Enhancing Maintenance Workers: A Controlled Field Experiment with Augmented Reality

Clay Greunke, James Fan, Perry McDowell
Quinn Kennedy, Imre Balogh, Chris Angelopoulos,
MOVES Institute/Naval Postgraduate School
Monterey, California
<larry.greunke/mcdowell/james.fan/ ilbalogh/mqkenned/cjangelo>@nps.edu

Matthew Stone
Naval Air Warfare Center/Aircraft Division
Patuxent River, Maryland
matthew.d.stone@navy.mil

ABSTRACT

Background: Maintenance is one of the most important aspects of any modern military's ability to conduct combat operations as reflected in the fact that the DoD allocated over $292 billion to operations and maintenance in 2020. The amount of time and resources required to perform maintenance has increased as weapons and equipment have become more advanced, therefore, leadership is motivated to find and implement streamlining methods. One technology which has significant capability to improve military maintenance speed and accuracy is augmented reality (AR). AR can function as a job aid to guide personnel as they conduct maintenance to improve performance.

Methods: We conducted a counterbalanced experiment to test the effect of using AR to assist personnel in a common aviation maintenance task under two different conditions: AR guidance or traditional desktop instructions. Completion time and accuracy were the main outcome measures. Subjects were personnel at the Center for Naval Aviation Technical Training who would conduct such maintenance in their regular jobs in squadrons and were grouped based upon their previous maintenance experience into two groups, novice ($n = 17$) and expert ($n = 17$). All subjects completed both conditions, with half completing the task first with AR guidance then with traditional desktop instructions. The other half completed the task in the opposite order. We predicted that novices would benefit more from AR guidance than experts.

Results: Results were consistent with our prediction. Novices showed statistically significant faster completion times under the AR condition than the traditional desktop instructions. Experts had a nonsignificant improvement in completion time under AR than the traditional condition. Because subjects rarely made errors, we were unable to assess accuracy. Additionally, novices reported that AR was easier to use than the traditional method; experts reported the two methods as equivalent in ease of use.

ABOUT THE AUTHORS

Clay Greunke is a Research Associate Faculty at the Naval Postgraduate. He completed his bachelor's in agricultural biotechnology at the University of Kentucky before being commissioned in the Navy. With a primary designator of a naval aviator, he served as an electronic attack pilot with VAQ-139 for his sea tour. For shore tour orders, he was accepted into the Naval Postgraduate School (NPS) and completed the requirements for a Master of Science in computer science in 2015. His thesis work in virtual environments and training received several national accolades. His last active-duty assignment he worked for the chief engineer at the Space and Naval Warfare Systems Command (SPAWAR), with the specific assignment of innovating processes within the Navy Enterprise. Clay joined the faculty at the NPS focusing his research efforts on the application of augmented reality within the Navy.

James Fan is an assistant professor of operations management at the Naval Postgraduate School. His research interests include behavioral economics and behavioral operations, laboratory and field experiments with decision-makers and organizations, and decision analysis for procurement and defense acquisitions.

Perry McDowell attended nuclear power training after commissioning as a naval officer. He served in USS VIRGINIA (CGN-38), USS ELROD (FFG-55), and USS ENTERPRISE (CVN-65). In 1995, Perry earned a Master of Science degree in computer science at the Naval Postgraduate School, where he was awarded the Grace Murray Hopper Award as outstanding computer science student. Upon leaving the Navy in 2000, he returned to NPS and joined the faculty. Although he has served as a principal investigator for a wide variety of projects in the MOVES Institute, from 2003 – 2012 he worked primarily as the executive director for the Delta3D open-source game engine. From 2012 to the present, he has taught courses in simulation for training and conducted research in the areas of training effectiveness and the creation of systems to improve warfighter performance.

Quinn Kennedy is a research associate professor in the Operations Research Department at the Naval Postgraduate School. Her work in behavioral science research focuses on optimizing human performance and decision-making and testing the effectiveness of new technologies on human performance and training. Dr. Kennedy received her PhD in Psychology and postdoctoral training from Stanford University.

Dr. Imre Balogh is the director of the Naval Postgraduate School's Modeling, Virtual Environments and Simulation (MOVES) Institute and has been working in the area of combat M&S for the past 23 years. He is currently doing research on dynamic behavior modeling in high resolution, agent-based combat simulations. Prior to joining the MOVES Institute, Dr. Balogh was the chief architect of the Combined Arms Analysis Tool for the 21st Century (COMBATXXI).

CDR Christopher Angelopoulos was commissioned as a naval officer at University of Florida with a bachelor's degree in computer engineering in 1994. Various operational F/A-18 tours supported global international and joint operations over the next 14 years. He served the NIOC Norfolk Cyber Readiness director in charge of red and blue team operations which included OPERATION ROLLING TIDE. As the first air operations officer for PCU GERALD R. FORD (CVN 78), he created a Mixed Reality Innovation Center generating interest from the Office of Naval Research, SPAWAR, and NAWC/TSD. He earned a master's degree in Modeling, Virtual Environment, and Simulation at the Naval Postgraduate School and received NPS's highest academic award. He is now faculty at NPS.

