2024 33rd IEEE International Conference on
Robot and Human Interactive Communication (RO-MAN)
August 26-30, 2024. Pasadena, California, USA.

2024 33rd IEEE International Conference on Robot and Human Interactive Communication (ROMAN) | 978-1-7281-5207-3/24/$31.00 ©2024 EU | DOI: 10.1109/RO-MAN60168.2024.10731196

Exploring Continuous Awareness Modelling for Improving Worker Safety and Trust*

Andrey Solovov[1,2], Mohammad Zarei[1,2], Vishal Gautam[1,2], Bruno Ferreira[1,2], Gustavo Assunção[1,2], Antonio Marin-Hernandez[3], Paulo Menezes[1,2]

***Abstract*— The Industry 5.0 Human-Robot Interaction (HRI) goals have the potential for significantly reducing production costs, but require factory workers to trust that robots will not cause them physical harm when sharing the workspace in order to be achieved. To this end, we propose a model that adjusts the robot's actions to an estimate of the workers' awareness of its behavior. Its premise is that mutual awareness should be continuous, and if no recent visual contact has been established, then the robot should adapt its range of motion and speed in order to minimize the intersection between the robot's space and the worker's space. On the other hand, if the worker steadily establishes visual contact from times to times, the robot operates under a normal or a faster configuration. After a period of habituation, this awareness model can be further tweaked to continuously improve production times, while improving worker's safety and trust. We present preliminary results of a pilot study with 32 participants, to asses if the existence of this awareness model impacts the outcomes of the combined task. The obtained tepidly optimistic results are the key to build upon in future designs.**

I. Introduction

The idea of humans sharing industrial spaces with robots is not novel, especially for improving production times. However, the current practices under the design of these Collaborative Robots (cobots) still lack human-centric analysis for safety concerns, namely for preventing hazards [1]. Human safety is usually ensured via robust protocols and standards applied to cobot design, implementation process, and behaviour [2]. The International Organization for Standardization (ISO) has played a central role in the development of guidelines for human-robot cooperative systems, cobots, and industrial robots. Namely, ISO-15066 is the standard that defines the currently accepted guidelines that HRI [1]. For instance, when humans are coexisting in the same workspace as N robots, safety mechanisms can take place and override the robot's behaviour as the intersection between both spaces increases. However, ISO-15066 imposes hard restrictions on the momentum of the robot, as the robot will decrease the speed of the task as required, and eventually, will stop, so that the human can conclude the assigned task [3]. This makes it impractical and increases the production cycle time drastically.

*This research has been financed by the PRR - Recovery and Resilience Plan and by the Next Generation EU European Funds, following NOTICE No. 02/C05-i01/2022, Component 5 - Capitalization and Business Innovation - Mobilizing Agendas for Business Innovation, under the project Greenauto - PPS10/PPS12/PPS13 with the reference 7255, C629367795-00464440, and by the project FCT/MCTES UIDB/00048/2020 DOI:10.54499/UIDB/00048/2020.

[1]Department of Electrical and Computer Engineering, University of Coimbra, Coimbra, Portugal
[2]Institute of Systems and Robotics, University of Coimbra, Coimbra, Portugal
[3]Universidad Veracruzana, Xalapa, Mexico

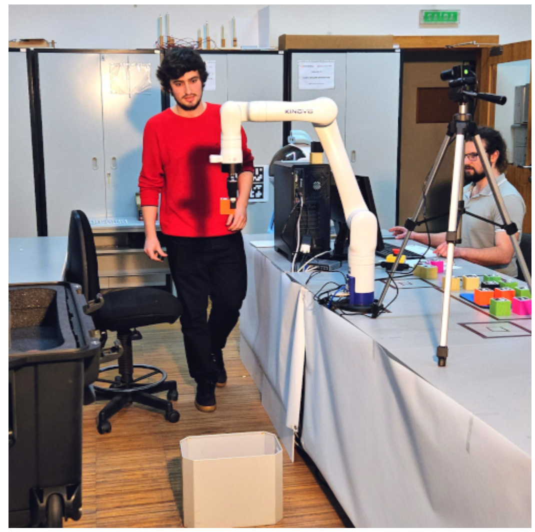

Fig. 1: Illustration of a participant performing the assigned task while sharing the same space with an industrial robot that is continuously operating.

Although these stringent practices are appropriate for the current state of the art, they can be alleviated upon the exploration of non-verbal human-robot communication methods. As a matter of fact, recent works on awareness during HRI are showcasing that this continuous problem is hard to model, and in most of the cases, impractical to directly deploy in many nowadays manufacturing facilities [4]. The focus under this line of work should reside in simple, but elegant solutions, that focus on key aspects of body language cues, such as eye gaze estimation.

On the other hand, factory employees have a learning curve regarding how they should work around/next to a cobot, and most importantly, the level of trust they should have with it [5]. This acquired trust comes along with the human's understanding of the cobot's capacity to read their body language, detect their acknowledgement of its presence, and comprehend the task they are performing. Once mutual awareness can be achieved, both assigned tasks can be carried out smoothly and according to expectations.

