Evaluating Mixed Reality and Tablet Technologies in Military Planning

Betsy Guzmán
Department of Industrial Engineering and Management Systems
Univ. of Central Florida
Orlando, FL
betsy.laxton@ucf.edu

Sarah Deresky, Sabrina Taylor, Hawk Wimmer, Ali Momen, Chad Tossell
Warfighter Effectiveness Research Center
U.S Air Force Academy
USAFA, CO
chad.tossell@usafa.edu

Michael Boyce, Joel Cartwright
Department of Behavioral Sciences and Leadership
U.S Military Academy
West Point, NY
michael.boyce@westpoint.edu

Charles Amburn
U.S. Army DEVCOM Simulation &Training Technology Center
Orlando, FL
charles.r.amburn.civ@army.mil

Ben Sawyer
Department of Industrial Engineering and Management Systems
University of Central Florida
Orlando, FL
sawyer@ucf.edu

2022 Systems and Information Engineering Design Symposium (SIEDS) | 978-1-6654-5111-6/22/$31.00 ©2022 IEEE | DOI: 10.1109/SIEDS55548.2022.9799294

Abstract – **Increasingly advanced technologies are penetrating military domains (e.g., air, land, sea, cyber, space) requiring more complex decision-making to support activities that apply across these domains (multi-domain planning & operations). These decisions often require humans to perceive, comprehend, project, and then communicate information in a timely and accurate manner, oftentimes with life-or-death consequences. To support these decisions, Department of Defense leaders are calling for more effective representations and displays of joint warfighting environments. This project addresses this requirement by examining novel technologies for integrating and displaying complex MDO plans for human decision-making. Using a mission planning scenario, we assessed situation awareness (SA), usability, cost, and overall effectiveness of a two-dimensional (2D) representation of a common joint warfighting display on a Samsung tablet against a three-dimensional (3D) display of the same information designed for use in a Microsoft HoloLens mixed reality system. A total of 22 U.S. Air Force Academy cadets were randomly assigned to either use the tablet or the HoloLens to develop and analyze a mission plan and assessed for situation awareness across two scenarios. Interestingly, the HoloLens did not provide any additional SA relative to the tablet. The tablet was also perceived as more usable and effective in terms of cost and overall performance. These results suggest more traditional technologies, such as a tablet, can provide SA at similar levels as more advanced technology with increased usability and affordability.**

Keywords—HoloLens, Human Factors, Military, Simulation

I. Introduction

As military planning, execution, and assessment extends beyond the air, land, and sea to domains such as cyber and space (i.e., multi-domain operations; MDO), displaying information to support situation awareness (SA) for decision-making is more challenging and important than ever. Newer technologies such as virtual reality (VR), augmented reality (AR), and mixed reality (MR) offer new ways to present information that could be easier to integrate MDO-specific information with the key aspects of the battlefield in both training and the real world. VR refers to an immersive experience where the head-mounted display blocks out reality, whereas AR refers to adding a layer of virtual elements over reality, and MR refers to a mixture of both [3]. Although these technologies purport to provide users with enhanced abilities and they are becoming ubiquitous in many military settings, very few empirical studies have assessed their effectiveness in military planning and operations.

We aim to fill this gap with the current study. Specifically, we compare a commercial-off-the-shelf (COTS) MR technology, the Microsoft HoloLens, with a more traditional tablet on providing cadets with SA in mission planning scenarios. The military may assume the former is more effective; however, our results aim to provide data-driven recommendations on how mission planning environments should be designed in the future.

II. Background

AR, VR, and MR technologies have already been shown to provide several benefits. For example, training for dangerous situations can be simulated while keeping the warfighter safe and presenting a situation that might not otherwise be possible [6]. Traditional training scenarios for dangerous situations can be costly to implement and can sometimes pose an opportunity for danger in the training situation. Using mixed realities can remove the danger but still present a faithful rendition of what a warfighter can expect in that situation. Additionally, use of these technologies in simulation can offer a more realistic experience than a classroom discussion [3]. Flight simulations also have a long record of accomplishment in providing transferable training skills that allow users to achieve a level of competency prior to real-life flying [7]. Other studies showed that using virtual interfaces improved SA without increasing the mental workload [8]. This demonstrates that it is possible to have users integrate surrounding elements and tactical information at a higher level of understanding to make better decisions than if presented with the information in a conventional manner. Similarly, Mitaritonna [9] found that there was high satisfaction using an augmented reality framework and higher SA with a more expert group of users. These studies point to the need for further investigation that can uncover the training and display combinations that would yield the greatest benefit to the warfighter.

