Evaluating the Use of Augmented Reality for Aircraft Maintenance Training

Charis Horner
Design Interactive Inc.
Orlando, FL
Charis.horner@designinteractive.net

Christina K. Padron
Dynepic Inc.
Charleston, SC
Christina@dynepic.com

Troy Westbrook
USAF AETC 367 TRSS
Hill AFB, UT
Troy.westbrook@us.af.mil

ABSTRACT

Maintaining aircraft within the Air Force remains key to being mission ready. However, as the Fleet continues to age, the average mission-capable rate across the Fleet has steadily declined. The older Fleet requires additional inspection and maintenance and the Air Force has named additional flight line maintenance as a potential solution. There remains, however, a shortage of experienced maintainers, as the recently closed manning gap resulted in an overabundance of 3-level maintainers and a shortage of 5-level and 7-level maintainers. Air Force Secretary Wilson said it herself: “Readiness is first and foremost about training people.” Air Force Training Support Squadrons (TRSS) are tasked with creating maintenance training. This paper is a follow up to previous work in which augmented reality (AR) was utilized for content creation of maintenance training and details next steps, which involved piloting training content on the flight line in conjunction with the 367 TRSS and C-5 maintainers out of Travis Air Force Base. The purpose of this paper is to detail the planning, execution, and results of this pilot program. Using surveys, interviews, and think-aloud protocol, the pilot program evaluated the potential for acceleration of maintainer task confidence through the use of an AR just-in-time training solution, as well as user acceptance of AR as a training medium. Results lead to analysis of potential confidence gains and user acceptance, suitability of AR training in the flight line environment, and transference to other maintenance applications and industries.

ABOUT THE AUTHORS

Charis Horner is a Product Engineer in the eXtended Reality Division at Design Interactive Inc. Her experience surrounds emerging technology delivery to diverse stakeholders, including the Office of Naval Research, the Air Force Research Laboratory, the Naval Air Systems Command, the Army Futures Command, and the Department of Homeland Security. Her current work is centered around the design and delivery of XR technologies across the Department of Defense (DoD) with a focus on driving user-centered design for scalable adoption of AR/VR/MR training and operational support solutions. She holds a Bachelor of Science in Industrial Engineering from the University of Central Florida with a minor in Engineering Leadership, where her focus included Human Factors.

Christina Padron is the Director of DoD Programs at Dynepic Inc. She has over 10 years of experience in the design, development, and evaluation of virtual assessment and training tools for a variety of DoD research agencies. Her work focuses on supporting the design, development, and usability of training and job aid solutions, specifically ensuring that the solutions are optimized for their specific users, tasks, and context of use. She has recently led efforts in which extended reality training systems were developed to train collective tasks in M1 Abrams tanks for the Army, network maintenance on Naval aircraft carriers, and Air Force aviation maintenance. She holds an MS from Penn State University in Industrial Engineering with a Human Factors focus, and a BS from Purdue University in Industrial Engineering.

Mr. Troy Westbrook is the Deputy Director for the 367 Training Support Squadron (TRSS). He entered the Air Force in 1983 and served over 21 years as a KC-135 Aircraft Crew Chief. He went on to serve as a Maintenance Qualification Training Program Instructor, retiring in the grade of Master Sergeant in 2005. He returned to the 367 TRSS as a civilian where he serves as the Deputy Director, continuing to support AETC’s Continuum of Learning and virtual training initiatives. Mr. Westbrook achieved a Bachelor of Science in Computer Science from Park University and an Applied Science degree in Aircraft Maintenance from the Community College of the Air Force.

Evaluating the Use of Augmented Reality for Aircraft Maintenance Training

Charis Horner
Design Interactive Inc.
Orlando, FL
Charis.horner@designinteractive.net

Christina K. Padron
Dynepic Inc.
Charleston, SC
Christina@dynepic.com

Troy Westbrook
USAF AETC 367 TRSS
Hill AFB, UT
Troy.westbrook@us.af.mil

INTRODUCTION

Maintaining aircraft within the Air Force remains key to being mission ready. However, as the Fleet continues to age, the average mission-capable rate across the Fleet has steadily declined, dipping below 70% in fiscal year 2018, the lowest in six years (Losey, 2019a). The older Fleet requires additional inspection and maintenance; the Air Force is seeking to reduce depot-based maintenance and more flight line maintenance as a potential solution. The flight line is defined as base locations where aircraft are stored, parked, loaded, offloaded, and serviced. This includes both within and outside of aircraft hangars. There also remains a shortage of experienced maintainers, as the recently closed manning gap resulted in an overabundance of 3-level maintainers, or apprentices, and a shortage of 5-level and 7-level maintainers, respectively known as journeymen and craftsmen. Specifically, the Air Force has been left with a 2,483-person shortfall of journeymen and craftsmen, while maintaining a surplus of 1,745 apprentice maintainers (Losey, 2019b). Thus, the Air Force is experiencing a skill level gap in existing maintainers, and the need exists to rapidly accelerate 3-level maintainers to become journeymen and craftsmen. In early 2019, Air Force Secretary Heather Wilson confirmed that while the manning gap has been closed, the challenge remains to "season those [new apprentices] and make them exceptional airmen on the flight line." She went on to state that "readiness is first and foremost about training people" (Pawlyk, 2019).

