Human-Robot Interaction with Drones and Drone Swarms in Law Enforcement Clearing Operations

Richard T. Stone[1], Thomas M. Schnieders[1], Kevin A. Push[1], Stephen Terry[2], Mary Truong[3], Inshira Seshie[4], & Kathryn Socha[1]
[1]Iowa State University
[2]Pennyslvania State University
[3]Northwestern University
[4]University of Maryland-Baltimore County

Police officers often must work alone in clearing operations, a procedure that involves surveying a building for threats and appropriately responding. A partnership between drone swarms and officers has potential to increase the safety of officers and civilians during these high-stress operations and reduce the risk of harm from hostile persons. This two part study examines aspects of trust, situational awareness, mental demand, performance, and human-robot interaction during law enforcement building clearing operations using either a single drone or a drone swarm. Results indicate that single drone use can increase time for operation, but accuracy and safety of clearing is enhanced. Single drone use saw increased situational awareness, a decrease in number of targets missed, and a moderate level of trust. For drone swarms, results indicate significant differences in mental workload from swarm data feeds compared to single drone feeds but no substantial difference in accuracy of finding targets.

INTRODUCTION

The Federal Aviation Administration defines drones as unmanned aircraft systems (UAS) (FAA, 2018). The earliest practical use of UAVs was during World War I where the US military used radio-controlled airplanes to attack German submarines (Keane & Carr, 2013). Drones are able to quickly cover large areas, process computational commands, and replace humans to reduce risk. With a century of technological advancements, drone interest has surged for practical applications reshaping industries in agriculture, film, and search and rescue operations (Cauchard, Kevin, Zhai, & Landay, 2015).

Drone swarms consist of two or more drones working in tandem to accomplish a goal. Some applications of drone swarms include construction (Parker, Zhang, & Kube, 2015), video conferencing (Jones et al., 2016), building inspection and surveillance (Cacace, Finzi, Lippiello, Loianno, & Sanzone, 2014), support in wilderness search and rescue missions (Jones, Berthouze, Bielski, & Julier, 2010), urban search and rescue (Gancet et al., 2010), natural disasters (Erat, Isop, Kalkofen, & Schmalstieg, 2018), and law enforcement (Schnieders et al., 2019).

Law enforcement is seeing a shift to using small drones to many tasks as an inexpensive substitute for traditional aviation units. Small drones are ideal for searching for missing people and plotting navigation for first responders during natural disasters due to their quick deployment time (Miller, 2016). An additional use of small drones in law enforcement is during room clearing operations – a task police officers perform on a regular basis where they physically rush into and check rooms for suspicious activity while trying to apprehend suspects. These close quarters room clearing operations are considered extremely dangerous and complicated (Greenwald, 2002). Due to personnel shortages, especially in rural areas, officers are often forced to handle these difficult situations on their own leaving them more vulnerable in dangerous environments. To mitigate these hazards, a trained officer could quickly deploy a drone or a drone swarm from a police vehicle to gather valuable information before performing a physical search. Maintaining distance away from hostiles while analyzing the situation is a necessity for safety (Miller, 2016). Having a drone enter a doorway first can mitigate the unknown variables of the clearing operation and can prepare an officer for a strategic and efficient cover of the building.

Drone swarms are advantageous in their ability to collect data from multiple vantage points simultaneously, saving time and resources (Jones, Berthouze, Bielski, & Julier, 2010). This is especially important during clearing operation where execution time needs to be short to reduce the possibility of new, complicating variables from emerging. Furthermore, drone swarms can clear buildings from multiple pathways decreasing the possibility of an evasive target avoiding detection. A drone swarm can better emulate the speed of a clearing team than a single drone, taking hostiles by surprise and allowing officers to take control of the situation (Miller, 2016).

This paper proposes two approaches to utilizing drones in law enforcement building clearing operations.

