Transparency in Multi-Human Multi-Robot Interaction

Jayam Patel, Tyagaraja Ramaswamy, Zhi Li and Carlo Pinciroli

Abstract—Transparency is a key factor in the performance of human-robot interaction. A transparent interface allows operators to be aware of the state of a robot and to assess the progress of the tasks at hand. When multi-robot systems are involved, transparency is a greater challenge, due to the larger number of variables affecting the behavior of the robots as a whole. Existing work studies transparency with single operators and multiple robots. Studies on transparency that focus on multiple operators interacting with a multi-robot systems are limited. This paper fills this gap by presenting a novel human-swarm interface for multiple operators. Through this interface, we study which graphical elements are contributing to multi-operator transparency by comparing four “transparency modes”: (i) no transparency (no operator receives information from the robots), (ii) central transparency (the operators receive information only relevant to their personal task), (iii) peripheral transparency (the operators share information on each others’ tasks), and (iv) mixed transparency (both central and peripheral). We report the results in terms of awareness, trust, and workload from a user study involving 18 participants engaged in a complex multi-robot task.

I. INTRODUCTION

Human-robot teams are often envisioned in complex scenarios [1], including humanitarian missions [2], [3], interplanetary exploration [4], ecosystem restoration [5], [6], mining [7], bridge inspection [8] and surgery [9]. The success of these missions depends on effective and efficient team interaction. One crucial requirement to make this vision a reality is making the multi-robot system more transparent [10], i.e., legible and interpretable, for the human operators.

Transparency is a key property of any human-machine interface. Transparent interfaces offer high usability and foster increased situational awareness [10], [11], [12], [13], [14]. Transparent interfaces limit or remove ambiguity, improve trust, and enhance decision-making [15], [16]. Lyon’s models of transparency [17] and the situational awareness-based transparency (SAT) [18] model provide guidelines for an effective interaction between an operator and a machine.

However, these models are designed and tested with a single operator in mind. The problem of designing a transparent interface intensifies when there are multiple human operators, or ‘the machine’ is, in fact, a multi-robot system. This is because the heterogeneous nature and sheer number of combined interactions among operators and robots affect the behavior and the performance of the entire system in non-trivial ways. As an example, imagine an automated warehouse in which hundreds of robots navigate and transport heavy objects. The robots might drop objects, experience hardware failure, or transport an incorrect object. To resolve these issues, the operators must collaborate and resolve the issues using information from robots and other operators.

In this paper, we study the information that each human operator must process and use, which affects cognitive load [19], [20]. To decrease cognitive load, a possible approach is to limit the amount and the type of information presented to the operator. However, this creates a trade-off with transparency, which intuitively suggests more information would be better.

We explore the design space of graphical user interfaces for human-robot interaction, focusing on the multi-human multi-robot scenario, which has, so far, received limited attention. We consider four types of interfaces, each presenting different amounts and types of information, and each corresponding to a specific ‘transparency mode’.

To characterize these modes, we study the effect of the field of view (FoV) an interface offers to the operators. We define the FoV as the observable area an operator can see through the interface camera. As shown in Fig. 1 we categorize the FoV into two regions, based on the distance from the center of the screen: central and peripheral. The central FoV is the region closest to the center; the peripheral FoV is the remaining region. Using this categorization, we produced four types of interfaces, each differing in the way information is displayed:

- No Transparency (NT): no information is available to the operator, used as a baseline for comparison;
- Central Transparency (CT): information is available at the center of the FoV and displayed directly on the robots;
- Peripheral Transparency (PT): information is shown at the boundaries of the FoV as dedicated widgets;
- Mixed Transparency (MT): a combination of central
and peripheral transparency.

We investigate the effects of the transparency modes on operator performance, awareness, task load, and trust in the system. This paper offers two main contributions:

1) A novel augmented-reality-based interface for multi-human multi-robot interaction with the mentioned transparency modes. This interface is an improvement of our mixed-granularity control interface [21], [22] for single operators;

2) A study, which to the best of our knowledge is the first, of the effects of transparency in multi-human multi-robot interaction. Our user study involved 18 participants in teams of 2, each team controlling 9 robots in an object transport scenario.

The paper is organized as follows. In Sec. II we discuss related work on transparency. In Sec. III we present our system and its design. In Sec. IV we report our user study procedures and results followed by analysis in Sec. V and summarize the paper in Sec. VI.

