2025 IEEE International Conference on Robotics and Automation (ICRA)
May 19-23, 2025. Atlanta, USA

Integrating Human-Robot Teaming Dynamics into Mission Planning Tools for Transparent Tactics in Multi-Robot Human Integrated Teams

Audrey L. Aldridge[1], Tyler Errico[2], Mitchell Morrell[2], Cindy L. Bethel[3], John James[2], Christa M. Chewar[2], and Michael Novitzky[2]

2025 IEEE International Conference on Robotics and Automation (ICRA) | 979-8-3315-4139-2/25/$31.00 ©2025 IEEE | DOI: 10.1109/ICRA55743.2025.11128530

Abstract— **This research aims to demonstrate how integrating human-robot teaming dynamics into mission planning tools impacts the abilities of robot operators as they coordinate multiple robot agents during a mission. This was investigated in a pilot study using two inter-robot collaboration modalities and interface tools, which required different human-robot interaction techniques to execute a mission with a team of four robots. In the first modality, the operator manually inserted waypoints for each robot, as they acted as individual agents. In the second modality, the operator used the Planning Execution to After-Action Review (PETAAR) toolset to plot a single waypoint for the team of robots, as the robots coordinated their movement as a group. One novel component of this study is the investigation of how human-robot teaming dynamics and the PETAAR toolset impacted robot operators' real-time situation awareness and perceived cognitive load as well as team performance. Although the teaming modalities differed greatly with respect to the level of operator input needed, the time required to complete the simulation, the participant's perceived cognitive load, and interface usability were very similar for both modalities. In contrast, the results revealed statistically significant differences between the two teaming modalities related to participants' abilities to maintain a wedge formation while remaining situationally aware. Results from this work will be used to guide development of PETAAR along with the design of future studies investigating more complex teaming scenarios and for creating a baseline for comparing future results.**

I. Introduction

In time-constrained situations, small-unit leaders must anticipate orders, prepare their units, and adapt to information updates from intelligence, surveillance, and reconnaissance assets. They create course of action (COA) sketches to outline mission plans, which human teammates can interpret to understand their roles, responsibilities, and actions to perform. However, integrating autonomous assets into human teams currently requires expert programmers to translate COAs into machine-executable instructions, a process that can take minutes to hours depending on complexity. Alternatively, the current state of the art for incorporating autonomous assets leverages a human operator teleoperating a single robot through a video view-screen, reducing a user's ability to remain situationally aware while consuming their cognitive load [1]–[4]. If any autonomy is included, it is considered a lower level of autonomy, such as GPS waypoint following and sometimes baseline obstacle avoidance, to achieve some desired trajectory. To bridge this gap, the Planning Execution to After-Action Review (PETAAR) human-robot interaction toolset aims to integrate multi-robot teams as organic assets into complex mission planning by enabling digital COA sketches to be directly converted into executable models for autonomous agents. As such, both human planners and robot operators should gain a clear understanding of the overall mission objectives, not just the local tasks and behaviors of the robots themselves.

Currently, most computerized mission planning, coordination, and control tools, like Ardupilot Mission Planner [5], [6], Starcraft [7], [8], and ROS Gazebo [9], are heavily localized to be robot-centric, lack widespread adoption of different types of team members, or are unable to portray the robot as an integrated member of a team. Additionally, when poorly designed or are insufficient given a task, control interfaces can worsen a user's situation awareness and overwhelm their cognitive load [4], [10]. In contrast to these interfaces, mission planning and monitoring tools, like the Android Team Awareness Kit (ATAK) [11], [12] and the Tactical Decision Kit (TDK) with its Interactive Tactical Decision Game (ITDG) System [13], are gaining traction in military operations because they are effective at painting the full picture and improving situation awareness [11]. These systems are well liked for their digitized course of action (COA) drawing capabilities, which uses common military graphics and control measures. However, autonomous systems have yet to be integrated into these tools.

