design to remove participants' personal biases in assessing such a novel task, while focusing on the differences between the self-reported measurements in the two conditions. The order of the conditions was randomised for each pair, also randomly matched. At the beginning of the experiment, each pair of participants was welcomed by the experimenter and introduced to the task. After having debriefed the instructions, each participant was randomly assigned to either location A or B, respectively. They were allowed to try out the related sequence of tasks; then, they were told to start performing the task as quickly as possible. They performed the task under one of the two conditions; after this, they switched locations and performed the task again. At this point, the participants were asked to fill the NASA-TLX and TiA questionnaires relatively to such experience. After this, they repeated the task under the other experimental condition. The repeated changes in their location were adopted to prevent adaptation to a specific sequence and carryover effects from one condition to the

[1]https://www.universal-robots.com/products/ur5-robot/

[2]https://robotiq.com/products/hand-e-adaptive-robot-gripper/

[3]The code repository is available at: https://github.com/francescosemeraro/multi_user_concurrent_hrc_task

[4]https://www.qualtrics.com/uk/

other. The participants filled the same questionnaires again relatively only to the last experienced condition, plus the post-experimental survey. The whole procedure lasted less than 1 hour per each pair of participants and a representation of it is summarised in Fig. 4.

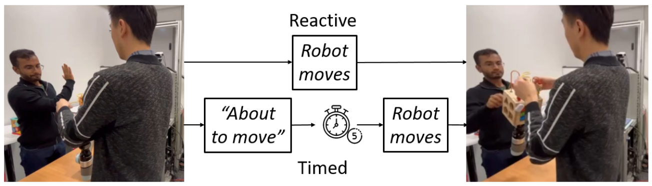

Fig. 5. Difference between the reactive and timed conditions in handling a conflict case between the two users.

The two experimental conditions, referred to as reactive and timed, differed in the way the robot behaved during the interaction with the users (see Fig. 5). In both conditions, the robot modified the pose of the target object to enable the users to perform the specific operation at hand properly. As the participants were performing their tasks concurrently, in some cases the robot had to mediate a position and an orientation of the object that were not ideal for either of the participants, but that would enable both of them to work at the same time with neither interruptions nor need to communicate with one another. In these cases, in the reactive condition, the robot moved the target object as soon as the request from the participant about to initiate a new operation was detected. In the timed condition, instead, the robot gave some time to the participant already working on the object to acknowledge the upcoming change in the configuration of the object. Specifically, when the request was detected, the system uttered, through a speaker, the sentence: *"About to move"*. This was done also to avoid that the inactivity of the robot could be misunderstood as failure to identify the situation. After 5 seconds from this event, the robot moved the object according to the specific case of the interaction.

IV. RESULTS

After having briefed the post-experimental surveys, 67 instances of participants were used for the analyses, granting effect size d=0.56, with 0.90 statistical power at an alpha level of 0.05. Their demographics and experience with robotic manipulators are reported in Table II.

Statistical tests were carried out to assess significant differences in the self-reported measurements of the participants regarding workload and trust between the reactive and timed conditions. From this point on, in the notation, the conditions will be referred to as R and T, respectively. Besides, $\tilde{X}$ will represent the median of a distribution. A null hypothesis was rejected when the p-value of the related statistical test was lower than 0.05.

A. Workload

Because of how workload is defined in the NASA-TLX [21], it is possible to derive a unique measurement of it, given by the combination of its components. As there is debate between using the raw workload score or weight it based on the ratings participants gave to the components of the workload [24], the analysis was carried out on both versions of the score. The raw and weighted workload scores in both conditions are reported in Fig. 6. The statistical hypothesis investigated regarding workload is stated below:

TABLE II

PARTICIPANTS' DEMOGRAPHICS AND EXPERIENCE WITH ROBOTS.