Still, it is unclear if these technologies can provide military decision-makers with more SA relative to previous technologies. SA has been defined in several ways; yet, Endsley's [10] is the most common and what is used in this study:

> The perception of the elements in the environment within a volume of time and space, the comprehension of their meaning and the projection of their status in the near future.

The three levels of SA from this definition are additive such that the first SA level (i.e., perception) is required for the second level of SA (i.e., comprehension) which is then required for the level three of SA (i.e., projection) which incorporates predictions of possible future outcomes [2]. SA has been assessed in a variety of ways in human factors. For example, situation awareness probes have been used in training to assess how well virtual reality technology provide users with adequate SA [1]. Because SA is so critical to effective military decision-making, it is frequently assessed within the context of military training and operations.

How do AR/VR/MR technologies influence SA in military planning compared to a more traditional technology? This study looks at the differences between a tablet display and a Microsoft HoloLens, as ways to display information to the warfighter. The tablet is a two-dimensional (2D) technology with which most people are familiar and would require minimal training to use. The Microsoft HoloLens used in this study is three-dimensional and expected to be novel to users. There is evidence that using an HMD can increase situation awareness [4] and that AR is more beneficial than traditional training for way-finding tasks [5]. Kaplan and colleagues showed that, in general, training using mixed reality methods were just as effective as traditional methods [3]. They also indicate that the transfer or knowledge is not the only factor to consider since things like cost of simulations can make MR preferable to a real-life simulation. Boyce and = colleagues [11] did an initial evaluation of these technologies in mission planning scenarios. Using the same scenarios, this study adds to the knowledge base regarding performance comparisons of these technologies in providing SA. We also evaluate system usability using the System Usability Scale (SUS) to increase the utility of mixed reality for training and operations [12].

III. METHODS

A. Participants

Data were collected from 22 cadets (*7 female*) at the United States Air Force Academy (USAFA) to this end. USAFA participants included cadets from across class years (freshman through senior), and between the ages of 18 and 23 ($M = 20.44$ *years*). All cadets receive training in military studies and have completed leadership experiences as part of their USAFA experience. Participants were recruited using the SONA system and given extra credit in exchange for their participation.

B. Materials

Two commercially available technologies were used to assess the differences between the tablet (2D) and MR (3D) technologies on SA and usability. The tablet used was a Samsung Galaxy Tablet S4 with standard functionality. A Microsoft HoloLens™ was used for the 3D condition. As shown in Figure 1, both technologies displayed a map of the United States Military Academy (USMA) area in West Point, New York. The 2D display allowed participants to move through the scenario and zoom in and out through tapping and/or sliding his or her fingers on the touchscreen. Conversely, the HoloLens™ shows the same strategic military scenario using the MR version of map on a projected Sandtable. This augmented reality headset displays holographic images that overlay onto the physical environment. The terrain and scenario shown in 3D emphasized the topography of the area relative to the tablet's 2-D view. Although the HoloLens can be controlled with voice and air tap gestures, only the latter was available for use with the custom Augmented Reality Sandtable (ARES) software used in the experiment. Participants assigned to the mixed-reality condition could physically walk around the scenario to see the different parts in different perspectives (see Figure 1C). Additionally, they were able to use hand gestures to navigate the map and adjust the display.

Both the HoloLens and Tablet used ARES software developed specifically for each of the technologies by the Army Research Laboratory. This software provided participants with a map of the terrain used for the scenario described below. Additionally, Qualtrics was used to record participants' perceptions of SA, usability (i.e., SUS), and overall effectiveness.

Figure 1. The two technologies used, the tablet with symbology (left) and the HoloLens (right) with a table displaying symbology (bottom).