Matthew Stone has B.S. degrees in aerospace and mechanical engineering from Missouri S&T, and a master's degree in Aerospace Engineering from the University of Maryland. He works for Naval Air Warfare Center Aircraft Division (NAWCAD) where he is the founder and team lead of the Human Systems Engineering Augmented Reality (HSEAR) group.

Enhancing Maintenance Workers: A Controlled Field Experiment with Augmented Reality

Clay Greunke, James Fan, Perry McDowell,
Quinn Kennedy, Imre Balogh, Chris Angelopoulos,
MOVES Institute/Naval Postgraduate School
Monterey, California
<larry.greunke/james.fan/ mcdowell
mqkenned/ilbalogh/cjangelo>@nps.edu

Matthew Stone
Naval Air Warfare Center/Aircraft Division
Patuxent River, Maryland
matthew.d.stone@navy.mil

INTRODUCTION

The United States Department of Defense allocated over $292 billion to operations and maintenance for the 2020 fiscal year (Office of the Under Secretary of Defense, 2019). Much of this budget supports the education, training, and on-site support of personnel responsible for conducting maintenance. The task is made more difficult because there is a large turnover of personnel at the junior level; the Navy's goal is to reenlist 57% of Sailors under six years of service (Maucione, 2020), while the Marines aim to keep 22% of Marines beyond the initial enlistment (United States Marine Corps, 2017). In the area of aviation maintenance, it is these junior maintainers who perform most of the required tasks. Obviously, training these junior personnel is critical for them to properly perform maintenance, but the turnover makes it difficult to preserve institutional knowledge crucial to a high performing workforce.

Therefore, either improving the training of junior personnel or reducing the difficulty of maintenance tasks is highly beneficial to the military by greatly reducing the cost and time of both training personnel and conducting the maintenance. One technology that has shown significant promise providing real-time visual support to complex tasks, thus meeting both goals, is augmented reality (AR). However, prior to implementing a new technology such as AR, it is necessary to test it to both verify its effectiveness and determine any unexpected complications that might arise.

To do this, we conducted a field-test of the efficacy of AR and its interaction with worker expertise. This paper describes a controlled experiment we conducted by inserting AR to provide the guidance of common maintenance tasks with naval maintenance workers at the Center for Naval Aviation Technical Training Unit (CNATTU) in MCAS Miramir. Based upon input from our sponsor, we chose as our task configuring a member of the Consolidated Automated Support System (CASS) family of testers (FoT). This task consists of connecting several wires between an electronic component removed from an aircraft to a test bench to perform troubleshooting. We employed a counterbalanced design in which subjects who varied in their degree of maintenance expertise (novice, expert) completed a maintenance task twice: once using AR and once using the traditional instruction method.

This study is important in the field and expands the field's knowledge base in two primary areas. First, many other studies examining the efficacy of AR utilize undergraduates, oftentimes engineering students, as research subjects (MacAllister et al., 2017) (Hoover et al., 2019). This choice of subject population can lead to findings that are not representative of how the new technology will perform when used by the actual user population, especially if the subjects are engineers (Tobias, 2016). We conducted our study using actual aviation maintenance workers, so the results are more representative of the actual user population. Second, because our subjects were actual maintainers and the experimental task simulated an actual maintenance duty some of them had performed previously, we were able to divide our pool into novices and experts. New technology must be used by both groups, and it is unlikely that both will interact with it similarly. By testing both groups, we can better predict how effective AR will be in the field.

Background

Consolidated Automated Support System

In 1990, a U.S. Navy aircraft carrier required twenty-four automated test systems and 300 aviation electronics technicians to support the electronics aboard the planes in the carrier's air wing (Meredith, 1990). Needing so many different systems created issues with shipboard space, maintaining multiple pieces of test equipment, and cost, so the Navy created the CASS FoTs in 1990 to reduce the number of test systems required to support squadrons on an aircraft carrier. The goal of the system was to significantly reduce the number of. Since its introduction, the CASS FoT has been modernized and expanded so that currently, "The CASS FoT provides fleet intermediate-level maintenance activities ashore and afloat with the capability to test and troubleshoot over 500 avionics Weapons Replaceable Assemblies and Shop Replaceable Assemblies across multiple Navy and Marine Corps type/model/series aircraft, avoiding repair costs of more than $1.1 billion per year… CASS stations [are] installed in Aviation Intermediate Maintenance Departments (AIMDs) ashore and afloat worldwide, as well as Reconfigurable Transportable CASS (RTCASS) stations in support of forward-deployed Marine Aviation Logistics Squadrons" (Naval Air Warfare Center Aircraft Division, 2021). Although the Navy expects the CASS systems to be replaced by the electronic CASS (eCASS) systems by 2025, CASS is still being used by large portions of the Naval Air enterprise (Office of Under Secretary of Defense for Sustainment, 2014).