Research, design, and implementation of augmented reality (AR) has occurred to determine the role this technology may have in supporting maintainer skill acquisition and task execution. In a systematic review of AR applications, aviation maintenance was tied with consumer technology as the third most common maintenance application subgroup found in the literature, after mechanical and plant maintenance (Palmarini, Erkoyuncu, Roy, & Torabmostaedi, 2018). AR aviation maintenance implementations were found in tasks such as assembly and disassembly, repair, inspection, and training (Palmarini et al., 2018). Previous research has found that AR should prove useful in supporting even highly complex maintenance activities involving many parts and materials (Ceruti, Liverani, & Marzocca, 2015). Thus, AR technology can be considered as a means to fill a wide array of aviation maintenance training needs (Haritos & Macchiarella, 2005), and aviation maintainer acceptance of AR as a training technology is expected to be high (Wang, Anne & Ropp, 2016). AR has also been evaluated as an operational support mechanism. While technical manuals provide standard operating and troubleshooting procedures, their use tends to be time-consuming; manuals are often in print form or hard-to-search online directories, slowing task execution and leading to maintainer frustration (Hincapie, Caponio, Rios, & Mendivil, 2011). As such, AR can provide a digital, contextually aware supplement to aviation maintenance technical manuals.

Padron, Horner, Westbrook, and Davidson (2019) built upon initial implementations of AR in industry, evaluating the suitability of AR to meet the needs of training content creation as carried out by Air Force Training Support Squadrons (TRSS). The use of AR in training content development was evaluated, specifically the ability to provide spatially contextualized three-dimensional (3D) training on actual equipment, as well as the ability to enable hands-free reference while practicing or completing maintenance tasks. An AR authoring tool was shown to support TRSS for centralized media capture, expediting the process of content authoring. This was done by allowing onsite maintainers to capture and store media within the centralized content creation platform, where TRSS content creators could directly access media to generate content within the platform. It was found that AR was particularly suitable for just-in-time training before scheduled or unscheduled maintenance and to provide spatial orientation around an airframe. It was posited that AR would accelerate maintainer proficiency on tasks as a form of just-in-time training, specifically in reference to en-route maintenance facilities. In these facilities, no aircraft are on base regularly, and maintainers may need to service any number of different models depending on Fleet needs. AR may be particularly well suited to meet such needs.

To expand this work, authored AR training content needed to be implemented and evaluated to determine suitability and effectiveness for just-in-time training on the flight line. Validation of this type of training could bolster the Air Force's ability to conduct less depot-based maintenance and more flight line maintenance, as well as provide a mechanism to accelerate skill acquisition of 3-level maintainers, thus enhancing fleet readiness. While Padron et al. (2019) demonstrated an AR authoring tool can effectively meet training content creation needs before having access to an aircraft on the flight line, evaluation of this authored content in context, as training, was still needed. This paper builds upon that work by completing an initial evaluation of AR training content with Air Force maintainers on the flight line to further validate the use of AR for aircraft maintenance training.

PILOT PROGRAM METHOD

The 367 TRSS out of Hill Air Force Base (AFB) is tasked to create training products for Air Force maintainers, following a robust process in which they elicit and validate training needs, analyze and establish training tasks, select the appropriate training medium, author content, iteratively test the product, and scale training roll out while monitoring validation (Padron et. al, 2019). In the iterative testing of the product, the 367 TRSS uses a small group tryout approach, which represents an initial pilot program with the intended end-user group to evaluate the product's effectiveness. The typical goal of this pilot program process is to validate the efficacy of authored content to meet the training need as identified by the 367 TRSS. To conduct this, representatives from the 367 TRSS travel to the maintainers' location, train the maintainers on the new product, and implement the product using a variety of analysis methods to test training gains. As such, the same process was followed to evaluate the authored AR training content.

Before this effort, the 367 TRSS had not implemented any AR products for maintenance training, so this evaluation represented a key innovation milestone. Because implementation of an AR training product introduced a novel technology, the main aim of this implementation process was to see AR authored content through the pilot program stage, determining relevance of its use in Air Force training. Additionally, pilot program goals for the AR implementation were to evaluate maintainer task confidence gains with introduction of an AR training product and determine end-user acceptance of AR as a flight line maintenance training technology.

The pilot program targeted the C-5 Super Galaxy, the largest aircraft in the fleet. The pilot program team included individuals from the 367 TRSS, Design Interactive Inc. (DI), and Air Force C-5 maintainers at Travis AFB. Prior to the start of the pilot program, the team worked to clarify assumptions, select a specific use case with specific tasks, and identify the appropriate AR hardware and software. Participants were recruited and a testing process established, along with the definition of measures for analysis.

