The Impact of an Instructional Manual's Spatial Location on User Experience and Task Completion in Industry Augmented Reality (IAR)

by

Blessing C. Leonard

May 2023

Presented to the

Division of Science, Information Arts, and Technologies

University of Baltimore

In Partial Fulfillment

of the Requirements for the Degree of

Master of Science

Approved by: DocuSigned by: Greg Walsh 5/9/2023

E147E72B25744E6...

Dr. Greg Walsh, Thesis Advisor

DocuSigned by: Giovanni Vincenti 5/5/2023

323C89E5A4C64F4...

Dr. Giovanni Vincenti, AIT Program Director

Abstract

Research has shown that Augmented Reality (AR) instruction manuals are more effective than their paper counterparts in reducing errors when performing manual tasks such as assembling objects. To further improve user experience with such tasks, this study compared two positionings of instructional materials in an AR environment. The goal was to identify the solution that offers a more pleasant user experience for fully immersed individuals, as they interact with the virtual manuals while their hands remain freed up to complete an assembly task. As the popularity of Industrial AR rises, this study seeks to equip future designers of AR and Industry 4.0 interfaces and content with best practices.

Acknowledgments

I would like to thank my family (Goodwill, Josephine, Uzoma, Adaobi, and Eunice) and friends (Victoria and Jane) for their support and prayers throughout the journey to this master's degree.

I also want to thank the Interaction Design and Information Architecture faculty at the University of Baltimore, especially Dr. Greg Walsh who inspired my journey into the field of designing for humans, and graciously provided expertise and guidance as my thesis advisor.

Lastly, I would like to thank Dr. Giovanni Vincenti and my thesis cohort members Angela and Stephanie for being my soundboards, and all participants for their time and contribution to the project.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

Industrial Augmented Reality (IAR) is the application of AR technology to operational activities in fields such as manufacturing. Over the years, this has been achieved through immersive (using special glasses or head mounted displays (HMDs)) and non-immersive methods such as smartphones and projectors. With benefits that include increased productivity, efficiency, safety, and lowered costs (Border States, 2021) & (PTC, n.d.), IAR continues to gain ground. (PTC, n.d.) believes that IAR can provide “Up to 40% improvement in new hire productivity, up to 30% improvement in first-time-fix-rate, up to 30% acceleration in sales cycle, up to 50% reduction in training costs, and up to 25% reduction in scrap and rework costs”.

Border States (2021), the sixth largest electrical distributor in the U.S., provided six application opportunities for IAR in today’s manufacturing world: product design and development, maintenance and repair, training, assembly, quality control, and order picking (inventory management). In addition to these, (PTC, n.d.), a computer software and services company that specializes in augmented reality believes that the adaptation of IAR can lead to “Better information delivery, faster knowledge transfer, modernized training methods, immediate access to remote expertise, and enhanced customer experiences.” With training, maintenance, and repairs being integral to a manufacturing plant’s operation, this research focused on the adaptation of AR for delivering task instructions.

In their 1999 and 2017 studies, researchers Baird & Barfield, and Blattgerste et al. showed that AR manuals are more effective than their paper counterparts in reducing errors when performing manual tasks such as assembling objects. To further improve user experience with assembly tasks, this research compared two positionings of instructional materials in an AR environment. The goal was to identify the solution that offers a more pleasant user experience for fully immersed individuals as they interact with the virtual manuals while their hands remain freed up to complete an assembly task. I’m hoping that this study can be instrumental to future designers of AR interfaces and manuals geared towards futuristic assembly tasks like those promised by Industry 4.0.

To help provide a scope for this study, the following research question was posed: “How are user experience and time to task completion impacted by the position of the AR assembly manual?” The goal of this question was to compare the effectiveness of two presentation styles for a virtual assembly manual. The two options being a stationary (in-view) AR manual and an in-situ AR manual, with the hypothesis being that the in-situ manual would result in less errors and a quicker task completion time.