The RTCASS was designed to operate the mainframe CASS legacy test program sets in a mobile version. Boeing originally designed the system to troubleshoot electronic systems on the F/A-18 Hornet, AV-8B Harrier and EA-6B Prowler aircraft platforms and it has been upgraded to support SOCOM V-22 variants (Office of Under Secretary of Defense for Sustainment, 2014). The RTCASS FoT's is shown in Figure 1.

Figure 1. RTCASS Family of Testers

The RTCASS system is operated by Marines with the 6469 military operational specialty: RTCASS Technician. These Marines "at the IMA, inspect, test, maintain, repair, and analyze airborne weapon replaceable assemblies, shop replaceable assemblies, automatic test equipment, and ancillary equipment failures, beyond normal fault isolation procedures" (United States Marine Corps, 2021).

Maintenance personnel use any variant of the CASS FoT in essentially the same way. The maintainers take a suspect piece of electronic equipment from an aircraft and bring it to the CASS bench. They then follow a series of instructions which are delivered via computer screen to connect the equipment to the CASS testbench via wires and cables. The required connections vary based upon aviation electronics part being tested as well as the type of test being performed. If the technician performs the pre-test wiring incorrectly, the test results will not be valid. As the test proceeds, the technician must reconfigure the connections based upon the findings of the system.

Augmented Reality

AR is a visual technology which overlays digital information or virtual objects over the user's view of the real world, ideally in a seamless manner requiring no user effort (Yuen, Yaoyuneyong, & Johnson, 2011).

AR has long been touted as having applications in a wide range of areas (Azuma, 1997). Using AR as a job aid to assist in manufacturing has long been envisioned. Boeing created an AR system in the early 1990's to guide the manufacturing process that is remarkably similar to the system that we used in our work (Figure 2). The designers wrote, "Our research and development project at Boeing is aimed at advancing the components of this technology to the point at which the use of AR in manufacturing applications is practical" (Caudell & Mizell, 1992, p. 662).

Unfortunately, reaching practicality took longer than many expected; until the last decade, AR's practical capabilities were more hype than reality. AR systems were very expensive and therefore primarily used in academic and industry labs or as a few prototypes to prove the concept. While this work was key to the development of AR and showed its

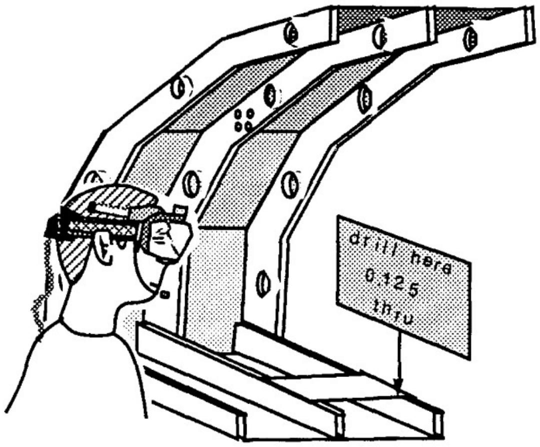

Figure 2. Early Boeing AR System for Manufacturing (Caudell & Mizell, 1992)

promise, it underscored that AR was not yet ready for widespread adaption. However, within the last decade advances in computing and peripheral devices led many to reexamine AR's potential for use in manufacturing and other areas, such as training, operations, and logistics (Donovan & Cimino, 2010). In the last five years, low-cost commercial AR products, such as Microsoft's HoloLens and Magic Leap's One, have become available and greatly expanded the opportunities for AR usage and many have begun to experiment with its capabilities.

The Virtual Reality Applications Center at Iowa State University worked with Boeing in conducting a series of experiments over several years where subjects assembled an aircraft mocked-up aircraft wing in two trials. The earlier experiments were intended to examine several factors, such as subjects' stress levels or the effects of various occlusion methods upon performance, they provided an excellent source of data on undergraduate engineering subjects completing the task using various modalities of instruction. (Richardson et al., 2014) (MacAllister et al., 2017). They later compared the earlier research to compare a HoloLens system similar to the one we used in our research to the earlier modalities: model-based instructions (MBI) delivered via a desktop system, MBI delivered via tablet, and AR instructions delivered via a tablet. They found that the HoloLens system outperformed the other three modalities in time and error rate on both trials, and performance improved from the first to the second trial. (Hoover et al., 2019). Boeing has implemented results from this work and has reported 90% improvement in first-time quality and a 30% reduction in time to complete tasks (Boeing, 2018).

One study compared performance of expert, intermediate practitioner, and novice neurosurgeons in identifying and classifying tumor samples in four modalities: 2-D images, orthogonal planes, 3-D images, and AR. Although it suffered from an extremely low sample size (11 total subjects across the 3 groups), it found that novices and experts performance was the same for one of the tasks using AR, and AR outperformed the other modalities for that task. In the other task, they found that experts outperformed the other groups in all modalities and that AR and 3D outperformed orthogonal planes (Abhari et al., 2013). This is important because it shows that even in similar tasks, AR might affect novice's performance in one but not the other.