Assumptions

Testing for this pilot program occurred on the flight line, both within a hangar and outside a hanger. The C-5 aircraft was chosen for this pilot program due to its immense size and the complexity of maintenance tasks, stretching the limits of AR use on the flight line, therefore ensuring that other flight line implementations would be feasible if the C-5 implementation was successful. As such, the team assumed flight line testing findings of the pilot program would apply to flight line implementations with other aircraft.

As previously discussed, this AR implementation aimed to evaluate the technology's suitability within the case of en-route maintenance for the C-5 Super Galaxy. En-route maintainer facilities, however, are typically in overseas locations difficult to access for an initial pilot program. As such, the pilot program was not able to be conducted in an en-route facility or with en-route maintainers. In lieu of this, pilot program testing occurred with maintainers who remain at the C-5 flight line and service only the C-5 aircraft. Therefore, it was assumed that results garnered by testing with these users would be sufficiently representative of the en-route user group to provide insight into the value proposition for AR use within Air Force maintenance.

Task Selection

There were two major task categories selected for training content creation and subsequent user testing on the C-5 Super Galaxy. The first was aircraft safe-for-maintenance, which involves confirming the aircraft is properly grounded, bonded, and safe so maintainers can carry out procedures. Safe-for-maintenance procedures span the entirety of the aircraft and include pre-flight, post-flight, and in-flight tasks. To enable optimal user testing, a subgroup of tasks was selected from the safe-for-maintenance category. Specifically, the pre-flight inspection procedure of the nose landing gear (NLG) and main landing gear (MLG) were chosen. The inspection of the landing gear is a critical pre-flight task, as it helps to ensure safety of the aircraft and crew each time the plane operates. These occur outside the physical aircraft, enabling testing of AR use and its suitability in an ambient aircraft storage environment. AR is suited to this task, as the spatial orientation of content to the landing gear itself allows maintainers to fully engage in physical actions necessary to complete the task within their environment, while still obtaining reference to training information provided in the AR headset.

The second category of selected tasks was the powering of electrical systems in the aircraft enabling the performance of maintenance. This involves procedures to apply, display, and control electrical power, as well as disconnecting the core processing power and external power upon completion of maintenance. Within this task category, the subtasks selected were those of Auxiliary Power Unit (APU) Operation and Shutdown, as well as Apply/Control and Display Power On/Off. Both tasks occur on the flight deck inside the aircraft and involve operating switches and checking displays. Therefore, selection of these two tasks allowed for evaluation of AR use both outside the aircraft at the landing gear and inside the aircraft on the flight deck.

Analysis and Implementation of Hardware and Software

In previous work (Padron et al., 2019), types of available AR hardware were evaluated for C-5 aircraft maintenance and the use of a spatially aware heads-up display (HUD) was selected. At the time of the study initiation, the two available spatially aware HUD hardware devices were the Microsoft HoloLens 1 and the Magic Leap 1. Each device was evaluated for its suitability to support task performance. In comparison to the Microsoft HoloLens 1, the Magic Leap 1 is superior in terms of field of view, ability for hand and finger tracking, eye tracking, and processing speed of the central processing unit. However, the Magic Leap 1 has drawbacks in terms of user experience, as users with corrective eyewear cannot use the device without custom optical lenses. By contrast, the Microsoft HoloLens 1 is compatible with the use of corrective eyewear without customization. The primary user interactions with the Microsoft HoloLens 1 are hands-free gestures carried out in space, while the primary form of interaction for the Magic Leap 1 is a controller, negating a key value of a HUD. From a software development perspective, the Mixed Reality Toolkit provided by Microsoft for the HoloLens 1 contains a vast toolset of developer workflows and functions allowing for rapid prototyping and quick use case implementation. This exceeded the development capabilities provided by Magic Leap, which require more time developing low-level sensor functionality and human interface interactions. As such, the use of the Magic Leap 1 proved less than ideal for the rapid nature of the training content development and delivery process for the pilot program. Therefore, the Microsoft HoloLens 1 was selected as the hardware of choice, which was also seen as being strategically sensible due to the proliferation of Microsoft product use in the Department of Defense (South, 2020) and team accessibility to these devices.

Four HoloLens devices were obtained for the pilot program. The HoloLens devices were provisioned with an AR authoring software created by DI, which enabled 367 TRSS content creators to develop training content without the need of software engineers. This authoring software was downloaded directly onto the HoloLens devices using the HoloLens Windows Device Portal. This software also served as the training platform, as authored content could be viewed and interacted with in a viewer mode intended to be used by maintainers. Additional software used was a web portal where the AR content was authored by the 367 TRSS before being finalized in the HoloLens. This web portal was accessible on laptop computers via URL and unique login credentials enabling both author and viewer roles. Other hardware used included personal laptop computers for web portal access and chargers for all devices.

Participant Selection and Demographics

A representative group of 37 stateside C-5 maintainers was chosen to participate in the pilot program. These C-5 maintainers were Crew Chiefs, of which the total population is 492. Crew Chief responsibilities include inspecting, troubleshooting, and maintaining aircraft structures, by performing tasks such as service operations, aircraft inspections, and maintenance problem diagnosis. The distribution of 3-level, 5-level, and 7-level maintainers within the user group is detailed in *Table 1*.

