Intelligent Human Systems Integration (IHSI 2025), Vol. 160, 2025, 398–405
https://doi.org/10.54941/ahfe1005889

Experimental Evaluation of Pilot Visual Response to Understand Situational Awareness While Controlling Multiple UAVs

Tyler Shillig, Karl Tschurtschenthaler, and Axel Schulte

Institute for Flight Systems, University of the Bundeswehr Munich, 85579 Neubiberg, Germany

ABSTRACT

This work presents the results of an experiment examining a pilot's situational awareness as they are tasked with controlling increasing numbers of unmanned vehicles. The primary metric for determining situational awareness was the pilot's gaze response time to on-line queries about the location of assets under control. The gaze response times are evaluated to determine the impact of additional drones on the pilot's situational awareness. We also document specific types of degraded situational awareness observed throughout the experiment. Finally, we identify challenges encountered and suggest future work that could provide deeper insight into the research questions addressed here.

Keywords: Human-autonomy teaming, Situational awareness, Manned-unmanned teaming

INTRODUCTION

The field of Human Autonomy Teaming (HAT) focuses on effective interactions between humans and unmanned intelligent agents. In the military domain, Manned-Unmanned Teaming (MUM-T) is an operational concept that specifically focuses on the teaming of manned and unmanned platforms to achieve a common objective. This mode of operations leverages HAT principles to improve the operator's effectiveness. A central concern in both fields is the operator's situational awareness—awareness of the present and future states of the unmanned vehicles under control (Endsley, 2023). A loss of situational awareness has been shown to increase accident rates and lead to non-optimal decisions.

Despite its importance, unobtrusively assessing a pilot's situational awareness remains challenging and has been the subject of extensive research. One reason for this is the difficulty of defining and measuring the operator's internal mental state throughout complex scenarios that involve numerous relevant aspects (Endsley, 1995; Uhlarik et al., 2002). For this work, awareness of the location of Unmanned Aerial Vehicles (UAVs) under control was the only attribute considered when evaluating situational awareness.

As the autonomy of unmanned platforms increases, these vehicles are anticipated to play an ever-larger role in military missions. In future air combat environments, it is likely that pilots will exert supervisory control over unmanned vehicles with a reduced emphasis on traditional piloting tasks. In this role as a battlefield manager, a single pilot may be responsible for multiple unmanned platforms simultaneously (Lewis et al., 2006). It is important to understand the impact that this scenario has on an operator's situational awareness. In addition to limiting the maximum number of platforms a pilot can effectively manage, this understanding will also help in the development of more effective strategies or systems to enhance pilot situational awareness.

The motivation for the experiment was a hypothesis that pilot visual response performance when instructed to fixate on an asset under control is correlated with basic situational awareness. In other words, if the pilot is unaware of the asset's location, their visual response to a fixation request will be inaccurate and a visual search will be necessary in order to find the asset. The additional time for the visual search can be measured and used to gain insight into how the number of assets under control impacts situational awareness.

EXPERIMENTAL SETUP

The experiments were conducted in a single-seat attack helicopter simulator designed as a research testbed for evaluating MUM-T technologies. The pilot was seated in a cockpit with standard helicopter controls. Mission relevant information was presented to the pilot on three touch-screen displays. A large central display measures approximately 29.5x52cm while left and right screens are approximately 13.5x21.5cm. The pilot used these screens to assign a variety of tasks to automated unmanned aerial vehicles (UAVs) in order to achieve mission objectives, including navigation, reconnaissance, and attack.

The simulator visual system is a dome providing a 210 degree horizontal field of view and an approximately 75 degree vertical field of view. The simulator is also equipped with a 6 camera eye tracking setup from SmartEye Inc. that provides the high-accuracy real-time gaze monitoring that was a key element of this experiment.

PARTICIPANTS

The data were collected from five professional military helicopter pilots from a variety of backgrounds. The average age of the participants was 43.4 years with an average of 2265 helicopter flight hours. Although flight tasks were not directly involved in this experiment, the military background of the pilots is relevant as a baseline that is representative of operators in a real-world MUM-T scenario. In fact, this experiment was part of the training regimen to prepare the pilots to fly full MUM-T missions in a larger experimental campaign.