II. BACKGROUND

Transparency is an important research topic in human-machine and human-robot interaction [23], [24]. Transparency affects usability [25], [26], [27], performance [28], [29], trust [30], [31] and explainability [32], [33]. The effect of these factors increases with the type and quantity of information provided to the operator [34], [35]. Coarse information often negatively affects decision time, trust, situational awareness, and performance; in contrast, detailed information typically results in higher cognitive load.

Ghiringhelli et al. [36] first proposed to graphically represent the actions of the robots using augmented reality for single operators. Chen et al. [13] and Mercado et al. [37] tested the impact of transparency on situational awareness, trust, and workload of an operator. Their work is based on simulated point-mass models of the robots, which lack important physical properties of mobile robots and create a ‘reality gap’ between results collected with a simulated environment and the results collected with physical environment [38].

With multiple operators, a novel problem arises: the need for operators to share robots and their information, to achieve a new form of transparency which we call **operator-level transparency**. To the best of our knowledge, there is no study on this topic in the context of multi-robot systems controlled through augmented reality (AR).

III. TRANSPARENCY-BASED INTERACTION SYSTEM

A. System Overview

Our system comprises four components (see Fig. 2):

1) A distributed AR interface implemented as an app for an Apple iPad;

2) A team of robots, pre-programmed with various behaviors to reach a defined point, recognize objects, and perform collective transport;

3) Vicon [39], a motion capture system that localizes the robots and the movable objects in the environment;

4) ARGoS [40], a multi-robot simulator modified to act as ‘software glue’ between the app, the robots, and the Vicon. We replaced the simulated physics engine shipped with ARGoS with a plug-in that receives positional data from the Vicon motion capture system, and developed new sensor and actuator plug-ins that interface with those on-board the robots. With these plug-ins, ARGoS acts as a middleware functionally similar to the **roscore** of ROS.

B. User Interface

Our interface integrates an AR software development kit, Vuforia [41], and the Unity [42] game engine. The interface detects robots and movable objects by their fiducial markers. The robots and objects recognized by the interface are overlaid by virtual objects. The operator can manipulate the virtual objects to send commands to the robots. For example, the operator can translate a virtual object with a one-finger swipe and rotate it with a two-finger twist. It is also possible to select a team of robots by drawing a closed path with a continuous one-finger swipe. Fig. 3 shows a screenshot of the default view of the application. The top-right corner shows the menu buttons to toggle the visibility of the transparency modes. The bottom-left corner shows the real-time global coordinate frame.
(a) Object recognition   (b) New Goal Defined  
(c) Robots approach and push  (d) Transport complete 

Fig. 4: Object manipulation by interaction with virtual objects. The overlaid dotted black arrow indicates the one-finger swipe gesture used to move the virtual object and the overlaid red dotted arrow indicates the two-finger rotation gesture.

(a) Object recognition   (b) New object position

Fig. 5: Robot manipulation by interaction with virtual robots. The overlaid dotted black arrow indicates the one-finger swipe gesture to move the virtual robot and the arrowhead color indicates the moved virtual robots.

C. Granularity of Control

In our previous work [21], [22], we proposed an interface capable of mixed granularity of control for single operators. The ‘granularity’ refers to the possibility to interact at the robot-, team-, and environment-level. Robot- and team-level control allow the operator to send direct commands to individual robots or groups of them. With environment-level granularity, the operator indicates the desired effect of a modification of the environment, e.g., moving an object, and the robots autonomously execute the action. As discussed in [21], [22], mixed granularity of control can outperform any individual level of control. We use these interface features in the present study.

Object Manipulation. The interface overlays virtual objects over the recognized objects (see Fig. 4). The user can move multiple virtual objects to define their respective desired poses, and teams of robots transport these objects to destination. If two or more operators simultaneously control the same object, the system processes the pose received last.

Robot Manipulation. The interface overlays virtual robots over the recognized robots (see Fig. 5). The color of the virtual robot resembles the color of the fiducial markers to differentiate between robots. The user can move multiple virtual robots to define their respective desired poses. If the robot is part of a team performing collective transport, the other robots in the same team pause until the selected robot reaches the desired pose. If the robot is part of a team not involved in collective transport, the selected robot overwrites the goal pose with the newly defined pose and does not affect its team members. If two or more operators simultaneously control the same robot, the system processes the last pose received. Fig. 5 shows how virtual robots look.

