Ethnographic Study of a Commercially available Augmented Reality HMD app for Industry Work Instruction

Andrew Pringle
School of Psychology
Trinity College Dublin
Dublin, Ireland
pringlea@tcd.ie

Stefanie Hutka
Adobe Design Research & Strategy
Adobe Inc.
San Francisco, CA
hutka@adobe.com

Jesse Mom
BAM Advies & Engineering
Royal BAM Group
Utrecht Area, Netherlands
jesse.mom@bam.com

Robin van Esch
BAM Advies & Engineering
Royal BAM Group
Utrecht Area, Netherlands
robin.van.esch@bam.com

Niall Heffernan
Sales
DAQRI & Curtiss-Wright
Dublin, Ireland
niall.heffernan@curtisswright.com

Paul Chen
Product Management
DAQRI & Komprise
Los Angeles, CA
paul.chen@komprise.com

ABSTRACT

Industrial applications of Augmented Reality (AR) are becoming increasingly commonplace but there are only a small number of published user studies examining the use of commercially available AR technologies, like AR HMDs, with real workers in real industry settings. This paper presents ethnographic research of an industry task that includes the context of the industry procedure, pain-points with current methods and a user experience study of an HMD-delivered AR application for delivering work instructions to support engineers performing the procedure. The AR application is delivered to engineers with different levels of experience through a commercially-available AR HMD (the DAQRI Smart Glasses®). Engineers (users) were observed and video recorded by researchers as they performed the procedure in the real-world setting of a sprinkler room of a hospital in the Netherlands. Engineers who used AR were found to deviate less from the correct procedure in comparison to an engineer who performed sprinkler maintenance using the current industry method, without AR instruction. Errors made by engineers on the procedure, together with semi-structured interview responses, shed light on customer pain points that AR can alleviate, useful UX/UI design considerations, barriers to adoption and insights for informing larger scale user evaluations of industry AR from maintenance to manufacturing.

CCS CONCEPTS

• Field studies • User studies

KEYWORDS

Augmented Reality, providing instructions, maintenance, head-mounted displays, ethnography, qualitative methods

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.

PETRA'19, June5–7, 2019, Rhodes, Greece
© 2019 Copyright is held by the owner/author(s). Publication rights licensed to ACM.
ACM ISBN 978-1-4503-6232-0/19/06…$15.00
https://doi.org/10.1145/3316782.3322752

1 INTRODUCTION

Industrial applications of Augmented Reality (AR) for the purpose of providing work instructions to guide employees in performing industrial procedures are becoming increasingly common. In 2011 Fite-Georgel [6] acknowledged that, despite AR's technical advancements real products using AR were not being applied in industry. There have been advancements since then in getting products using AR into industrial settings but Kim et al. [13] noted that there is now a need for evaluations of such commercial systems for delivering AR in real-world settings. Evaluations of AR's effectiveness compared to current non-AR methods of work instruction are required to demonstrate the worth of AR to industry and identify the current barriers to its adoption. In this paper we present an ethnographic user experience study of an HMD delivered AR application for supporting service engineers in performing a fifty-nine-step maintenance procedure on a hospital sprinkler system – a procedure that they currently perform as part of their workflow.

2 RELATED WORK

2.1 User studies of AR for Work Instruction

AR technologies for delivering Industrial work-instructions [6] are currently available in the form-factor of head mounted displays (HMDs) (e.g., custom built prototypes [10, 11], commercially available HMDs such as the Microsoft Hololens and DAQRI Smart Glasses®), in custom applications for tablet devices [5, 18, 20] and *in situ* projection [7, 8, 9]. Indeed, many user studies have been performed to evaluate the benefits of AR for work instruction delivery using these different AR technologies, compared to currently-used non-AR methods of work instruction delivery (usually instruction manuals) [2, 3, 7, 8, 9, 10, 11, 12, 15, 16, 17, 18, 19, 20, 21, 22, 23]. A large proportion of these user studies have demonstrated that users perform procedural tasks more accurately and more quickly when guided by AR, compared to paper- or electronic-document-based task guidance using a tablet. Though these studies are suggestive about the benefits of AR for industrial procedures, their real world value is limited because most of the procedural tasks used are not "real" industrial procedures but proxies for them; for instance, studies have used Lego assembly tasks [12, 21] or aircraft-wing assembly using wooden components in place of real aircraft wing parts [20]. It is difficult to see the industry sector being convinced to adopt such technologies based on the "toy" nature of these tasks.

A few highly cited studies have examined the benefits of using AR HMDs for guiding users in performing "real" industry procedures, for instance, in the assembly and disassembly of an engine combustion chamber [10] and in maintenance of an armoured personnel carrier turret [9]. However, the HMDs used in these studies were custom prototypes not commercially available products.

Two recent user studies have evaluated commercially available technologies for AR work instruction on real industry tasks with real industry employees [7, 23]. In-situ projection systems for delivering AR have been deployed in the workplace and are commercially available [7]. Funk et al. [7] performed a user study with an in-situ projection system on an industrial assembly task. The authors found that untrained workers benefitted from AR, assembling products faster and without making errors following three days of learning the task. However, the benefits did not persist after the initial learning phase, and expert workers were slowed down by the AR instructions. Commercially available and portable AR HMDs provide an opportunity to deliver AR at scale across different industry sites. Werrlich et al. [23] evaluated an AR application for an engine assembly task delivered to trainee automotive technicians using the Microsoft Hololens AR HMD. Trainees made fewer errors when using the AR application compared to a paper instruction manual, but performed the task more slowly using AR.