METHOD

The screen layout and positions of experiment-relevant areas are shown in Figure 1. The pilot receives text instructions on the right screen (Area 1) that detail a specific mission task that should be assigned to a designated UAV, for example 'White Attack V9' or 'Pink Fly P3'. The location of friendly and enemy units is presented on the tactical map in Area 2, also shown in Figure 2 with more detail. The tactical map is also where the pilot assigned tasks to UAVs. Preset navigation points (the blue dots P1-P12) and targets of interest (V1-V12) are named and arranged on the tactical map to match a clock face in order to eliminate the need for the pilot to memorize their locations. On the mission timeline (Area 3), the pilot confirmed the task assignment and visualized the UAV task execution order. For certain tasks, the pilot used the sensor page (Area 4) to investigate targets and identify them as friendly or hostile.

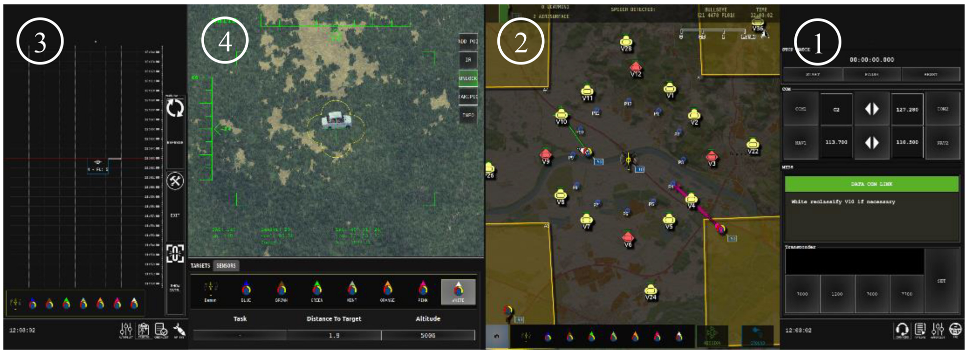

Figure 1: Simulator tasking interface.

Figure 2: Tactical map displaying UAVs, targets, and waypoints.

At the beginning of the experiment the pilot received instructions that involved only a single UAV, e.g. the UAV named "White." As the pilot assigned tasks, additional UAVs up to a maximum of 7 were gradually introduced. A number of instructions for the pilot to look at a specific asset under control (location queries) were also presented to the pilot throughout the tasking sequence.

For each UAV location query, the time it took for a pilot to find the UAV of interest was measured. This was done post-experiment by replaying the mission at.1x speed and visually identifying the time that the saccade from the right screen to the circle of objects was initiated, as well as the time that the pilot's gaze fixated on the target. This duration was the primary metric of interest. The pilot gaze tracks were also qualitatively evaluated to identify general trends associated with degraded situational awareness. The total tasking sequence involved 44 tasks that the pilot assigned to UAVs with 8 UAV position queries directed at the pilot.

RESULTS

Visual inspection of the experimental results in Figure 3 seem to indicate a difference in the median tasking performance when different numbers of unmanned vehicles are involved, with a small but clear reduction in performance between the 2-UAV and 3-UAV tasking conditions.