C. Procedure

Participants who were randomly assigned to the HoloLens (3D) condition completed the scenario in a laboratory within a 10 ft. by 10 ft. area marked off similar to previous studies using this technology (e.g., [11]). Participants assigned to the tablet condition sat at a desk within the same laboratory.

After signing an informed consent form, participants filled out a short survey to collect demographic data and became familiar with the software and scenario. Participants were given training slides with an overview of military symbols along with the Warning Order pertaining to the mission. Once oriented, participants were introduced to the technology and given a one-minute familiarization session to learn the technology. familiarization and early was consistent across technologies.

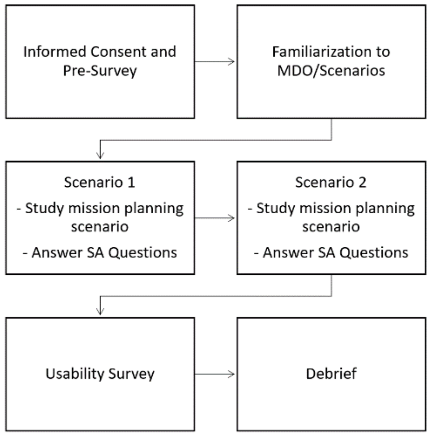

Figure 2. Procedure flow. Participants used either a tablet or HoloLens. Scenario 1 and 2 counterbalanced the presentation of the easy and hard scenario and accompanying Situation Awareness questions.

As shown in Figure 2, each participant completed two scenarios, one easy and one hard, given in a randomized order. A map of terrain surrounding West Point, as well as icons of the relevant military symbols laid out within the terrain, was provided, and the map remained the same for both scenarios while the symbology and described tactical situation changed. The symbols include friendly forces (blue icons) and enemy forces (red icons) as well as different checkpoints and landmarks. The participant was given five minutes to look at the initial scenario before answering the first set of questions regarding perception. The participant was then given three minutes to look at the same scenario before answering the second set of questions, which tested comprehension. Finally, the participant was allowed to look at the same scenario for three more minutes before answering the last set of questions regarding projection. This process was then repeated for the next scenario. After completing all questions and surveys regarding both the hard and easy scenario, the SUS was given to the participant to complete.

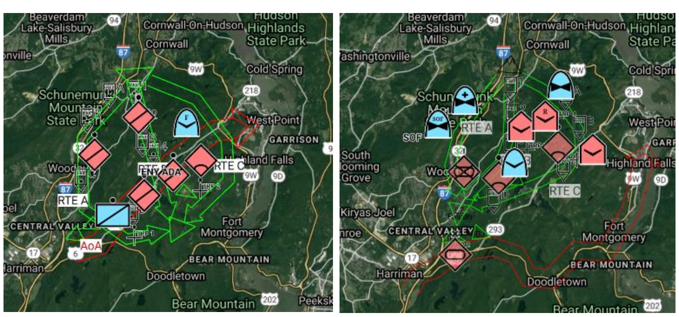

Figure 3. The easy scenario (left) and hard scenario (right) were counterbalanced across Scenario 1 and Scenario 2.

IV. Results

A. *Situation Awareness*

Data from nineteen participants (n = 19) are included in the present analysis, while data from three additional participants were discarded due to data quality issues. Hard and easy SA test questions were scored in terms of percentage correct and then averaged across perception (SA Level 1), comprehension (SA Level 2), and projection (Level 3) categories. Data were analyzed using a within-between subjects MANOVA, to assess the within-subjects impact of two levels of SA test difficulty 2 (difficulty: easy vs. hard, and the between-subjects impact of two different technologies 2 (technology: HoloLens vs Tablet) used by participants during their entire session. The order of receiving the hard and easy SA test was counterbalanced, but imperfectly in this pilot sample: among 9 tablet users, easy SA tests were presented first to 7, and among 10 HoloLens users, easy SA tests were presented first to 6. As such, these data are not analyzed. Further, our data discarded left an uneven sample size for the present analysis. The general linear model is robust to such issues, but out of an abundance of caution we additionally report Wilk's Lambda, in order to quantify how well each function separates cases into groups.