2.2 Ethnography in HCI and the present study

In this paper we conducted a rapid ethnographic study of AR on an industry task with real industry users (engineers). Ethnographic research methods are commonly used in the field of human computer interaction [4, 14]. Typically, ethnographic research involves field work done in natural settings with a focus on understanding activities from the user's perspective and it allows for a richer understanding of the context of an activity than would be possible in a controlled laboratory study [14]. Another benefit of ethnographic methods is that they may provide ways to discover user requirements that are difficult for users to articulate themselves [14]. Ethnographic research is commonly used to provide requirements for the future design of interactive systems but Dourish [4] has argued it should go beyond this to provide models for thinking about the settings in which these systems are used and shaping future research or corporate strategy [4]. Ethnographic methods include participant observation, activity walkthroughs and semi-structured interviews that can focus on specific benefits, bottlenecks or problems [14]. There is recognition that traditional ethnography can be too demanding in terms of the time taken to collect data. Rapid ethnography has been proposed as an alternative where fieldworkers conduct short focused studies to quickly gain an understanding of the work setting [14].

In this paper we present a rapid ethnographic study of AR on an industry task that includes the context of the industry procedure and current pain-points together with a user experience study of an HMD-delivered AR application for supporting real industry employees in their performance of a fifty-nine-step maintenance procedure on a hospital sprinkler system. We combined three of the important facets of the user studies conducted by Funk et al. [7] and Werrlich et al. [23] reviewed in section 2.1, that is, the inclusion of real employees with different levels of experience with the industry procedure, a commercially available AR HMD and an AR application designed specifically to alleviate customer pain points with current methods [24] that were identified from initial discussions with management.

Discussions with management informed the development of quantitative measures for evaluating performance on the industry procedure and quantitative data on these measures was collected from users performing the entire industry procedure of interest. Reasons for users deviating from the correct way of performing the industry procedure were explored using semi-structured interviews to provide qualitative insights on how AR was supporting or hindering users in performing the industry procedure correctly. We conclude by discussing how the findings from this study can inform the design and evaluation of future AR applications in industry, from maintenance to manufacturing.

3 INDUSTRY PROCEDURE: SPRINKLER MAINTENANCE

Information on the industry procedure presented in this paper was obtained from initial interviews conducted by the third and fourth authors, employees of the construction company Royal BAM Group. They interviewed both an expert on the procedure and the facility manager responsible for overseeing that it was performed correctly in order to gain a deep understanding of the procedure (the Sprinkler Test) and the context for performing it.

In the Netherlands, the law requires that all public buildings fire suppression sprinkler systems be tested every two weeks, by conducting the Sprinkler Test in the Sprinkler Room. Royal BAM group are responsible for the management of the Sprinkler Room of a particular hospital near Amsterdam (in Zaandam). The Sprinkler Room contains an installation that will pump water from a water basin into the Sprinkler pipe system in the hospital when the pressure in the system drops (Figure 1). Such a pressure drop is very likely the result of a fire, causing the sprinkler head to discharge. To make sure that this system works under all circumstances, a number of checks and balances have been installed in the room that will notify the facility manager through the facility management system when the pump is compromised. By law, the facility manager has to verify every two weeks that these checks and balances are still working properly through the Sprinkler Test procedure.

3.1 Current Methods

The Sprinkler Test procedure is conducted by Royal BAM group service engineers. Service engineers perform the fifty-nine-step procedure entirely from their memory of being shown the procedure by a colleague who has performed it before. There is no instruction manual for the procedure. Although the procedure is

Figure 1: Sprinkler Room with Sprinkler Installation.

performed every two weeks, different service engineers perform the Sprinkler Test, meaning that there is often a long gap in time between a particular service engineer performing the Sprinkler Test procedure. Different service engineers also have different levels of experience in performing the task and have been taught the procedure by different colleagues. Each Sprinkler Test is currently performed by two service engineers who check each other's work. Below we summarize the main steps of the Sprinkler Test:

- **Step 1:** The Sprinkler Test begins with the service engineer setting the Siemens Sprinkler Management system to "test mode".
- **Step 2:** The service engineer walks down to the Sprinkler Room.
- **Step 3:** The service engineer then begins a fifty-nine-step procedure that involves pressing and turning switches on various control boxes and rotating multiple valves on the Sprinkler Installation in a specific direction (Figure 2, left).
- **Step 4:** The service engineer must inspect the control panel on the sprinkler management system (Figure 2, right) to ensure the correct configuration of indicators are lit in response to performing the actions listed in Step 3 (e.g. pressing and turning switches, rotating valves).

2.2 Pain Points in Current Methods

Werrlich et al. [24] recommend that AR support systems alleviate customer pains. To this end, we identify pain points in the current methods proposed by the facility manager and expert in initial interviews. As described earlier, current methods require the service engineer to perform the Sprinkler Test from memory. Given that fifty-nine steps are required, there is significant risk of mistakes being made. Hospital alarms have been set off as a result of such mistakes. Current methods don't require the steps to be performed in a set sequence, resulting in different service engineers completing the steps in different sequences. This inconsistency can have an impact on new service engineers learning the

Figure 2: Left: a section of valves on the sprinkler installation. Right: control panel with indicator lights lit green.

Sprinkler Test procedure, specifically, they may find it harder to learn the procedure if they are trained by multiple different expert service engineers. Current methods lack any documentation detailing the steps to be performed and the order in which to perform them, so that steps may regularly be missed, mistakes made unbeknownst to facility management, and incorrect procedures passed down to new service engineer trainees.