SINGLE DRONE STUDY

Materials and Methods

A total of 14 officers, age 22-63 (Median = 26.5 years, SD = 10.8 years, 2 females, 12 males) with experience in law enforcement (mean = 5.4 years) and building sweeping

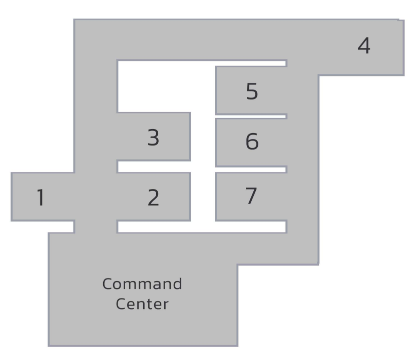

Figure 1: Map of Building

(mean = 3 years) participated in the study. Participants completed both the control and experimental conditions for a total of 28 trials. The order of starting with the control or experimental condition was randomized. Upon completion of a control or experimental condition, each participant completed a post-study survey, a SART-10D Rating Sheet, a modified NASA-TLX, and a Trust in Huma-Robotic Interaction Survey. The control conditions had a single officer performing a clearing operation of a one-story building without the use of a drone. The experimental condition performed a clearing operation of the same building using the drone as assistance. There were eight total scenarios (four control, four experimental) that were randomly assigned. These scenarios indicated if there was no target, an active shooter target, and a civilian target and the scenarios randomized where the potential target may be in the building.

The drone followed directions from the on-the-ground officer. The drone would only enter rooms when instructed and scanned each room from left to right spending no more than 40 seconds in each room unless a target has been located. If a target is identified, the drone would remain in the room and the drone operator would report "Target spotted in room. Target on left/right side of room." Officers than proceeded with regular room clearing operations. Building clearing resumed until the operation had been fully completed.

RESULTS

Performance

Performance was measured by two factors; completion time and target miss rate. A paired-samples t-test was conducted to compare time of operation in drone assistance and no drone assistance conditions. There was a significant difference in time take for drone assistance (M = 215.6s, Sd = 50.79s) and no drone assistance (M = 109.5s, SD = 26.05s) conditions; $t(13) = 11.713$, $p = 2.794e\text{-}08$.

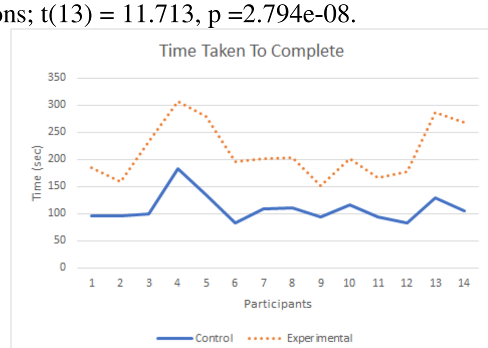

Figure 2: Time take to complete floor

Of the 14 scenarios, it was possible for each group to encounter a civilian or active shooter nine times. The control gropu missed finding the target a total of three out of ninne times, yielding an average miss rate of 33.33% while the experimental group only missed finding the target once out of nine times, yielding an average miss rate of 11.11%.

SART-10D Rating Sheet

Officers had their situational awareness measured using the Situational Awareness Rating Technique known as SART (Taylor, 1989). This questionnaire proposed by Endsley (1988) measures three levels of situational awareness, namely attentional demand, attentional supply, and understanding.

A paired-samples t-test was conducted to compare the results of each metric of the SART. There was no significant difference in situational awareness metrics except for information quality (see Table 1).