No significant interaction was seen; Wilks' Lambda = .78, $F(3, 15) = .1.47$, $p = .31$, η^2p = .227. There was, however, a significant main effect of difficulty, Wilks' Lambda = 0.59, F(3, 15) = 3.42, $p = .04$, η^2p = .406. There was no significant main effect of technology, Wilks' Lambda = 0.59, $F(3, 15) = 3.42$, $p = .04$, η^2p = .406

B. *System Usability Scale (SUS)*

The SUS was calculated for each technology used between participants. The tablet was scored more usable (*M*=73) than the HoloLens (*M*=60). No statistical comparison of these scores is offered, as the SUS is usually analyzed in terms of the categories of above average and below average, with a cutoff score of 68 [12]. As such, per the SUS, our participants rate the tablet as having above average usability, and the HoloLens as having below average usability.

V. Discussion

Mixed, augmented, and virtual reality technologies are penetrating military contexts to help decision makers gain and maintain SA [13]. Surprisingly, there is little evidence to support these technologies' effectiveness on SA and decision-making. Our results suggest 3D displays provided by AR/VR technologies do not provide greater SA for users relative to 2D technologies. Participants responded to queries with the same level of accuracy regardless of what technology they used. Additionally, the tablet technology provided a similar level of capability at reduced cost and scored higher on perceived usability. Taken together, 2D tablet technologies seem to provide a usable system for users to gain and maintain SA in a complex military scenario.

Still, our findings should be interpreted with caution for several reasons. First, our sample size was low and did not allow for more nuanced analyses of additional categories. Second, familiarity with 2D tablet systems was likely higher. The official statistics collected through the American Community Survey reports that 62.5% of US households had tablets and 84.4% of US households had smartphones in 2018 [14], and the percentages are likely to have grown leading up to 2022. Sites tracking market growth show that VR and AR headsets have grown tremendously over the 2020 to 2021 period [15]. Finally, the complexity of scenarios such as air, land, cyber, space, and other domains may introduce important variables which require 3D visualizations.

This notwithstanding, our results should inform policy and acquisitions such that new technology is well-aligned for the task. Kaplan et al. [3] notes that training in augmented reality was not better or worse than conventional methods but has some advantages in terms of cost and safety. Specifically, the cost of creating complex training scenarios when a VR environment can provide similar outcomes would make the VR technology potentially more cost-effective. Once created, a VR environment could potentially allow for multitudes of trainees whereas a physical training space limits the number of people who can use it. There may also be time costs if the physical training space must be reset or repaired between uses. If the task has similar outcomes regardless of technology, it may be more cost-effective to use the tablet. For reference, the Samsung Tab S4 tablet used in the study was $300 while the Microsoft HoloLens was $3000, an order of magnitude of difference. However, the tablet would only prove more cost-effective if there is no reduction of information by using the 2D rendition.

A. Limitations and Strengths

This is a pilot study within a larger body of programmatic research. As such, a number of minor irregularities with data were expected, and did in fact occur as reported. We do hope that future efforts will improve upon the easy and hard maps, as we believe that in going from one scenario to the next, regardless of “easy” or “hard” designation, the participant must necessarily learn something about the map that could be used for the next scenario. Similarly, the questions used to measure the situation awareness were close enough in nature that the participant was likely to look for similar information when reviewing the map between question periods. For example, by asking about airborne assets in one section, the participant looks for the airborne assets in the next scenario, anticipating that it will be one of the questions asked. Future studies with balanced samples would allow for additional analyses of how to parse the effects of using different scenarios. Asking different questions to measure SA would allow for a greater distinction between the information gathered to minimize the effects of familiarity when the participant reaches the second scenario.

The present study was conducted within a room, over a neutral background. While the results are important evidence as noted, further research in more diverse environments would be necessary to determine applied potential to the warfighter. A proposed strength of the HMD technology is that it would overlay information over the reality of what the user was seeing, making it ideal for use while out in the field.

One strength of the study is the participant population. As cadets at USAFA, the participants have received training on the symbology used as well as in tactical decision-making. In addition, this is the population that will be the recipients of the newest technology in the field and their input will be invaluable to maximizing the technology for their greatest benefit. Another strength is the applied and environmentally valid nature of the SA tests. The questions were adapted from work done by Strater and colleagues [1] for the Army Research Institute. They have been tested on and used by the military to measure situation awareness. The results have been shown to reflect SA in a military setting.