The methodology applied to this study was human-centered design, which uses the observed needs of a target audience as inspiration for designing what will be the solution to the observed problem. IDEO defines this method as "cultivating deep empathy with the people you're designing with; generating ideas…and eventually, putting your innovative new solution out in the world." ("What's the difference", n.d.). RubyGarage, a European software development and consulting company, suggests three principles to be followed when adapting the human-centered design method: collaboration, empathy, and experimentation. Collaboration was applied by adapting the prototype to include feedback from the participants, empathy was embodied by using personas to inspire the design process, and experimentation was incorporated through user tests and design iterations. A combination of questionnaires, interviews, surveys, and observations were used for data collection.

There are many concerns when testing with human subjects. Some of which include ensuring that neurodivergent and at-risk populations are not taken advantage of; participants are not deceived; and privacy measures are implemented to protect the participants. To ensure these are active considerations for researchers, schools have implemented Institutional Review Boards (IRB) that perform an objective analysis of research topics, methods, and procedures before the human subjects are involved. In addition to what is required by the institution, this study also applied the following ethical principles suggested by the Association of Computing Machinery (ACM) (Association for Computing Machinery, 2018):

Contribute to society and to human well-being, acknowledging that all people are stakeholders:

This principle requires that the study is not conducted just because, but rather that there is a goal of potentially improving lives with its findings. It also urges the researcher to actively think of the effect the final solution will have on the general population which is diverse in so many ways. In this case, the goal is to improve the efficiency of assembly tasks in industrial organizations regardless of the job experience level.

Avoid harm:

While this study is not expected to have negative effects on the participants, the University of Baltimore's IRB evaluated communications and study methods to ensure avoidable harms (physical and otherwise) are not overlooked. Should unforeseen issues have arisen during the testing phase, the individual would have been addressed, and the study revised to prevent further harm.

Be honest and trustworthy:

Participants were provided both verbal and written overviews of what the study entails, what data will be recorded, and their rights to leave the study at any point. The data analysis, discussion, and results sections of the final paper will also be free of manipulation should they not support the hypothesis.

Respect privacy and Honor confidentiality:

Save for photos/videos which could be used to illustrate the test environment set up or observed nuances, participants' personally identifiable information will not be stored for this study; they will be identified with terms such as Participant 1 or P1. The photos/videos were only taken at the consent of the participant, stored on the university's cloud storage medium, and deleted at the end of the study.

Chapter 2: Previous Work

Over the years researchers have compared the effectiveness of paper instructions to those provided via AR devices. Baird & Barfield (1999) compared four modes of presenting the same assembly steps for installing cards and chips onto a motherboard. The four modes evaluated were paper, computer-aided, opaque AR, and see-through AR displays. The sample set included 15 male graduate and undergraduate students (aged 20 – 37) and 20/20 vision was required. As there were 7 steps in total, the paper manual had images and instructions printed out on 5 total pages, the computer-aided manual used 7 slides in PowerPoint to present each step (instructions and photos were identical to that of the paper manual), and the AR HMDs (which were monocular for both the opaque and see-through devices) both had the instructions all displayed at once across a 35-degree horizontal field of view. The instructions for the AR HMDs were created in WordPad. Each participant tested all 4 manual modes randomly. For usability analysis, after each condition, the same 4 questions were asked and answered using a 7-point Likert-type scale (With 1 being the best and 7 the worst).

They observed that the participants made less body movements between the task and the manual when using the AR displays. They explained this as "…displaying the assembly instructions in the subject's work field[-]of-view may have reduced the amount of information the operator needed to store in working memory and may also have reduced the time needed to store assembly instructions in working memory."