Robot Team Selection and Manipulation. The user can draw a closed path with a one-finger continuous swipe to select all the robots in the enclosed region (see Fig. 6). A contour-shaped virtual object with a virtual cube at its centroid appears in the graphical view. The user can manipulate this cube to define the desired pose for the selected team of robots. The user can handle only one team at a time. If two or more operators have the same robot in their team, the robot processes the pose received last.

D. Collective Transport

We employ a collective transport behavior based on the finite state machine (FSM) shown in Fig. 7. This behavior is identical to the one presented in our previous work [21]. The states in the FSM are explained next.

Reach Object. Upon receiving the desired goal pose for the object, the robots organize themselves around the object in a circular manner. These poses are decided based on the number of robots in the team and their distance from the object. This state comes to an end when all the robots reach their designated poses.

Approach Object. The robots move towards the centroid
Fig. 7: Collective transport state machine.

TABLE I: Analogy between the features in our interface and the levels of the SAT model.

| Transparency mode       | Our feature       | SAT level          |
|-------------------------|-------------------|--------------------|
| Central Transparency    | On-robot status   | Level 1 + 2        |
|                         | Robot Direction Pointer | Level 2           |
|                         | Shared Awareness  | Level 3            |
| Peripheral Transparency | On-robot status   | Level 1 + 2        |
|                         | Object Panel      | Level 2            |
|                         | Text-based Log    | Level 3            |

of the object. This state is completed when all the robots are in contact with the object.

**Push Object.** The robots first rotate in-place facing the direction of the goal. The robots then move towards the goal. The robots modulate their speeds to maintain a set distance from the centroid of the object and keep their formation. If a robot breaks the formation, the team switches back to Approach Object, waits for its completion, and subsequently resumes the transport behavior. The state comes to an end once the object reaches the goal position.

**Rotate Object.** The robots rearrange around the object and move along a circular path, thereby rotating the object in place. If any robot breaks the formation, the team rearranges and resumes object rotation. The state ends when the object reaches the desired orientation.

E. Transparency Modes

We present different transparency modes based on the visual FoV of our interface. The interface provides an option to switch between modes. The modes incorporate transparency features that reflect an operator’s perception, comprehension, and projection, i.e., in terms of the three levels described in the SAT model [18]. Table I lists the features, the transparency modes, and the corresponding SAT levels of information.

**No Transparency (NT).** Operators can send control commands, but without access to any feedback information.

**Central Transparency (CT).** The interface overlays each robot with a direction vector and text to report the current task (see Fig. 8). The direction vector indicates the heading of the robot. The color of the vectors resemble the color of the fiducial markers to differentiate between vectors when the robots are close to each other. The interface updates the information 10 times per second. The displayed states are: **Idle, Reach, and Error.** The interface also reports the commands of other operators in real time, to foster collaboration and shared awareness, and to minimize (ideally avoid) conflicting control of the same robots and objects. This information is only visible if an operator is focusing the tablet camera on a specific robot or object, i.e., at the center of the FoV.

**Peripheral Transparency (PT).** The interface displays a robot panel, an object panel, and a text-based log at the edges of the screen (see Fig. 9). The robot panel shows the robots as icons. The highlighted icons correspond to the robots that are moving or performing operator-defined actions. The interface conveys error conditions as blinking red exclamation points. Analogously, the object panel shows the objects as icons. The interface highlights the icons that correspond to an object manipulated by the robots. The interface also offers the option to select an object icon to lock it for future use. By locking, an operator indicates that they intend to work with that object. The interface of other operators highlights the lock with a red icon. An operator can lock only one object at a time, removing past locks when a new one is requested. The text-based log reports the last 3 control actions taken by other operators.

**Mixed Transparency (MT).** This mode offers the features of both central and peripheral transparency.
IV. USER STUDY

A. Hypotheses

The primary purpose of this work is to investigate the effect of different transparency modes on the operators’ awareness, workload, trust, interaction, and performance in a multi-human multi-robot scenario. We based our experiments on three hypotheses:

H1: Mixed transparency (MT) has the best outcome as compared to other modes, in terms of the mentioned metrics.