FEATURE	GROUPING	PERCENTAGE
Gender	Male	78%
	Female	22%
	Non-Binary/ Preferred not to Say	0%
Age	Young (16-35)	88%
	Middle-Aged (36-55)	9%
	Senior (56-)	3%
Experience with Robotic Manipulators	None	76%
	One	10.5%
	Few	10.5%
	Regular	3%

H1: There is a statistically significant difference in the distribution of differences of the perceived workload between the reactive and timed conditions.

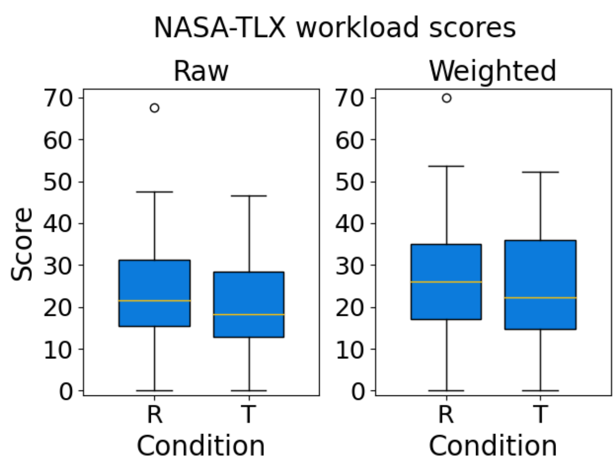

Fig. 6. NASA-TLX workload scores in the two conditions (R = Reactive, T = Timed).

Thanks to the high amount of samples collected and because the score distribution from aggregated Likert scales can be considered continuous [25], Shapiro-Wilk normality tests [26] were carried out on the distribution of differences of the paired samples. This resulted in proving the distributions to not be normally distributed for the raw scores (W=0.954, p=0.0152) and the weighted scores (W=0.941, p=0.0032). Consequently, a Wilcoxon signed-rank test [27] was run on both scores. This rejected the corresponding null hypothesis for the weighted scores (Z=763, p=0.0431) and failed to reject it for the raw scores (Z=701.5, p=0.0538), therefore proving *H1* when using the weighted scores. More precisely, the weighted workload in the timed condition ($\tilde{X}_T$=22.33) was significantly lower than the one measured in the reactive condition ($\tilde{X}_R$=26). A similar result was obtained with the raw scores ($\tilde{X}_T$=18.33 and $\tilde{X}_R$=21.67), but not at a significant level.

B. Trust

As trust is a complex phenomenon that is not advised to be expressed by means of a unique variable [23], analyses on trust were performed by considering the scores of the six aforementioned components (see Subsection III-B.1) singularly. Their scores are reported in Fig. 7. The following statistical hypothesis is applied to each component:

H2: There is a statistically significant difference in the distribution of differences of the perceived trust component towards the robotic system between the reactive and timed conditions.

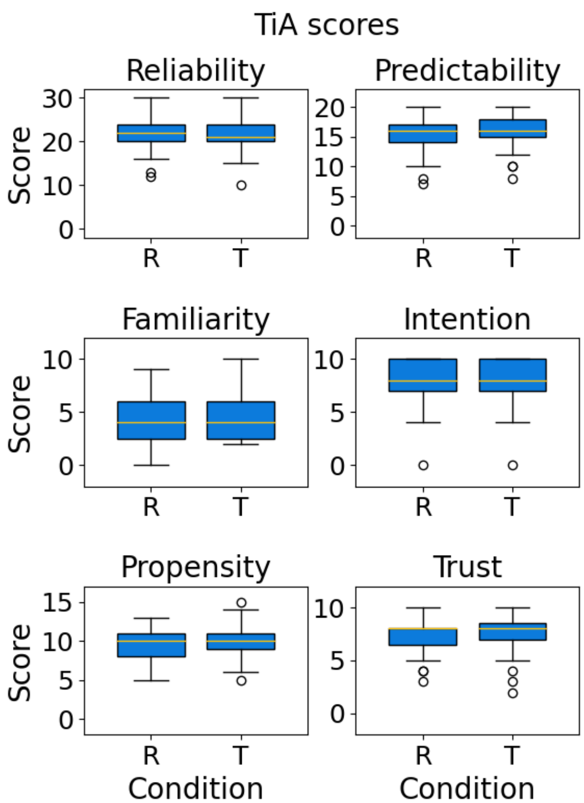

Fig. 7. TiA trust component scores in the two conditions (R = Reactive, T = Timed).