We propose the use of a lightweight eye contact estimation method that serves as a backbone for a novel awareness model. The proposed awareness model is meant to be deployed in industrial settings, in order to dynamically adapt the robot's behaviour over time in accordance to whether the person has paid attention to the task that it is performing in the same environment. We sustain the idea that the attention that we pay to dynamic/moving objects is primarily continuous, yet highly dependent to whether we have looked at them or not.

This paper is organised as follows: Section II overviews the state-of-the-art regarding trust in HRI, and the correlation between human gaze and awareness, shortly exploring how this has been measured. Section III details the proposed methodology for robot behavioural adaptation based on perceived human awareness. Subsequently, Section IV presents the experimental scenario and protocol used to evaluate our methodology, followed by Section V where results of this experiment are listed and discussed. Section VI adjourns the paper and advances potential paths for future research.

II. State of the Art

The research community has been working over the recent years on different methods that focus on ensuring workers' safety while decreasing the production cycle time. Of those, safety will always be the crucial factor to improve trust of the workers in robotics applied to the industry [5]. Below, one can delve into whether using different actions by cobots, mostly based on user attention levels, will improve the factor of perceived safety by humans and the overall level of trust.

The interdisciplinary nature of HRI drives researchers to establish multiple definitions of trust. As trust plays such an important role in building and maintaining a proper social relationship between humans and robots, in industry, it is considered one of the basic requirements for the development of a cooperative environment [6]. Several literature reviews and meta-analyses exist on this matter. One of them is Lee and See [7], which proposes a three-dimensional approach that defines human trust based on the following factors: *purpose* - the user's knowledge on what the robot is supposed to do; *process* - how the task will be conducted; and *performance* - the reliability and correlation between used questionnaires. Additionally, Hancock et al. presented a meta-analysis on factors affecting trust in HRI, highlighting its complexity [8]; while Mittu et al. [9] tried to propose a trust measurement scale to evaluate the trust factor in HRI. Tenhundfeld et al. have also reviewed and compared multiple methods for measuring trust in automation [10]. These authors have discussed the requirements for these measures, noting the importance that they should be concise, unintrusive, interaction-based, context-specific, dynamic, and that they take into account task dependency and risk level.

Schroepfer et al. proposed a framework for representing trust as a system represented by a meta model in [11]; they divided the research around trust into components while modelling the relationship between them, becoming an important foundation for the following works. For optimal performance, human trust should be well calibrated as both situations of under- and over-trust can lead to accidents, due to inexperience and complacency respectively.

In terms of practical results, Hald et al. [12] showed the eminent correlation between trust and operator movement patterns. The work of Hopko et al. [13] assessed the use of human factors (mental state of the user, difficulty level of task, and experience) for evaluating trust in Human-Robot Coexistence (HRCx) scenarios; while Lawrence et al. [14] evaluated trust in robots during a competitive game scenario. Others, such as Goubard et al. [15] present their research on a trust-aware policy for Human-Robot Collaboration (HRC), by using gaze and posture to estimate the trust level of the user. Giulio Campagna et al. [16] study the relationship between the proximity of robot on trust of the user, by exploring a chemical industry scenario, where a robot assists a human in mixing chemicals.

From the above, one can understand that there is no unified framework, tool, or approach for addressing the HRI trust problem in a way that is applicable to all scenarios.

A. Attention and Awareness using Visual Contact

Delving even more deeper into the scope of this work, trust points to the human's perception of the robot's competence at keeping them safe: they will (rightly) distrust a robot that performs an action with the potential of causing them harm; or one that endangers them by not changing behaviors when it should. Thus, the level of trust that the human exhibits in HRI scenarios is predicated on maintaining and acting upon a good understanding of the mutual awareness between the human and the robot over time [17]; namely, the robot should act more cautiously around a human that displays insufficient awareness of its behavior; but it is also free to continue performing its task as before upon the approach of a human exhibiting sufficient awareness.

To this end, Noton and Stark's influential Scanpath Theory of Attention [18], [19] provides a connection between eye gaze and awareness via attention, while simultaneously making a case for the quantifiability of awareness: scene appraisal through successive foveal fixations to relevant stimuli occurs over time, building up semantic understanding of the scene (i.e. awareness), thus making it attractive to associate a measure of the awareness level: it would represent that the semantic understanding has overall increased as the person kept "scanning" the scene. Similarly, as the scene changes while the person's attention is directed elsewhere, it is reasonable to suppose that their awareness will involuntarily decay. The timewise increase and decrease of awareness have commonly agreed-upon figures, referred to in [20] and [21] respectively: they are $200 - 300ms$ for the increase in awareness from inexistent to "full-blown"; and $15s$ for an involuntary attention shift to occur, representing awareness decay. Considering that this work concerns an industrial environment, it was chosen to err on the side of safety by underestimating the level of awareness, imposing a longer *scene appraisal time*, $t_S = 0.5$s, and a shorter *distraction time*, $\tau_h = 10$s.