They found no effect of order on time to complete the tasks. Based on the subjective feedback after each condition, the paper manual was rated least favored. The opaque HMD was voted most favorable for making the task easiest to perform; CAI was favored for clarity of instructions; opaque for easiest image and instruction readings; and opaque for perceived most effective. The order of manual favorability was opaque AR, CAI, see-through AR, and paper. The AR modes were attributed with faster completion times (approximately 50% faster) and fewer errors than the paper manual. They also suggested that the reason the opaque AR display fared better than the see-through was because the instructions "floated" around the motherboard for the see-through mode which could have been distracting to the subjects. This thesis is also hoping to verify if floating see-through (which I am interpreting to mean in-situ instructions) is still an issue, and if after 23 years, static remains the preferred display mode for see-through HMDs.

Blattgerste et al.'s (2017) sample size of 24 were slightly younger (aged 20 – 33) than (Baird & Barfield, 1999)'s and included both male and female participants. While Blattgerste et al. (2017) like Baird & Barfield (1999) also studied the impact of multiple instructional modes, Blattgerste and his co-researchers did not use monocles for the AR devices, and they compared two AR methods (in-situ and in-view). The 3D in-situ instructions were delivered via an Android smartphone and a Microsoft HoloLens 2, while the in-view or pictorial instructions were delivered via paper and an Epson Moverio BT-200 smart glasses. The in-view instructions provided by the Epson smart glasses were a replica of the paper instructions while the in-situ version of the manual was converted to 3D and featured assistive features such as a location marker. Users were given the authority to advance to the next step by the press of a button for both views.

The task was assembling a Lego set and they found that paper caused less cognitive load than the others. Paper was the fastest in locating the accurate brick container, picking the brick, and performing the assembly action of implanting the brick (followed closely by the HoloLens). The overall least number of errors occurred with the HoloLens, followed by paper, which begs the question: *in IAR, will less errors compensate for longer completion times*? A key discovery was that while the HoloLens recorded the least number of errors for participants when trying to identify the bin with the needed assembly pieces, the in-situ devices (one of which is the HoloLens) were responsible for more of the errors that occurred when identifying where each piece should go on the assembly field. Paper had lesser errors with assembly positions.

The inferred cause of the higher error rate for assembly positions with in-situ devices is that the instruction projection did not consider the placement of other objects in the environment such as a brick that had been assembled before the current one that needs to be placed behind it. Hence, the obstruction may have led to initial confusion for the participants. In their words: “Tang et al. [15] have shown that AR instructions reduced errors and cognitive load of participants. However, they also found that occlusions of target objects by AR content or presenting information over a cluttered background can decrease task performance.”

Lampen et al. (2019) made a similar discovery with their study to introduce a novel way of illustrating task execution to fully AR-immersed individuals. They used a human avatar (depiction of a life-sized virtual human) to walk users through the steps associated with assembling a car passenger door. They compared this method to the 3D in-situ and paper instructional methods and found that the simulated method reduced cognitive load and time to task completion. They however found that while the 3D in-situ approach fared best with error reduction, it didn’t do well with changes in the real-world environment that may occur after the initial calibration at the beginning of the task. They suggest continuous calibrations of the real space as a countermeasure for potential spatial changes that may occur after the initial mapping. While I did not apply the human avatar simulation to improve cognitive load and time to completion, continuous recalibration of the test environment to further improve error rates was considered.

Blattgerste and his co-researchers conducted another study in 2018 where they compared an improved version of their 3D in-situ manual to an in-view/side-by-side counterpart (Blattgerste et al., 2018). This improved version accounted for the environmental occlusions they hypothesized contributed to participants' confusion in 2017. They tested with a 24-participant sample size aged 18 – 31 and found that the in-situ instructions had better results than the side-by-side instructions for errors, task completion time, and perceived task load. "The correct handling of occlusions by real objects seems to improve the error rate, which went down from 2.875% per placement in our previous work [4] to only 0.13% in this study." The 3D in-situ instruction only had one person make one error, while the side-by-side had the slowest assembly time and highest error rate. The NASA Task Load Index (TLX) questionnaire was used to measure cognitive load and the 3D in-situ manual also dominated as the option that generated the least load.