EXPERIMENTAL DESIGN

Overview

Our experimental goal was to determine whether presenting instructions via AR improved performance compared to the current method of delivering instructions via text and schematic on a computer screen while completing the task of connecting wiring to electronic equipment on an RTCASS workbench. We used aggregate time to complete the multi-step maintenance action to measure performance. We chose a maintenance task on the RTCASS workbench that was known to be problematic in the Marine maintenance community and reproduced that procedure, to best of our ability, with the analog bench that we created. The experiment lasted approximately one and a half hours per subject. Performance was measured by speed (aggregate time to complete the multi-step maintenance action) and accuracy (number of errors).

Subjects

Our subjects included students and instructors at CNATT. We grouped subjects into two categories, novice (n=17) and expert (n=17), based upon their level of maintenance experience. Demographic values for the groups are shown in Table 1.

Table 1. Subject Demographics

	Novice	Expert
n	17	17
Mean Age	23.0	31.5
Male/Female	13/4	14/3
Mean Years of Service	2.2	8.5
Rank: E1 – E-3	15	0
Rank: E-4 – E-6	2	15
Rank: > E-6	1	2
# Wearing Corrective Lenses	5	8
Mean Yrs. of Military Maintenance	0.4	6.8
Mean Yrs. Performing Assembly Tasks	1.5	6.3
# w/ Previous VR Experience	8	9
Mean # of VR Experiences	5.6	6.2
# Prone to Motion Sickness	0	1

The RTCASS Task

To alleviate the need for subjects to use an actual RTCASS during the experiment, we built a functional analog and transported it to CNATTU Miramar. While the experimental version was less complex than the actual RTCASS, it was spatially accurate to where each of the connections were (see Figure 3). Of note, we chose to only have one type of connector for the experiment, where the real machine has multiple types of connectors. Subject matter experts (SMEs) reported that the analog created was a good representation of the actual system in size and complexity.

For the experiment, we reproduced the 46-step procedure that was detailed as being very problematic for the maintainers. This procedure consisted of a series of common actions while connecting electronic equipment to the RTCASS: connecting wires, disconnecting wires, repositioning wires, and moving plugs. SMEs judged the procedure to be of an equivalent difficulty to the real process.

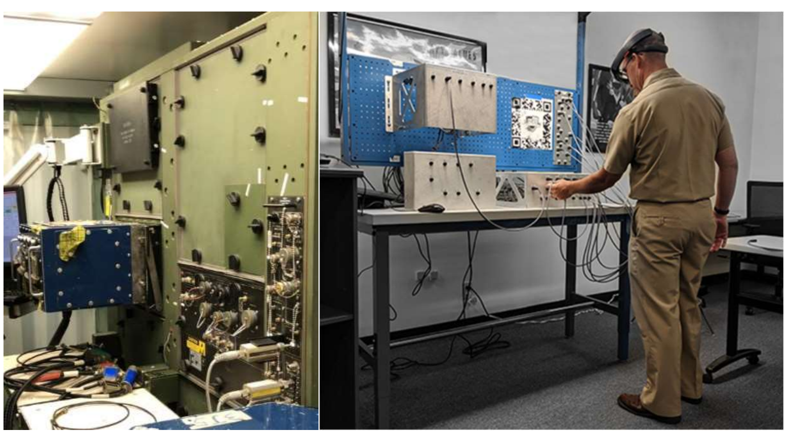

Figure 3. Deployed RTCASS (left) and Replica Created for Experiment

We found the current instruction manual for the actual RTCASS to be difficult to use and it often contained multiple actions within the same steps. We wanted to ensure that any experimental differences between the AR and control conditions were due to the improved capabilities of AR rather than problems arising from a non-intuitive user manual. Additionally, we felt it would be easier to measure and compare the subjects' performance if there were only one action per step. Therefore, we recreated the instruction manual desktop application for RTCASS maintenance and improved the user interface for ease of use. Figure 4 shows the actual instructions used in the desktop guidance and our modified versions.