H2: Operators prefer mixed transparency (MT) to the other modes.

H3: Operators prefer central transparency (CT) to peripheral transparency (PT).

B. Gamified User Study

We devised a gamified scenario in which the operators must use the robots to perform object transport. Teams of two participants had to move 6 objects (2 big and 4 small) from their initial position to a goal region. Big objects were worth 2 points each, and small objects were worth 1 point each. The operators had to work collaboratively to gain as many points as possible (out of a maximum of 8) in a fixed time limit of 8 minutes. The operators could move the big objects using the collective transport behavior, or using the robot or robot-team manipulation modalities at will. Small objects could only be transported with the robot and team control modalities. The operators were given 9 robots to complete the game. Fig. 10 shows the initial positions of the robots, the objects, and the goal region.

C. Participant Sample

We recruited 18 university students (10 female, 8 male) with ages ranging from 19 to 41 (23.78 ± 5.08) in accordance with protocols approved by WPI’s IRB1. No participant had any prior experience with the system.

D. Procedures

Each session of the study approximately took 105 minutes and involved four games. After signing the consent form, we explained the scenario and gave the participants 10 minutes to play with the system. After each game, the participants had to answer a subjective questionnaire. All the participants played the game with all transparency modes (NT, CT, PT, MT) once. We randomized the order of the modes to reduce learning effects.

E. Metrics

We recorded subjective and objective measures for each participant for each task. We used the following measures:

Situational Awareness. We used the Situational Awareness Rating Technique (SART) [43] on a 10-point Likert scale [44] to assess the situational awareness after each game.

Task Workload. We used the NASA TLX [45] scale on a 4-point Likert scale to compare the perceived workload in each game.

Trust. We used the trust questionnaire [46] on a 10-point Likert scale to compare the trust in the interface affected by each transparency mode.

Interaction. We used a custom questionnaire (see Fig. 11) on a 5-point Likert scale to assess the operator-level and robot-level interaction.

Performance. We used the points earned in each game as a metric to scale the performance achieved with each transparency mode.

Usability. We asked the participants to select the features (Log, Robot Panel, Object Panel, and On-Robot Status) they used during the study. Additionally, we asked them to rank the transparency modes from 1 to 4, 1 being the highest rank.

---

1https://www.wpi.edu/research/resources/compliance/institutional-review-board
TABLE II: Results with relationships between transparency modes. The relationship are based on mean ranks obtained through Friedman’s Test. The symbol * denotes significant difference (p < 0.05) and the symbol ** denotes marginally significant difference (p < 0.10). The symbol − denotes negative scales and lower ranking is a good ranking.

| Attributes                  | Relationship | \( \chi^2(3) \) | p-value |
|-----------------------------|--------------|-----------------|---------|
| SART SUBJECTIVE SCALE       |              |                 |         |
| Instability of Situation    | not significant | 4.192          | 0.241  |
| Complexity of Situation     | NT>MT>PT>CT* | 6.435          | 0.092  |
| Variability of Situation    | not significant | 4.192          | 0.241  |
| Arousal                     | NT>MT>PT>CT* | 7.093          | 0.069  |
| Concentration of Attention   | not significant | 4.664          | 0.198  |
| Spare Mental Capacity        | not significant | 3.526          | 0.317  |
| Information Quantity         | MT>CT=PT>NT* | 16.160         | 0.001  |
| Information Quality          | MT>CT>PT>NT* | 11.351         | 0.010  |
| Familiarity with Situation   | not significant | 1.911          | 0.591  |
| NASA TLX SUBJECTIVE SCALE   |              |                 |         |
| Mental Demand                | not significant | 6.169          | 0.104  |
| Physical Demand              | not significant | 3.526          | 0.317  |
| Temporal Demand              | not significant | 0.564          | 0.903  |
| Performance                  | not significant | 4.573          | 0.206  |
| Effort                       | NT>PT>CT>MT* | 9.203          | 0.027  |
| Frustration                  | NT>CT>MT>PT* | 9.205          | 0.027  |
| TRUST SUBJECTIVE SCALE       |              |                 |         |
| Competence                   | not significant | 3.703          | 0.295  |
| Predictability               | PT>CT>MT>NT** | 6.359          | 0.095  |
| Reliability                  | not significant | 4.338          | 0.227  |
| Faith                        | not significant | 1.891          | 0.595  |
| Overall Trust                | PT>MT>CT=NT* | 12.607         | 0.005  |
| Accuracy                     | PT>MT=CT=NT* | 12.214         | 0.007  |
| PERFORMANCE OBJECTIVE SCALE  |              |                 |         |
| Points Scoreed               | not significant | 5.554          | 0.135  |

