Human Non-Compliance with Robot Spatial Ownership Communicated via Augmented Reality: Implications for Human-Robot Teaming Safety

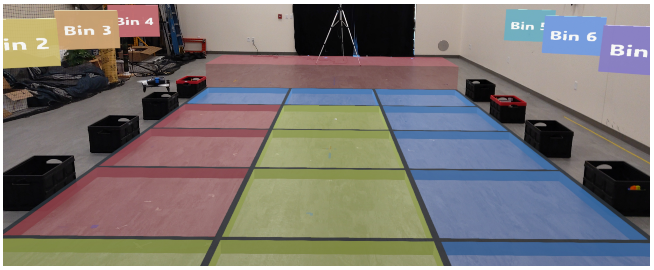
Christine T. Chang[1,2], Matthew B. Luebbers[1], Mitchell Hebert[2], and Bradley Hayes[1]

***Abstract*— Ensuring the safety and efficiency of human workers in environments shared with autonomous robots is of paramount importance. In this work we examine the behavior and attitudes of participants performing tasks in a noisy environment collocated with an autonomous quadcopter robot. Visual communication of spatial ownership and nonverbal (deictic gesture) requests for changes in spatial ownership are facilitated using an augmented reality (AR) head-mounted device that renders a color-keyed grid on the floor. After a request, the robot can alter floor ownership to provide participants with a safe path to complete their work. Participants (*n*=20) in a between-subjects study took part in either a shared space condition (concurrently occupying the work floor with the robot, with obvious rationale for floor ownership) or a turn-taking condition (alternating excursions onto the grid with the robot, without apparent rationale for the floor grid colors). We find consistent evidence of potentially dangerous over-trust in the system that led to non-compliance; notably, 25% of participants intentionally walked across forbidden floor regions during the experiment. We identify design considerations and a variety of user-borne rationale for committing safety violations that designers will need to explicitly take measures to remedy in production AR safety systems.**

I. INTRODUCTION

Due to a confluence of technological availability and utility, humans and robots are increasingly operating in close proximity to each other. The current state of safety in human-robot collaborative and cooperative collocated tasks generally revolves around protecting the human from any contact with the robot, using physical barriers and sensors to pause robot operation in the vicinity of humans. This is oft realized as robots installed within physical cages or within fences in a manufacturing environment, or as ground robots in a well-structured warehouse environment that stop when a human approaches wearing specially instrumented clothing. While effective at preventing negative interactions, these approaches tend to be inefficient and cause frustration.

Research on increasing predictability and interpretability of quadcopter robots by collocated humans [1, 2, 3, 4], in addition to work that predicts human movement [5], tends to assume that a robot should always defer or conform to human preferences independent of the rationale behind them. However, the practical alternative of expecting the human to conform to the robot's movements or demands is less explored. With increased deployment of robots in established processes within warehouse, manufacturing, and even space environments, we must find safe, efficient, and robust ways of collaborating with them.



(a) View through the HoloLens from Home. Eight bins contain task components, a 3x5 grid indicates spatial ownership, and bin labels are selected to request access. Shown is a path to Bin 4.

(b) Looking towards Home in the ARHMD during the trap scenario (no return path) in the Shared Space condition.

Fig. 1: Example views through the HoloLens ARHMD.

This work surfaces insights about human compliance and non-compliance with robot instructions for spatial ownership as delivered via augmented reality in a collocated environment with important safety implications. These insights are gathered from an experiment where human and robotic agents held ownership over different areas of a warehouse floor. We designed and implemented the FENCES (Facilitation of Efficient Nonverbal Collocated Environment Safety) System to enable this interaction. FENCES enables a user to request permission from an autonomous robot to traverse the work floor to reach bins containing parts needed for an assembly task. The robot, an autonomous free-flying quadcopter that is conducting an inventory task, gives permission by giving the human temporary ownership of parts of the floor indicated by hologram coloration (see Figure 1a).

We investigated user behavior and compliance with respect to the FENCES system through an Institutional Review Board-approved, between-subjects study with two conditions: (1) a shared-space condition where the human and robot occupied the floor concurrently, and (2) a turn-taking condition where the human and robot performed their tasks sequentially, with only one of them allowed on the work floor at a time. The main contributions of this work are our findings surrounding human compliance and the justifications

[1] Authors affiliated with University of Colorado Boulder
[2] Authors affiliated with Draper Laboratories

they provide for non-compliance and the subsequent identification of critical design considerations for future AR-based safety systems to incorporate, with implications for safety, trust, and cognitive load.

II. RELATED WORK

The FENCES system and the experimental design in this work are based on insights synthesized from collections of research within the multiple interconnected themes of communication, safety, augmented reality, and human-robot interaction, expanded upon in the subsections that follow.