2 INDUSTRY-READY AR HMD

The DSG is a head mounted display (HMD) for delivering high-resolution 3D AR content. The field of view is 44 degrees per eye. The DSG weighs 335g and a powerpack (compute pack) weighing 496 g is connected to the HMD via a cable. The compute pack houses the battery which the user wears on their person, attaching it to their clothing using a clip. The DSG is certified eye protection and is designed for both indoor and outdoor use. Full specifications for the DSG can be found on the DAQRI website (https://daqri.com/products/smart-glasses/). The DSG has a hands-free user interface with a cursor termed "the reticle", controlled by the user moving their head. The reticle allows the user to navigate through the interface, "dwelling" on components for a few seconds to select them. This method of interface control is known as "gaze and dwell". The AR app in this user study used marker-based tracking where the AR content in the sprinkler room is loaded and tracked by the user scanning a physical marker with the AR HMD. The marker was positioned in a pre-determined location in the sprinkler room.

2.2 AR application

Two employees from Royal BAM group, the 3rd and 4th authors, worked closely with DAQRI to design the AR application for Sprinkler maintenance in Unity 3D. This application consisted of:

- Work instructions for all fifty-nine steps of the Sprinkler Test procedure delivered in a fixed, linear flow to app users
- Directional arrows to guide app users to the correct position in the Sprinkler Room to perform each step
- AR annotations overlaid onto their physical counterparts in 3D space dynamically showing the correct direction to turn valves/switches (see figure 3, top right)
- Photographic displays showing the correct configuration of indicator lights that should be lit at each step on the

control panel (figure 2, right) or other panels (such as figure 3, bottom right)

- AR displays of written instructions for each step above the physical components (i.e., valves/switches) that app users were required to interact with (such as in figure 3, bottom right)
- Functionality that enabled app users to document the completion of each step

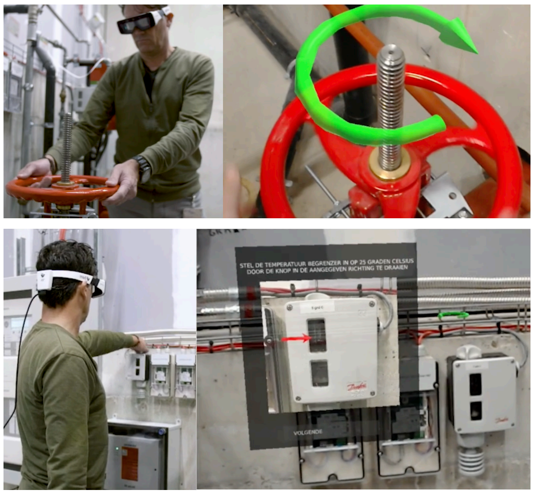

Figure 3: Top panel (right) displays dynamic AR annotations seen through the DSG. Bottom panel (right) displays the required temperature setting on a temperature sensor & dynamic green AR annotation above the button on the far right, © BAM Netherlands [1]

3. USER EXPERIENCE STUDY

A user experience study was designed and conducted to provide quantitative and qualitative insights on how AR was supporting or hindering users in performing the industry procedure correctly. Since our focus was on users experience of the AR HMD app for supporting task performance, the majority of engineers (*N*=6) performed the procedure while wearing the optical see through AR HMD (AR-instruction). One engineer completed the procedure using the current industry method (performing the fifty-nine-step procedure from memory) to provide a baseline against which AR instruction could be compared. We then conducted semi-structured interviews with all users. Key themes were noted along with relevant quotations.

3.1 Participants

Seven engineers from Royal BAM group participated in the user study. Four of these were service engineers who currently perform the Sprinkler maintenance procedure in their workflow. Two of these service engineers were classed as "somewhat experienced" having completed the sprinkler test 1 or 2 times prior to the user study, with the previous time they performed it being 2 to 6 months ago. Two were classed as "experienced" having completed the sprinkler test 10 to 20 times previously and having performed it 2 weeks to 1 month ago and the final service engineer classed as an "expert" having completed the sprinkler test thousands of times previously and last performed it the day before the study. The two remaining engineers had never previously performed the sprinkler test and were classed as "novices". User profiles showing the age ranges, experience level of engineers and whether they completed the "dry" or "live" test using the AR app or current method (see section 3.2. below) are shown in table 1. All users gave informed consent to participate in the research and permission for DAQRI to use images from the user study in external publications.

3.2 Design

Six engineers completed the sprinkler test wearing the AR HMD and following the work instructions on the AR app. This group, termed the AR-instruction group, included the "expert", two "somewhat experienced" service engineers, one "experienced" service engineer and all "novices". One "experienced" service engineer completed the sprinkler Test using the current method (i.e. performing it from memory without wearing the AR HMD). This provided a baseline against which to evaluate the effectiveness of the AR app for sprinkler maintenance.

There were two variants of the sprinkler test procedure. Both variants included the fifty-nine steps but in one termed the "dry test" users followed the sprinkler test procedure and indicated the actions they should perform in each step but *did not* actually physically perform actions (e.g., rotating valves, turning/pressing switches) on the Sprinkler installation. This differed from the "live test" where users did physically perform actions on the Sprinkler installation. The reason for the two variants of the Sprinkler test was because Royal BAM group only had permission from the hospital to perform the "live test" on one day. We included the "dry-test" so as to include more users in the user study. Two users took part in the "dry-test" on day 1; the "experienced" service engineer and one "novice" engineer, who both completed it wearing the DAQRI Smart Glasses® and using the AR app. Five users took part in the "live-test" on day 2; one "novice", two "somewhat experienced", one "experienced" service engineer and the "expert". All participants except for the second "experienced" service engineer completed the procedure wearing the AR HMD and using the AR app. The second "experienced" (baseline) service engineer completed the procedure using the current method.