Table 1: SART-10D Non-Significant Results

SART-10D Results				
Instability of the Situation				
Control		Experimental		p-Value
Mean	Standard Deviation	Mean	Standard Deviation	
4.9	2.3	5.6	1.6	0.058
Complexity of the Situation				
Control		Experimental		p-Value
Mean	Standard Deviation	Mean	Standard Deviation	
4.4	2.4	4.36	2.3	0.45
Variability of the Situation				
Control		Experimental		p-Value
Mean	Standard Deviation	Mean	Standard Deviation	
4.6	2.3	4.36	2.3	0.27
Alertness				
Control		Experimental		p-Value
Mean	Standard Deviation	Mean	Standard Deviation	
6.5	0.65	6.4	0.85	0.34
Concentration of Attention				
Control		Experimental		p-Value
Mean	Standard Deviation	Mean	Standard Deviation	
6.5	0.65	6.5	0.85	0.5
Division of Attention				
Control		Experimental		p-Value
Mean	Standard Deviation	Mean	Standard Deviation	
5.4	2.03	5.9	1.6	0.055
Spare Mental Capacity				
Control		Experimental		p-Value
Mean	Standard Deviation	Mean	Standard Deviation	
5.2	1.97	4.9	1.99	0.25
Familiarity with Situation				
Control		Experimental		p-Value
Mean	Standard Deviation	Mean	Standard Deviation	
5.9	1.59	6.07	1.14	0.17

A significant difference was found in information quality between the control (M = 5.29, SD = 1.98) and experimental (M = 6.36, SD= 0.63); $t(13) = 2.11$, $p = 0.027$. Information quality is described as 'How good is the information you have gained about the situation? Is the knowledge communicated very useful (high) or is it a new situation (low)?'.

NASA-TLX

Workload is defined in the context of the NASA-TLX as the cost of the user to finish a task, that is, how much mental, physical, or temporal demand was there to complete the building operation. The Raw TLX score was used to eliminate the pairwise comparison between the six subjective metrics to increase experimental validity (Bustamante, E. A., 2008).

A paired-samples t-test was conducted to compare the results of each metric of the NASA-TLX. There was no significant difference in mental workload metrics except for temporal demand (see Table 2).

Table 2: NASA-TLX Non-Significant Results

NASA-TLX Results				
Mental Demand				
Control		Experimental		p-Value
Mean	Standard Deviation	Mean	Standard Deviation	
12.6	5.80	13.2	5.62	0.13
Physical Demand				
Control		Experimental		p-Value
Mean	Standard Deviation	Mean	Standard Deviation	
7.79	5.13	8.36	4.94	0.195
Performance				
Control		Experimental		p-Value
Mean	Standard Deviation	Mean	Standard Deviation	
9.00	7.15	9.07	5.17	0.48
Effort				
Control		Experimental		p-Value
Mean	Standard Deviation	Mean	Standard Deviation	
11.0	5.17	9.29	5.85	0.08
Frustration				
Control		Experimental		p-Value
Mean	Standard Deviation	Mean	Standard Deviation	
4.57	4.22	4.79	3.96	0.34

A significant difference was found in temporal demand between the control (M = 10.14, SD = 5.53) and experimental (M = 8.14, SD = 5.20); p = 0.04.

Trust in Human-Robotic Interaction

Trust is important in any partnership, without trust, police officers may never attempt to use drones in their tasks, and the drone can become a liability in room clearing operations. Twelve trust related questions on a 1 to 7 scale were asked. A score of indicates 'not at all' and a score of 7 indicates 'extremely'.

Table 3: Trust in HRI Results

	Mean	SD
The system is deceptive	1.64	1.08
The system behaves in an underhanded manner	1.50	0.94
I am suspicious of the system's intent, action, or outputs	1.46	0.84
I am wary of the system	2.86	1.51
The system's actions will have a harmful or injurious outcome	1.75	1.01
I am confident in the system	4.75	1.42
The system provides security	5.39	1.60
The system has integrity	4.79	1.67
The system is dependable	4.57	1.34
The system is reliable	4.29	1.27
I can trust the system	4.68	1.23
I am familiar with the system	2.50	1.29