Prior to testing, each participant was trained on the basics of operating a HoloLens device, including the use of gestures such as tap, pinch, and bloom. They were also introduced to the AR training software and how to navigate through the training procedure, specifically how to use the menu and in-space steps to access the training content.

Table 1. Maintainer Participant Distribution

Experience Level	Count
3-level	8
5-level	15
7-level	11
Unidentified	3
Total	***37***

Authoring Process and Schedule

The process began with authoring just-in-time training content for the selected tasks: 1) NLG Pre-Flight Check, 2) MLG Pre-Flight Check, 3) APU Operation and Shutdown, and 4) Apply Power and Control Display. The 367 TRSS team, with support from researchers, followed their aforementioned TRSS development process using the AR authoring software to create a training product for the tasks. Prior to the time of the pilot program, authors at the 367 TRSS created content using the web portal and AR authoring tool. The training was authored using information from technical orders, as well as media associated with carrying out the tasks (e.g., a video of how to conduct an inspection). The content was first authored in the web portal, then spatially situated in the environment. The environment included the flight deck, MLG, and NLG. Media, specifically pictures, videos, and holograms, were added to some steps of the training. The content was initially authored during a trip to Dover AFB in Delaware and at the 367 TRSS at Hill AFB in Utah. At Travis AFB, the content was refined using the AR authoring tool in the HoloLens on the first day of the pilot program. Refinement included spatially placing training content in relevant locations and adding additional media.

The schedule for the pilot program included authoring of content, creation of surveys and evaluation methods, and conducting user testing. Initial content authoring occurred on a C-5 located at Dover AFB in Delaware in early fall of 2019. This is where media capture occurred, as pictures and videos were retrieved from onsite expert maintainers at Dover AFB. Content authoring refinement, in which content authors edited media, incorporated information from technical orders, and generated the structure of the content, continued at Hill AFB over the next couple of months. During this time, a pre-course survey and a post-course survey were developed. The pre-course survey included asking participants to assess their confidence level on each of the four tasks they would be completing on a Likert Scale from 1-5, with 1 being “not confident at all” and 5 being “very confident”. The post-course survey included a confidence assessment of the same type, intended to measure self-reported task efficacy. The post-course survey also included a system evaluation intended to determine user acceptance. The training content was tested with end-users in the late fall of 2019. When the team arrived at Travis AFB to set up for the pilot program, the AR content was reviewed and refined to correct its spatial location. Over the next few days, 37 participants reviewed the training content within a Microsoft HoloLens while on a C-5 parked in a hangar. On the final day, environmental suitability testing with the team occurred on a C-5 parked outside a hangar.

Participant Testing Procedure

Each participant began by taking a pre-course survey, which assessed their confidence and understanding of the tasks to be completed. They then followed one of the two participant procedures shown in *Table 2*. Which participant procedure each individual followed was determined by whether more participants were currently in the flight deck or at one of the two landing gear locations. The less populated location is where participants would begin.

As participants completed the just-in-time training, they were asked to speak aloud, noting anything they liked or did not like, anything helpful or confusing. After completing the training, all participants completed the post-course survey, reevaluating their task confidence and providing a review of the system. Participants were also interviewed to collect anecdotal data, including what they thought was useful or unnecessary, and in what types of environments, if any, they could see the AR training content being valuable.

Table 2. Participant Procedures

Step	Participant Procedure 1	Participant Procedure 2
1	Travel to the MLG	Travel to the flight deck
2	Receive training on how to use the HoloLens and navigate the AR software	Receive training on how to use the HoloLens and navigate the AR software
3	Complete MLG Pre-flight Check training	Complete APU Operation and Shutdown training
4	Travel to the NLG	Complete Apply Display and Control Power training
5	Complete NLG Pre-flight Check training	Travel to the MLG
6	Travel to the flight deck	Complete MLG Pre-flight Check training
7	Complete APU Operation and Shutdown training	Travel to the NLG
8	Complete Apply Display and Control Power training	Complete NLG Pre-flight Check training

Definition and Collection of Measures

Participant knowledge of targeted tasks was measured and assessed through surveys as described above. Participant performance in this context was defined as the ability to navigate through the AR system successfully while comprehending training content as delivered. This performance was measured through interviews, observations, and surveys. Methods used to collect data were participant face-to-face interviews, collection of participant observations while thinking aloud during training, and participant feedback surveys including a pre-course and post-course survey. Metrics collected included percent increase in task confidence directly after being trained in AR, reported likelihood to adopt or recommend the AR training modality, and percent increase in training satisfaction.