The use of human awareness to adapt robot behavior is novel within the state of the art. Previous works in a similar vein have focused on HRC tasks: examples include Dufour et al. [22], where a cobot's end-effector was only allowed to move while within the visual-spatial attention of the person; or Prajod et al. [23], where a person's gaze was used to initiate a cobot's action. The circumstances studied so far have thus been ones where the person must always pay attention to the robot for it to be working, imposing it as insufficient for the human to be merely aware of it. This limitation is reasonable for the HRC interactions that were of interest for those works, but too strict for the HRCx and Human-Robot Cooperation (HRCp) tasks that are also common in industry: for HRCx and HRCp, it is assumed that the person cannot dedicate their full attention to the robot, but may still exhibit sufficient awareness of it for crossing paths to be safe.

III. Methodology

This section presents the different modules created to support the awareness estimation, as it requires a synergy between different topics.

A. Visual Contact Classification

The backbone for this work is a visual contact classification model that leverages RetinaFace [24] to efficiently detect and crop the subject's face from each video frame. This cropped face is then resized to a standardised format of 224x224 pixels. A modified SENet50 architecture [25] acts as the core component of the model as it analyses the facial features within the cropped image to determine if the person is looking at the robot. Consequently, the model outputs a boolean value for "looking" or "not looking" for each processed frame. The model achieves approximately 10 frames per second (FPS) on a GPU, enabling real-time analysis of the video stream with input frames sized at 1024x768 pixels. Notably, the pixel area of the subject's face is also extracted to infer walking direction.

B. Continuous Awareness Model

The awareness model implemented for this work is based on the one proposed in Zarei et al. [26]. The basic structure is the same: a human *awareness measure* $\mathcal{A}$ and an *Awareness Sufficiency Threshold (AST)* Θ are calculated independently of each-other based on data extracted from camera input. If $\mathcal{A} \geq \Theta$, then the interaction is deemed safe, and the robot can continue working as normally, else it must adjust its behavior so that the interaction will be less hazardous. The assumptions from Zarei et al. are also mostly preserved, particularly the main one: that slow and predictable robots require less awareness for a safe interaction than fast robots with complex tasks. It is assumed like in the aforementioned work that awareness is additive upon continuous appraisal, with t_S setting the rate of increase, and decays exponentially (Fig. 3), scaled upon τ_h and the solid angle covered by the robot in the human's field-of-view. This work also adheres to calibrating $\mathcal{A}$ and Θ so that the least extreme values for both are achieved at the Habitual Workspace Conditions (HWCs), taken to correspond to a human-to-robot distance of $r_{\mathrm{h}} = 2m$, with the human stationary and facing the robot.

Besides this, the model's expressions for $\mathcal{A}$ and Θ were simplified in order to allow for a bufferless implementation, thus allowing for the omission of an explicit time coordinate. With these adjustments, the awareness measure took on the form shown in Eq. 1, where $\mathcal{D}$ is a user-inputted dimensionless parameter, Δt is the time interval since the previous calculation of $\mathcal{A}$, LR^+ is the positive likelihood ratio of the gaze detector, `int(looking)` is the latest gaze detector-outputted *looking flag* (Fig. 2), with average precision of 98.74% [26], converted into an integer, and r is the current distance of the person to the robot, calculated using the area of the bounding box of the person's face:

$$\mathcal{A} \leftarrow \mathcal{A} \times \exp\left[-\frac{\ln 2\,\mathcal{D}}{r_h^2 \tau_h} r^2 \Delta t\right] + \mathrm{LR}^+ \times \texttt{int(looking)} \tag{1}$$

The *AST* took on the form in Eq. 2, where q and $\mathcal{V}$ are user-inputted dimensionless calibration parameters and v is the scalar approach speed of the human toward the robot.

$$\Theta \leftarrow q \times \exp\left[\frac{v}{\mathcal{V}} + \frac{r_h}{r} - 1\right] \tag{2}$$

In practice, the only difference between the current simplified design of the awareness calculation and the originally proposed one is that the decay rate is now tied to the current distance to the robot, rather than the distance during the time of the hit. Due to the slow robot and human speeds in this work, the discrepancies due to this change were negligible.

The awareness model's user-inputted parameters were calibrated AST-side first: q was set so that, under HWCs, $\mathcal{A}$ would rise from zero to Θ in t_S time if a person kept looking at the robot. Then $\mathcal{D}$ was adjusted so that $\mathcal{A}$ would decrease from its maximum possible value to Θ in τ_h, under HWCs. These two factors were alternately adjusted until an optimal value was reached. This left $\mathcal{V}$ as a free parameter: ultimately, it was chosen to set $\mathcal{V}$ to a high value, as v was over oscillating due to a simple calculation based on r.