PETAAR aims to expand upon the functionality being used in current tools to provide more freedom of control to robot operators and (autonomous) robot assets. The specific portion of PETAAR introduced in this work is the intersection of planning and execution (PE). In order to coordinate large amounts of resources, military planning examples that utilize common military graphics and control measures such as phase lines are leveraged. This is critical because using famil-

*We gratefully acknowledge the support of ARL DCIST CRA W911NF-17-2-0181 for this work. This material is based on Cindy Bethel's work supported while serving at the National Science Foundation. Any opinion, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

[1]Audrey L. Aldridge is with the Robotics Research Center (RRC) at the U.S. Military Academy and the Dept. of Computer Science and Engineering, Mississippi State University, Mississippi State, MS 39762 USA ala214@msstate.edu

[2]Tyler Errico, Mitchell Morrell, John James, Christa Chewar, and Michael Novitzky are Faculty with the RRC, Dept. of Electrical Engineering and Computer Science, United States Military Academy, West Point, NY 10996, USA {Tyler.Errico, Mitchell.Morrell, John.James, Christa.Chewar, Michael.Novitzky}@westpoint.edu

[3]Cindy L. Bethel is a Professor of Computer Science and Engineering, Mississippi State University, Mississippi State, MS 39762 USA cbethel@cse.msstate.edu

iar military symbols eliminates the need for a new language, allowing seamless coordination of human and human-robot teams. Mission planners can manage robot teams without technical/robot expertise, and all levels - from robot operators to company commanders - can quickly understand mission plans.

PETAAR will undergo multiple rounds of development to enhance multi-robot team integration for complex missions. Key experimental measures include situation awareness, time, cognitive load, and mission success. For this pilot study, the "PE" portion of PETAAR was adjusted to enable robot autonomy to coordinate using phase lines (Fig. 1), map-based markers that guide unit movements across time and space. Phase lines were chosen for this first set of studies because they embody a basic level of control in tactical graphics, include both spatial and temporal constraints, and demonstrate the utility of a multi-robot team employing the PETAAR toolset over those requiring human operators to balance these requirements themselves.

The goal of the pilot study was to learn how leveraging PETAAR with a multi-robot team impacted mission success, situation awareness, and cognitive load in comparison to the state-of-the-art of waypoint selection on a graphical user interface (GUI). This was done by investigating how teaming dynamics, like robot-robot collaboration and human-robot interaction, impacted a human user's ability to use the PETAAR toolset and work with a team of robots to complete a mission. It was hypothesized that with PETAAR, participant-robot teams would have greater mission success (shorter completion times, higher robot formation scores), increased situation awareness, and reduced cognitive load.

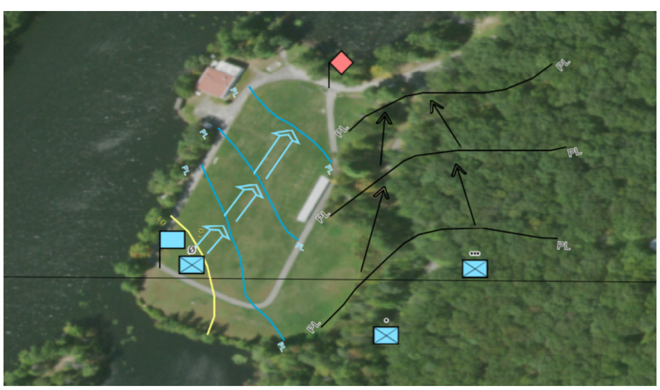

Fig. 1: Mission plan (COA sketch) provided to participants.

II. Method

To demonstrate how different teaming dynamics impact the abilities of robot operators, 22 participants were asked to complete a mission with a team of robots under two experiment conditions, following a within-subjects study design. Participants were provided with similar equipment as to what expert autonomy operators would use to perform this mission in the field. This included a digital COA sketch (mission plan), a computer to interact with their team of robots, and a radio to communicate movement with a higher-ranking officer, referred to as 'higher' for the scope of this work. Participants were evaluated on their performance, situation awareness, cognitive load, and system usability. Prior to beginning the study, approval from USMA's Human Research Protection Program was given under protocol reference number CA-2024-122.

A. Participants

From an *a priori* power analysis, the total sample size was calculated to be n = 36. However, with this being a pilot study aimed to guide future research, the goal sample size was reduced to be n = 20 with 10 participants per condition. To account for unforeseen issues with data collection, 22 participants were recruited from cadets, staff, and faculty at USMA in West Point, NY. There was no compensation or exclusion criteria for the sample population, as a diverse mix of civilians and Army personnel were of interest for this study. Participants ranged from 19 to 77 years of age and included eight higher ranking Army personnel, eight civilians (two former Army), and six USMA cadets.