B. Future Directions

Future studies will address the training component to determine if there is a training modality that works best. The tablet technology is so ubiquitous that training on its use is not necessary. As a new technology, the ability to use the HoloLens seamlessly will take time. Previous research supports the possibility that additional training may have an impact in future studies since the HoloLens is a new technology for most users [11].

The display should maximize the capability of the technology. There needs to be a closer look at the impact of using flat symbols, originally created for use on two-dimensional maps, in a three-dimensional environment. There may be more intuitive displays that would increase SA while allowing the warfighter to remain cognizant of the surroundings.

Research to improve the capability of the HoloLens in training and, eventually, in the battlefield can benefit MDO missions by providing simultaneous communication to various units. With an increased understanding of current status and tactical options, decisions can be made quickly to maintain the advantage in battle but, more importantly, to increase the safety of the warfighters.

ACKNOWLEDGMENTS

The views expressed in this paper are those of the authors and do not reflect those of the U.S. Air Force, Department of Defense, or U.S. Government.

REFERENCES

[1] Strater, Laura D., et al. *Measures of platoon leader situation awareness in virtual decision-making exercises*. TRW INC FAIRFAX VA SYSTEMS AND INFORMATION TECHNOLOGY GROUP, 2001.

[2] Endsley, Mica R. "A taxonomy of situation awareness errors." *Human factors in aviation operations* 3.2 (1995): 287-292.

[3] Kaplan, Alexandra D., et al. "The effects of virtual reality, augmented reality, and mixed reality as training enhancement methods: A meta-analysis." *Human factors* 63.4 (2021): 706-726.

[4] Wickens, Chris, et al. "Developing and evaluating an augmented reality interface to assist the joint tactical air controller by applying human performance models." *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. Vol. 62. No. 1. Sage CA: Los Angeles, CA: SAGE Publications, 2018.

[5] Goldiez, Brian F., Ali M. Ahmad, and Peter A. Hancock. "Effects of augmented reality display settings on human wayfinding performance." *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)* 37.5 (2007): 839-845.

[6] Salas, Eduardo, Clint A. Bowers, and Lori Rhodenizer. "It is not how much you have but how you use it: Toward a rational use of simulation to support aviation training." *The international journal of aviation psychology* 8.3 (1998): 197-208.

[7] Rantanen, Esa M., and Donald A. Talleur. "Incremental transfer and cost effectiveness of groundbased flight trainers in university aviation programs." *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. Vol. 49. No. 7. Sage CA: Los Angeles, CA: SAGE Publications, 2005.

[8] Roldán, Juan Jesús, et al. "Multi-robot interfaces and operator situational awareness: Study of the impact of immersion and prediction." *Sensors* 17.8 (2017): 1720.

[9] Mitaritonna, Alejandro, María José Abásolo, and Francisco Montero. "An augmented reality-based software architecture to support military situational awareness." *2020 International Conference on Electrical, Communication, and Computer Engineering (ICECCE)*. IEEE, 2020.

[10] Endsley, M. R. *The functioning and evaluation of pilot situation awareness*. Northrop Technical Report: NOR DOC 88–30, 1988.

[11] Boyce, M. W., Rowan, C. P., Shorter, P. L., Moss, J. D., Amburn, C. R., Garneau, C. J., & Sottilare, R. A. (2018). The impact of surface projection on military tactics comprehension. *Military Psychology*, *31*(1), 45–59.

[12] Brooke, John. "Sus: a 'quick and dirty' usability." *Usability evaluation in industry* 189.3 (1996).

[13] Saunders , Karen, and John M Murray. "Joint Asa Alt and AFC Statement on the Integrated Visual Augmentation System." *www.army.mil*, 18 Oct. 2021.

[14] Martin, Michael. "Computer and internet use in the United States: 2018." *States* (2021).

[15] "AR/VR Headset Shipments Grew Dramatically in 2021, Thanks Largely to Meta's Strong Quest 2 Volumes, with Growth Forecast to Continue, According to IDC." *IDC*, Mar. 2022.