The awareness model was implemented in Python over three objects asynchronously infinitely looping one function each: one for $\mathcal{A}$; one for Θ; and one for yielding a safety flag (Fig. 2) based on the output of the other two. It achieved a slightly faster cycle rate than the gaze detector module, so that it easily worked in tandem with it: an example of the awareness model in action is depicted in Fig. 3. The graph is timewise-divided into four regions from A to D. In region A, the person is approaching the robot, causing Θ to increase. During this movement, they briefly glance at the robot, causing $\mathcal{A}$ to increase to where the interaction is deemed safe, but as the person looks away from the robot, $\mathcal{A}$ begins to decay. After reaching the closest distance to the robot at the peak Θ value, region B shows Θ decreasing as the person moves away from the robot. In region C, the person is outside of the frame, so that the gaze detector continues publishing the last known distance of the person to Robot Operating System (ROS). Finally, in region D, the $\mathcal{A}$ value has decayed to where $\mathcal{A} < \Theta$, and thus the interaction is no longer deemed safe, so the robot slows down.

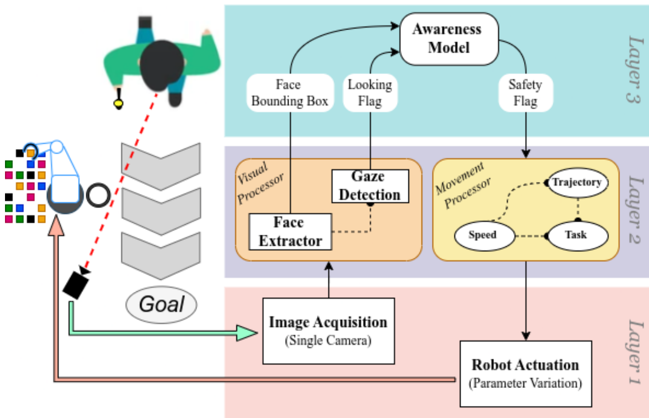

Fig. 2: A layered overview of the system architecture, sided next to a top view of the experimental scenario.

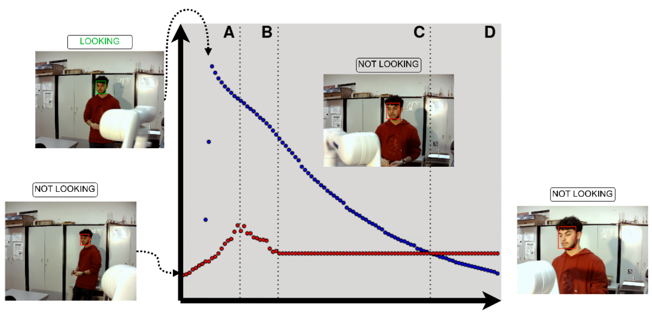

Fig. 3: An example of the awareness model working in tandem with the gaze detection. The blue points represent the $\mathcal{A}$ values and the red points represent the Θ values.

C. Integration with an Industrial Robot

In this experiment, a collaborative Kinova Gen3 with 7 Degrees of Freedom (DoF) served as the hardware platform. Additionally, we utilised the ROS on the software side, operating at a frequency of 10 Hz matching the speed of our eye-contact detection model, to facilitate smooth communication between the robot and our models. The experiment framework consists of 3 layers, acquisition and actuation (low-level functionalities), detection and movement (mid-level functionalities), and awareness estimation and safety classification layer (high-level functionalities). As shown in Fig.2, the awareness model sends a signal to the lower layer in order to modify the speed of the robot according to the safety threshold. Then, the relevant speed will be sent to the robot's controller. Nevertheless, the maximum speed of the robot was set at 50 cm/s and with a safe speed of 5 cm/s.

IV. Experiment

In order to evaluate the architecture described in the previous section, a collaborative environment was simulated for experimental purposes, wherein a robotic arm and user work in tandem, with variable awareness. As depicted in Fig. 2, the setup encompasses object carrying by both parties. The robot arm was tasked with package filling, moving cubes from a designated area and into a fixed basket. Simultaneously, each user was asked to carry token items to and from a goal location, traversing a path which the robot arm could likewise cross. With the integration of our awareness model, we aimed at evaluating how the user would respond to the robot's adaptive behaviour in terms of trust and sense of safety, by premising two research questions:

RQ1: Is the user aware of the robot's behavioural change?
RQ2: Is the awareness model behavioural modification enough to foment trust in the user?

At the start of the experiment, participants were asked to fill out a socio-demographic questionnaire, compiling information related to age, gender, occupation, as well as previous experience with robots. Subsequently, each participant received a short briefing explaining their task of carrying an object along the straight path, and was asked to complete two separate trials of one minute each, in tandem with the robot. In these, the awareness model was either active or inactive, setting up an ablation study. Plus, the participant pool was divided, with the first half engaging with the robot in a inactive (dummy) to active (model) trial order, and the second half doing the opposite. Information regarding model activation or any other mention to perceived danger or fomented trust was not disclosed to the users. Post both trials, each participant was asked to complete a short questionnaire pertaining to their perceived experience.