B. Mission Planning Task

The goal of this study's task was for participants to quickly move their team of four robots in wedge formation across a field, by following a mission plan, while also remaining situationally aware of their surroundings. To establish a baseline for evaluating a robot operator's abilities in more complex scenarios, a simple mission in a static environment was used. By limiting the task to moving in a straight line to three equally spaced waypoints across flat terrain, differences in interface interactions were minimized. This approach reduced confounding factors while preserving participants' ability to assess interface usability. The idea was to create comparable human-robot interaction data for future analysis.

Fig. 1 displays the mission plan that participants executed with their team of four robots. In this COA sketch, the two blue envelope-style symbols on the right were included to show the participants that they were part of a bigger unit, all working toward the same goal (i.e., reaching the red flag). The black lines show the (other) unit's movement, however participants were informed that they were not responsible for executing that part of the mission plan and that they would not see the unit's movement. The blue envelope-style symbol next to the blue flag shows the starting point for the team of four robots. The yellow line signifies the Line of Departure, i.e. the robots' starting place. Participants were told not to plot a waypoint for the Line of Departure or for the red flag. Instead, participants were instructed to focus on the three phase lines, shown as blue lines crossing the open field. The arrows outlined in blue indicate the direction of movement through each phase line.

To reach the objective (red flag), participants had to move their team of robots, following the instructions in the COA sketch (Fig. 1), while keeping the robots in a wedge formation, shown in Fig. 2. At each phase line, participants used the radio to "call to higher", informed higher (one of

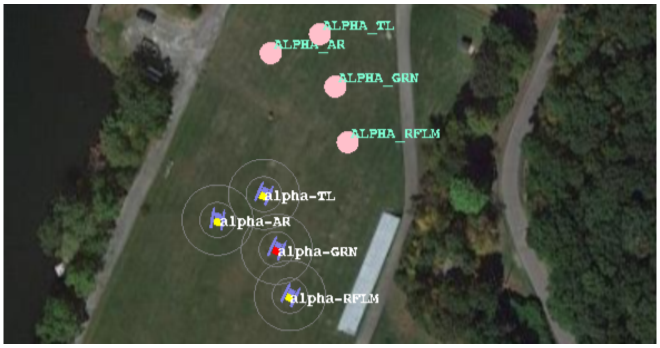

Fig. 2: Robot team in a wedge formation, waiting to proceed to the next set of waypoints.

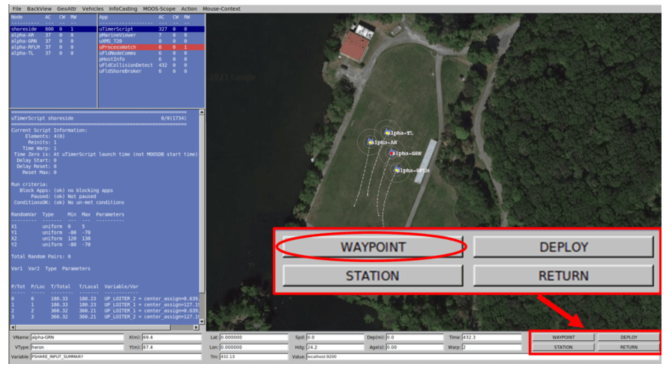

Fig. 3: The pMarineViewer interface with zoomed in section showing waypoint button for robot navigation.

the researchers) of their team's movements, and waited for confirmation from 'higher' to proceed to the next phase line.

C. Interface Tools and Teaming Dynamics

The underlying robot autonomy for these experiments is the Mission Oriented Operating Suite-Interval Programming (MOOS-IvP) toolbox [14], [15]. The mission planning and monitoring systems of interest for this study include 1) pMarineViewer, a GUI for mission monitoring and control that is well-established in the robotics literature [16]–[19] and 2) ATAK, which is currently used in the field by military groups to remain situationally aware of teammates' movements and maneuvers [11], [12]. For this study, pMarineViewer was used as a waypoint point-click interface and for mission monitoring, while the PETAAR toolset leveraged ATAK for its native support of military graphics.

1) Individual-Agent: In the pMarineViewer interface (Fig. 3), each robot in the team acted as an individual agent. Therefore, the 'Point-and-Click' method was used to maneuver the robots individually while maintaining a wedge formation. This meant that when a participant was ready to move his or her team of robots to a phase line, the participant selected the desired robot from a drop-down menu, moved the mouse pointer to the location on the screen where a waypoint was desired, and clicked on that position, dropping a waypoint for the selected robot. This was repeated for each robot in the team for each phase line, although waypoints were only plotted for one phase line at a time. When participants were ready for the robots to move, they clicked the 'Deploy' button. Although the robots could take any position in the wedge formation, participants were informed to plot the robots' waypoints such that the robots' initial alignment was preserved from the starting order through each waypoint sequence, shown in Fig. 2.