V. ANALYSIS AND DISCUSSION

Table II shows the results for all the subjective scales and the objective metrics. We used the Friedman test [47] to analyze the data and to assess the significance among different games. We formed rankings based on the mean ranks for all the attributes that showed statistical significance (set to p < 0.05) or marginal significance (set to p < 0.10). Fig. 12 shows the percentage of operators using a particular feature. Fig. 13 reports how participants ranked the transparency modes.

We used the Borda count [48] to calculate the rankings. We inverted the ranking of the negative scales when calculating the Borda count scores. Table III shows the results for each category. This table indicates mixed transparency (MT) as the overall winner in terms of performance, as well as the preferred mode across participants, in accordance with hypotheses H1 and H2. The data suggests that central transparency is better than peripheral transparency, confirming hypothesis H3.

Mixed Transparency. This mode is the overall best choice for operators. The data suggests that this mode has the best information quality and quantity. The operators could pick the information they wanted from central and peripheral regions. However, the operators reported higher perceived complexity and higher arousal than with other transparency modes. The usability tests suggested this mode as the best to understand the teammate’s intent and actions. This justifies mixed transparency as the first choice.

Central Transperancy. This mode has the lowest complexity and arousal. The users found it easier to focus on the center of the screen, and use on-robot status over the side panels. The users reported better information quality and clarity w.r.t to peripheral transparency. 55.55% of the operators preferred this mode over peripheral transparency.

Peripheral Transparency. The operators found the information displayed at periphery of the screen hard to parse and access. This led to increased effort, complexity, and arousal w.r.t the central transparency mode. However, as the information was available on-demand and was not constantly displayed in the FoV, the users reported the lowest amount of frustration. Operators also preferred the icon panels over the text-based log. Additionally, the operators preferred PT over CT to gain awareness of their teammate’s intention.

Performance. Our experiments did not report a substantial difference in performance across transparency modes. We hypothesize that this lack of difference is due to a learning effect across the four runs that each team had to perform.

![Fig. 13: Game Preference.](image-url)

TABLE III: Ranking scores based on the Borda count. The gray cells indicate the leading scenario for each type of ranking.

| Attributes                  | Based on Collected Data Ranking (Table II) | Based on Preference Data Ranking (Fig. 13) |
|-----------------------------|------------------------------------------|------------------------------------------|
| Borda Count                 | NT CT PT MT                               | 17.5 40 39 43.5                           |
|                            |                                          | 18 46 45 72                               |
could not avoid this learning effect through randomization of the transparency modes or pre-study training. The training sessions improved the participants’ understanding of the interface and its features, but did not improve operation proficiency. We attribute this issue to the fact that our study was conducted with real robots, exposing the participants to real-world issues with robots they never encountered before (e.g. noise, failures).

Fig. 14 reports the performance recorded in each game. Fig. 15 shows the increase in performance sorted by game performed (learning effect). Task performance dropped or stayed the same for teams that used no transparency after (learning effect). Task performance dropped or stayed the same for teams that used no transparency after (learning effect). Task performance dropped or stayed the same for teams that used no transparency after (learning effect). Task performance dropped or stayed the same for teams that used no transparency after (learning effect). Task performance dropped or stayed the same for teams that used no transparency after (learning effect). Task performance dropped or stayed the same for teams that used no transparency after (learning effect). Task performance dropped or stayed the same for teams that used no transparency after (learning effect).

VI. CONCLUSION AND FUTURE WORK

In this paper, we studied the effects of different transparency modes in multi-human multi-robot interaction. We classified transparency based on visual FoV. We demonstrated the design of a novel augmented-reality interface that supports different modes of transparency and provides both operator-level and robot-level information.

We performed a user study with 18 operators to assess the effects of these modes of transparency on awareness, workload, trust, and interaction. Mixed transparency outperformed other modes in terms of overall effect and usability, and the participants chose mixed transparency as the best mode. We also compared central transparency with peripheral transparency. More operators preferred central transparency (55.55%) over peripheral transparency (45.45%). Although the difference between the central and peripheral transparency is small, these modes of transparency have their respective benefits. Central transparency offers better robot-level information, while peripheral transparency provides better operator-level information.