Table 1: User study participant profiles

Experience Level	Novice	Somewhat experienced	Experienced	Expert
Number of participants	2	2	2	1
Ages	28, 45	47, 51	55, 46	60
Number of times performed sprinkler test	Never	1, 2	20, 10	1, 2
When last performed sprinkler test	N/A	6 months, 2 months ago	2 weeks, 1 month ago	Yesterday
Dry or Live test?	Dry, Live	Live, Live	Dry, Live	Live
Instruction method (AR or current)	AR	AR, AR	AR, Current	AR

3.3 Procedure

All users completed the Sprinkler Test procedure on the real Sprinkler Installation housed in the Sprinkler Room of the hospital in Zaandam (Netherlands) on the 24th and 25th October 2017. There were three individuals present on-site to oversee the study and record data from all users. One of these was the first author, the other was the third author who was involved in developing the AR app and the final individual was the expert user who in addition to participating in the study themselves was also present to tell the 1st author if a user deviated on any step from the correct procedure. Indeed, it was the expert who defined the 'correct' procedure. The third author acted as a translator to translate between the expert who spoke Dutch and the first author who only spoke English.

All participants received instruction on how to use the AR HMD and AR App, with the exception of the "baseline" user, who completed the Sprinkler Test procedure without wearing the DAQRI Smart Glasses®. Specifically, training consisted of the third author, narrating a video that showed the user scanning a marker to open the AR app, navigating the menu, and interacting with text-based and graphical AR content. The video also guided users through use of the gaze-and-dwell function for moving between work instruction steps. Users watched this video on a tablet. Prior to beginning the Sprinkler Test procedure, users were instructed to complete the procedure as accurately as they could at a normal pace. Users then completed the fifty-nine step Sprinkler maintenance procedure wearing either the AR HMD and guided by the AR app or in the case of the baseline user, performed the fifty-nine steps from memory. All users were video-recorded in order to double check the quantitative measures were recorded accurately and to enable accurate recording of task completion time. Following completion of the Sprinkler Test, all users, completed semi-structured interviews (see sections 3.4 for details).

3.3 Quantitative Measures of Task Performance

Data on the following quantitative measures of task performance was collected based on interviews that the co-authors from Royal BAM group had conducted with the facility manager for the Sprinkler Installation based on key performance indicators (KPIs) important to them:

Accuracy is defined as the correct procedure being performed at each step. In total there are fifty-nine steps, so the maximum possible accuracy score is fifty-nine for each participant. Accuracy was only calculated based on steps that participants performed. Steps that participants missed were captured by "task completion" (below).

Task completion is defined as the lack of missed steps. A missed step is defined as any one of the fifty-nine steps that a participant did not perform, until prompted by the expert.

Consistency is defined as the extent to which the performed procedure followed the correct linear sequence that was defined in consultation with experts in Sprinkler maintenance, and which was previously captured by Royal BAM group and programmed as the linear sequence of steps in the AR app. All steps that a user performed out of sync with the correct procedure were summed to provide a score capturing a lack of consistency.

Time taken is defined as the time to complete all fifty-nine steps of the Sprinkler Test Procedure

3.4 Semi-structured interviews

Semi-structured interviews were conducted where users were asked what they thought the main benefit of the app to their job was and was there anything that the app could better address. Any other comments users made were also recorded and the interviewer (the first author) probed answers to zero in on specific benefits or problems conferred by use of the AR app. Responses to these questions and other comments freely offered by users were recorded as quotes.

4. USER STUDY RESULTS

4.1 Accuracy across fifty-nine step procedure

All users were able to perform the Sprinkler Test procedure with a high level of accuracy using the AR application (all >90%) (see table 2). The level of accuracy using the AR app was very similar to that achieved in the baseline live test, where an experienced user performed the Sprinkler Test using the current method. Accuracy was similar for both live and dry tests. Using the AR app, the two novices were able to perform the Sprinkler Test procedure with greater than 90% accuracy.

4.2 Time taken to complete fifty-nine step procedure

All users were able to complete the Sprinkler Test procedure in approximately 30 minutes. The experienced user who used the current method completed it faster (10 minutes) than the experienced (31 minutes) or expert (23 minutes) users who used the AR app. Novices completed Sprinkler Maintenance in a similar time to the expert and experienced users who used the AR app. User ID 3 was unable to complete the entire procedure due to an interruption caused by a hardware issue with the AR HMD, namely the power cord disconnecting mid-procedure. As such that user's completion time could not be recorded and is listed as "n/a" in table 2.