If two or more robots got too close to each other (potentially by crossing paths), a yellow circle appeared on the screen indicating a near-collision. For crashes, a red circle appeared on the screen at the site of the collision. It is important to note that near-collisions and collisions did not impact the robots' abilities to complete the mission; they were visual occurrences to inform participants. Once the robots reached the phase line and permission to proceed was received from 'higher', the participant could then plot the waypoints for the next phase line. To reduce user error, a mouse-click added a red dot to the environment. These red dots, as seen in Fig. 2, which signified a dropped waypoint, disappeared after 20 seconds to ensure participants could clearly see the robots' movements.

2) Groupthink: In the ATAK interface, the team of four robots acted as a group with one leader and three followers. For this behavior, participants plotted a single waypoint per phase line. Plotting a single waypoint for a 'Groupthink' team by using this 'Auto' interaction style involved most of the same steps in ATAK as it did in pMarineViewer. The major differences between the two interfaces has to do with the vast array of mission planning tools and symbols provided in ATAK. Mission planning in ATAK took place before the robots were deployed. Referring to the mission plan (COA sketch) provided by the researchers, a participant would approximate the position of each phase line in the ATAK environment, drop a single waypoint on each phase line, and rename the waypoints to 'RP1', 'RP2', and 'RP3'. Realistically, this mission file would then be sent to the robots for them to execute, but because this study was computer-based, the mission file was exported to pMarineViewer for participants to watch as the robots executed mission. When the mission plan from ATAK was viewed in pMarineViewer, the robots still paused at each phase line to wait for the participant to call to 'higher', receive permission from 'higher' to proceed, and then command the robots to move to the next phase line by pressing the "enter" key twice. When in motion, the robots handled maintaining a perfect wedge formation without needing help from participants. Fig. 4 shows three waypoints, one for each phase line, as they should have been plotted in ATAK.

D. Conditions

The conditions for this study were broken down into two sets of teaming dynamics, Individual-Agent and Groupthink. Both of these teaming dynamics included a human-robot interaction style ('Point-and-Click', 'Auto'), a robot-robot collaboration modality (individual-agent, groupthink), and an interface tool (pMarineViewer, ATAK). As a randomized

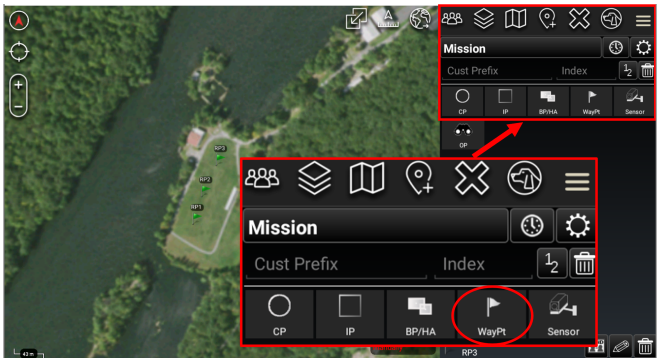

Fig. 4: The ATAK interface with zoomed in section showing waypoint button for robot navigation.

crossover study, counterbalanced randomization was used to assign the condition order in which participants worked with their robot teammates. The condition order assigned to a participant was one of the following: Individual-Agent → Groupthink, Groupthink → Individual-Agent.

E. Assessments

During the mission, participants were evaluated on their situation awareness and performance. Unaware of the timing between questions, participants verbally answered “yes” or “no” to a set of Situation Awareness Global Assessment Technique (SAGAT) probes [20]–[23], which were verbally asked every 30 seconds without freezing or blanking the screen. Examples of these questions for the three dimensions of situation awareness [20] include, “Do you see a body of water on the map?” (perception), “Are there 5 robots on your team?” (comprehension), and “Could the robots crash into each other?” (projection). The idea behind verbally answering real-time SAGAT probes is that, in addition to having to ‘call to higher’, this verbal secondary task resembles a soldier having to maintain radio communication with other teammates. For performance, two researchers scored how well participants were able to keep their robots in a wedge formation. As part of this score, near collisions and actual collisions were considered. Duration, or the recorded time to complete a mission, was also used to evaluate performance. After completing a mission, participants’ cognitive loads were analyzed using the NASA Task Load Index (NASA-TLX) [24]. Interface usability was evaluated using the System Usability Scale (SUS) [25] to determine the “ease of use” equivalency between the two interfaces.