DRONE SWARM STUDY

The drone swarm study used the same participants and building layout as in the single drone study. Each participant watched four different pre-recorded drone video feeds. There were no targets present during the control scenario. In the experimental scenario, a target was always hidden in a random room. Participants were instructed to call out when they say a possible target on the video feed and to mark where they believed they saw the target on a printed map of the testing area. All videos were watched to completion; if the target did not identify a target after watching the entire video, they marked 'No Target" on the post-study survey. All participants first watched a control and experimental feed from a single drone on a single monitor in a random order. They were then randomly assigned to watch footage from a three-drone swarm on either a single monitor or three monitors. In each trail, participants watched a control and experimental scenario in random order (see Figure 3). All monitors had a 24" diagonal viewable size and all drone feeds were displayed at 50% screen size.

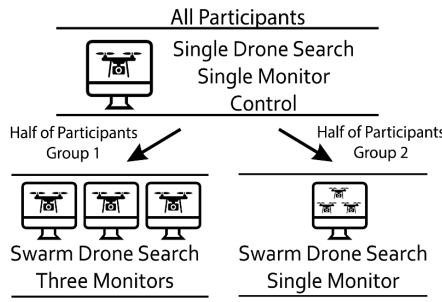

Figure 3: Experimental participant trial group (swarms)

After finishing four runs, the participants completed a NASA-TLX, SART-10D, Trust in HRI survey, and informal interview. Of the 40 total runs, participants correctly identified if there was a target and the target's location in 35 instances. Of the 20 trials involving a single drone, participants always correctly identified the presence of a target, but two misidentified the room in which the hostile was located. Out of the ten trials displaying a swarm feed on multiple monitors, one misidentified the room that the hostile was located, and one misidentified the room in which the target was located and completely missed the target.

NASA-TLX

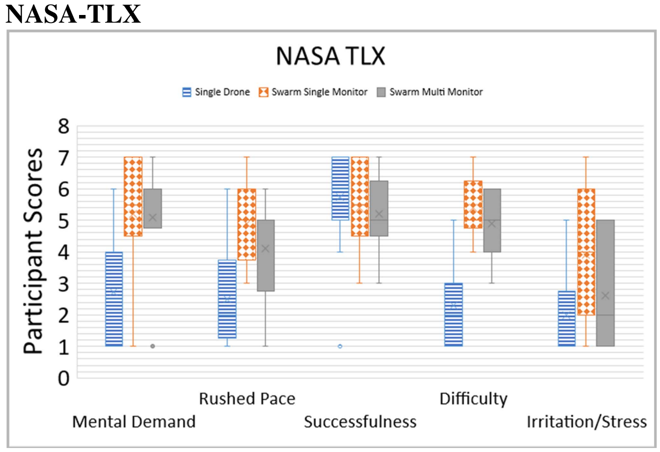

Figure 4: Disitrbution of NASA-TLX scores

A one-way, single factor ANOVA was condcuted. Results suggest that a single monitor feed is more mentally demanding than watching a single drone feed [F(3,16) = 3.966, p = 0.0273]; looking at a multiple monitor swarm feed is more difficult than watching a single drone feed [F(3,16) = 6.020, p = 0.0060]; the pace of clearing opeartion was perceived as more rushed than in the multiple monitor swarm feed [F(3,16) = 3.636, p = 0.0357]. Results for insecurity and stress showed no signficance in either single monitor drone swarm compared to single monitor single drone [F(3,16) = 1.972, p =0.1589] and multiple monitor swarm compare to multiple monitor single drone [F(3,16) = 0.177, p = 0.9106].