Table 2: Percentage accuracy of participants across 59 step procedure

ID	Participant	% Accuracy (across all 59 steps)	Time to complete procedure (minutes)
1	Novice (AR, Dry test)	98	30
2	Novice (AR, Live test)	91	30
3	Somewhat experienced (AR, Live test)	97	n/a
4	Somewhat experienced (AR, Live test)	95	25
5	Experienced (AR, Dry test)	98	31
6	Experienced (Current Method, Live test)	100	10
7	Expert (AR, Live test)	98	23

Table 3: Number of Task completion and Consistency errors across participants. Measures for the baseline participant (using the current, non-AR method) are shown in bold

ID	Participant	Task Completion (steps missed)	Consistency (steps out of sequence)
1	Novice (AR, Dry test)	0	0
2	Novice (AR, Live test)	1	0
3	Somewhat experienced (AR, Live test)	0	0
4	Somewhat experienced (AR, Live test)	0	0
5	Experienced (AR, Dry test)	0	0
6	**Experienced (Current Method, Live test)**	**5**	**3**
7	Expert (AR, Live test)	0	1

4.3 Task completion (i.e. lack of missed steps)

Task completion errors in the form of number of missed steps (out of a total of 59 steps) are shown in table 3. Out of the six users who used the AR app to complete the Sprinkler Maintenance procedure, only one novice missed one step (miss rate of 0.3% across users). In contrast, the experienced user who performed the Sprinkler Maintenance procedure using the current, non-AR, method missed five steps (miss rate of 6%). The AR app thus resulted in a 20 times lower miss rate versus the current method.

Somewhat experienced, experienced and expert users demonstrated 100% task completion using AR, as not one of them missed any of the fifty-nine steps.

The expert who observed the performance of all other users commented during the semi-structured interview that a step of critical importance was missed by the experienced engineer (ID 6) who followed the current (i.e. non-AR) method. The step of turning the flow switch was forgotten because the experienced user said that *"a different procedure was taught to them in training, specifically that turning a sprinkler head (another valve on the installation) is equal to turning the flow switch".* The expert however disagreed stating that *"turning the flow switch is however an important step in the procedure because it is necessary for giving the hospital information on the location of the alarm".*

4.4 Consistency (i.e. steps performed out of linear sequence)

Consistency errors in the form of number of steps performed out of correct linear sequence (out of a total of 59 steps) are shown in table 3. Out of the seven participants who used the AR app, only one, the expert, performed a step out of sequence because they were trying to go through the steps too quickly (see section 4.6). All other participants correctly followed a consistent linear sequence of steps that were presented in the AR app. In contrast, the experienced user who performed the Sprinkler Maintenance procedure using the current method performed three steps out of sequence.

4.5 Types of Errors Made

When users made errors, they were asked why they had made them. One novice user and one somewhat experienced user made errors on the step shown in the left panel of figure 4. This step required them to 'rotate the switch in the direction indicated', with a dynamic green arrow showing the correct direction of rotation. However, as figure 4 shows, the arrows were not tracked correctly to align with the correct switch. The arrow should be located above the black switch but instead it was located to the left and above it. The users indicated that this made them unsure what switch to turn and as such they made the error of turning the grey switch instead. A similar error was made on the step shown in the right panel of figure 4 where the user should rotate the top valve in the direction indicated but again the green arrow was not tracked accurately to overlay the top valve. One experienced user and one novice user made errors on this step and indicated that the tracking problem was the reason. The steps shown in figure 4 show multiple switches or valves in close proximity to one another and this may be why tracking problems contributed to errors here more than elsewhere in the installation where components were spaced further apart.

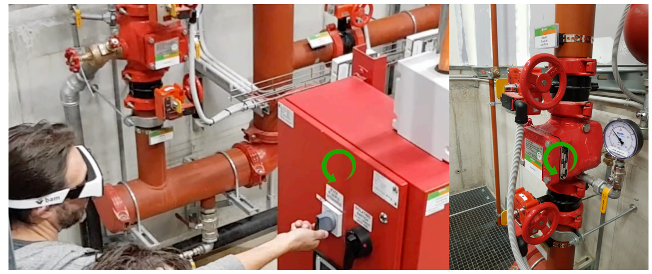

Figure 4: Displays errors made due to tracking problems at two different steps (left & right panels) in the Sprinkler Test procedure. Note that the arrows are only to illustrate what was seen through the AR HMD, they are not from HMD point-of-view recordings.

4.6 Semi-structured interview responses

We found that user's responses in semi-structured interviews could be summarized into nine themes. These themes are accuracy & independence, learning, confidence, tracking problems, perceptual difficulties in AR, positive experience with AR app annotations, user interface improvements, a hardware issue and the expert being slowed down by AR. These themes are listed and discussed below with participant quotes for illustration.

(1) AR instructions allowed users to be more accurate, independent and avoid erroneously missing task steps

One experienced user answered in response to the question "what is the main benefit of the app to my job" that the AR app *"will ensure I do all the steps. You have to do the procedure by protocol. If you do it another way it is easy to miss some steps."* Another answered in response to the same question that *"using the AR app will be quicker because you can do it on your own. You sometimes need two engineers to help each other through the procedure."*

(2) AR instructions helped users to learn the sprinkler system

One somewhat experienced user answered in response to the question "what is the main benefit of the app to my job" that you can *"Learn from interacting with the AR app how the installation works. The AR app makes it easier to work with the installation components."* One novice user commented that *"for someone who's never done it before its just easy to go through"* and the other said in response to the question "what is the main benefit of the app to my job" that *"You can perform sprinkler maintenance as a novice, you don't have to be an expert."*. The expert concurred also saying he thought the main benefit of the app was *"Novices learning the procedure for the first time in a new environment (i.e., an installation they've never interacted with before)"*.

(3) AR instructions build confidence

One novice said that the main benefit of the app to their job was *"you know you are doing the right operations on the installation".*

(4) Tracking problems

All users of all experience levels reported tracking problems when asked what the app could better address. One experienced user stated that *"AR overlay annotations are not well aligned. If the AR overlay annotations fall between two valves, you question yourself on which one to rotate"*. This point clearly relates to the types of errors made shown in figure 4. Another user mentioned the problems with tracking the marker which often had to be intermittently rescanned a few times during the study to improve tracking. They elaborated on the tracking problem saying, "*The AR instructions are sometimes running away from you"*. One user suggested a solution to this problem would be *"to have certain valves that are in close proximity to one another numbered so that one can more easily differentiate between them"*.