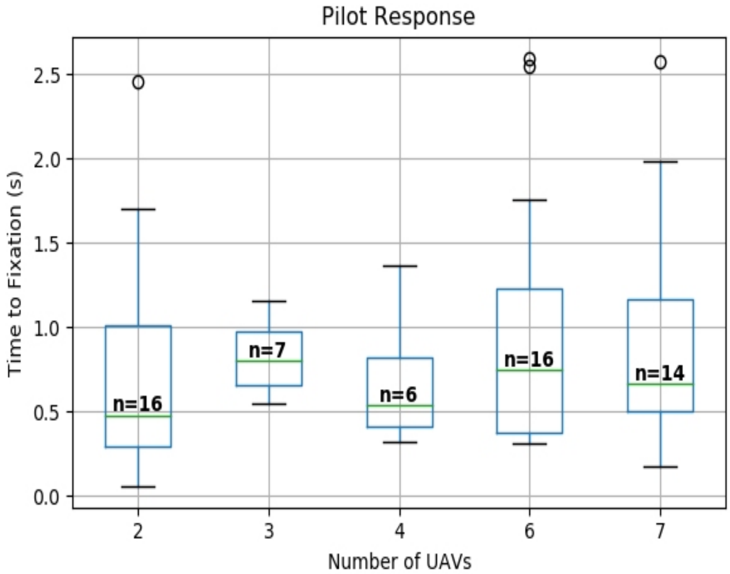

Figure 3: Visual response times as pilot interacts with UAVs.

The raw data is also presented as a scatter plot in Figure 4. Large outliers in the data can clearly be seen in the 2-UAV tasking condition which is likely the result of inadequate preliminary training. This is particularly relevant when interpreting the reduced performance associated with the transition from the 2-UAV to 3-UAV tasking condition. Without the training-induced outliers in the 2-UAV condition, the change in performance would be significant and substantial.

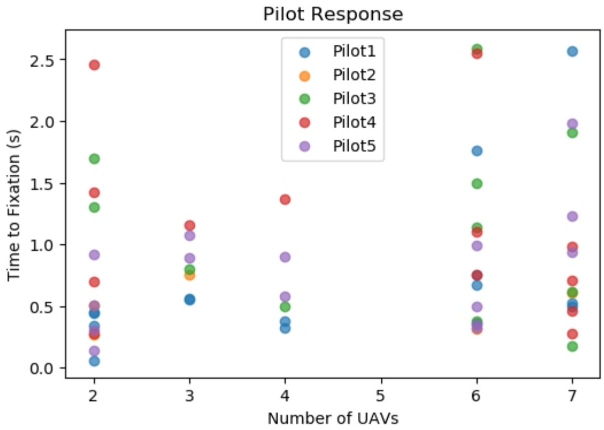

Figure 4: Response time observations.

When tasking three to seven UAVs, the median pilot performance is relatively constant, but the variance within the data increases substantially, as shown in Table 1. The sample size was too small for the results to be significantly different using a Mann-Whitney U test (Table 2).

Table 1. Gaze response data.

UAVs	Samples	Median (s)	MAD (s)	Mean (s)	Std. Dev. (s)
2	16	.48	.22	.74	.66
3	7	.80	.24	.83	.23
4	6	.54	.19	.68	.40
6	16	.75	.39	1.00	.75
7	14	.67	.30	.96	.71

Table 2. Mann-Whitney U Test - statistical significance between datasets.

UAV Dataset	2	3	4	6	7
2	1.00	.17	.74	.17	.18
3		1.00	.37	.81	.74
4			1.00	.51	.41
6				1.00	.9
7					1.00

DISCUSSION

Perhaps the most relevant conclusion to draw from the experimental results is that the operator's capacity to track vehicles under their control is significantly worse than the 5–7 object limit that was anticipated. This assertion is due to the reduced performance when moving from the 2-UAV to the 3-UAV condition. This is likely due to the visually complex environment and diversity of tasks that the pilot was asked to perform. The degree to which the pilot awareness can be improved through additional training and human factors considerations is highly application-dependent. This question is the subject of an extensive body of research (Chen and Barnes, 2014). The ability to track the UAVs may also be improved when they are placed into a more realistic operational environment. Numerous studies have shown the importance of context for the ability of experts to recall the position of objects (Robbins et al., 1996).

A second observation from the experiment relates to the numerous ways in which the gaze track changes as situational awareness degrades. When the operator had a clear understanding of an object's position, the gaze saccade jumps almost immediately (<~100ms) to the vicinity of the queried UAV. As this spatial clarity is reduced, one common failure mode involved switching the object of interest with another. In this case, the gaze saccade initially jumped to the object of confusion and then immediately to the correct one. The initial object of fixation also appeared to display recency bias, i.e. objects that were most recently interacted with tended to be the initial saccade target more frequently. As situational awareness degraded further, saccade accuracy grew increasingly inaccurate.