RESULTS

Content Authoring Results

Leading up to the pilot program, the majority of content authoring was conducted by two training content authors at the 367 TRSS. These two authors used the web portal to enter information from the technical orders, traveled to Dover AFB to capture media, and put content together into training modules for the four tasks selected. For the pilot program, these two authors planned to be on-site to verify the content at Travis AFB and complete any necessary modifications. Unfortunately, due to extenuating circumstances, these two authors could not attend the pilot program and two other TRSS members traveled to Travis AFB as authors in their stead. These authors had received minimal training on the software. Once onsite, the new authors were able to quickly learn to author and edit previously made content. As relative novices for authoring, they became proficient with the software in a matter of hours. Throughout the four days of the pilot program, the 367 TRSS authors reacted to how participants were perceiving the content and software and adjusted the content to mediate user concerns and garner more complete feedback. These adjustments included refining the spatial location of content and adding media such as technical order specification images.

These experiences illuminated key benefits of the AR authoring tool. Even as novice users in the authoring software and AR hardware, the 367 TRSS team could adapt content in a short time frame, tailoring it to the needs of the specific environment. They could move virtual content within the tool, placing it in relevant locations both on the flight deck and at both sets of landing gear, highlighting the flexibility and efficiency of the software. The usability and intuitiveness of the authoring tool were also demonstrated here, as the procedure for developing content in both the web portal and AR device took little time to explain and the new authors were able to autonomously carry out authoring with cues from the software. Both of the new content authors considered themselves expert AR authors by the end of the four-day pilot program, highlighting the learnability of the software. The content authors also found the experience of using the software to be satisfying, as it featured an aesthetic and minimalist design. Flexibility of the software was also identified as a key benefit. The option to author AR training content using existing media like pictures, videos, and holograms allowed augmentation of training content, as media could be captured in or added to the tool within minutes. Authors could capture new media in the tool or add existing content via the web portal. This way, AR could coexist with previous training mediums, as existing media could be populated within the AR software tool to be accessible and viewable in a hands-free environment. One highlighted component was the ability to add spatially relevant holograms to direct users' attention to important points within the training environment, supporting navigation and wayfinding. Authors found the use of this component to be both satisfying and robust.

Additional needs for the hardware and software of the AR authoring tool were also identified through the pilot program authoring stage. Although content authors understood how to create content, they did not have guidance as to functions of the AR authoring tool which could be used to enhance content quality. For example, participants responded well to the introduction of media to steps in the training content, but not all steps contained media beyond the step description and its location in the environment. If the content had been authored to contain more descriptive media, value derived from the tool may have increased. As such, there is a need for additional guidance on how to fully utilize an AR authoring tool in tandem with principles of instructional design and contextualized learning.

Improving the spatial location mapping of the HoloLens device was also identified as an authoring need. The utilized AR authoring software took a spatial anchor approach to orient content. Similar to a quick response (QR) code, the anchor would be scanned and virtual content affixed to the world relative to the anchor location would appear. When content was first authored at Dover AFB, virtual content was placed throughout the aircraft relative to a single spatial anchor. This was done to determine if the AR hardware and software could support training across the entirety of the aircraft without needing multiple spatial anchors. The aircraft has a length of 247 feet 10 inches, and a height of 65 feet 1 inch (U.S. Air Force, 2006). This stretched the Simultaneous Localization and Mapping (SLAM) capabilities of the HoloLens, which recommends virtual content be placed with 9.8 feet of the spatial anchor (Microsoft, 2019). Content did appear correctly relative to the single spatial anchor when first authored at Dover AFB. However, it did not appear correctly on a different C-5 at Travis AFB. This was the main impetus for the editing of the authored procedures once on location for the pilot program. The team had to reorient the training content to compensate, adding a spatial anchor on the flight deck and spatial anchor near each landing gear.

The most frequent participant feedback, which will be further explored in the subsequent section, was that spatial locations of the training content relative to the physical environment were inaccurate. After the pilot program, an AR researcher determined the content location was likely affected by the spatial anchor placement. The anchor was suspended from the ceiling of the flight deck and placed on the curved surface of the airframe at the landing gear. Slight changes in the angle of the anchor led to warping of the environment mesh and misalignment of the content. Therefore, the importance of placing the spatial anchor on a stable, flat surface was discovered. Further inconsistencies in the placement of the virtual content can be tied to the built-in hardware SLAM system. Therefore, enhancement of the SLAM system in the AR hardware, guidance on anchor placement, and refinement of spatial placement strategies within the software are identified as areas in need of additional research.

Training Survey Results

The first survey was a participant self-assessment. Participants were asked to rate their knowledge on the tasks before and after taking the AR training course on a Likert scale from one to five, one being "not confident at all", and five being "very confident". Responses to each item were averaged across participants. On all courses, participants showed improved average confidence in task execution after taking the course in AR. The most significant increase was with the Pre-flight NLG course. Notably, this course was adapted prior to the second day of testing, to include more holograms, images from the technical orders, videos, and other media content. This suggests that increased use of media authoring tools will increase confidence in successfully performing tasks after completing the AR course.