SART-10D Rating Sheet

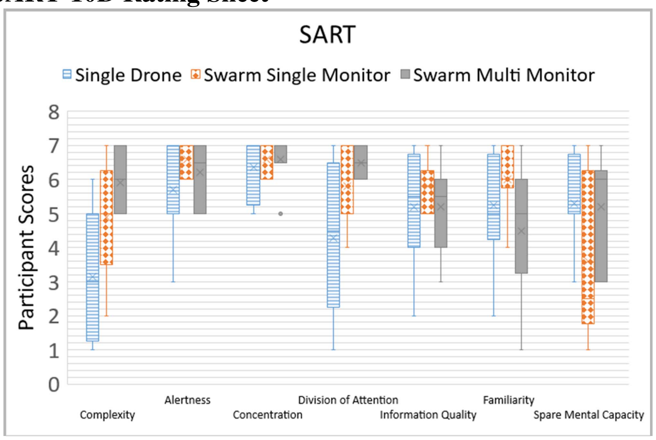

Figure 5: Distribution of SART scores (swarm)

There were no significant differences in situational awareness except for complexity. The results suggest that looking at a multi-monitor swarm feed is more complex than watching a single drone feed [F(2,16 = 5.529, p = 0.0085].

Trust in Human-Robotic Interaction

There was no statistically significant difference in trust metrics for the single drone trial compared to the multi-monitor or single monitor swarm.

Informal Interview

Based on the informal interviews, law enforcement like the idea of using drones to assist in room clearing operations but are currently reluctant to trust the drones due to technology limitations. Participants indicated distrust of automated drones primarily due to the lack of drone control. They indicated having a trained officer with more control and knowledge of the situation would be better than not having another person.

DISCUSSION AND CONCLUSION

The aim of these two studies was to investigate if the use of a single drone during room clearing operations can increase the officers' situational awareness, decrease cognitive load on officers, increase their efficiency during room clearing operation, as well as measure the level of trust between officers and the UAV used in the study. These metrics were extended for use in drone swarms with single monitor and multiple monitor scenarios.

As the results section of this document suggest, the use of a drone during room clearing exercises increased the information quality, increased cognitive demand in the area of time pressure during the task, and the efficiency of room clearing was increased with the use of a single drone. It is also very important to note that none of the areas assessed were negatively impacted using drones. In the drone swarm scenario officers were accurately able to identify hostiles in the drone feed 35 out of 40 runs. There was no statistical significance in officers' accuracy in finding the hostile in any of the different trials.

Task performance is always a critical factor when conducting studies in the workplace, even more so when the task involves potentially life-threatening scenarios. The results of these studies indicate that every officer took significantly longer to complete their clearing task when utilizing the drone. This was impacted by several factors. Firstly, there is a delay between when the officer completes the room to when they communicate the room is clear to when that information is processed by the drone operator and reacted to. In addition, the drone itself slowed progress because movement through doorways is a relatively difficult operation due to the narrow size of the doorframes. This is a common issue for both airborne robots as well as land-based robots that has been well-documented in the literature (Stone, Schnieders, & Zhong, 2017).

The second study looked at alleviating the increase of time used to complete a clearing operation by utilizing a drone swarm. Before practical application of drone swarms in clearing operations could be conducted, an analysis of how to display this information to officers needed to be done. Results of the second study indicated viewing a single drone feed was easier than viewing the swarm feed due to the increased mental demand of tracking multiple drones' search paths. In addition, using multiple monitors to view swarm footage was less stressful than multiple feeds on a single monitor. Participants indicated that this allowed them to more easily distinguish the paths and locations of each drone in the swarm. This potentially poses an issue if larger drone swarms are to be used.

In conclusion, use of a single drone during law enforcement clearing operation is a viable option that does not significantly increase mental demand or decrease situational awareness, while also improving target identification. Using a drone swarm does not impact the accuracy of officers, compared to single drone operations, and significantly decreases the time to complete the operation. However, more work needs to be done to find an ideal, manageable number of drones in a swarm to not significantly increase cognitive demand. Additional work should investigate drone swarm orientation awareness, heading direction, and relative location to be deployed successfully and the drone swarms should be tested in further practical applications.

ACKNOWLEDGEMENTS

This research was supported by the National Science Foundation Grant 1757900. The researchers would like to acknowledge Zhonglun Wang for assistance in running part of the study.