Communication of Information in AR. McIntire et al. [6] find that stereoscopic 3D displays have equal or superior information communication performance as compared to non-stereo (2D) displays the majority of the time. Augmented reality (one form of stereoscopic 3D display) is a preferable option due to its dynamic visualization capabilities, non-obstruction of the visual field, and relative ease of use. Szafir and Szafir [7] indicate that most past research on human-robot interface design has centered around situational awareness and user control. While our system provides situational awareness in terms of spatial ownership, we look beyond control and towards back-and-forth communication between the human and the robot.

AR for Human-Robot Communication. A rich corpus of work on use cases and experiments exists regarding using AR for human-robot communication [1, 2, 3, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17]. Many systems are designed to improve communication **from the robot to the human**, such as providing insight into motion intent [1, 2, 8], assistive control predictability and legibility [10], aiding teleoperation [11], improving control handovers for autonomous vehicles [9], and using AR-assisted robot gestures [17]. Other systems exist that facilitate communication **from the human to the robot**, including programming or otherwise adjusting the system [12, 13, 14, 15], teleoperation [11], providing boundaries to the robot [4, 16], or functioning as a team [18, 19]. While our work builds on this growing body of research, we specifically address human ***compliance*** with a communicative system as it relates to safety.

AR and Safety. AR is increasingly used to improve worker safety in a variety of environments [20, 21]. Tatić and Tešić [22] presented a case study using AR to improve safety in an industrial environment by providing virtual safety instructions and other information. AR-equipped hard hats are also increasing in prevalence, indicating there is growing acceptance of using AR in high-risk environments [23, 24]. Our work leverages these findings and techniques in ***spaces containing humans and robots***.

AR for Human Safety in Shared Spaces with Robots. A system from Choi et al. [25] provided safety signals in the form of a green, yellow, or red dot for low, medium, and high risk of danger in the corner of the user's field of view. Makris et al. [26] also shaded regions of the workspace in red to denote the robot's space or green to indicate the operator's safe working area. In practice, for our system we found that users had difficulty distinguishing between yellow and green holograms, leading to our use of blue instead of green, but maintaining the overall principle of using color to denote ownership or imply safety.

Some primary applications for our findings include manufacturing and fulfillment centers. There are indications that humans working in close proximity to robots at Amazon Fulfillment Centers might alter their workflow to accommodate or support the work of their robot teammates [27], prompting the authors to ask how AR can further facilitate these human-robot teams. Amazon has already initiated work on this front, as evidenced by the existence of a patent on an AR display for fulfillment center workers [28, 29].

In this work, we utilize augmented reality to provide both a communication modality and spatial ownership information for a person working collocated with an aerial robot and draw conclusions related to human compliance and safety.

III. THE FENCES SYSTEM

The FENCES system includes a Microsoft HoloLens 2 augmented reality head-mounted display (ARHMD), a Parrot Bebop 2 quadcopter robot, a Vicon Tracker motion capture camera system for tracking the robot and the user, and a computer performing sensor fusion, state management, and robot control. In the component descriptions below, the term "user" refers to the human participant.

FENCES was designed as a test bed for analyzing human behavior while interacting with AR and a collocated robot. Within the system, a user can request permission to traverse a controlled space in order to reach a specific goal location. Through the ARHMD, the user can see who has ownership of the spaces on the floor: the robot, themselves, or no one.

A. Microsoft HoloLens 2 ARHMD and User Interface

The Microsoft HoloLens 2 is capable of projecting images and text in the wearer's field of view. The user interface was designed in Unity [30] and consists of the following features, some of which can also be seen in Figures 1a and 1b: (1) A large 3-by-5 grid on the floor, with the 8 bins and table serving as boundaries. (2) The 1.5 x 1.5 meter grid squares are colored red, yellow, or blue, depending on whether they are "owned" by the robot, no one, or the human, respectively. (3) Billboards above each bin are labeled with a corresponding number and always face the user. They can be selected using a HoloLens "air tap" to indicate a user request. Audio feedback is provided when a bin/billboard is selected ("Bin [number] selected."). (4) A "Home" billboard hovers above and behind the home base table. (5) Text confirming completion appears when the experiment has ended. The ARHMD is the sole mode of communication between the user and the system. The user initiates a request to approach a bin by selecting its billboard, and the system may give permission to traverse the floor, indicated by shading the grid squares in blue that the user is permitted to enter.

B. Experiment Manager and Experimenter Interface

All of the robot goal locations, floor color configurations (and thus user access routes), and anticipated bin selections are predetermined by the experimenters and implemented as sequentially reachable states in the system. The states have transition criteria based on specific conditions being met: user location, robot location, and bin request.