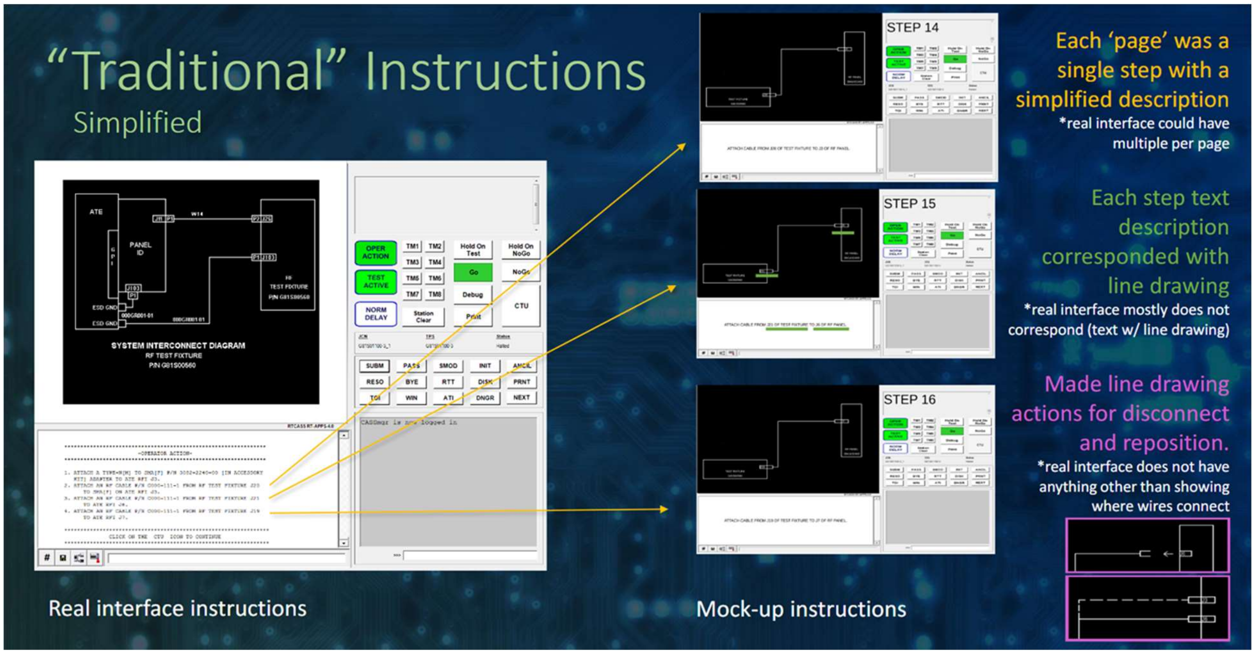

Figure 4. Actual Instructions (on left) and Modified Versions Used in Experiment (on right)

AR Equipment

To present the AR instructions to the subjects, we used a Microsoft HoloLens providing optical see-through AR. The HoloLens' video cameras provided registration of the RTCASS position and was used to overlay virtual objects atop real-world objects. We did this by creating a virtual environment of the replica test bench using the Unity game engine. The instructions were programmed into Unity using its scripting language. In both treatments, subjects clicked on the mouse to advance to the next step. We used a fiducial to orient the view. When the fiducial was not visible in the camera view, the system relied upon the HoloLens motion system to orient the user's position and direction and would recalibrate when the fiducial was again visible.

The AR system highlighted the locations of wires and plugs to be manipulated and demonstrated what actions the subject needed to perform to successfully complete the step. Figure 5 shows the AR guidance for connecting a wire from one location to another. Video of the task being performed from the user's viewpoint in AR are available at https://www.youtube.com/watch?v=3dAOjO_IUzk&feature=youtu.be and from an onlooker's viewpoint at https://www.youtube.com/watch?v=PkH8xBB4BZY.

Experimental Procedure

We implemented a 2x2 experiment design to study the guidance method (AR vs. traditional desktop instructions) and the maintenance expertise of workers (novice and expert). We recruited maintenance workers at CNATT, separated them into those with naval maintenance experience (experts) and those with no naval maintenance experience (novices). Each subject completed both the AR and desktop treatments to produce higher statistical power, and we alternated whether a subject received the AR treatment or desktop treatment first to control for learning effects. We collected demographic data and background history (such as prior non-military maintenance experience and experience with AR).

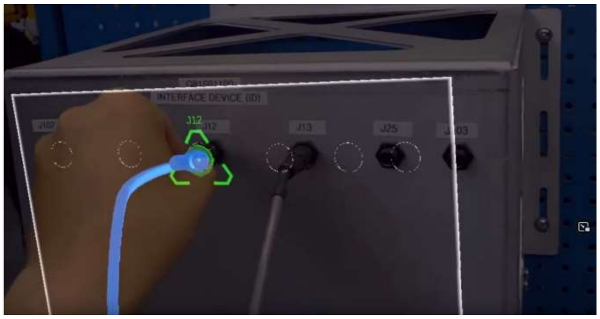

Figure 5. User's Viewpoint in AR

In the AR treatment, subjects used the HoloLens system descried above. Each of the 46 procedural steps of the task is timed. The first four steps served as practice to familiarize subjects with the task and were not included in the analyses. We characterize each step as one of four types of steps: connecting wires, disconnecting wires, repositioning wires, or moving plugs.

In the desktop treatment, subjects used the written instructions we had adapted from the RTCASS maintenance manual on a computer screen (see Figure 4) while they completed the task list to service the RTCASS until completion of all 46 steps. The first four steps again served as practice and were not analyzed.

RESULTS

We were unable to analyze accuracy because the number of errors both groups committed was so low (<1% of all steps) with no commonality in either the step, modality, or grouping. Speed was measured by both aggregate time to complete the task and time per step.

Aggregate Task Completion Times

Our results confirmed our predictions that using AR would affect the total task completion time of novices more than that of experts. Figure 6 summarizes aggregate task completion times. Because the data did not meet the normality requirements, we used the Wilcoxon signed rank test to compare results. Overall, there was a trend for subjects to be faster using AR by average of 89 seconds (std error = 45.34), but it did not reach the level of statistical significance ($p = 0.07$). This result was driven by the novices who reduced task completion time by an average of 136 seconds (std error = 60.10) ($p = 0.04$). Expert performance using AR was not better than using the traditional method at a statistically significant level ($p = 0.68$).