F. Study Protocol

Upon arrival to the Robotics Research Center at the USMA in West Point, NY, participants reviewed and verbally provided consent to participate. Next, participants watched orientation videos on the primary task (mission planning), human-robot interface tools, and the secondary task (verbally answering situation awareness questions). In learning about the primary task and their performance criteria, participants also learned about phase lines and maintaining a wedge formation. Following orientation, participants were asked to provide a back-brief on the task to ensure they learned all necessary information required for the task. Finally, participants watched a short conclusion video wrapping up the orientation and were given the opportunity to ask questions.

Next, participants were given the mission plan (COA sketch overlayed on the environment map), seen in Fig. 1, a radio to call to ’higher’ upon reaching a phase line, and were instructed to use the first assigned teaming dynamic in the respective interface tool to plan and coordinate (in real-time) their team of robot agents in accordance with the mission plan. During this planning task, the participants were asked SAGAT questions for which they responded with a verbal “yes” or “no”. Once the participants and robot team completed the mission, the participants answered a series of questionnaires about their perceived cognitive load and system usability. After completing the questionnaires, participants planned and coordinated (in real-time) their team of robot agents using the second assigned teaming dynamic in the respective interface tool. The planning task in the second interface tool followed the same mission plan as the first. A slightly different set of SAGAT questions were asked and verbally answered “yes” or “no”. The same series of questionnaires followed mission completion. After all questions were answered, participants answered demographics questions, which concluded the study.

III. Results and Data Analysis

This section contains a mixed-methods analysis of results to determine how two conditions of teaming dynamics (Groupthink and Individual-Agent) differed in performance (wedge formation and duration), situation awareness, cognitive load, and usability. Table I displays the average values for duration, maintained wedge formation score, situation awareness, cognitive load, and interface usability before normalization. The Shapiro-Wilk test for normality concluded that the wedge formation scores, situation awareness, and cognitive load data were not normally distributed, although cognitive load was successfully normalized for analysis. Multiple attempts were made to normalize the wedge formation scores and situation awareness data.

Due to only having two groups in the independent variable, a t-test was used to analyze factors individually. The Wilcoxon signed-rank test, a non-parametric alternative to the dependent samples t-test, was used to test for significant differences in central tendency for situation awareness between interfaces and respective teaming dynamics. The Wilcoxon one-sample signed-rank test was used to test for significant differences in the individual-agent’s wedge formation scores versus 12 points, the perfect wedge formation score from the group-think condition. Effect size (r) was calculated for significant results.

A. Performance

Performance was evaluated based on a participant’s ability to keep four robots in a wedge formation, to follow a provided mission plan, to avoid near collisions and actual collisions between robots, and to complete quickly. In

TABLE I: Results: raw mean values for each study measure per scenario and condition

TEAM DYNAMICS	Metric	Duration	Wedge Formation	Situation Awareness	Cognitive Load	Usability
Individual-Agent	avg st dev	4.05 min 1.20	9.45 3.10	87.61 % 6.90	14.17 7.42	68.86 16.69
Groupthink	avg st dev	4.32 min 1.26	12 -	92.78 % 8.83	12.50 9.89	69.43 17.89

scoring the wedge formations, near collisions and actual collisions were taken into consideration. With there being four robots, a participant was awarded one point for each robot that was considered 'in a wedge formation', such as the formation shown in Fig. 2. The side of the formation which contained the longer tail was unimportant in scoring. This meant that for each phase line, participants could receive four points for successfully maintaining all four robots in a wedge formation and a total of 12 points for successfully maintaining a wedge formation for all three phase lines. Near collisions resulted in a reduction of 0.5 points, and actual collisions resulted in a reduction of one point.

1) Wedge Formations: Fig. 5 shows example wedge formations and plotted waypoints (red dots) created by participants. The shape of the plotted waypoints as well as the robots' movements were both observed for wedge performance scores. Following the rules for scoring, the plotted wedge formation in image (a) in Fig. 5 received a score of 3 for having three robots correctly in formation. In image (b), the plotted formation appears to have three robots in a wedge, however this wedge formation received a score of 2.5 for having two robots nearly collide (yellow circle) on their way to the waypoints. Lastly, the plotted wedge formation in image (c) would have received a score of 4, but due to two near collisions (yellow circles), the formation resulted in a score of 3.