A gaze track reflecting degraded situational awareness is shown in Figure 5 below, where the gaze history is indicated by the light blue line. The UAV of interest for this query was "White". It can be seen in the figure that the gaze initially starts at the lower right side of the screen where the pilot had just been reading the location query. Instead of accurately jumping directly to White, it jumped to the vicinity of V6 and briefly hesitated there before moving up to V4, then P3, and finally perceiving and fixating on the target. The total time for this gaze sequence was .5 seconds.

Figure 5: Gaze sequence with degraded situational awareness.

FUTURE RESEARCH

One suggested area for further exploration is the impact of tasking modality (voice commands vs. touch) on the pilot's performance. It is unclear how different tasking modalities could alter a pilot's spatial awareness capacity. Engaging with the UAVs via voice could potentially reduce performance due to the reduced engagement of the visual cortex when compared with the coordinated effort required for tasking through touch interfaces. On the other hand, mentally separating the visual observation of the situation from the spoken tasking mechanics could potentially improve performance due to the separation of processing resources (Dixon and Wickens, 2003).

A second suggested area of research is having the UAVs provide verbal feedback when arriving at their destination. It is proposed that this would provide an additional stimulus for the brain, which can increase active memory performance. The increased feedback could also increase system transparency, which has been shown to increase the operator performance and situational awareness (Stowers et al., 2020).

CONCLUSION

This work presented the results from an experiment that assessed an operator's situational awareness as they interacted with increasing numbers of autonomous agents. When controlling more than two UAVs, the pilots appeared to have a reduced situational awareness concerning the spatial location of the vehicles. As additional agents came under control, this situational awareness was further compromised, with increased variance in pilot performance as the primary statistical indicator of this decline. The results could be improved by significantly increasing the number of observations and increasing preliminary training to eliminate performance variation as a result of system familiarization. The experiment serves as a foundation for future research because of the lessons learned and the empirical data that was collected.

REFERENCES

Chen, J. Y. C., Barnes, M. J., 2014. Human–Agent Teaming for Multirobot Control: A Review of Human Factors Issues. IEEE Transactions on Human-Machine Systems 44, 13–29. https://doi.org/10.1109/THMS.2013.2293535

Dixon, S., Wickens, C., 2003. Control of Multiple-UAVs: A Workload Analysis. Presented at the 12th International Symposium on Aviation Psychology, 2003, Dayton OH, United States, p. 6.

Endsley, M. R., 2023. Supporting Human-AI Teams: Transparency, explainability, and situation awareness. Computers in Human Behavior 140, 107574. https://doi.org/10.1016/j.chb.2022.107574

Endsley, M. R., 1995. Toward a Theory of Situation Awareness in Dynamic Systems. Hum Factors 37, 32–64. https://doi.org/10.1518/001872095779049543

Lewis, M., Wang, J., Scerri, P., 2006. Teamwork Coordination for Realistically Complex Multi Robot Systems. RTO-MP-HFM-135. https://doi.org/10.14339/RTO-MP-HFM-135

Robbins, T. W., Anderson, E. J., Barker, D. R., Bradley, A. C., Fearnyhough, C., Henson, R., Hudson, S. R., Baddeley, A. D., 1996. Working memory in chess. Memory & Cognition 24, 83–93. https://doi.org/10.3758/BF03197274

Stowers, K., Kasdaglis, N., Rupp, M. A., Newton, O. B., Chen, J. Y. C., Barnes, M. J., 2020. The IMPACT of Agent Transparency on Human Performance. IEEE Transactions on Human-Machine Systems 50, 245–253. https://doi.org/10.1109/THMS.2020.2978041

Uhlarik, J., Comerford, D. A., Kansas State University, M., 2002. A Review of Situation Awareness Literature Relevant to Pilot Surveillance Functions (No. DOT/FAA/AM-02/3).