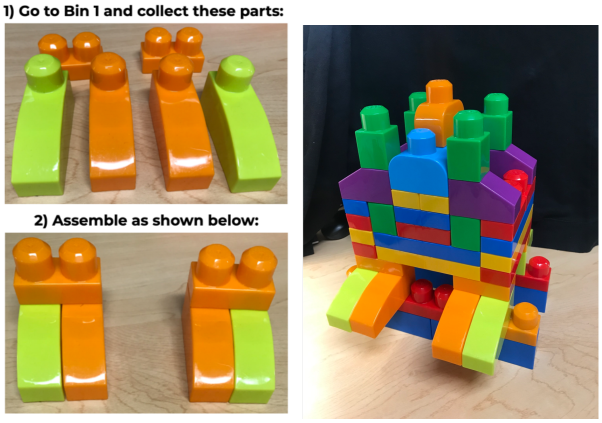
IV. EXPERIMENT DESIGN

We designed this IRB-approved experiment (*n=20*) as a between-subjects study with two conditions. Participants were assigned pairwise randomly to conditions: odd numbered participants were randomly assigned a condition and the following even numbered participant received the opposite condition. Pairwise randomization is an unbiased assignment mechanism to ensure balanced cases when there is a guaranteed pair [31]. We recruited 22 participants, but two trials were discarded due to issues with the motion capture system. The participant population drew from students at our university and was 25% female, 5% nonbinary, and 70% male. On a scale from 1 ("Never interacted with") to 5 ("Extensive experience with"), average experience across participants was 3.1 for robots and 2.4 for AR.

We deployed FENCES in an experimental flight space lab arranged to replicate an assembly environment with eight distributed parts bins along east and west sides (Figure 1a). The task space was approximately *8m x 12m*. A table for the user's workspace was at the south end, deemed "Home Base" for the human. Participants received an orientation at this table, which also contained the assembly workspace and instruction booklet. The experimenter and control equipment were behind protective netting to the west of the table.

After signing the consent form, participants read one page of instructions describing the experimental task. The activity involved constructing a small assembly with Mega Bloks according to a printed booklet of step-by-step instructions with words and photos (see Figure 2). They were instructed to collect the blocks from the bins in a strict order from the bins and told that they should only walk on the blue areas in the grid. They wore the Microsoft HoloLens 2 ARHMD described in Section III-A, which provided the interface for users to request permission from the robot to traverse the space and obtain access to particular parts bins (see Figures 1a and 1b). Simultaneously, the quadcopter robot flew about the room, stopping at bins to simulate inventory checks.

The two conditions were designated "Shared Space" (SS) and "Turn-Taking" (TT). In the **SS condition**, the participant and the robot were permitted to work in the grid area simultaneously, in non-overlapping regions of the space. The robot never returned to Robot Home, a red, robot-only location at the north end of the grid analogous to the human's "home base". The entire task took approximately 15 minutes to complete in the SS condition. In the **TT condition**, the participant alternated with the robot occupying the floor space; while the robot conducted its inventory route, the participant was required to stay in their respective home base, and while the participant was collecting items from a bin and traversing the grid, the robot hovered at Robot Home. After each inventory excursion, the quadcopter returned to Robot Home via the same general path by which it had departed. Since the robot and the participant were never on the grid at the same time, the duration for the entire task increased to approximately 30 minutes. In both scenarios, the "ownership" of the grid squares (robot, human, or neutral/unowned) was communicated to the participant using the virtual grid described in Section III-A and pictured in Figures 1a and 1b.



(a) An example of the instructions provided. Each appeared on separate pages for clarity.

(b) The completed assembly of multicolored MegaBloks.

Fig. 2: The task (a) instructions and (b) final assembly.

These conditions were chosen to investigate behavior in two different yet equally relevant situations: one where the spatial ownership rationale was more recognizable (Shared Space) and one where the spatial ownership rationale and associated safety concerns were less obvious (Turn-Taking). Participants were not provided explicit explanations in either condition about why certain regions were permitted or prohibited, only what the colors denoted. Because we were investigating behavior with respect to the floor ownership as designated in AR, we do not compare their behavior to an AR-free condition. Further, without any indicator of spatial ownership or a significant deviation of the quadcopter's behavior, travel through the space would have been prohibitively unsafe for participants.

Immediately after the task ended, participants answered verbal questions about their experience in an interview with an experimenter. They were asked about their thoughts and behavior during the experiment, as well as whether they perceived any inefficiencies and whether they felt unsafe. Finally, they responded to a survey consisting of Likert (5-point scale) and free response questions.

A. *Land Scenario*

In both conditions, the robot landed on the workspace floor approximately 60% of the way through the experiment. This scenario was designed to reduce the perceived risk involved in the shared space condition, since the robot was not currently flying, potentially tempting the participant to disregard the floor ownership indicators and to return home via a more direct route. In the TT condition, this also served to allow the experimenter to quickly replace the robot's battery with minimal disruption to the experimental timing of quadcopter behaviors between conditions.

B. Trap Scenario

Partway through the experiment, the participant requests access to Bin 4 and access is granted (Figure 1a). Once the participant arrives at Bin 4, only the grid squares along the northern edge remain blue while the rest of the workspace floor turns red, effectively eliminating their route back to Home Base (Figure 1b). The system then begins a 60-second timer, after which the path back to Home Base will reappear. The quadcopter hovers adjacent to Bin 1 in the SS condition and hovers at the Robot Home in the TT condition.