From the Wilcoxon one-sample signed-rank test, there exists a significant difference (W = 0, $P < .0001$, $r = 0.71$ (large effect)) in wedge formation scores between the two sets of teaming dynamics. In the Individual-Agent condition, six participants successfully maintained a wedge formation across all three phase lines without near collisions or actual collisions. Only one participant received a score of zero for their total wedge formation score. Eight participants experienced one or more near collisions, and four participants had actual collisions. Additionally in the Individual-Agent condition, the average wedge formation score was 9.45; the median was 10; and the mode was 12. In the Groupthink condition, participants were not tasked with maintaining a wedge formation when plotting waypoints. Instead, the robots handled maintaining a wedge formation based on a single waypoint plotted by a participant for the robot team. Consequently, all participants received the maximum score possible (12) for maintaining a wedge formation.

2) Duration: Duration was measured as the time it took a participant to move all four robots from the starting point to the third phase line. The result of a two means t-test indicates no significant difference between the average duration times when participants used either set of teaming dynamics (Groupthink = 4.315 min, Individual-Agent = 4.046 min). Additionally, the median and mode duration times seen for the two conditions are 3.72 min and 3.28 min, respectively, for Individual-Agent and 4.26 min and 3.32 min, respectively, for Groupthink.

B. Situation Awareness

Situation awareness via a set of real-time SAGAT questions were evaluated for accuracy as a percentage. The Wilcoxon signed-rank test shows a significant difference (W = 103, $P = .02$, $r = 0.53$ (large effect)) in situation awareness when participants used Groupthink teaming dynamics versus when they used Individual-Agent teaming dynamics. Fig. 6 displays the difference in situation awareness scores between the two conditions. In the violin/box plots, the red circles show the average scores. The thicker bold line in each box plot shows the median values (Groupthink = 100%, Individual-Agent = 85.71%). Additionally, the mode for situation awareness scores in the two conditions are 100% for Groupthink and 80% for Individual-Agent.

C. Cognitive Load

Self-reported cognitive load was collected using the NASA-TLX after each mission. The NASA-TLX assesses cognitive load on a scale from 0 to 100, with 100 being the highest score. Cognitive load data was normalized using a log transformation. A two means t-test indicated no significant difference between average cognitive loads when participants used the Groupthink teaming dynamic (2.42) versus the Individual-Agent teaming dynamic (2.68). The median and mode for cognitive load under the Groupthink teaming dynamic are 2.52 and 2.12 and for Individual-Agent are 2.65 and 2.46.

D. Usability

The interfaces' usability was evaluated via the SUS questionnaire, which provides a score from 0 to 100 with 100 being the highest usability score. A two means t-test indicated no significant difference between the average usability for ATAK (69.43) and for pMarineViewer (68.86). In addition to these scores, the median and mode scores for ATAK are 75.0 and 80.0, while the median and mode scores for pMarineViewer are both 67.5.

(a) (b) (c)

Fig. 5: Examples of misshapen wedge formations and robot collisions.

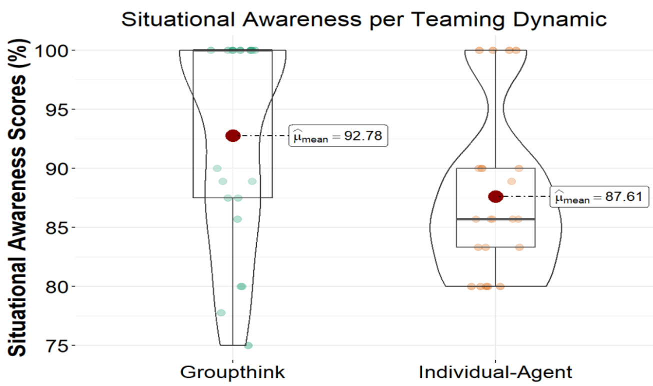

Fig. 6: Real-time situation awareness scores per condition of teaming dynamics.