It should also be noted that pre-course confidence self-assessment for each of the tasks was greater than three, indicating participants on average had some level of confidence in completing these tasks before the AR training. Because these participants are Crew Chiefs for the C-5 and this aircraft is the only one most are tasked to service, it is unsurprising that most participants would be confident on how to carry out these C-5 servicing tasks prior to additional AR training. Therefore, the maximum potential for confidence increase was low for each task, ranging from 22% to 47.1%, and survey results should be regarded in light of this. Even with the high level of confidence participants had coming into the pilot program, participants' confidence still increased with the use of the just-in-time AR training for each task. The results of this survey speak to the original pilot program goal of determining maintainer task confidence gains with the use of AR. Findings display gains of up to 11.8% with the adoption of an AR training tool for maintainers who are already confident in their abilities. *Table 3* details the outcomes of the self-assessment**.**

Table 3. Self-Assessment Survey Results

Course	Pre-course confidence self-assessment	Post-course confidence self-assessment	Average Delta	Self-assessed confidence learning increase
Pre-Flight NLG	3.4	3.8	0.4	11.8%
Pre-Flight MLG	4.0	4.3	0.3	7.5%
APU Operation & Shutdown	4.1	4.2	0.1	2.4%
Apply/control & display power on/off	4.1	4.3	0.2	4.9%

Participants were asked to state what they liked about the AR training modality and provide suggested improvements as a part of the survey. On the first day of testing, participants called out the spatialized media displayed in the course as a key feature they liked. They also enjoyed the addition of another avenue for training in comparison to traditional methods. Others believed the AR content was an easier way to learn tasks than following the technical order alone. Key areas of AR improvement noted by participants included the need to make the AR content location more accurate, desire to see more media, and need for inclusion of more technical order specific information (e.g., specification cards).

At the start of day two of data collection, AR content was modified by content authors to include more media. On this day, survey feedback changed slightly, reflecting the modified content. Key elements of the system participants reported enjoying were the step-by-step spatialized location of training content, ease of interaction and use of virtual content, and ability to see technical content from manuals in space while carrying out tasks. Potential improvements noted once again referred to the accuracy of AR content location in space and better-quality media (e.g., pictures with higher resolution).

The final element of the survey was evaluation of the AR training system itself. Participants were asked to rate each statement on a Likert scale of one to five (one being "strongly disagree" and five being "strongly agree"). For all statements save one, the average response was greater than three, indicating that participants had a neutral or positive response to the question. Average participant responses are shown in *Table 4*.

Table 4. System Evaluation Survey Results

Statement	Average Participant Response
I understand the purpose of the training provided.	4.5
The AR/VR method was a beneficial approach to maintenance training.	4.1
The software helped me improve my overall understanding of the maintenance procedures.	3.2
The software was easy to use.	3.9
I thought the course flowed smoothly.	3.8
I believe that most people would learn to use this system quickly.	4.2
I felt comfortable using the AR/VR system.	4.2
Using AR/VR will improve my job performance.	3.3
I was able to see the virtual content easily while using the AR/VR system.	3.8
I found the headset comfortable to wear.	3.4
I found the interaction with the software easy to use (hand gestures, taps, etc...)	3.7
I would like to continue using AR/VR for training.	3.7
I prefer using AR/VR in comparison to traditional training.	2.7
I would recommend the continued use of AR/VR for maintenance training.	3.9

The one statement to which participants reacted with an average response of less than three was "I prefer using AR/VR in comparison to traditional training." In this context, traditional training would include schoolhouse instruction with guided instructors, computer-based training, and interactive multimedia instruction. This type of training is not all representative of just-in-time training, as much of this training is preloaded prior to completing tasking. It is possible that participants when responding to this statement believed the just-in-time AR training would replace the robust schoolhouse instruction, which lays the groundwork for skill acquisition. The just-in-time training content tested as a part of the pilot program was created for maintainers with experience, rather than new maintainers. As such, this flight line AR training is meant to complement, not replace, schoolhouse training.

Four statements had a mean response of greater than four, indicating responses of agree or strongly agree. These four statements are as follows: "I understand the purpose of the training provided," "The AR/VR method was a beneficial approach to maintenance training," "I believe that most people would learn to use this system quickly," and "I felt comfortable using the AR/VR system." The first statement, understanding the purpose of the training provided, speaks to the clarity of the presented training content and ability of the AR tool to communicate meaning to the participants. The remainder of the statements to which participants responded favorably speak to one of the initial pilot program goals, determining end-user acceptance of AR as a flight line maintenance training technology. Participants believed the AR training method was beneficial, easy-to-learn, and comfortable, suggesting they would be likely to adopt such a capability in the future.

Training Interview Results

The anecdotal responses from participants, both as they were thinking aloud while using the AR software and during the interview once they completed training, were recorded and coalesced. A significant perception was that it was helpful to see, resize, and manipulate media in space. Participants also believed the AR training tool would be useful for new airmen unfamiliar with an airframe or tasks. When asked where they saw AR most valuable, eleven airmen commented that trainees coming out of maintenance school would benefit from this system. The system can provide new airmen orientation to the aircraft, attention direction for important tasks, and step-by-step task walkthrough in a heads-up spatialized way. Participants also believed it was helpful to be navigated in space from one step to the next. Ten airmen commented on the navigation system included in the AR software, which directed them to the next piece of content, noting it to be helpful for task wayfinding. Potential enhancements identified through interviews included ensuring easy stepwise transition between pieces of AR content in space; eight participants commented on this function. Additionally, seven participants commented on the display of media within the tool, expressing desire for the media to be more easily located in space. This can be achieved through improved AR design paradigms. Overall, participants remarked that the software is innovative and engaging, and they liked being able to see practical training content in a new and exciting way. More testing, verification in a variety of environments, and system hardening would allow the identified opportunities to be realized in the Air Force environment.