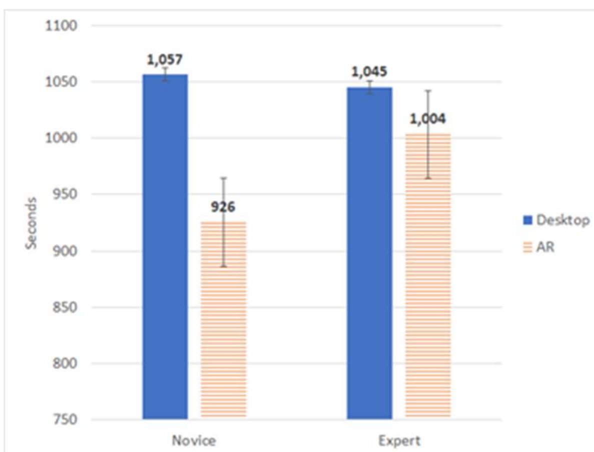

Figure 6. Aggregate Task Completion Times

Ordering Effects and AR

Because we used a counterbalanced design, we next analyzed order effects of giving the desktop treatment vs. AR. Completion time based upon initial modality used by subjects is shown in Figure 7. An order effect was found such that both novice and experts improved from their first attempt to their second. Table 2 shows the average completion times by condition and expertise level.

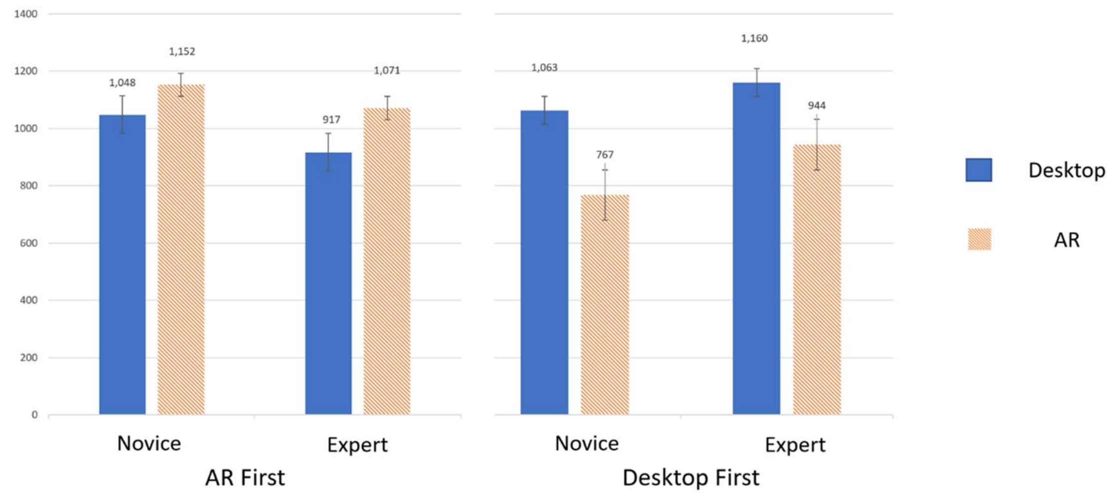

Figure 7. Average Aggregate Time Based upon Initial Modality

Table 2. Average Time to Complete Tasks (sec.)

	Desktop First		AR First	
	Novice	**Expert**	**Novice**	**Expert**
Both	914.98 (128.15)	1051.64 (236.84)	1099.79 (236.17)	993.85 (151.59)
Desktop	1062.76 (163.39)	1159.65 (305.75)	1047.77 (193.47)	916.80 (217.39)
AR	767.21 (122.59)	943.62 (185.87)	1151.82 (293.59)	1070.90 (121.99)

Post-Experiment Questionnaire Results

The subjects completed a post experiment questionnaire in which they rated the difficulty of completing the task with each modality on a linear analog scale from 1 (not difficult) to 100 (extremely difficult). The novices reported that AR was almost 75% less difficult than the desktop guidance, but the experts reported higher difficulty with AR by over 25%. To determine whether these results were statistically significant, we performed Wilcoxon signed rank tests on the two groups' results. Only the novices' results were statistically significant. Table 3 shows the numeric results.

Table 3. Self-reported Difficulty (Lower is Better)

	Desktop Guidance	AR Guidance	Standard Error	Wilcoxon Signed Rank Result
Novice	23.75	6.17	5.06	Prob < |S| = 0.0002
Expert	10.89	13.74	4.04	Prob > |S| = 0.34

The survey also asked subjects whether they had any difficulties seeing either the virtual or actual objects. Only three of the 34 subjects reported difficulty seeing the virtual objects, and only 4 of the 34 had difficulty seeing the actual objects. Of these, one subject who had trouble with seeing both virtual and actual objects was wearing corrective lenses, and one subject who had difficulty seeing actual objects was wearing corrective lenses. The remainder of those with difficulties did not require corrective lenses.