C. Hypotheses

Through the system and experiment described above, we test the following hypotheses: **H1:** Participants will feel safer in TT than in SS due to the reduced proximity to the quadcopter. This will lead to increased deviations in TT, as participants will rely on potentially faulty reasoning (i.e., based only on priors and directly observable features) when determining whether to follow the system guidance. They will also spend more time on the grid in SS due to increased caution near the robot. **H2:** Longer or less direct routes will invoke more deviations from the blue path than shorter or more direct routes. Thus, the land scenario and trap scenario will also invoke deviations that participants will self-justify.

V. RESULTS

A. Mixed Methods in HRI

For a model of mixed methods analysis, we consulted Veling and McGinn's [32] recent survey of 73 qualitative research papers in human-robot interaction, specifically the categories of insights-driven, design, and hypothesis-driven studies. There is a substantial history of prior work in HRI that use qualitative and mixed methods [33, 34, 35]. Using widely accepted qualitative methods we gathered data in semi-structured interviews as well as textual analyses [32], and coded the responses for repeated key words and themes.

B. Trap Scenario

A striking 25% of participants chose to walk through the red and yellow regions to return to Home, disregarding the instructions they had received at the start to only walk through blue regions. Three were in the TT condition while two were in the SS condition, showing similar rates of non-compliance regardless of robot proximity.

- "I knew I was faster than it, so [wherever] it was gonna go I was gonna get out of dodge before it could get there." (TT)

In fact, in one case it seemed that *because* a participant had high trust in the robot's consistency, they disobeyed the floor colors to return to Home Base.

- "I can see that it's safe, so [walked through the red]." (TT)

Eleven of the 20 participants became impatient or assumed a malfunction when the trap scenario began and selected the "Home" button as a solution; 7 participants considered requesting another bin to generate a path, such as one close to Home, or Bin 4 again (the bin where they were trapped); 2 participants admitted that they considered going around the experiment area, outside the grid entirely.

- "I did come close to wondering whether [to walk] around the outside because...nothing will be there..." (TT)

When asked why their path back to Home disappeared, participants generally thought that there was a software issue (n=7) or that the robot was claiming the area (n=9).

- "It seemed like there was a glitch so I broke the rule [and] went straight through." (SS)

However, there was no significant correlation between participants' reasoning about why the path disappeared and their decisions about what to do, suggesting that all of the reasons provided warrant consideration. Furthermore, we can see that when an autonomous system lags, users will not wait patiently; instead they desire ways to work around the lag.

C. Safety, Efficiency, and Trust

One of the most remarkable results from this study was that **all participants felt safe during the experiment, with the exact same distribution across both conditions**. Given the statement, "I felt safe throughout the exercise," all responses were either 4 or 5, with an average of 4.7 (see Table I). Furthermore, 7 participants mentioned the word "safe" in the interview *before* they were asked whether anything felt unsafe about the experience. Two participants used the word "safe" in their response to the question, "Did you find anything inefficient about this process?" One participant (SS condition) believed the system was too safe:

- "It's overly safe...there's not enough risk involved." (SS)

When asked if they thought anything was inefficient about the system, of the 20 participants, 17 identified inefficiencies, while 3 did not. One TT participant described a SS environment that would be more efficient, but SS participants had suggestions as well:

- "There were...times where there was a yellow part that didn't belong to anyone, and it still made me go around." (SS)

Participants also volunteered their thoughts about trust, sometimes combined with issues of safety and efficiency:

- "...I trusted the robot to stay in its red areas." (TT)
- "I trusted it. I think it was very safe at the cost of efficiency, I'd be comfortable with less safety if possible." (SS)

As presented in Table II, participants in the SS condition felt that the robot was more fair than those in the TT condition, suggesting a willingness to sacrifice safety for a perception of fairness. As expected and required by the experiment design, participants were closer on average to the robot in the SS condition (3.7 m) than in the TT condition (4.4 m), with $p<0.0001$ (Figure 3). However, the self-reported feelings of safety showed identical data for the two conditions (Table I), **suggesting that participants felt as safe nearly 3 meters from the robot as they did when it was waiting predictably in Robot Home**.

Participants in the SS condition responded statistically significantly more positively to the statement, "I thought the robot was fair," (see Table II), suggesting that the longer wait time in TT implied a level of unfairness.

TABLE I: Summary responses to select survey items. Parentheses indicate SS responses. 1 = Strongly disagree, 5 = Strongly agree.