DISCUSSION AND CONCLUSION

The initial aim of this effort was to determine if AR can bolster the Air Force's ability to conduct flight line maintenance, providing a mechanism to accelerate the skill level of maintainers. To accomplish this, the pilot program conducted sought to determine the suitability of AR for use in the flight line environment, to get a glimpse of task confidence gains achievable with AR training software, and to understand end-user acceptance of AR as a flight line maintenance training technology. The pilot program results speak to the ability of an AR tool to meet these Air Force maintenance training needs.

Suitability for Maintainer Training

Participant survey data collected during the pilot suggests task confidence gains may be achievable with AR training software in a flight line environment. Each tested course led to an increase in participant confidence in carrying out the maintenance task. Though increased confidence in performing a task does not necessarily correlate with improved task performance, it provides evidence further testing focused on learning gains would be useful. Further, many experienced participants commented this type of just-in-time AR training could be greatly beneficial to newer maintainers who recently completed schoolhouse training and for en-route maintainers who have less opportunity for familiarity with a specific aircraft. Notably, many participants commented that content displayed as a part of the training would also be helpful as an operational support aid, specifically for newer maintainers who are less comfortable carrying out a task. This is sensible, as training displayed as a part of the pilot program was intended to be used by an experienced technician as a refresher just prior to executing a task. As such, training content closely mirrored operational steps conducted as a part of the task, with additional guiding media to explain how to complete steps. Within just-in-time training, each step was affixed in a relevant location within the environment. For example, if a step within APU startup operation and shutdown occurred on a switchboard panel in the flight deck, training content would be affixed to the specific switchboard where the step occurred. This could be easily translated for operational support, as type of displayed content closely mirrors actual tasks to be performed. Finally, results of the system analysis survey showed participants believe AR is a beneficial approach for maintenance training, anticipating most maintainers could quickly learn to use AR and feel comfortable with the technology. Thus, it was confirmed end-users are likely to accept AR as a training technology if fully integrated.

Environmental Suitability

To investigate the suitability of AR for use in a flight line environment, the pilot program was conducted on the flight line at Travis AFB, both inside and outside a hangar. In both instances, participants were able to use AR training to access and progress through content. Previous work has suggested use of HUD AR is suited exclusively to indoor environments with consistent lighting (Sailer, Rudi, Kurzhals, & Raubal, 2019). On the last day of the pilot program, testing occurred to assess whether content could be used outside of a hangar in the outdoor flight line environment. While in this outdoor environment, content was evaluated both at landing gear underneath the plane, which is shaded by the plane itself, and within the flight deck inside the plane. In both cases, content was able to be viewed and accessed with similar suitability as in hangar testing. This suggests use of an AR heads-up device may be suitable for outdoor flight line training if at least one of the following conditions is met: 1) area of training content is inside the aircraft or 2) area for training is shaded, either by the aircraft or by other means.

While testing inside the hangar, digital AR content was perceived well at landing gear and trainees had no difficulty accessing and navigating through content. In the flight deck, however, some participants had difficulty seeing the actual environment through brightness of digital content within the AR device. During testing, the flight deck environment was darker than the landing gear environment, as landing gear had ambient lighting from sunlight and hangar, while the flight deck was shaded. Additionally, much of the physical environment of the flight deck was dark. For example, the switchboards used to power on the aircraft were black. Due to these environmental factors, the digital content partially obscured the physical environment, making it difficult to see digital content overlaid on actual equipment. This was mitigated, however, by adjusting built-in brightness setting on the AR device, allowing digital content to be dimmed to suit the environment. Therefore, it is important when testing in a dim environment or with darker equipment to ensure brightness of the digital content is adjusted to suit the physical environment.

Applicability to Other Industries

The aircraft maintenance environment is comparable to other maintenance applications across the Department of Defense and industry. Similar AR training software has been implemented on the CVN-78 Gerald R. Ford aircraft carrier. Preliminary findings suggest that AR is also suitable to this environment, can help sailors achieve training gains, and is likely to be accepted by end-users. Vehicle maintenance, such as Army Joint Light Tactical Vehicle maintenance or commercial semi-trailer truck maintenance, is also a key area primed to benefit from this type of technology, as equipment similarly requires hands-on maintenance activities which would benefit from a heads-up display suitable for both indoor and outdoor environments. In fact, findings of AR suitability through this pilot program can be applied to any domain requiring hands-on interaction with physical environments.