The post-experiment questionnaire issued to the participants consisted of ten purpose-built questions aimed at HRCx scenarios, as listed in the left-hand side of Fig. 4 in the same order presented. It consisted of a Google Form where the participants assigned a value to each question, within a 5-point Likert scale where 1 represents strong disagreement and 5 indicates the opposite. In addition, three of the questions explicitly request that the participants compare some aspect of the first and second trials, namely Q2, Q4 and Q7, to understand if the behavioural difference is perceptible. The remaining questions address other factors such as the user's mental state and level of trust given to the robot.

The remaining seven questions are independent of trial order: they serve to address different factors affecting user trust. Namely, Q5 and Q6 evaluate the participants' mental state; Q1, Q8 and Q9 evaluate perceived safety [13]; and Q3 and Q10 evaluate the trust of the participant.

V. Results

In total, 32 users took part in this experiment: 9 female and 23 male. All of them except 2 were bachelors, masters or PhD students and post-doc researchers from the University of Coimbra, the remainder being university staff. Half of the participants were in the 18-25 year age range, 2 were over 40, and the rest were in between. Finally, half of the participants reported no prior HRI experience.

A. Analysis

The right-hand side of Fig. 4 summarizes the post-experiment questionnaire results as obtained for the two populations with opposite trial orders (trial with a dummy awareness model first, then a trial with the above-described

Question
Q1. Was the environment / task challenging?
Q2. Did you notice a difference in the behaviour of the robot?
Q3. Did you feel like you could trust the robot to not harm you?
Q4. Did you trust the robot more this last trial?
Q5. Did you feel comfortable working alongside the robot?
Q6. Did you feel like you needed to pay attention to the robot often?
Q7. Did you feel safer working alongside the robot this last trial?
Q8. Did you feel that the robot took your position into account?
Q9. Did you feel like the robot could hit you?
Q10. Would you be willing to collaborate with the robot on a task?

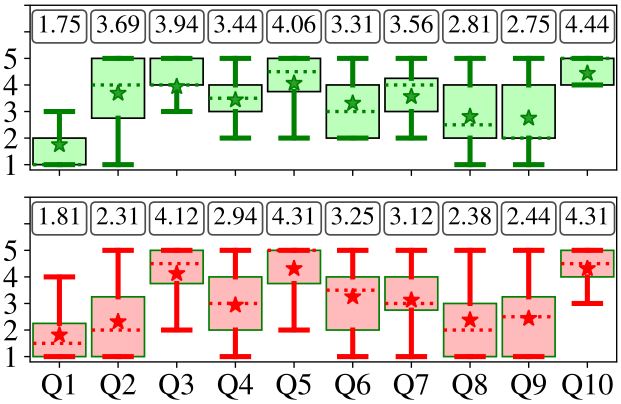

Fig. 4: Questions making up the post-experiment questionnaire issued to the participants. *Right:* box plots for the questionnaire answers for the "dummy → model" population (top) and the "model → dummy" population (bottom), with numbers above indicating the average answer values, indicated on the box plots with stars. The horizontal dashed lines represent the median.

awareness model, and vice-versa). The expectation was that, if there existed a preference for the presence or absence of the awareness model, then it would manifest in the person first noticing a difference in the robot's behavior in the two trials (Q2); and then expressing a preference for the trial that contained the awareness model (Q4 and Q7). The null hypothesis would thus be that there existed no preference at all: it would manifest as a statically negligible difference in the Q2, Q4, and Q7 averages between the two populations.

The obtained results conflict with the null hypothesis much more than they conform with it: they showed a statistically significant increased preference for the second trial when it contained the awareness model (average higher by 0.50 and 0.44 on Q4 and Q7 respectively), but although participants tended to notice a difference in behavior when the second trial contained the awareness model (average of 3.69 and median of 4 for Q2), they usually did not notice a difference when an awareness model trial was followed by a dummy model trial (average of 2.31 and median of 2 for Q2).

Besides displaying the highest average differences, the statistical significance of the answers to Q2, Q4, and Q7 was also assessed by generating random sample bisections into two populations of 16 participants each, in order to construct a distribution of per-population average differences assuming that the null hypothesis is true. In this manner, 10^4 random sample bisections were performed per question (Fig. 5), yielding a p-value of 1.5% for Q2's average difference of 1.38, and p-values of 21.9 % and 23.8% for the Q4 and Q7 average differences respectively. Finally, supposing that the null hypothesis is true, then the Q4 and Q7 average differences are independent of each-other, so the p-value for Q4 and Q7 together is the product of both p-values: that is 5.2%. Thus, the Q2 average difference is statistically significant, as are the joint average differences for Q4 and Q7, and so the null hypothesis can be rejected.

Consequently, regarding RQ1, it was determined that the user usually somewhat agreed that they noticed a behavioral change in the robot if the second trial was with the awareness model, and that this change was deemed as preferable. Regarding RQ2, the statistical significance of the joint Q4 and Q7 average differences, as well as the generally high score obtained for Q3 (average of 4.04, median of 4), suggest that the inclusion of the model aided in further increasing users' initially high trust.