(5) Perceptual difficulties in AR

One experienced user commented on the text instructions in AR saying, *"that I also have difficulty reading and I'm not used to reading while doing a task like this"*. Similarly, a novice user commented, *"The text is sometimes a bit unclear, it is hard to focus on text that is moving around"*. The text was moving due to the tracking being imperfect. One participant said the photographic images could be clearer and one experienced user said that on the step, shown in the bottom panel of figure 3, *"on the temperature dial it was hard to read the temperature value through the Smart Glasses"* indicating a reduction in real-world vision while wearing the AR HMD.

(6) Positive Experience with AR app annotations

One experienced user was positive about the green directional arrows shown in AR in the step shown in figure 3 (bottom right panel) stating that *"they are useful to show which direction you should turn the button on the temperature dial"*. Another stated more generally that using the AR app brought them satisfaction saying, *"I really like it, it brings an element of joy to performing the task"*. Two users commented positively on the futuristic aspect of the technology with one saying, *"It is a good system, I think it is the future"*.

(7) User Interface improvements

The expert suggested that the app could be improved by the addition of a back button. The expert made an error because they were trying to go through the steps too quickly and hit the next button too many times. An experienced user similarly commented that it would be better *"to go to the next step more easily, dwelling to select it is tricky"*.

(8) Hardware issue

For one participant, ID 3, the cable connecting the AR HMD to the powerpack got pulled out and they commented that *"The cable is problematic as it can be pulled out as happened to me"*.

(9) Expert was slowed down by AR

The expert said they were slowed down using the app giving the reason *"I'm already an expert I know what to do"*.

5. DISCUSSION

This paper presents ethnographic research of Augmented Reality (AR) in industry that includes the context for the industry procedure, pain-points with current methods and a user experience study of engineers using an HMD-delivered AR app to guide them through the industry procedure in a natural setting; a hospital sprinkler installation. Overall, users who wore the HMD and followed AR instruction were found to deviate less from the correct procedure in comparison to the baseline user who performed sprinkler maintenance using the current industry method, that is performing the procedure from memory. Errors made by users on the procedure together with subsequent semi-structured interview responses shed light on customer pain points that AR can alleviate, useful UX/UI design considerations, current barriers to adoption and insights informing future larger scale user evaluations of industry AR. We now discuss how findings relate to each of these in turn.

5.1. Customer Pain-points alleviated by AR

AR instruction increased performance on two out of four key performance indicators (KPI) identified by sprinkler facility management. These were 'consistency' where all engineers using AR instruction completed all steps of the fifty-nine-step procedure in the same correct linear sequence, with the exception of one step performed out of sequence by the expert, and 'task completion' where only one out of six engineers using AR instruction missed a single task step. This contrasts with an experienced engineer who, using the current industry method, performed three task steps out of sequence and missed five task steps. A third KPI 'accuracy' was similarly high for both AR and the current method, with greater than ninety-percent accuracy evidenced by engineers across both. The only KPI on which performance was reduced by AR instruction relative to the current method was task completion time, which was slowed by AR.

Errors made by the experienced engineer using the current, non-AR method illustrate two pain-points identified by facility management that appear to be alleviated by AR. A first pain-point identified problems with consistency and missed steps in the procedure which was borne out by the experienced engineer performing the procedure in a different sequence from the standard linear sequence of steps *and* missing steps using the current method. The interview response of another experienced engineer, that you have to do the procedure by "protocol" to avoid missing steps, supports the link observed between inconsistency and missing steps in the baseline user performing the current method. A second pain-point identified incorrect procedures passed down to trainees. The experienced user missed a crucial step, turning a flow switch, because they said that an incorrect procedure was passed down to them during training; namely that turning the valve they turned would do the same job as the flow switch valve they missed. The expert engineer who knew the procedure more deeply however pointed out the importance of turning the flow switch for letting the hospital know the location of the alarm. The AR instructions that were carefully designed with guidance from the deep expert and given to all new trainees would alleviate this pain-point of a lack of precision in training.

5.2. Informing future larger scale evaluations of industry AR

Examining the types of errors made using the current method that were alleviated by AR would appear of value more broadly to inform larger scale user evaluations of industry AR. Consistency may be a less obvious metric to evaluate performance than accuracy or task completion time but, as was the case in this study, it could represent a valuable KPI on which AR can evidence benefits relative to non-AR methods of work instruction. The potential of AR to confer greater precision in the transfer of procedural knowledge to new trainees could be evaluated in future longitudinal studies, such as that previously conducted by Funk et al. [7], that capture the transfer of knowledge between employees over time, comparing precision of knowledge transfer using AR to precision of knowledge transfer using face to face instruction or manuals.

Differences in the benefits conferred by users of different experience levels are similar to those found by Funk et al. [7] and Werrlich et al. [23], namely that the expert user did not benefit from AR and was slowed down by using it. However, findings suggest a nuance to this issue not noted previously in prior work, namely that an experienced user who had performed the procedure twenty times in the past stated that they would benefit from using the AR app and enjoyed using it. Our finding that the other experienced user missed steps and wasn't consistent when performing the procedure using the current method suggests that there is a need to support more experienced users on this procedure and thus, potentially in industry procedures more generally, value in AR support for more experienced users. These findings suggest that future empirical work examining the benefits of AR shouldn't exclude experienced users based on prior findings that they did not benefit from AR [7, 23], it should instead differentiate experienced users from deep experts who know the procedure 'inside-out' and are not likely not to benefit from AR support.