The results of the statistical tests are reported in Table III. Every distribution of differences did not pass the normality test. All the related Wilcoxon signed-rank tests failed to reject the null hypothesis, except for the predictability dimension. In this case, *H2* was proven: there is a statistically significant difference in terms of predictability between the two conditions. Although the medians are equal in both conditions ($\tilde{X}_R$=$\tilde{X}_T$=16), it is possible to visually appreciate from the top right plot of Fig. 7 that predictability is generally higher in the timed condition. Besides, both first and third quartiles are higher in the timed condition than in the reactive condition (Q_{1R}=14 and Q_{1T}=15, Q_{3R}=17 and Q_{3T}=18).

TABLE III
OUTCOMES OF THE STATISTICAL TESTS ON TRUST.

TRUST DIMENSION	SHAPIRO-WILK TEST		WILCOXON SIGNED-RANK TEST	
	W	*p*	*Z*	*p*
Reliability	0.929	<0.001	608.5	0.778
Predictability	0.954	0.0145	**423**	**0.00896**
Familiarity	0.722	<0.001	186.5	0.133
Intention	0.848	<0.001	264	0.76
Propensity	0.809	<0.001	398.5	0.102
Trust	0.885	<0.001	299	0.183

C. Covariate analysis

Because of the results reported in Subsections IV-A and IV-B, additional analyses were carried out on the weighted workload and predictability scores. Specifically, it was ascertained whether other independent variables have not influenced the achieved results. For this purpose, a Generalised Estimating Equation (GEE) model [28] was fitted to the two dependent variables separately. This was not done to find a multi-linear model of the workload and predictability, but rather to understand through the statistics of the model whether other independent factors could have heavily altered the outcome of the experiment. Alongside the experimental condition chosen as independent variable, covariates considered in the models were the experimental condition, the demographic factors of the participants and their previous experience with robotic manipulators (see Table II).

TABLE IV
z AND p-VALUES OF THE GENERALISED ESTIMATING EQUATION MODEL.

VARIABLE	PREDICTABILITY		WEIGHTED WORKLOAD	
	z	*p*	*z*	*p*
Condition	**2.452**	**0.014**	**-1.42**	**0.157**
Gender	-0.5	0.617	-0.343	0.732
Age	1.39	0.164	0.218	0.827
Experience	1.56	0.118	-0.886	0.376

The result of the fittings, reported in Table IV, allow us to conclude that the experimental condition was the most influential factor determining the outcome of the statistical analysis. Indeed, looking at the statistics of the predictability model, the experimental condition has the highest z-value, indicating the highest effect in the GEE model. Besides, its p-value is lower than 0.05, indicating statistical significance of the related coefficient. Regarding the weighted workload, the p-value of all estimated coefficients was above the threshold for every independent variable. However, the experimental condition was still the most determining variable among the chosen ones to describe the weighted workload. Indeed, its p-value was the lowest while its z-value was the highest in absolute value, proving largest evidence against the corresponding null hypothesis of the model, albeit not at a statistically significant level.

D. Evaluation of the collaborative task

The use of the weightings proved to be crucial to ascertain the validity of the hypothesis on workload at a statistically significant level (see Subsection IV-A). Therefore, further investigation on the weightings of the task was carried out. Indeed, they provide us with an overview of what type of workload the task was more suited to elicit. Fig. 8 reports the distribution of weights given by the participants in the NASA-TLX (see Subsection III-B.1) [21].