To further improve the natural quality of virtual objects when immersed, Nijholt (2021) suggests that the virtual objects in the AR environment be impacted by changes in the real world. For instance, he suggests that a virtual creature should become wet if the AR setting is outdoors and it starts raining. This resonates with the earlier suggestions by Blattgerste et al. (2017) to keep calibrating, so that the assembly task is true to time and not impacted by steps that have been completed since the initial space calibration (like with the Legos). The "ENHANCED WALKING EXPERIENCE IN AR" portion of Nijholt's paper suggests that instructions can be geotagged such that they become available/launch based on the location of the user. For instance, the tire mounting manual will be restricted to the tire department of a car factory, while the package assembly manual will be limited to the assembly floor of a warehouse. Applying geotags to the manuals will eliminate the noise generated by irrelevant manuals and reduce user interaction with the AR software, thereby allowing more time to focus on the main task.

In proposing a method for converting traditional manuals to 2D graphic symbols that can be used in AR, Gattullo et al. (2019) tested their proposed solution for a maintenance manual for hydraulic breakers with 22 people. Given the need to convert traditional paper manuals into AR-friendly formats, this source provided some *Industry 4.0* principles that could be applied to my process. They advocated for the application of an ever-present *table of content* placed off to the side so users can easily pick and choose what portion of the manual they want to access. Their conversion of pdf to AR-friendly manuals involved using less texts and more pictures, using universal 2D icons to depict tasks such as pushing, and using a yellow circle to highlight the portion of the physical object that needed attention. They suggest grouping instructional contents into three categories (concept, task, & reference) and then suggest ways to reduce text for each category. They found that "with visual manual, organization of information is clearer than with PDF version of the manual. However, users felt there was missing information.

Given a substantial amount of academic research has been done on applying AR to improve industry processes, Fite-Georgel (2011) conducted a survey to identify why there weren't many IAR solutions in the market. After collating his data, he identified scalability to be one of the major issues plaguing the adaptation of IAR. He recommends working with industry representatives during design and test iterations to ensure the finished product can be applied to the market. His suggestion for incorporating IAR solutions with existing workflow is "to understand the reason for using a specific type of data, why the process is performed and the expected output (quality and data).". He also re-iterated the benefits of an IAR system: "performance enhancement (support the worker's task), saving material and resources (by replacing some real resource by simulated ones or offering remote access to experts), and improving service (automatize some less rewarding but important task)." (Fite-Georgel, 2011)'s survey reinforces the initial drive to test the proposed solution from this study with individuals already in the industry. He has also inspired thinking of the solution from the perspective of improving the existing system vs introducing a solution that can only function in a new tailored environment.

Uva et al. (2017) compared projected instructions to immersive instructions given through smart glasses. While their lit review concluded that projected instructions fared better than smart glasses, this study will be sticking to smart glasses because NASA astronauts (the inspiring audience) will be fully immersed at time of use. They also suggest that switching back and forth between the paper manual & workspace can increase cognitive load; and like (Gattullo et al., 2019), they used a *tree-like/table of contents* structure to provide instructions on-demand.

Funk et al. (2016) in their paper "*Interactive Worker Assistance: Comparing the Effects of In-Situ Projection, Head-Mounted Displays, Tablet, and Paper Instructions*", infer that paper instructions can fare better if they are tailored better than just having images on pages. They believe their validity can also vary based on the industry. Their study involved paper, tablet, projected in-situ, and HMD center-aligned instructions. They found that paper and in-situ projection allowed for quicker location of assembly parts and their rightful positions, while both tasks were slower using the HMD. This made me wonder if the completion time would be impacted since the method applied to my study is in-situ via HMD.