We recognize that the sample size of our study is limited, making our study in some ways exploratory from a statistical standpoint. However, the complexity of the task we studied is compelling, especially when compared with existing literature. The next iteration of our work will focus on expanding the user study in two directions. First, understanding the effects of learning and training on transparency, i.e., comparing the need of information for a novice user with the needs of an expert user. Second, studying the effects of our transparency features on the operator’s reaction time, i.e., the time taken to resolve a problem.

ACKNOWLEDGEMENTS

This work was funded by an Amazon Research Award.

REFERENCES

[1] M. Brambilla, E. Ferrante, M. Birattari, and M. Dorigo, “Swarm robotics: A review from the swarm engineering perspective,” Swarm Intelligence, vol. 7, no. 1, pp. 1–41, 2013.
[2] R. R. Murphy, Disaster robotics. MIT press, 2014.
[3] A. P. Hamins, C. Grant, N. P. Bryner, A. W. Jones, and G. H. Koepke, “Research roadmap for smart fire fighting,” NIST, Tech. Rep. 1191, 2015.
[4] D. Goldsmith, “Book review: Voyage to the milky way: the future of space exploration/tv books, 1999,” Sky and Telescope, vol. 98, no. 5, p. 81, 1999.
[5] J. S. Denkenberger, C. T. Driscoll, S. W. Effler, D. M. O’Donnell, and D. A. Matthews, “Comparison of an urban lake targeted for rehabilitation and a reference lake based on robotic monitoring,” Lake and Reservoir Management, vol. 23, no. 1, pp. 11–26, 2007.
[6] T. M. Buters, P. W. Bateman, T. Robinson, D. Belton, K. W. Dixon, and A. T. Cross, “Methodological ambiguity and inconsistency constrain unmanned aerial vehicles as a silver bullet for monitoring ecological restoration,” Remote Sensing, vol. 11, no. 10, p. 1180, 2019.
[7] R. F. Rubio, “Mining: the challenge knocks on our door,” Mine Water and the Environment, vol. 31, no. 1, pp. 69–73, 2012.
[8] J.-K. Oh, G. Jang, S. Oh, J. H. Lee, B.-J. Yi, Y. S. Moon, J. S. Lee, and Y. Choi, “Bridge inspection robot system with machine vision,” Automation in Construction, vol. 18, no. 7, pp. 929–941, 2009.
[9] S. Sirouspour and P. Setoodeh, “Multi-operator/multi-robot teleoperation: an adaptive nonlinear control approach,” in 2005 IEEE/RSJ International Conference on Intelligent Robots and Systems. IEEE, 2005, pp. 1576–1581.
[10] K. A. Roundtree, M. A. Goodrich, and J. A. Adams, “Transparency: Transitioning From Human–Machine Systems to Human-Swarm Systems,” Journal of Cognitive Engineering and Decision Making, p. 1555343419842776, Apr. 2019. [Online]. Available: http://journals.sagepub.com/doi/10.1177/1555343419841276.
[11] A. Bhaskara, M. Skinner, and S. Loft, “Agent Transparency: A Review of Current Theory and Evidence,” IEEE Transactions on Human-Machine Systems, pp. 1–10, 2020. [Online]. Available: https://ieeexplore.ieee.org/document/8982042/
[12] I. Chakraborti, A. Kulkarni, S. Sreedharan, D. E. Smith, and S. Kambhampti, “Explicability? Legibility? Predictability? Transparency? Privacy? Security? The Emerging Landscape of Interpretable Agent Behavior,” Proceedings of the international conference on automated planning and scheduling, vol. 29, pp. 86–96, 2019.
[13] J. Y. C. Chen, S. G. Lakhmani, K. Stowers, A. R. Selkowitz, J. L. Wright, and M. Barnes, “Situation awareness-based agent transparency and human-autonomy teaming effectiveness,” Theoretical Issues in Ergonomics Science, vol. 19, no. 3, pp. 259–282, May 2018. [Online]. Available: https://www.tandfonline.com/doi/full/10.1080/1463922X.2017.1315750