5.3. UX/UI design considerations for industry AR

The types of errors made by users who used AR and their interview responses are valuable more broadly for providing user-experience (UX) and user-interface (UI) design considerations for future industry AR apps. Engineers were observed in the user study making errors due to problems with tracking that resulted in AR annotations being incorrectly aligned with the components (e.g., switches, valves) on which they indicated the correct actions to be performed. An experienced user commented that this made them uncertain of the correct actions to perform and indeed the common errors when tracking problems were reported by users as the cause were where switches and valves were in close proximity to one another. A UI improvement to label components in close proximity was suggested by one of the users. This problem and the proposed solution should be readily transferable to other industry AR apps and could be summarized as a way of designing around imperfect tracking. Tracking issues together with the need for periodic rescanning of the AR marker were identified as the main barriers to adoption of AR for this procedure by Royal BAM group at the time this study was conducted.

Perceptual difficulties that users had using AR are also instructive for future UI design. Difficulties reading text particularly if it is moving in AR suggest the importance of minimizing text content in future AR apps or at least locking text in world space, in other words so virtual AR content remains fixed to a real-world location and doesn't move when the user moves their head. The comment from one user that they couldn't read a temperature dial while wearing the AR HMD represents a recognized downside of superimposed AR imagery, in that it can mask information in the real world, referred to as "overlay clutter" [25].

The dynamic green AR arrows were viewed positively by users and could be a UI component included in a wide range of industry AR apps where users have to be directed around a workspace and instructed on which components to interact with, particularly if these involve actions like turning valves or switches.

Interaction difficulties where a user commented on the trickiness of using the gaze and dwell AR app interface could be tested in longitudinal studies. It is the first author's intuition that this difficulty would decrease as a user becomes more familiar with the AR interface that they were only using for the first time in this study, but this needs to be tested empirically.

5.3. Limitations

Contributions from the work reported here should be considered in the context of several limitations. Firstly, the sample size of users in the study was small although this did include all the engineers who perform sprinkler maintenance in the sprinkler room of the hospital in which the study was conducted. Evidently the quantitative differences in user performance between AR instruction and the current method are only suggestive of the benefits of AR. Only a large-scale user study with many more users performing the procedure using AR instruction and the current method from memory (i.e., baseline) would be able to establish if the suggested benefits of AR on this industry procedure are statistically robust. Further, it would be valuable to compare AR instruction to the same instructions delivered in another medium, such as on paper or digital non-AR instruction on a tablet to provide a more precise test of the specific benefits of AR in this industry context.

However, the purpose of the quantitative measures in this study was not to establish robust benefits of AR over current methods of work instruction, or over other instruction mediums, but instead to combine quantitative with qualitative data to better understand the benefits and limitations of AR from the user's perspective. For instance, recording missed steps in the procedure resulted in the researchers questioning the user who missed these steps (i.e. the experienced user who used the current method) on why they did so which led to the important insight about a lack of precision in training mentioned in section 5.1, a pain-point that AR instruction appears ideally suited to alleviating. This may have been difficult for the experienced user to articulate themselves given they were not aware, prior to the missed steps being pointed out by the expert, that they were not performing the procedure correctly, since they were performing it as they were trained to do. This is a benefit of ethnographic methods, that they may provide ways to discover user requirements (i.e. the need for more precise training) that users would otherwise have difficulty articulating themselves [14].

A second limitation is that the present work used rapid ethnography to quickly gain an understanding of the work setting and the benefits and limitations of AR for the industry procedure [14]. The hospital only permitted the sprinkler room to be used for the user study for two days so a full ethnographic study of AR over time was not possible in this context. A longer ethnographic study examining AR use over time and obtaining interviews from facility management would be beneficial for more deeply understanding issues around adoption of AR in this context.

5.4. Conclusion

This paper presents a different approach to typical user experience and usability studies of Augmented Reality, by reporting an ethnographic study of commercially available AR in a sample of real users using it within their workflow in industry. Findings reveal real customer pain-points in industry alleviated by AR, key performance indicators and user profiles that could be incorporated in future controlled large-scale laboratory evaluations of industry AR, AR UX/UI design considerations and current barriers to AR adoption. The aims of the authors are that these findings will prove useful in informing and stimulating the design of future AR systems and user evaluations of AR across a range of industries, from maintenance to manufacturing.

ACKNOWLEDGMENTS

The authors wish to thank DAQRI, Royal BAM group and the engineers who participated in the ethnographic study. The work of the first author was supported by a Science Foundation Ireland (SFI) Industry Fellowship, Grant #16/IFA/4331.