Step-time Analysis

We defined each of the steps into one of four types of actions: connecting wires, disconnecting wires, repositioning wires, and moving plugs. We conducted a regression analysis to determine how each instruction modality affected the time to complete each type of steps, as well as interaction effects between instruction modality and task type. We summarize the key results below in Table 4.

Table 4. Variable Effects Upon Step Performance Time

	Effect	Std. Error
AR Treatment (d)	-1.515*	0.791
Expert (d)	-0.244	1.881
Connection	0	0.0
Disconnect	-15.01****	0.925
Move Plug	-15.45****	0.748
Reposition	-13.44****	0.704
Connection * AR	-3.011***	1.113
Move Plug * AR	-1.918***	0.712
First time	4.625****	0.791
Age	0.137	0.183
Gender (F=0; M = 1)	2.903**	1.320
Observations	2772	
* $p < 0.10$ ** $p < 0.05$, *** $p < 0.01$ **** $p < 0.001$		

We find that workers complete tasks 1.5 seconds faster on average under AR instruction. This is consistent with our aggregate results. Broken down by task type, disconnecting wires, moving plugs, and repositioning wires are significantly faster to complete compared to connecting wires, which we use as the baseline for comparison. Furthermore, the interaction effects of connecting wires and moving plugs with AR instruction further decreases task times by 3 seconds and 1.9 seconds, respectively. This provides preliminary evidence that AR instructions may be even more beneficial for certain maintenance tasks, particularly tasks with high cognitive load.

DISCUSSION

Results

As expected, novices completed the RTCASS significantly faster with the use of AR, whereas experts did not significantly differ in task completion time.

We were surprised that there was no statistically significant difference between the two groups when receiving desktop guidance first, as our intuition was that the experts would perform significantly better. We posit that this may have been because we simplified the written instructions of the actual task in the desktop condition. Thus, the task was easier than what the experts encounter when actually performing the task, and these slight differences may reduced their performance.

We did not have any expectations as to which type of tasks would be most affected by using AR guidance. A specific advantage of AR was found for connecting wires and moving plugs, and on average providing AR instructions decreases task times by 1.5 seconds. We believe that further research is needed to both validate these results as well as gain insight into the underlying causes if they are validated.

We believe that cognitive load theory (CLT) might explain why novices did statistically better using AR while there was no statistical difference for experts. We posit that the AR interface produced additional extraneous load in the experts who did not receive as much benefit from its guidance as the novices. Kalyuga (2005) states that "design principles that help low-knowledge learners may not help or even hinder high-knowledge learners (pg. 325)." "The expertise reversal effect [of CLT] generally occurs when high-information instruction is beneficial to novices when compared with the performance of novices who receive a low-information format, but is disadvantageous for more expert learners when compared with the performance of experts who receive a low-information format" (pg. 327-328). This effect would be similar to many expert computer users preferring to use a command line interface rather than a graphical user interface. While most research into the expertise reversal effect has been done in the area of learning, we believe that it is likely to apply in other areas such as operations and maintenance.

FUTURE WORK

Our original idea was to perform this experiment with more graded levels of expertise (e.g., novice, competent, expert) in order to get a better idea of how AR affects maintenance at various levels across the spectrum of maintainer levels. Unfortunately, constraints due to COVID and time limitations forced us to revise our experimental design to only two levels. However, we believe that future work should be performed to look at AR's effect upon performance with a greater level of granularity of skill levels.

One potential downside to implementing an AR solution will be a reduction in learning from completing tasks. We would like to perform additional experiments to compare the long-term knowledge retention of novices who perform tasks using assistance such as AR and those who perform them using more traditional instructions.

Additionally, as in most studies with a limited number of subjects, confirmation with a larger sample size is always appropriate to validate the results.

IMPLICATIONS

Several results from study are important for planning how to implement advanced modalities to improve performance and perhaps reduce the amount of training they require before they can perform advanced maintenance:

- Because AR is significantly more effective for novices and does not negatively affect expert performance, changing to AR solutions for delivering instructions for assembly tasks will produce improvements in performance in the field.
- Novices' performance is similar to experts without significant training or experience. Therefore, novices may not require as much training before reporting to commands and still perform at acceptable levels.
- Although there was a large overall effect, not all types of actions the subjects performed showed an improvement in AR. When determining what tasks should be implemented in AR in the future, it is important to choose tasks that have been demonstrated to produce results.
- Experts do not necessarily see AR as being beneficial to performing the tasks. Whenever a new technology is implemented, it is critical to get buy-in from those who will be using it. Therefore, prior to implementing any AR solutions, leaders must ensure that it is implemented in such a way that all users, especially those senior ones in leadership positions, will accept it.

ACKNOWLEDGEMENTS

The authors wish to acknowledge the support provided by LtCol George Roberts of PMA-257 who provided funding for this research. We also would like to thank the personnel of CNATT who facilitated and participated in this research.