REFERENCES

Bustamante, E. A. (2008). Measurement Invariance of the NASA-TLX". *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 52(19), pp. 1522-1526, DOI:10.1177/154193120805201946.

Cacace, J., Finzi, A., Lippiello, V., Loianno, G., & Sanzone, D. (2014). Aerial service vehicles for industrial inspection: task decomposition and plan execution. *Applied Intelligence*, *42*(1), 49–62. https://doi.org/10.1007/s10489-014-0542-0

Cauchard, J. R., Kevin, J. L. E., Zhai, Y., & Landay, J. A. (2015). Drone & Me: An Exploration Into Natural Human-Drone Interaction. *UbiComp* '15, pp. 361-365.

Endsley, M.R. (1988). Design and evaluation for situational awareness enhancement, *Proceedings of the Human Factors Society 32nd Annual Meeting* (pp. 97–101). Santa Monica, CA: Human Factors and Ergonomics Society.

Erat, O., Isop, W. A., Kalkofen, D., & Schmalstieg, D. (2018). Drone-Augmented human vision: Exocentric control for drones exploring hidden areas. *IEEE Transactions on Visualization and Computer Graphics*, *24*(4), 1437–1446. https://doi.org/10.1109/TVCG.2018.2794058

FAA. (2018, May 25). Unmanned Aircraft Systems. Retrieved from https://www.faa.gov/uas/

Gancet, J., Motard, E., Naghsh, A., Roast, C., Arancon, M. M., & Marques, L. (2010). User interfaces for human robot interactions with a swarm of robots in support to firefighters. *Proceedings - IEEE International Conference on Robotics and Automation*, 2846–2851. https://doi.org/10.1109/ROBOT.2010.5509890

Jones, B., Dillman, K., Tang, R., Tang, A., Sharlin, E., Oehlberg, L., Neustaedter C., & Bateman, S. (2016). Elevating Communication, Collaboration, and Shared Experiences in Mobile Video through Drones. *Proceedings of the 2016 ACM Conference on Designing Interactive Systems - DIS '16*, 1123–1135. https://doi.org/10.1145/2901790.2901847

Jones, G., Berthouze, N., Bielski, R., & Julier, S. (2010). Towards a situated, multimodal interface for multiple UAV control. *Proceedings - IEEE International Conference on Robotics and Automation*, *i*, 1739–1744. https://doi.org/10.1109/ROBOT.2010.5509960

Keane, J. F. & Carr, S. S. (2013). A brief history of early unmanned aircraft. Johns Hopkins APL Technical Digest, 32(3), pp. 558-571.

Miller, W. M. (2016). The use of small unmanned aircraft systems (sUAS) in Law Enforcement.

Parker, C. A. C., Zhang, H., & Kube, C. R. (2015). Blind bulldozing: multiple robot nest construction. In *Proceedings 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2003) (Cat. No.03CH37453)* (Vol. 2, pp. 2010–2015). IEEE. https://doi.org/10.1109/IROS.2003.1248950

Schnieders, T. M., Wang, Z., Stone, R. T., Backous, G., & Danford-Klein, E. (2019). The Effect of Human-Robot Interaction on Trust, Situational Awareness, and Performance in Drone Clearing Operations. *International Journal of Human Factors and Ergonomics*, in press.

Stone, R. T., Schnieders, T. M., & Zhong, P. (2017). Wayfinding in human-robot collaborated exploration: Orientation awareness and its enhancement. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. Austin TX, October 9-13, Vol. 61, pp. 1257-1261 DOI: 10.1177/1541931213601795.

Taylor, R.M. Situational awareness rating technique (SART): the development of a tool for aircrew systems design. *Proceedings of the NATO Advisory Group for Aerospace Research and Development (AGARD) Situational Awareness in Aerospace Operations Symposium (AGARD-CP-478)*; October 1989, p. 17.