REFERENCES

[1] Bambouwentechniek.nl. 2017. Themas, BIM. (November 2017). Retrieved March 29, 2019 from http://www.bambouwentechniek.nl/themas/bim
[2] Sebastian Büttner, Markus Funk, Oliver Sand and Carsten Röcker. 2016. Using Head-Mounted Displays and In-Situ Projection for Assistive Systems: A comparison. *Proc. PETRA '16*. 1-8.
[3] Thomas P Caudell and David W Mizell. 1992. Augmented reality: An application of heads-up display technology to manual manufacturing processes. *Proc HICSS'92, Vol.2., IEEE*: 659–669.
[4] Paul Dourish. 2006. Implications for Design. In *Proc.* CHI'06. 541-550.
[5] Nirrit Gavish, Teresa Gutierrez, Sabine Webel, Jorge Rodrıguez, Matteo Peveri, Uli Bockholt, and Franco Tecchia. 2015. Evaluating Virtual reality and Augmented Reality Training for Industrial Maintenance and Assembly tasks. *Interactive Learning Environments*, 23(6): 778–798.
[6] Pierre Fite-Georgel. 2011. Is there a reality in Industrial Augmented Reality. In *Proc. ISMAR '11*. IEEE. 201-210.
[7] Markus Funk, Andreas Bächler, Liane Bächler, Thomas Kosch, Thomas Heidenreich, and Albrecht Schmidt. 2017. Working with Augmented Reality? A Long-Term Analysis of In-Situ Instructions at the Assembly Workplace. *Proc. PETRA '17*. 222-229.
[8] Markus Funk, Thomas Kosch, Scott W. Greenwald, and Albrecht Schmidt. 2015. A benchmark for interactive augmented reality instructions for assembly tasks. *Proc. MUM '15*. 253–257.
[9] Markus Funk, Thomas Kosch, and Albrecht Schmidt. 2016. Interactive worker assistance: comparing the effects of in-situ projection, head-mounted displays, tablet, and paper instructions. *Proc PETRA '17*. 222-229.
[10] Steven J. Henderson and Steven K. Feiner. 2009. Evaluating the benefits of augmented reality for task localization in maintenance of an armored personnel carrier turret. *Proc*. ISMAR '09, IEEE. 135–144.
[11] Steven J. Henderson and Steven.K. Feiner. 2011. Augmented reality in the psychomotor phase of a procedural task. *Proc*. ISMAR '11, IEEE. 191–200.
[12] Lei Hou, Xiangyu Wang, Leonhard Bernold, and Peter.E.D. Love. Using Animated Augmented Reality to Cognitively Guide Assembly. 2013. *Journal of Computing in Civil Engineering*. 27 (5): 439-451.
[13] Kangsoo Kim, Mark Billinghurst, Gerd Bruder, Henry Been-Lirn Duh, Gregory F. Welch. 2018. Revisiting Trends in Augmented Reality Research: A Review of the Second Decade of ISMAR (2008-2017). *IEEE Transactions on Visualization and Computer Graphics*. 24 (11): 2947-2962.
[14] David R Millen. 2006. Rapid Ethnography: Time Deepening Strategies for HCI Field Research. In *Proc DIS'00*. 280-286.
[15] Michael R. Marner, Andrew. Irlitti, and Bruce H. Thomas. 2013. Improving Procedural Task Performance with Augmented Reality Annotations. *Proc. ISMAR '13*. 39–48.
[16] Miwa Nakanishi, Mugihiko Ozeki, Toshiya Akasaka and Yusaku Okada. 2007. Human factor requirements for Applying Augmented reality to manuals in actual work situations. *Proc.* ISIC '07. *IEEE.* 2650–2655.
[17] Nattapol Pathomaree and Siam Charoenseang. Augmented reality for skill transfer in assembly task. *ROMAN '05, IEEE International Workshop on Robot and Human Interactive Communication.* 1–25.
[18] Jarkko Polvi, Takafumi Taketomi, Atsunori Moteki, Toshiyuki Yoshitake, Goshiro Yamamoto, Christian Sandor and Hirokazu Kato. Handheld Guides in Inspection Tasks: Augmented Reality versus Picture. *IEEE Transactions on Visualization and Computer Graphics.* 24 (7): 2118-2128.
[19] Andrew Pringle, Abraham Campbell, Stefanie Hutka, Alberto Torrasso, Colin Couper, Fabian Strunden, Jan Banjana, Kamil Jasztrab, Ralph Croly, Rob Quigley, Ross McKiernan, Paul Sweeney and Mark T Keane. 2018. Using an Industry Ready AR HMD on a Real Industry Maintenance Task: AR benefits performance on certain task steps more than others. *Adjunct Proc ISMAR'18. IEEE* (in press).
[20] Trevor Richardson, Stephen A. Gilbert, Joseph. Holub, F. Christian Thompson, Anastacia MacAllister, Rafael Radkowski, Eliot Winer, Paul Davies and Scott Terry. 2014. Fusing self-reported and sensor data from mixed-reality training. *Proc.* I/ITSEC '14. 1–12.
[21] Arthur. Tang, Charles Owen, Frank Biocca, and Weimin Mou. 2003. Comparative effectiveness of augmented reality in object assembly. In *Proc.CHI '03.* 73–80.
[22] Antonio E. Uva, Michele Gatullo, Vito M. Manghisi, Daniele Spagnulo, Giuseppe L. Cascella, Michele Fiorentino. 2018. Evaluating the effectiveness of spatial augmented reality in smart manufacturing: a solution for manual working stations. 94 (1–4): 509–52.
[23] Stefan Werrlich, Austino Daniel, Alexandra Ginger, Phuc-Anh Nguyen, Gunther Notni. 2018. Comparing HMD-based and Paper-based Training. *Proc. ISMAR '18, IEEE*. 134–142.
[24] Stefan Werrlich, Kai Nitsche, Gunther Notni. 2017. Demand Analysis for an Augmented Reality based Assembly Training. *Proc PETRA '17*. 416-422.
[25] Christopher D. Wickens, Justin G. Hollands, Simon Banbury, Raja Parasuraman. 2016.Engineering Psychology and Human Performance (4th edition). Routledge, New York.