Regarding the questions independent of trial order, the mental state ones yielded 4.18 mean and median 5 for Q5; and 3.21 mean and median 3 for Q6: these results suggest that most participants felt comfortable and attentive during the task. The perceived safety questions all gave low Likert scale answers: 1.78 mean and median 1 for Q1; and 2.50 mean and median 2 for both Q8 and Q9. Thus, the task was deemed easy, participants felt like there was a reasonably low likelihood of the robot hitting them, but mostly did not judge the robot as accounting for their position. The trust questions both yielded high results, 4.03 mean and median 4 for Q3; 4.30 mean and median 5 for Q10, indicating a generally high trust for the robot.

Finally, it should be noted that participants' experience with robots or lack thereof did not affect their answer to Q2 in a statistically significant way: the average difference for Q2 between the populations with and without prior experience was 0.25, which amounted to a p-value of 65.9%.

B. Discussion

Regarding the RQ1 result, it is notable how much more the participants noticed a difference in the robot's behavior when the first trial was with the dummy model and the second was with the awareness model, rather than vice-versa. Under the dummy model, the robot only exhibits one mode of behavior, whereas under the awareness model it exhibits two, so the large difference in Q2 results for the two populations may be an indication that the participants noticed the introduction of a new behavior more easily than the ommission of a preexisting one. This is not surprising, as the appearance of novel stimuli tends to stand out compared to the ommission of old stimuli, when attention is directed to them involuntarily.

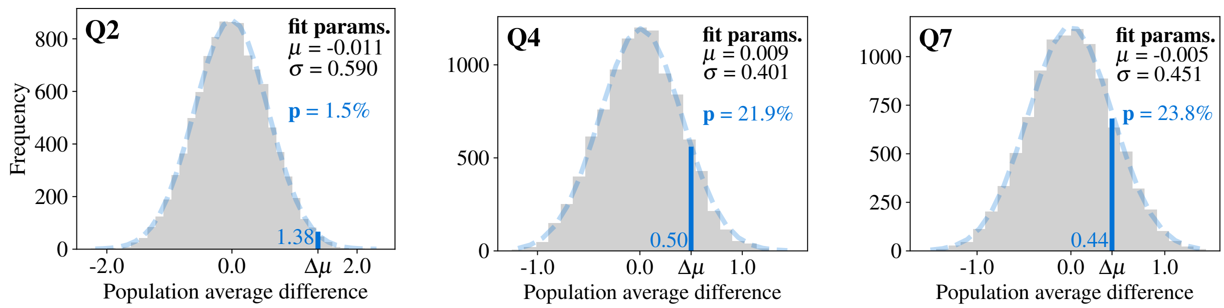

Fig. 5: Average difference distributions for 10^4 random sample bisections into two populations of 16 participants each, with resulting histograms and corresponding Gaussian fits, for questions Q2, Q4, and Q7 from Figure 4. μ and σ are the average and standard deviation obtained for the fits, respectively. The $\Delta\mu$ values represented by the blue vertical bars are the average differences for the corresponding questions. Finally, the resulting p-values are also displayed.

This is the case, given how the participants' main task is not related to the robot. Thus, the RQ1 result suggests that the participants had gotten used to the robot's behavior by the end of the first trial, and felt no need to reappraise its behavior on the second trial if it did not become different.

Although the RQ2 result may suggest that the participants preferred the awareness model over the dummy model, the extracted data do not conclusively show that this is not due merely to it being more comfortable to work alongside the robot when it is moving slowly. If it were, then the results would become ambivalent in regard to the contribution of the awareness model's inclusion. Still, arguments can be made for why it may not be so likely as it first seems that the RQ2 result merely points to a slow working-speed preference.

The first argument begins by considering that the Q3, Q5, and Q9 answers pointed to the participants feeling generally safe sharing the workspace with the robot, regardless of trial order: in other words, this was true even for the case where the second trial was with the dummy model. Given that most participants in the "Model $\rightarrow$ Dummy" population did not notice a difference, the implication is that the participants were already comfortable with the higher speed of the robot, since they would have assumed that it remained unchanged in the second trial. Thus, decreasing the speed in this case would in fact risk *increasing* discomfort, as the robot would risk being in the way of the participant for a longer time, rendering interactions with it annoying.

The second argument hinges on the fact that participants did not notice a change in behaviour when the dummy model was the second trial. If the high speed were uncomfortable, participants would have had a propensity to seek out or grow affinity for the slower behaviour mode available to the awareness model. Noticing a difference in the robot behaviour between trials in the "Model $\rightarrow$ Dummy" case would then be more likely. The lack of support of this hypothesis by our results suggests that the high speed was not uncomfortable. Therefore, the RQ2 is likely not an indication of preference for a slower robot working behaviour.