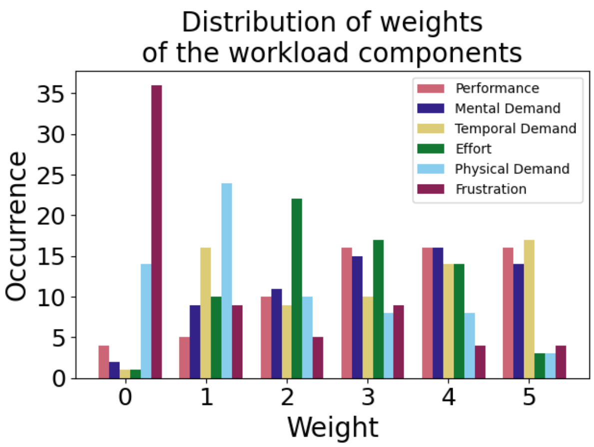

Fig. 8. Distribution of weights of the workload components of the NASA-TLX. The components were sorted in the legend by decreasing order of the total number of times the component was chosen throughout the user study.

Frustration, Physical Demand and Effort were the components chosen less frequently by the participants. This is a good indicator that the task was well designed from an ethical point of view to not cause discomfort to the participants during the experiment. Moreover, little physical demand is consistent with a manufacturing operation. On the other hand, Performance was the component chosen most frequently by the participants. As Performance is the only reverse item of the NASA-TLX [21], this means that the participants were satisfied with the way they performed. This result indicates that the task was easy for the participants to execute. Mental Demand was the main indicator of workload. This is consistent with the situation the participants faced. Most of them were not familiar with robotic manipulators (see Table II), so it was demanding for them to understand how to interact with it. Besides, they were asked to carry out the collaborative task without any previous knowledge about it. Furthermore, it is interesting to point out that they felt slightly lower Temporal Demand than Mental Demand. This means that they did not feel an excessive time pressure during the experiment, despite the task being new to them and being asked to perform it as quickly as possible.

V. DISCUSSIONS

The statistical tests carried out (see Section IV) allowed us to draw important conclusions regarding the design of the behaviour of a human-human-robot collaboration. It was assessed whether a different timing of the robot response regarding conflictual situations between the two users (see Section I) induced a different workload and trust on the users. It was found that the timed condition resulted in lower workload of the task and higher predictability of the system with respect to the reactive condition. This was tested out by having both participants experiencing all possible cases in which this design choice could have affected them (see Subsection III-B.2). Indeed, they could benefit or be negatively affected by the timed response, according to whether they were querying the robot or working in the scenario (see Section I). However, on the whole, having pauses and receiving messages from the robot during the task helped the users keep track of what was happening in the task and feel less stressed. Moreover, this result was proved not to be influenced by other variables of the design (see Subsection IV-C). Finally, the ratings on the workload components of the NASA-TLX showed that the novel interaction scenario is clear to understand and keeps the participants generally comfortable while doing the experiment (see Subsection IV-D). Having gathered these insights from participants involved in the field of manufacturing increases the validity of the reported findings.

Furthermore, it is interesting to point out that drawing out the conclusions regarding workload at a statistically significant level was possible thanks to the weighting procedure used in the NASA-TLX [21]. This suggests that the use of weightings can be useful to strengthen the significance of the information given by the workload scores. Besides, a procedure of this kind allows the whole task to be sensitive to the users' own ratings of the experience, favouring a design that further incorporates aspects of personalisation in human-robot interaction [29].

This work could be enhanced in the future by exploring the impact of the response of the robots on other human factors. First, it would be interesting to explore how these design choices impact on other features of HRC, such as fluency of the collaboration [30]. Secondly, exploring different experimental setups of this kind would further increase the generalisibility of the achieved results to different situations. Finally, an explainability component was introduced in the design by having the robotic system uttering a vocal message to notify the users about the intended delay in the response (see Subsection III-B.2). Exploring different modalities of conveying information from users to robots and viceversa [31], [32] is another potential direction of further investigation.

VI. CONCLUSIONS

This work contributed to enrich the underexplored research domain of multi-user human-robot collaboration. A physical collaborative task with two users and one robotic manipulator was built. The task was designed to include concurrency in the interaction, by having the users working on different operations on the same target object at the same time. The concurrency was addressed by the robot acting in the scenario by changing the configuration of the target object to meet the requirements of both users at the same time.