IV. Discussion

The results indicate PETAAR helped participants remain more situationally aware during the mission. No significant difference was expected in duration and usability; however, cognitive load and situation awareness were predicted to be significantly different between the two conditions. Since there was a significant difference in wedge formation scores but no significant difference in usability scores to suggest a difficulty effect of the ATAK interface (with its full set of military symbols), it is unexpected to see that participants felt their cognitive load was nearly equivalent under both conditions. This brings forth the question of whether a real-time measure of cognitive load would have revealed significant differences between the two conditions, similar to how real-time situation awareness questions resulted in significantly different SAGAT scores.

An important aspect of this experimental method was having the situation awareness questions asked in real-time to mimic the radio communication a robot operator experiences in the field added a complexity to the task by involving and potentially overloading more of the inputs and outputs of a human's senses. Participants were not informed of the timing between situation awareness questions and, therefore, could not anticipate or ignore external sounds. However, one common inexplicable difficulty participants had involved noticing a body of water in the environment; if that SAGAT question is removed, situation awareness would remain statistically significantly better in the Groupthink condition.

Using a more complex mission plan with the same study design would likely result in greater differences in cognitive load, situation awareness, and performance, as this initial pilot study's mission plan was very simple. Additionally, in a more real-world task with more sensory inputs (external sounds) and real-time adjustments to the initial mission plan, the results suggest that using PETAAR may provide a greater situational awareness advantage to robot operators. Although applied here in a military setting, study of PETAAR should expand to broader applications areas for human-robot teaming, such as search and rescue and exploration in dangerous or unknown environments.

V. Conclusion

Although the results of this work were unexpected, they show that with the help of the PETAAR toolset and increased coordination among robots, participants were better able to remain situationally aware of their surroundings. This pilot study not only demonstrates the utility of the experimental methodology and its applicability to future research, but it also provides a foundation for assessing how the incorporation of human-robot teaming dynamics into mission planning tools affects the abilities of robot operators. As military groups more regularly adopt the use of mission planning tools and robot assets, a greater divide in information dissemination to human teammates versus to robot assets may become apparent. As such, there will emerge a need to update the control of, interaction with, and communication to robot assets. The PETAAR toolset is one approach at solving a lot of the complications that come with integrating autonomous robots into human teams.

The ultimate goal of PETAAR is that multi-robot teams are easier to incorporate into large complex mission plans and that the autonomy is given sufficient information about a mission plan so it can execute the mission but also intelligently adapt should drastic changes occur to neighboring units or mission updates. In the future, the PETAAR toolset will incorporate other artifacts for planning and execution like a synchronization matrix that affords agents an automated understanding of their timing among the larger movements. This allows for easier translation from current planning artifacts into an executable robot mission and requires less ad-hoc methods from humans adapting plans.

References

[1] J. E. Cossitt, "Developing a model of driver performance, situation awareness, and cognitive load considering different levels of partial vehicle autonomy," Ph.D. dissertation, Department of Computer Science and Engineering, Mississippi State University, 2022. [Online]. Available: https://scholarsjunction.msstate.edu/td/5466

[2] Y. Yang, B. Karakaya, G. C. Dominioni, K. Kawabe, and K. Bengler, "An hmi concept to improve driver's visual behavior and situation awareness in automated vehicle," *21st International Conference on Intelligent Transportation Systems (ITSC)*, pp. 650–655, 2018.

[3] B. Tam, N. Kottege, and B. Kusy, "Augmented telepresence for remote inspection with legged robots," 12 2017.

[4] J.-M. Georg, E. Putz, and F. Diermeyer, "Longtime effects of video-quality, videocanvases and displays on situation awareness during teleoperation of automated vehicles," in *2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 2020, pp. 248–255.

[5] W. Suparta, A. Basuki, and M. Arsyad, "The development of quadcopter using arducopter apm 2.8 with autopilot for tracking point of drop-off goods," *IOP Conference Series: Earth and Environmental Science*, vol. 1151, no. 1, p. 012032, mar 2023. [Online]. Available: https://dx.doi.org/10.1088/1755-1315/1151/1/012032

[6] M. Oborne. (2024) Ardupilot mission planner. [Online]. Available: https://ardupilot.org/planner/docs/mission-planner-overview.html

[7] N. Othman, J. Decraene, W. Cai, N. Hu, M. Y. H. Low, and A. Gouaillard, "Simulation-based optimization of starcraft tactical ai through evolutionary computation," in *2012 IEEE Conference on Computational Intelligence and Games (CIG)*, 2012, pp. 394–401.