Another possibility allowed by the RQ2 result is that the awareness model caused the robot to obstruct the participants' path less often, leading to a more comfortable work experience. This possibility is less damning of the awareness model than the former, but points to a reduced ability of the model to foment trust. The ideal alternative would be that participants' trust increases due to their realisation that the robot responds to their awareness of it, but besides there existing several implementation and conveyance difficulties associated to teaching this quickly enough through the robot's behaviour alone, it is not supported by evidence, as can be seen by the responses to Q8. Thus, it is quite likely that the positive reaction to the model was mainly due to the robot being out of the way of the participants more often.

As referred in Sec. II trust is depends on a number of factors: for this experiment, we were only interested in the mental state and perceived safety of the participant. The mental state results were optimistic, but the perceived safety ones were conflicting due to the unfavorable Q8 value despite favorable Q1 and Q9 values.

Generally speaking, the obtained results point to the implementation of awareness models for industrial robot behavior adjustment as a viable avenue for improving worker safety, comfort and trust under HRI scenarios, especially provided that further honing of techniques is performed. Participants generally felt safe and willing to perform HRC with the robot, and reported increased feelings of safety and trust due to the implemented model, which was the desired outcome.

On the other hand, the implementation and experiment displayed some limitations, most notably evidenced by the RQ2 result: the difference in opinion between the two populations clearly did not go beyond being statistically significant. Although it suggests that this avenue of research has potential, it also points to our implementation requiring significant improvements. A few of the required improvements were also noted during the experiment, namely a need to reduce the delay between the generation of a safety flag and subsequent behavior alteration; and improved speed data via the use of a more sophisticated speed calculation algorithm allied with smoothed bounding box data.

VI. Conclusion and Future work

In this paper, we integrated a novel modelling framework for human awareness in robot behavioural adaptation, integrated in industrial collaborative scenarios. A simulated environment was used to test this framework with a set of voluntary participants, tasked with object carrying while traversing a path also crossed by a robotic arm doing its own placement task. Level of trust and perceived behavioural difference were evaluated via questionnaires, post experimental sessions. Statistical evaluation of obtained results revealed a preference for the awareness model being active, albeit by a small margin, even despite users not being informed about its existence in either trial. Results also revealed users may not demonstrate this preference, depending if rapid functioning speeds are already habitual or found to be comfortable, in which case users have become themselves used to adapting their behaviour and slower robot speeds would instead negatively affect their trust.

Naturally our study leaves some matters open-ended, which we intend to address as future work. Methodologically, several avenues can be taken to make the results more robust: the use of trajectory changes as an objective trust measure complementing the questionnaire; use of active perception for better gaze classification; or adjusting the AST to the robot's current pace. It is also worthwhile to determine how efficiency is affected by the awareness-based behavioural change: given the necessity of industrial settings to sustain a certain throughput, negatively affecting this could invalidate the integration of an awareness model in collaborative robots, regardless of its usefulness for trust and safety. Additionally, it is unclear whether user preference was necessarily tied with perceived human awareness by the robot, or simply related to its slower functioning speed. It would also be interesting to exploit motion planning dependent on human awareness and motion, for instance with the robot actively moving away from the user in addition to adapting its speed, as well as assess production efficiency here. Besides more closely simulating an industrial environment, to mitigate these ambiguities, we intend to improve speed calculation for a faster and more fluid change, as well as adapt actuation to the angle between the robot and the person facing their trajectory. Reducing gaze detection variability could also be useful, and may be achieved through a multi-camera setup. Finally, it will be necessary to test our approach with multiple simultaneous users, to further approach a real scenario.

References

[1] W. Li, Y. Hu, Y. Zhou, and D. T. Pham, "Safe human–robot collaboration for industrial settings: a survey," *Journal of Intelligent Manufacturing*, 2023.

[2] Z. Bi, C. Luo, Z. Miao, B. Zhang, W. Zhang, and L. Wang, "Safety assurance mechanisms of collaborative robotic systems in manufacturing," *Robotics and Computer-Integrated Manufacturing*, vol. 67, p. 102022, 2021.

[3] P. Chemweno, L. Pintelon, and W. Decre, "Orienting safety assurance with outcomes of hazard analysis and risk assessment: A review of the ISO 15066 standard for collaborative robot systems," 2020.

[4] F. K. Konstantinidis, N. Myrillas, S. G. Mouroutsos, D. Koulouriotis, and A. Gasteratos, "Assessment of industry 4.0 for modern manufacturing ecosystem: A systematic survey of surveys," *Machines*, vol. 10, no. 9, 2022.

[5] D. S. Valentina, V. D. Pasquale, V. Giubileo, and S. Miranda, "Human-Robot Collaboration: An analysis of worker's performance," *Procedia Computer Science*, vol. 200, no. 2019, pp. 1540–1549, 2022.

[6] M. Lewis, K. Sycara, and P. Walker, "The Role of Trust in Human-Robot Interaction," *Studies in Systems, Decision and Control*, vol. 117, pp. 135–159, 2018.