A year and half later, Funk et al. (2017) studied the long-term effects of in-situ instructions at an assembly workplace, and they found that in-situ was better for novice users, but after a while it became a distraction. Similarly, expert users found it annoying. They suggest using in-situ instructions for teaching a new process then dropping it at some point; however, they haven't figured out the sweet spot. One of their sources found that users complained of headaches, focus fatigue, and needing a 15-minute break while using the HMD. Yet, another source did not find a difference in the users' fatigue between HMD & paper instructions.

There seems to be a consensus that although the paper manual allows for a shorter completion time, the 3D in-situ manuals offer less errors. Given the limitation of full immersion being required, and the aim of the study to reduce erroneous actions, the 3D in-situ approach was one of the modes applied to this study. Part of the study included measuring subjective feedback on if fewer errors can offset the time cost.

Industry 4.0

In their paper that evaluated the expected influence Industry 4.0 technology will have on industrial performances, Dalenogare et al. (2018) mentioned that "Industry 4.0 concepts are proposed to enable companies to have flexible manufacturing processes and to analyze large amounts of data in real time, improving strategic and operational decision-making". While Frank et al. (2019) agrees that Industry 4.0 adopts digital technology to collect and analyze data in real-time for the overall improvement of the manufacturing system, they go further to define Industry 4.0 as the fourth industrial revolution with smart manufacturing as its core. Hence, all the talk about application of technology to collate and analyze data in real-time.

Referred to as the Industrial Internet of Things by Gilchrist (2016), this fourth industrial revolution that originated in Germany continues to see worldwide adaptations. According to Carvalho & Cazarini (2020), Industry 4.0 champions flexibility in the manufacturing process through the adaptation of smart technologies such that production behaviors can account for change in demand, technological availability, and labor expertise. It aims to reduce manual tasks through automation and leverage artificial intelligence to improve labor-intensive tasks such as monitoring seasonal demands and device uptime. The adaptation of Industry 4.0 methods also supports the customization of industrial processes and products on-demand. A major component of Industry 4.0 are Cyber-Physical Systems (CPS) which the National Science Foundation (NSF) defines as “integrate sensing, computation, control and networking into physical objects and infrastructure, connecting them to the Internet and to each other.” (National Science Foundation, n.d.).

Cognitive Load

Given this study is focused on improving the experience of workers in environments that require physical awareness, I was intent on not recommending a solution that would increase their cognitive load as that could lead to frustrations and more user errors. Essmiller et al. (2020) studied how to reduce cognitive load while encouraging the use of immersive Mixed Reality (MR) for providing users with instructions for simple procedural tasks. They also used Microsoft HoloLens and found that people were more motivated to engage with an MR app when they were tasked with either playing a game or freely exploring the HoloLens features. However, people felt more self-efficacy after going through a tutorial on how to use the HoloLens.

Paas et al. (2003) talked about how element interactivity during a task is directly proportional to cognitive load intensity. This is also referred to as intrinsic cognitive load because the demand it places on the working memory is needed for learning to occur. They also mentioned that the way to reduce this load will be to implement simpler tasks that will require fewer interactive elements. Another type of cognitive load is the extraneous/ineffective cognitive load which results from unnecessary demands being placed on the working memory. Extraneous cognitive load interferes with forming and using schemas and can result from both the information delivery method and the type of learning tasks presented to the user. It is suggested that tasks geared at forming schemas may help reduce cognitive load since schemas can consolidate multiple steps into one automated step thereby reducing the processing power needed to complete a task.

While intrinsic and extraneous cognitive load are additive (they add up to increase the load on working memory), modifying instructional designs to reduce total load will only be effective with materials that have high element interactivity. As there is little to no effect on materials with low element interactivity that have been designed to reduce load. Germane cognitive load boosts learning as the working memory is devoted to schema acquisition and automation. For learning to occur, the total of these three cognitive loads must not exceed the working memory's capacity. "Note that increases in effort or motivation can increase the cognitive resources devoted to a task." It is suggested that baseline/foundational knowledge is first presented so it can be transferred into a schema in the long-term memory and used throughout the task, then more specific information can be provided when it is needed. Finally, it was mentioned that the level of instructor control over the learning material should decrease as the level of student expertise increases.