Future Work

Future work is needed to continue evaluation of AR for Air Force maintenance training. Testing should be conducted within en-route facilities to ensure confidence gains and user acceptance are similarly applicable to en-route maintainers. Additionally, this pilot program utilized AR tied to a physical environment where equipment was present. A key benefit of AR is ability to provide contextualized, embodied training when equipment is not present by using virtual holograms. Therefore, future work should evaluate implementation of virtual equipment training. This pilot program evaluated self-assessed confidence gains as reported by participants, showing promise with improved results. It is recommended future work utilize empirical formative evaluations to assess participant learning gains to be compared to self-assessed confidence increases. Finally, findings of this pilot program call for future improvements to commercially available AR hardware devices, specifically improving SLAM quality, along with previously identified needs to improve field of view and storage limitations (Padron et al., 2019). Expansion of analysis to other benefit areas such as ground vehicle maintenance and shipboard maintenance operations is recommended.

This paper intended to evaluate AR in use within the Air Force training environment, laying groundwork for AR to assist with acceleration of maintainer task proficiency. Preliminary analysis included evaluation of maintainer task confidence increases and determination of maintainer acceptance of AR as a training tool. Through implementation of a just-in-time AR training product, the pilot program illuminated benefits provided by AR training, along with potential enhancements to more acutely meet needs of the Air Force. Findings show AR training can assist with preparation for and execution of flight line maintenance, enabling the Air Force to decrease reliance on depot-based maintenance. It was shown that AR training can increase maintainer confidence in task execution. This is the first step in evaluating ability of AR to increase proficiency of maintainers, assisting with the need to accelerate 3-level apprentices to 5- and 7-level journeymen and craftsmen. Thus, AR can provide an effective new mechanism to train maintainers, bolstering the ability of the Air Force to fly, fight, and win.

REFERENCES

Ceruti A., Liverani, A., & Marzocca, P. (2015). A 3D User and Maintenance Manual for UAVs and Commercial Aircrafts Based on Augmented Reality. *Proceedings of the SAE 2015 Aerotech Congress and Exhibition*, *1*, 1-8.

Haritos, T., & Macchiarella, N. (2005). A Mobile Application of Augmented Reality for Aerospace Maintenance Training. *24th Digital Avionics Systems Conference*, *1*, 1–9.

Hincapie, M., Caponio, A., Rios, H., & Mendivil, E. G. (2011). An introduction to Augmented Reality with applications in aeronautical maintenance. *2011 13th International Conference on Transparent Optical Networks*, 1–4.

Losey, S. (2019, August 1). Aircraft mission-capable rates hit new low in Air Force, despite efforts to improve. Retrieved May 6, 2020, from https://www.airforcetimes.com/news/your-air-force/2019/07/26/aircraft-mission-capable-rates-hit-new-low-in-air-force-despite-efforts-to-improve/

Losey, S. (2019, February 8). The Air Force still has a serious maintainer staffing problem, GAO says - but no strategy to fix it. Retrieved May 6, 2020, from https://www.airforcetimes.com/news/your-air-force/2019/02/08/the-air-force-still-has-a-serious-maintainer-staffing-problem-gao-says-but-no-strategy-to-fix-it/

Microsoft. (2019, February 24). Spatial anchors. Retrieved May 12, 2020, from https://docs.microsoft.com/en-us/windows/mixed-reality/spatial-anchors

Padron, C. K., Horner, C., Westbrook, T., Davidson, J. (2019). Utilizing Augmented Reality for Air Force Maintenance Training. In *Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC),* Orlando, FL.

Palmarini, R., Erkoyuncu, J. A., Roy, R., & Torabmostaedi, H. (2018). A systematic review of augmented reality applications in maintenance. *Robotics and Computer-Integrated Manufacturing*, *49*, 215–228.

Pawlyk, O. (2019, February 8). The Air Force Has Fixed Its Active-Duty Maintainer Shortage, SecAF Says. Retrieved May 6, 2020, from https://www.military.com/dodbuzz/2019/02/08/air-force-has-fixed-its-active-duty-maintainer-shortage-secaf-says.html

Sailer, C., Rudi, D., Kurzhals, K. & Raubal, M. (2019). Towards Seamless Mobile Learning with Mixed Reality on Head-Mounted Displays. In *Proceedings of World Conference on Mobile and Contextual Learning 2019*, 69-76.

South, T. (2020, February 11). The Army wants to buy 40,000 ‘mixed reality’ goggles. Retrieved May 12, 2020, from https://www.armytimes.com/news/your-army/2020/02/11/the-army-wants-to-buy-40000-mixed-reality-goggles/

U.S. Air Force. (2006, May 15). C-5M Super Galaxy. Retrieved May 6, 2020, from https://www.af.mil/About-Us/Fact-Sheets/Display/Article/104492/c-5-abc-galaxy-c-5m-super-galaxy/

Wang, Y., Anne, A., & Ropp, T. (2016). Applying the Technology Acceptance Model to Understand Aviation Students’ Perceptions toward Augmented Reality Maintenance Training Instruction. *International Journal of Aviation, Aeronautics, and Aerospace, 3*(4).