Statement	**1**	**2**	**3**	**4**	**5**
I felt safe throughout the exercise.	0	0	0	6(3)	14(7)
I deviated from the given path during the exercise.	15(8)	0	0	1(1)	4(1)
I felt informed throughout the exercise.	1	1(1)	4(1)	7(3)	7(5)

TABLE II: Mean responses, by condition, to select survey items. *$p<0.05$

Statement	**Shared Space**	**Turn-Taking**
I thought the robot was fair.*	4.0	3.1
I liked the way I interacted with the AR device.*	4.7	3.8
I thought the robot was very responsive to my requests.	3.7	3.0
I thought the robot was intelligent.	3.4	2.7

- 'The robot thought its priorities were more important." (SS)

Participants further personified the robot and the system in some of their interview responses:

- "Sometimes you...had to...wait a little bit for it [the robot] to realize, 'Wait, I don't need that square, I can give it up."' (SS)
- "It knew when I was on the field and when I wasn't." (TT)

We also noted how many times each participant checked the robot's position by looking at it while they were on the grid. Data shown in Figure 4 indicate with statistical significance that the higher they perceived its intelligence, the fewer location checks a participant made. Repeated checks for the robot suggest that the human is engaged in tracking the robot. As multitasking increases cognitive load [36], **this suggests that increasing the perception of intelligence can be a powerful way to reduce cognitive load.**

During the pre-experiment briefing, the experimenter interacted with participants on the south side of the table located at Home Base, facing the grid. However, 4 participants chose to work from the north side of the table with their backs to the robot and grid as they were constructing the assembly. One participant chose to work from the west end of the table. This behavior (working without view of the robot) is possibly another indicator of participant trust in the system.

A number of other interesting behaviors were observed. Despite being told to conduct the tasks in the order provided in the instruction booklet, one participant in the TT condition attempted to increase efficiency by gathering blocks from more than one bin per excursion, for example if the following bin was also in the given path, as well as by trying to select future bins while he was on the grid. Other participants tried to anticipate what side of the table the path would start from, waiting on their predicted side, though frequently the path that appeared started from the opposite side as predicted.

D. AR Interface

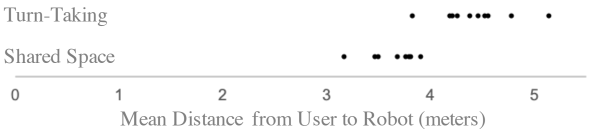


Fig. 3: Mean distance from the robot sorted by condition. Across all participants, mean distance in SS = 3.7 meters, while mean distance for TT = 4.4 meters, $p<0.0001$.

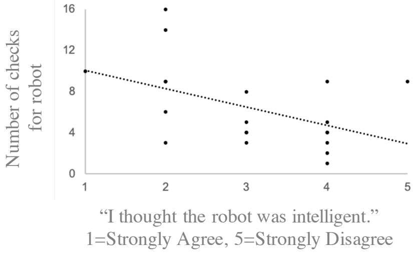


Fig. 4: Relationship between participant response to the item, "I thought the robot was intelligent" and how frequently they looked for the robot while on the grid ($p<0.05$).

Despite having the same user interface across both conditions, participants in the SS condition responded statistically significantly more positively to the statement, "I liked the way I interacted with the AR device" (Table II). They generally liked seeing everything that was in the AR view, except for 1 participant who stated that the grid hologram obscured the robot, making it difficult to see where the drone was, which he also said made him feel less safe. (This participant still responded with 4/5 to "I felt safe throughout the exercise.") Participants consistently made the following suggestions for other information to share in AR: robot intent or priorities (n=8), a timer showing remaining wait time (n=4), task instructions (n=4), and an indicator for the robot location (n=2) were some of the most popular responses.

Participants had a number of suggestions for additional information they would like to see in the display. By showing the red robot-owned regions, we intended to convey the current and near-future movements of the quadcopter. However, over half the participants (n=11) desired even more insight into the robot's intent, priorities, and planning, with which they felt that they could make their own decisions about how to move about the space. However it is unclear whether, with this additional insight, they would continue to stay within the blue grid squares or feel empowered to make their own, potentially deviant, choices for movement around the space. **This information could be useful when designing such systems to know what kind of deviations to expect and how to prime users to use the systems as intended.**

E. Support for Hypotheses

The first hypothesis addresses efficiency between the two conditions. However we found no significant difference after comparing the time SS participants spend on the grid to the TT participants' time. We also analyzed the mean participant distance from the robot as compared to participants' perceptions of safety. Looking at these data in concert, we see that despite SS participants being closer on average to the robot throughout the experiment (Figure 3), they were just as likely to report that they felt safe throughout (Figure I). While most participants (n=15, or 75%) stayed within the blue regions throughout the experiment, the other 25% deviated by walking back through the red and yellow regions during the trap scenario. Considering this is a safety-critical system, we view 25% non-compliance as an alarming result. Of those 25%, one participant also cut corners when taking circuitous routes and took a more direct route back to Home Base in the land scenario. Participant deviations occurred in both SS and TT conditions, and some participants felt the grid ownership guidance was unnecessary. The data partially support H1 in that participants use faulty priors to justify feelings of safety, but there were no differences in perceived safety across the conditions. The data support H2.

F. Limitations

Experimenters were present in the same room as the participant for reasons of safety and practicality, enabling participants to communicate with the experimenters at will, which happened on three occasions. In those instances, a preplanned response was given that did not offer any information about the task or system. Additionally, the motion capture capability varied. Two participants were more difficult to track than others, requiring experimenter intervention to advance the system to its next state. This induced a level of variability in responsiveness to built-in triggers, such as floor colors changing upon the participant's return to Home Base. Participants also had mixed success learning the HoloLens "air tap" gesture, possibly affecting their impressions of the system. This work was also limited by the participant population: all were STEM majors; 70% identified as male.