This setup was used to run a user study that investigated whether people could benefit more from a reactive or timed response of the robot during conflictual situations. The experiment showed that a timed response decreases the perceived workload of the users and increases the predictability of the robotic system.

ACKNOWLEDGMENT

We acknowledge Norman Suchorzewski for helping with the recruitment of participants and Stephen Rhodes for the modifications to the prop to make it suitable for the experiment.

REFERENCES

[1] E. Hornecker, A. Krummheuer, A. Bischof, and M. Rehm, "Beyond dyadic hri: Building robots for society," Interactions, 2022.

[2] E. Schneiders, E. J. Cheon, J. Kjeldskov, M. Rehm, and M. B. Skov, "Non-dyadic interaction: A literature review of 15 years of human-robot interaction conference publications," ACM Transactions on Human-Robot Interaction (THRI), vol. 11, p. 13, 2 Feb. 2022.

[3] F. Semeraro, A. Griffiths, and A. Cangelosi, "Human–robot collaboration and machine learning: A systematic review of recent research," Robotics and Computer-Integrated Manufacturing, vol. 79, pp. 1–16, Feb. 2023.

[4] F. Semeraro, J. Carberry, and A. Cangelosi, "Simpler rather than challenging: Design of non-dyadic human-robot collaboration to mediate human-human concurrent tasks," Proceedings of the 22nd International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS), pp. 2451–2453, 2023.

[5] "QS-004CS Quality Standard for Sealant Application" Accessed on July 26, 2024. (2024), [Online]. Available:https://supplier.aero.bombardier.com/A220-SQA/sqa-docs/quality_docs/.

[6] H. Wali, "Human Robot Collaboration in Assembly Processes", 2018.

[7] F. Semeraro, J. Carberry, and A. Cangelosi, "Towards multi-user activity recognition through facilitated training data and deep learning for human-robot collaboration applications," 2023 International Joint Conference on Neural Networks (IJCNN 2023), pp. 01–09, Jun. 2023.

[8] E. Schneiders, "Non-dyadic human-robot interaction: Concepts and interaction techniques," ACM/IEEE International Conference on Human-Robot Interaction, vol. 2022-March, pp. 1176–1178, 2022.

[9] E. Schneiders, C. Fourie, S. Celestin, J. Shah, and M. Jung, "Understanding entrainment in human groups: Optimising human-robot collaboration from lessons learned during human-human collaboration", Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '24), 2024, Honolulu, HI, USA, vol. 1, 2024.

[10] E. Merlo, E. Lamon, F. Fusaro, M. Lorenzini, A. Carfì, F. Mastrogiovanni and A. Ajoudani, "An ergonomic role allocation framework for dynamic human–robot collaborative tasks", Journal of Manufacturing Systems, vol. 67, pp. 111–121, 2023.

[11] D. Kourosh, E. Simetti, F. Mastrogiovanni and G. Casalino, "A hierarchical architecture for human–robot cooperation processes", IEEE Transactions on Robotics, vol. 37, no. 2, pp. 567-586, 2020.

[12] M. Story, P. Webb, S. R. Fletcher, ·. G. Tang, C. Jaksic, and J. Carberry, "Do speed and proximity affect human-robot collaboration with an industrial robot arm?" International Journal of Social Robotics, vol. 14, pp. 1087–1102, 2022.

[13] M. G. Collins, I. Juvina, and K. A. Gluck, "Cognitive model of trust dynamics predicts human behavior within and between two games of strategic interaction with computerized confederate agents," Frontiers in Psychology, vol. 7, pp. 1–17, 49 Feb. 2016.

[14] G. Tsamis, G. Chantziaras, D. Giakoumis, et al., "Intuitive and safe interaction in multi-user human robot collaboration environments through augmented reality displays," 30th IEEE International Conference on Robot and Human Interactive Communication, RO-MAN 2021, pp. 520–526, Aug. 2021.