[8] P. Garćıa-Sánchez, A. Tonda, A. M. Mora, G. Squillero, and J. Merelo, "Towards automatic starcraft strategy generation using genetic programming," in *2015 IEEE Conference on Computational Intelligence and Games (CIG)*, 2015, pp. 284–291.

[9] S. Kalluraya, G. Pappas, and Y. Kantaros, "Multi-robot mission planning in dynamic semantic environments," *2023 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 1630–1637, 2022. [Online]. Available: https://api.semanticscholar.org/CorpusID:252222587

[10] M. R. Endsley and D. G. Jones, "Situation awareness oriented design: Review and future directions," *International Journal of Human–Computer Interaction*, vol. 40, no. 7, pp. 1487–1504, 2024.

[11] K. Usbeck, M. Gillen, J. Loyall, A. Gronosky, J. Sterling, R. Kohler, K. Hanlon, A. Scally, R. Newkirk, and D. Canestrare, "Improving situation awareness with the android team awareness kit (atak)," in *Proc. SPIE 9456, Sensors, and Command, Control, Communications, and Intelligence (C3I) Technologies for Homeland Security, Defense, and Law Enforcement XIV*, 2015, p. 295–305.

[12] J. Kim, T. Gregory, J. Freeman, and C. M. Korpela, "System-of-systems for remote situational awareness: Integrating unattended ground sensor systems with autonomous unmanned aerial system and android team awareness kit," in *2022 International Conference on Unmanned Aircraft Systems (ICUAS)*, 2022, pp. 1418–1423.

[13] "Tactical decision kit (tdk) system overview," *Marine Corps Rapid Capabilities Office*, [Presentation].

[14] M. R. Benjamin, H. Schmidt, P. M. Newman, and J. J. Leonard, "Nested autonomy for unmanned marine vehicles with moos-ivp," *Journal of Field Robotics*, vol. 27, no. 6, pp. 834–875, 2010.

[15] M. R. Benjamin. (2024) Moos-ivp. [Online]. Available: https://oceanai.mit.edu/moos-ivp/

[16] M. Novitzky, C. Pippin, T. R. Collins, T. R. Balch, and M. E. West, "Bio-inspired multi-robot communication through behavior recognition," in *2012 IEEE International Conference on Robotics and Biomimetics (ROBIO)*, 2012, pp. 771–776.

[17] A. Spears, M. West, and T. Collins, "Autonomous control and simulation of the videoray pro iii vehicle using moos and ivp helm," in *2013 OCEANS - San Diego*, 2013, pp. 1–10.

[18] L. Paull, S. S. G., M. Seto, and H. Li, "A multi-agent framework with moos-ivp for autonomous underwater vehicles with sidescan sonar sensors," *Autonomous and Intelligent Systems. AIS 2011. Lecture Notes in Computer Science*, vol. 6752, 2011.

[19] R. Mabry, J. Ardonne, J. N. Weaver, D. Lucas, and M. J. Bays, "Maritime autonomy in a box: Building a quickly-deployable autonomy solution using the docker container environment," in *OCEANS 2016 MTS/IEEE Monterey*, 2016, pp. 1–6.

[20] M. R. Endsley, "Measurement of situation awareness in dynamic systems," *Human Factors*, vol. 37, no. 1, pp. 65–84, 1995.

[21] ——, "Theoretical underpinnings of situation awareness: A critical review," *Situation awareness analysis and measurement*, pp. 3–32, 2000.

[22] D. G. Jones and M. R. Endsley, "Can real-time probes provide a valid measure of situation awareness?" in *Proceedings of the Human Performance, Situation Awareness and Automation: User Centered Design for the New Millennium Conference*, 2000.

[23] ——, "Use of real-time probes for measuring situation awareness," *The International Journal of Aviation Psychology*, vol. 14, no. 4, pp. 343–367, 2004.

[24] S. G. Hart and L. E. Staveland, "Development of nasa-tlx (task load index): Results of empirical and theoretical research," in *Human Mental Workload*, ser. Advances in Psychology, P. A. Hancock and N. Meshkati, Eds. North-Holland, 1988, vol. 52, pp. 139–183.

[25] J. Brooke, "Sus - a quick and dirty usability scale," in *Usability Evaluation in Industry*, P. W. Jordan, B. Thomas, B. A. Weerdmeester, and I. L. McClelland, Eds. Taylor & Francis, 1996.