REFERENCES

Abhari, K., Baxter, J. S., Chen, E. S., Khan, A. R., Wedlake, C., Peters, T., Eagleson, R., & di Ribaupierre, S. (2013). The role of augmented reality in training the planning of brain tumor resection. In H. Liao, C. L. Linte, K. Masamune, T. M. Peters, & G. Zheng (Ed.), *Proceedings of Augmented Reality Environments for Medical Imaging and Computer-Assisted Interventions* (pp. 241-248). Nagoya, Japan: Springer.

Azuma, R. T. (1997). Survey of Augmented Reality. *Presence: Teleoperators and Virtual Environments, 6*(4), 355-385.

Boeing. (2018, January 19). *Boeing Tests Augmented Reality in the Factory*. Retrieved June 22, 2021, from Boeing.com: https://www.boeing.com/features/2018/01/augmented-reality-01-18.page

Caudell, T. P., & Mizell, D. W. (1992). Augmented reality: an application of heads-up display technology to manual. *Proceedings of the Twenty-Fifth Hawaii International Conference on System Sciences* (pp. 659-669). IEEE. doi:https://ieeexplore.ieee.org/document/183317

Donovan, D., & Cimino, J. (2010). Augmented reality as an emerging military training technology. *Proceedings of the Interservice/Industry Training Simulation and Education Conference (I/ITSEC).* Orlando, FL: Nationtal Training and Simulation Association.

Hoover, M., Miller, J., Gilbert, S., Winer, E., & Davies, P. (2019). Evaluating augmented reality assembly instructions delivered via microsoft hololens. *Proceedings of the Interservice/Industry Training, Simulation and Education Conference (I/ITSEC).* Orlando, FL: National Training and Simulation Association.

Kalyuga, S. (2005). Prior knowledge principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning* (pp. 325-337). Cambridge, UK: Cambridge University Press.

MacAllister, A., Hoover, M., Gibert, S., Oliver, J., Radkowski, R., Garrett, T., Holub, J., Winer, E., Terry, S., & Davies, P. (2017). Comparing visual assembly aids for augmented reality work instructions. *Proceedings of the Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC).* Orlando, FL: National Training and Simulation Association.

Maucione, S. (2020, December 30). After a year of high retention, the Navy is thinking about force balance. *Federal News Network*. Retrieved June 22, 2021, from https://federalnewsnetwork.com/navy/2020/12/after-a-year-of-high-retention-the-navy-is-thinking-about-force-balance/

Meredith, M. (1990). *The consolidated automated support system (CASS): a comparative evaluation* [Master's Thesis : Naval Postgraduate School]. Monterey, CA: Naval Postgraduate School. Retrieved from https://calhoun.nps.edu/bitstream/handle/10945/34811/90Jun_Meredith.pdf

Naval Air Warfare Center Aircraft Division. (2021). C*onsolidated automated support system family of testers (CASS FoT)*. Retrieved May 29, 2021, from NAWCAD Lakehurst: https://www.navair.navy.mil/lakehurst/product/consolidated-automated-support-system-cass-family-of-testers

Office of the Under Secretary of Defense. (2019). O*peration and maintenance overview fiscal year 2020.* Washington D.C.: U.S. Department of Defense. Retrieved January 25, 2021, from https://comptroller.defense.gov/Portals/45/Documents/defbudget/fy2020/fy2020_OM_Overview.pdf

Office of Under Secretary of Defense for Sustainment. (2014). *Consolidated automated support system (CASS)*. Retrieved May 29, 2021, from Office of Under Secretary of Defense for Sustainment: https://www.acq.osd.mil/log/MPP/ats_CASS.html

Richardson, T., Gilbert, S., Holub, J., Thompson, F., MacAllister, A., & Radkowski, R. (2014). Fusing self-reported and sensor data from mixed-reality training. I/ITSEC. *Proceeedings of the Interservice/Industry Training,* *Simulation, and Education Conference (I/ITSEC)*. Orlando, FL: National Training and Simulation Association.

Tobias, S. (2016, June 19). 'Silicon Valley' season 3, episode 9: What exactly is pied piper? *New York Times*. Retrieved June 15, 2021, from https://www.nytimes.com/2016/06/19/arts/television/silicon-valley-season-3-episode-9-recap.html

United States Marine Corps. (2017, June 30). *FY18 enlisted retention goals*. Retrieved June 10, 2021, from official website of the Marine Corps: https://www.marines.mil/News/Messages/Messages-Display/Article/1235588/fy18-enlisted-retention-goals/

United States Marine Corps. (2021, April 30). *MOS 6469 reconfigurable transportable consolidated automated support system (RTCASS) technician, IMA*. Retrieved May 30, 2021, from Marine Corps Credentialing Opportunities On-Line: https://www.cool.navy.mil/usmc/enlisted/6469.htm

Yuen, S. C., Yaoyuneyong, G., & Johnson, E. (2011). Augmented reality: An overview and five directions for AR in education. *Journal of Educational Technology Development and Exchange, 4*(1), 11.