[15] E. Matheson, R. Minto, E. G. G. Zampieri, M. Faccio, and G. Rosati, "Human–robot collaboration in manufacturing applications: A review," Robotics, vol. 8, pp. 1–25, 2019.

[16] D. Fu, F. Abawi, and S. Wermter, "The robot in the room: Influence of robot facial expressions and gaze on humanhuman- robot collaboration," RO-MAN 2023 - 32nd IEEE International Conference on Robot and Human Interactive Communication, pp. 85–91, Nov. 2023.

[17] M. Faria, R. Silva, P. Alves-Oliveira, F. S. Melo, and A. Paiva, ""me and you together" movement impact in multi-user collaboration tasks," 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), vol. 2017- September, pp. 2793–2798, Sep. 2017.

[18] E. Schneiders, C. Fourie, J. Shah, and M. Jung, "Humanhuman collaboration: What can we learn from human group collaboration to improve human-robot collaboration?" ACM/IEEE International Conference on Human-Robot Interaction (HRI'23 WYSD Workshop), pp. 1–5, 2023.

[19] E. Schneiders and S. Celestin, "Non-dyadic entrainment for industrial tasks," Workshop on Joint Action, Adaptation, and Entrainment in Human-Robot Interaction at the HRI'22 conference, 2022.

[20] Y. Maruyama, S. Kato, and T. Azumi, "Exploring the performance of ros2," Proceedings of the 13th International Conference on Embedded Software, EMSOFT 2016, Oct. 2016.

[21] NASA, "Task load index (nasa-tlx) v 1.0 manual," pp. 0-26, 1986.

[22] M. Korber, "Theoretical considerations and development of a questionnaire to measure trust in automation," Advances in Intelligent Systems and Computing, vol. 823, pp. 13–30, 2019.

[23] M. Romeo, P. E. McKenna, D. A. Robb, et al., "Exploring theory of mind for human-robot collaboration," RO-MAN 2022 - 31st IEEE International Conference on Robot and Human Interactive Communication, pp. 461–468, 2022.

[24] S. G. Hart, "Nasa-task load index (nasa-tlx); 20 years later," Proceedings of the Human Factors and Ergonomics Society, pp. 904–908, 2006.

[25] S. E. Harpe, "How to analyze likert and other rating scale data," Currents in Pharmacy Teaching and Learning, vol. 7, pp. 836–850, 6 Nov. 2015.

[26] S. S. Shapiro and A. M. B. Wilk, "An analysis of variance test for normality (complete samples)," Biometrika, vol. 52, pp. 591–611, 3-4 Dec. 1965.

[27] F. Wilcoxon, "Individual comparisons by ranking methods," Biometrics Bulletin, vol. 1, pp. 80–83, 6 1945.

[28] K.-Y. Liang and S. L. Zeger, "Longitudinal data analysis using generalized linear models," Biometrika, vol. 73, pp. 13–22, 1986.

[29] I. Tarakli, G. Angelopoulos, M. Hellou, et al., "Social robots personalisation: At the crossroads between engineering and humanities (concatenate)," ACM/IEEE International Conference on Human-Robot Interaction, pp. 920–922, Mar. 2023.

[30] G. Hoffman, "Evaluating fluency in human-robot collaboration," IEEE Transactions on Human-Machine Systems, vol. 49, pp. 209–218, 3 Jun. 2019.

[31] F. Semeraro, L. Fiorini, S. Betti, G. Mancioppi, L. Santarelli, and F. Cavallo, "Physiological wireless sensor network for the detection of human moods to enhance human-robot interaction," Lecture Notes in Electrical Engineering, vol. 544, pp. 361–376, 2018.

[32] L. Fiorini, F. Semeraro, G. Mancioppi, S. Betti, L. Santarelli, and F. Cavallo, "Physiological sensor system for the detection of human moods towards internet of robotic things applications," 17th International Conference on New Trends in Intelligent Software Methodology Tools and Techniques (SoMeT), pp. 967–980, 2018.