[7] J. D. Lee and K. A. See, "Trust in automation: Designing for appropriate reliance," *Human Factors: The Journal of Human Factors and Ergonomics Society*, vol. 46, pp. 50 – 80, 2004.

[8] P. A. Hancock, D. R. Billings, K. E. Schaefer, J. Y. Chen, E. de Visser, and R. Parasuraman, "A meta-analysis of factors affecting trust in human-robot interaction," *Human Factors: The Journal of Human Factors and Ergonomics Society*, vol. 53, pp. 517 – 527, 2011.

[9] R. Mittu, D. Sofge, A. Wagner, and W. F. Lawless, "Robust intelligence and trust in autonomous systems," *Robust Intelligence and Trust in Autonomous Systems*, no. April 2016, pp. 1–270, 2016.

[10] N. Tenhundfeld, M. Demir, and E. de Visser, "Assessment of Trust in Automation in the "Real World": Requirements for New Trust in Automation Measurement Techniques for Use by Practitioners," *Journal of Cognitive Engineering and Decision Making*, vol. 16, no. 2, pp. 101–118, 2022.

[11] P. Schroepfer and C. Pradalier, "Why There is No Definition of Trust: A Systems Approach With a Metamodel Representation," *2023 32nd IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*, pp. 1245–1251, 2023.

[12] K. Hald, M. Rehm, and T. B. Moeslund, "Human-robot trust assessment using top-down visual tracking after robot task execution mistakes," *2021 30th IEEE International Conference on Robot and Human Interactive Communication, RO-MAN 2021*, pp. 892–898, 2021.

[13] S. Hopko, J. Wang, and R. Mehta, "Human Factors Considerations and Metrics in Shared Space Human-Robot Collaboration: A Systematic Review," *Frontiers in Robotics and AI*, vol. 9, no. February, pp. 1–15, 2022.

[14] S. Lawrence, N. Azizi, K. Fan, M. Jouaiti, J. Hoey, C. L. Nehaniv, and K. Dautenhahn, "The Impact of Social Norm Violations on Participants' Perception of and Trust in a Robot during a Competitive Game Scenario," *2023 32nd IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*, pp. 1287–1294, 2023.

[15] C. Goubard and Y. Demiris, "Cooking Up Trust: Eye Gaze and Posture for Trust-Aware Action Selection in Human-Robot Collaboration," *ACM International Conference Proceeding Series*, 2023.

[16] G. Campagna and M. Rehm, " Analysis of Proximity and Risk for Trust Evaluation in Human-Robot Collaboration * ," *2023 32nd IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*, pp. 2191–2196, 2023.

[17] S. Li, P. Zheng, S. Liu, Z. Wang, X. V. Wang, L. Zheng, and L. Wang, "Proactive human–robot collaboration: Mutual-cognitive, predictable, and self-organising perspectives," *Robotics and Computer-Integrated Manufacturing*, vol. 81, p. 102510, 2023.

[18] D. Noton and L. Stark, "Eye movements and visual perception," *Scientific American*, vol. 224, no. 6, pp. 34–43, 1971. Place: US Publisher: Scientific American, Inc.

[19] D. Noton and L. Stark, "Scanpaths in eye movements during pattern perception," *Science*, vol. 171, no. 3968, pp. 308–311, 1971.

[20] M. D. Anderson, J. H. Elder, E. W. Graf, and W. J. Adams, "The time-course of real-world scene perception: Spatial and semantic processing," *iScience*, vol. 25, p. 105633, Dec. 2022.

[21] L. Gmeindl, Y.-C. Chiu, M. S. Esterman, A. S. Greenberg, S. M. Courtney, and S. Yantis, "Tracking the will to attend: Cortical activity indexes self-generated, voluntary shifts of attention," *Attention, Perception, & Psychophysics*, vol. 78, pp. 2176–2184, Oct. 2016.

[22] K. Dufour, J. Ocampo-Jimenez, and W. Suleiman, "Visual–spatial attention as a comfort measure in human–robot collaborative tasks," *Robotics and Autonomous Systems*, vol. 133, p. 103626, Nov. 2020.

[23] P. Prajod, M. L. Nicora, M. Mondellini, G. Tauro, R. Vertechy, M. Malosio, and E. André, "Gaze Detection and Analysis for Initiating Joint Activity in Industrial Human-Robot Collaboration," 2023.

[24] J. Deng, J. Guo, Y. Zhou, J. Yu, I. Kotsia, and S. Zafeiriou, "Retinaface: Single-stage dense face localisation in the wild," *CoRR*, vol. abs/1905.00641, 2019.

[25] J. Hu, L. Shen, and G. Sun, "Squeeze-and-excitation networks," *CoRR*, vol. abs/1709.01507, 2017.

[26] M. Zarei, A. Solovov, V. Gautam, B. Ferreira, G. M. S. Assunção, A. Marin-Hernandez, and P. Menezes, "Transforming Collaboration: A Vision for Human-Aware Robots to Enhance Worker Safety and Boost Production Efficiency," Feb. 2024. [In Press].