Like Paas et al. (2003), Kirschner (2002) explains that because all the processing that needs to occur on the elements in working memory is also done by working memory, only two to three items (out of its maximum capacity of seven items) can be truly held by working memory. Kirschner also mentions that having complex processes involving unknown knowledge can be problematic. To combat this, they advise considering how information is stored in the long-term memory and designing informational materials that will leverage this (schema formation and automation) so that the needed baseline information can be presented when needed, thereby reducing cognitive load on the working memory. This Echoes what Paas et al. (2003) said about increasing germane cognitive load by designing the material in a way that supports schema formation and ensuring all three cognitive loads do not exceed the working memory's capacity.

Fast-Berglund et al. (2013) found a direct correlation between an increase in user choices and assembly errors, and that the errors were more likely in manual tasks than in tasks that offered cognitive support. Thus, they advise using cognitive automation/support to reduce the negative effect of increased choices in final assembly. They record Thurman et al. as defining cognitive automation to be the use of technology to compute information such as situational awareness, monitoring and fault management that would otherwise have been calculated by humans before commencing a task. They then define cognitive automation as "the amount of technique and information provided to the operator in order to know what, how and when to do a specific task in the most efficient way.". In cases where the task is performed by a machine, cognitive automation is used mostly to control and supervise said machine.

Watson et al.'s 2010 study compared 3 instruction models (text, static diagrams, & animations) over five build tasks. They found that there was no significant difference between animation and static diagrams, however, both yielded significantly faster completion times than the text model. They also observed that the text and static diagram groups required less manual referencing at Build 5 than the animation group. They suggest this may be due to the participants paying passive attention to the animated manual since it was easiest to follow thus less learning was happening. Unlike with the text format that required more concentration from the inception which could have in turn led to more retention. They also think that the quicker time and less error with animation and static diagram models may have resulted from lesser cognitive load than the text model.

The NASA TLX questionnaire is a tool for measuring subjective perceptions of workload in scenarios that involve human/machine interactions. Originally developed as a paper and pencil version by Sandra Hart in the 1980s, it has since been adapted into a mobile app; however, the original paper version was used for this study. To measure the workload required by a task, the questionnaire asks participants to rate six categories on a 21-point Likert-type scale that ranges from ‘very low’ on the left to ‘very high’ on the right. The six categories measured are: Mental Demand, Physical Demand, Temporal Demand, Performance, Effort, and Frustration.

Head Mounted Displays (HMDs)

So far head mounted displays (HMDs) have been mentioned several times, thus it is only fair that an analysis of their capabilities is also discussed. When considering HMDs, the two common methods for establishing an AR environment are video and optical see-through. According to ThirdEye Gen’s (2018) article, optical see-through uses semi-transparent mirrors to project computer-generated virtual objects in a way that it merges with the user’s view of the real world. They suggest optical see-through is better for mission-critical applications, however it may contribute to darkening the field of view (as it offers the sensation of wearing sunglasses). Medeiros et al. (2016) agree that optical see-through (OST) HMDs use semi-transparent mirrors. However, they go a step further to explain that these mirrors are “used to reflect the computer-generated images into the user’s eyes”.

Video see-through on the other hand uses the device’s cameras to provide a live video feed to be augmented with virtual objects. ThirdEye Gen (2018) suggests an advantage for video see-through is that the AR objects’ mapping to real world are more accurate than optical. While there is no latency with optical, there could be a mismatch of the graphics and what is seen through the glasses. Video on the other hand has some latency, but this latency allows the video and graphics to sync up and provide a less distracting finished product. This project used an optical see-through HMD (Microsoft HoloLens 2) since one had been previously purchased by the supporting institution.