VI. DISCUSSION AND FINDINGS

In our collocated, physically unprotected environment, participants had to rapidly draw conclusions about the robot's current state, its intentions for the future, and the trustworthiness of its communications.

One of the most surprising results was the demonstrated and reported **overwhelming feelings of safety by all participants**. As explored in Section V, all participants shared that they felt **safe** throughout the experiment, some explicitly stated that they **trusted** the robot to stay in its red areas, and they generally felt **informed** throughout the exercise (Table I). All of this resulted despite not being provided any explanations about the robot's trustworthiness or reliability. Prior work has shown that humans tend to over-trust robots, even in high-risk situations and when they have experience with the robot misleading them [37]. Our work provides further evidence of the potential to over-trust autonomous systems and leads to **Finding 1: Humans working in close proximity to robots appear willing to sacrifice some amount of safety to achieve increased efficiency**.

Lee and See [38] reported that written descriptions induce high levels of initial trust, and that trust in automation begins with faith, then dependability, and finally predictability. Our system initialized trust with the written task description and demonstrated dependability with its consistency until the trap scenario; participants were building their levels of trust in the robot as the task progressed.

Research on trust and safety in high-risk situations contain some key ideas that are useful for understanding the behavior of our participants. Although much of that work relates to trust in people, we observed evidence that participants were personifying the robot. Furthermore, some even viewed it as intelligent (Figure 4). Pidgeon et al. [39] define ***critical trust*** as a "practical reliance on other people combined with a skepticism of the system" [40]. Prior work also demonstrates that it is possible to trust people but not trust dangerous situations; in our experiment, as the trap condition occurred, participants had established some level of trust with the robot, however the system behaved unexpectedly. Five participants then trusted the robot to continue behaving as it has been, simultaneously distrusting the floor colors, ignoring them to return home. Four other participants trusted the system to allow them a path back eventually and waited for this to occur. Further evidence indicates that "trust and distrust are unlikely to lie on the same dimension" [40]. We can conclude that an optimal model of safety requires both critical trust and distrust, leading to **Finding 2: Users desire insight into the decisions and priorities of an autonomous system to help them understand the reasoning behind its actions, decrease frustration, and help them make their own decisions about how to act during uncertainty.**

In human-machine interactions that are facilitated by an interface, it is the interface that establishes shared expectations and trust [41]. The ARHMD and the AR visualizations play a crucial role in the participants' trust development. The virtual images and text are the only methods the system possesses to communicate any information to the user; aside from the actual robot behavior and any prior experience, almost all trust is derived via the ARHMD. By incorporating the suggested features from the participant responses - such as robot intent, prioritization, or wait time - trust and safety can be increased, which informs **Finding 3: Increasing the perception of a collocated robot's intelligence could significantly decrease a worker's cognitive load.**

VII. DESIGN RECOMMENDATIONS

Placing autonomous robots into a shared environment with humans introduces risks and safety considerations. Our study has demonstrated that augmented reality is not necessarily a clear solution to those problems; simply displaying spatial ownership does not dictate safety nor compliance, especially when unexpected events occur. We conclude with recommendations for collocated human-robot systems utilizing AR to aid communication, informed by our results and findings:

Recommendation 1: Provide deviation warnings to deter self-justified rule-breaking that could result in additional risk. **Recommendation 2**: Brief people about the robot's abilities and limitations as part of system training to mitigate intelligence and over-trust perceptions. **Recommendation 3**: Include live visual information to improve real-time understanding of system operation. **Recommendation 4**: Provide training on actions to take during uncertainty; enable the system with corresponding capabilities.

References

[1] D. Szafir, B. Mutlu, and T. Fong, "Communicating Directionality in Flying Robots," in *2015 10th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, Mar. 2015, pp. 19–26.

[2] M. Walker, H. Hedayati, J. Lee, and D. Szafir, "Communicating robot motion intent with augmented reality," in *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*, ser. HRI '18. New York, NY, USA: ACM, 2018, pp. 316–324, event-place: Chicago, IL, USA. [Online]. Available: http://doi.acm.org/10.1145/3171221.3171253

[3] D. Szafir, "Mediating Human-Robot Interactions with Virtual, Augmented, and Mixed Reality," in *Virtual, Augmented and Mixed Reality. Applications and Case Studies*, ser. Lecture Notes in Computer Science, J. Y. Chen and G. Fragomeni, Eds. Cham: Springer International Publishing, 2019, pp. 124–149.

[4] D. Sprute, K. Tönnies, and M. König, "Virtual Borders: Accurate Definition of a Mobile Robot's Workspace Using Augmented Reality," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Oct. 2018, pp. 8574–8581, iSSN: 2153-0866.

[5] J. F. Fisac, A. Bajcsy, S. L. Herbert, D. Fridovich-Keil, S. Wang, C. J. Tomlin, and A. D. Dragan, "Probabilistically Safe Robot Planning with Confidence-Based Human Predictions," *arXiv:1806.00109 [cs]*, May 2018, arXiv: 1806.00109. [Online]. Available: http://arxiv.org/abs/1806.00109

[6] J. P. McIntire, P. R. Havig, and E. E. Geiselman, "Stereoscopic 3D displays and human performance: A comprehensive review," *Displays*, vol. 35, no. 1, pp. 18–26, Jan. 2014. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0141938213000929

[7] D. Szafir and D. A. Szafir, "Connecting Human-Robot Interaction and Data Visualization," in *Proceedings of the 2021 ACM/IEEE International Conference on Human-Robot Interaction*, ser. HRI '21. New York, NY, USA: Association for Computing Machinery, Mar. 2021, pp. 281–292. [Online]. Available: https://doi.org/10.1145/3434073.3444683

[8] M. D. Coovert, T. Lee, I. Shindev, and Y. Sun, "Spatial augmented reality as a method for a mobile robot to communicate intended movement," *Computers in Human Behavior*, vol. 34, pp. 241–248, 2014. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0747563214000612

[9] M. Colley, S. Krauss, M. Lanzer, and E. Rukzio, "How Should Automated Vehicles Communicate Critical Situations? A Comparative Analysis of Visualization Concepts," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 5, no. 3, pp. 94:1–94:23, Sep. 2021. [Online]. Available: https://doi.org/10.1145/3478111

[10] C. Brooks and D. Szafir, "Visualization of Intended Assistance for Acceptance of Shared Control," in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Oct. 2020, pp. 11 425–11 430, iSSN: 2153-0866.

[11] H. Hedayati, M. Walker, and D. Szafir, "Improving Collocated Robot Teleoperation with Augmented Reality," in *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*, ser. HRI '18. Chicago, IL, USA: Association for Computing Machinery, Feb. 2018, pp. 78–86. [Online]. Available: https://doi.org/10.1145/3171221.3171251

[12] M. Luebbers, C. Brooks, C. L. Mueller, D. J. Szafir, and B. Hayes, "ARC-LfD: Using Augmented Reality for Interactive Long-Term Robot Skill Maintenance Via Constrained Learning from Demonstration," in *ICRA 21*, Xi'an China, 2021, p. 7.

[13] S. M. Chacko and V. Kapila, "An augmented reality interface for human-robot interaction in unconstrained environments," in *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2019, pp. 3222–3228, ISSN: 2153-0866.

[14] C. P. Quintero, S. Li, M. K. Pan, W. P. Chan, H. Machiel Van der Loos, and E. Croft, "Robot programming through augmented trajectories in augmented reality," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2018, pp. 1838–1844, ISSN: 2153-0866.

[15] D. Krupke, F. Steinicke, P. Lubos, Y. Jonetzko, M. Görner, and J. Zhang, "Comparison of Multimodal Heading and Pointing Gestures for Co-Located Mixed Reality Human-Robot Interaction," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Oct. 2018, pp. 1–9, iSSN: 2153-0866.

[16] D. Sprute, P. Viertel, K. Tönnies, and M. König, "Learning Virtual Borders through Semantic Scene Understanding and Augmented Reality," in *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Nov. 2019, pp. 4607–4614, iSSN: 2153-0866.

[17] T. Williams, M. Bussing, S. Cabrol, E. Boyle, and N. Tran, "Mixed reality deictic gesture for multimodal robot communication," in *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 2019, pp. 191–201, ISSN: 2167-2148.

[18] E. Rosen, D. Whitney, M. Fishman, D. Ullman, and S. Tellex, "Mixed Reality as a Bidirectional Communication Interface for Human-Robot Interaction," in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Oct. 2020, pp. 11 431–11 438, iSSN: 2153-0866.

[19] K. Chandan, V. Kudalkar, X. Li, and S. Zhang, "Negotiation-based Human-Robot Collaboration via Augmented Reality," *arXiv:1909.11227 [cs]*, Mar. 2020, arXiv: 1909.11227. [Online]. Available: http://arxiv.org/abs/1909.11227

[20] T. A. Sitompul and M. Wallmyr, "Using augmented reality to improve productivity and safety for heavy

machinery operators: State of the art," in *The 17th International Conference on Virtual-Reality Continuum and its Applications in Industry*, ser. VRCAI '19. Association for Computing Machinery, 2019, pp. 1–9. [Online]. Available: https://doi.org/10.1145/3359997.3365689

[21] X. Li, W. Yi, H.-L. Chi, X. Wang, and A. P. C. Chan, "A critical review of virtual and augmented reality (VR/AR) applications in construction safety," *Automation in Construction*, vol. 86, pp. 150–162, Feb. 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0926580517309962

[22] D. Tatić and B. Tešić, "The application of augmented reality technologies for the improvement of occupational safety in an industrial environment," *Computers in Industry*, vol. 85, pp. 1–10, Feb. 2017. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0166361516302718

[23] "Hard Hat for HoloLens 2 Solution," running Time: 23. [Online]. Available: https://visuallive.com/hard-hat-for-hololens-2-system/

[24] "Trimble XR10 with HoloLens 2." [Online]. Available: https://www.microsoft.com/en-us/d/trimble-xr10-with-hololens-2/8smjj5mx7zt7

[25] S. H. Choi, K.-B. Park, D. H. Roh, J. Y. Lee, M. Mohammed, Y. Ghasemi, and H. Jeong, "An integrated mixed reality system for safety-aware human-robot collaboration using deep learning and digital twin generation," *Robotics and Computer-Integrated Manufacturing*, vol. 73, p. 102258, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0736584521001381

[26] S. Makris, P. Karagiannis, S. Koukas, and A.-S. Matthaiakis, "Augmented reality system for operator support in human–robot collaborative assembly," *CIRP Annals*, vol. 65, no. 1, pp. 61–64, 2016. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0007850616300385

[27] N. Scheiber, "Inside an amazon warehouse, robots' ways rub off on humans," *The New York Times*, 2019. [Online]. Available: https://www.nytimes.com/2019/07/03/business/economy/amazon-warehouse-labor-robots.html

[28] A. Delfanti and B. Frey, "Humanly extended automation or the future of work seen through amazon patents," *Science, Technology, & Human Values*, vol. 46, no. 3, pp. 655–682, 2021, publisher: SAGE Publications Inc. [Online]. Available: https://doi.org/10.1177/0162243920943665

[29] U. Madan, M. E. Bundy, D. D. Glick, and J. E. Darrow, "Augmented reality user interface facilitating fulfillment," USA patentus 10,055,645 B1, 2019.

[30] U. Technologies, "Unity Real-Time Development Platform | 3D, 2D VR & AR Engine." [Online]. Available: https://unity.com/

[31] C. Fairhurst, C. E. Hewitt, and D. J. Torgerson, "Using pairwise randomisation to reduce the risk of bias," *Research Methods in Medicine & Health Sciences*, vol. 1, no. 1, pp. 2–6, Sep. 2020, publisher: SAGE Publications Ltd STM. [Online]. Available: https://doi.org/10.1177/2632084319884178

[32] L. Veling and C. McGinn, "Qualitative Research in HRI: A Review and Taxonomy," *International Journal of Social Robotics*, Feb. 2021. [Online]. Available: https://doi.org/10.1007/s12369-020-00723-z

[33] M. Luria, J. Forlizzi, and J. Hodgins, "The Effects of Eye Design on the Perception of Social Robots," in *2018 27th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, Aug. 2018, pp. 1032–1037, iSSN: 1944-9437.

[34] K. S. Welfare, M. R. Hallowell, J. A. Shah, and L. D. Riek, "Consider the Human Work Experience When Integrating Robotics in the Workplace," in *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, Mar. 2019, pp. 75–84, iSSN: 2167-2148.

[35] D. Silvera-Tawil, D. Bradford, and C. Roberts-Yates, "Talk to Me: The Role of Human-Robot Interaction in Improving Verbal Communication Skills in Students with Autism or Intellectual Disability," in *2018 27th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, Aug. 2018, pp. 1–6, iSSN: 1944-9437.

[36] O. Örün and Y. Akbulut, "Effect of multitasking, physical environment and electroencephalography use on cognitive load and retention," *Computers in Human Behavior*, vol. 92, pp. 216–229, Mar. 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0747563218305624

[37] P. Robinette, W. Li, R. Allen, A. M. Howard, and A. R. Wagner, "Overtrust of robots in emergency evacuation scenarios," in *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, Mar. 2016, pp. 101–108, iSSN: 2167-2148.

[38] J. D. Lee and K. A. See, "Trust in Automation: Designing for Appropriate Reliance," *Human Factors*, vol. 46, no. 1, pp. 50–80, Mar. 2004, publisher: SAGE Publications Inc. [Online]. Available: https://journals.sagepub.com/doi/abs/10.1518/hfes.46.1.5030392

[39] N. Pidgeon, J. Walls, A. Weyman, and T. Horlick-Jones, *Perceptions of and Trust in the Health and Safety Executive as a Risk Regulator*. Health and Safety Executive, 2003.

[40] B. C. Gunia, S. H. Kim, and K. M. Sutcliffe, "Trust in Safety-Critical Contexts," in *The Routledge Companion to Trust*. Routledge, 2018, num Pages: 15.

[41] B. Hayes and M. Moniz, "Trustworthy Human-Centered Automation Through Explainable AI and High-Fidelity Simulation," in *Advances in Simulation and Digital Human Modeling*, ser. Advances in Intelligent Systems and Computing, D. N. Cassenti, S. Scataglini, S. L. Rajulu, and J. L. Wright, Eds. Cham: Springer International Publishing, 2021, pp. 3–9.