Chapter 3: Methods

Given one of the main goals of this study is to improve user experience with AR-based assembly manuals, surveys and questionnaires were used to elicit subjective feedback based on participants' experiences. The sources mentioned during the literature review used demographic, subject-matter-familiarity, and the NASA TLX questionnaires to generate a baseline for result comparisons, however, this study maintained only the subject-matter familiarity and NASA TLX questionnaires (NASA TLX Paper and Pencil Version, 2022), as demographic influences were not studied.

The primary investigator was both the facilitator and the observer; hence supporting tools like filming and taking pictures were used for more accurate observable data collection. The assembly manual presentation modes were assigned randomly, and regardless of experience with an AR HMD or assembling tasks, each participant received a demonstration before commencing the task for each mode. The NASA TLX questionnaire was completed at the end of each task for each medium, and a final survey was given at the end of the test session. The post-test survey allowed participants to compare and rate the different presentation modes.

The two manual presentation modes that were compared were the static/in-view and the in-situ modes. The static mode displayed each assembly step as a series of 2D pictures that can be anchored outside the user's view of the work area, while the in-situ mode displayed the assembly steps as 3D holograms. Both modes gave the user the option to anchor the guides wherever they chose. However, because of the structuring of Microsoft Dynamic 365 Guides (the tool used for this study), the in-view guides were typically placed either at eye level or above eye level and never stood the chance of interacting with the work area, while the holograms could have been anchored in the field of view over the physical workspace.

As the research question seeks to identify the impact that an AR assembly manual's presentation style may have on user experience and time to task completion, Blattgerste et al. (2017 & 2018)'s proposed benchmark value of time to assembly (time to task completion) was measured. It should be noted that task completion here refers solely to successful assembly for each task.

Designing the Test

To prepare for the testing phase, personas were created as soon as enough knowledge was gathered to support the direction for the study. Five personas were created to include occupations like factory worker, and educational backgrounds that ranged from high school dropout to graduate degree holder. Familiarity and comfortability with technology were also included with the persona backgrounds. Including these details helped ensure the design was implemented from the perspective of potential users versus that of the primary investigator. After the personas were created, the first iteration of the design sketches was done. The sketch included some basic features that previous studies recommended such as a table of contents to offer the flexibility to move around the guide.

Due to time constraints, job descriptions were reviewed (Deaton, 2009) in lieu of performing an inquiry via an interview with an assembly line worker to understand the job description and requirements. The review focused on grasping an understanding of daily tasks and expectations to better inform what needs to be displayed in the field of view during and outside task completion. Once the Microsoft Dynamic 365 Guides was identified as the tool of choice for the study, the initial sketches were adapted to fit the features offered by the tool. Also due to time constraints, an official pilot test was not held, instead the in-situ prototype was updated before the third (out of four) test day to incorporate the feedback received from the previous two days. The second iteration was adopted to the HMD and used to test on days three and four. At the end of the entire test session, participants from the latter days were asked to look through the first iteration of the in-situ guides and give comparative feedback based on their experience with the newer version. The snowball method was employed for participant recruitment with the testers comprising of those approached on the campus of the host institution and the primary investigator’s contacts. While the preferred target testers would have been those with factory/warehouse/assembly line experience, I believe the credibility of the study remains, as the resulting solution is expected to be intuitive.

The test environment and gear included a worktable, an office chair, a subset of Lego pieces from the Classic Legos 11008 set, and a Microsoft HoloLens 2 loaded with Microsoft Dynamics 365 Guides (Figure 1). All testing took place in a lab with the same equipment and gear to offer a controlled environment for all participants. The instructions were offered through the HoloLens, and the participants were expected to locate all needed parts and assemble the items in the physical world based on the information provided by the manual. For each assembly task, the manual was presented in one of the two modes (in-view or in-situ) with the order of presentation randomized.

Figure 1

Test Environment

The in-situ method made use of 3D models that were found online (Yauhen, 2015 & 2018), downloaded, and then converted to SLR files so they could be imported into the Microsoft Dynamics 365 Guides tool that was eventually used to present the guides to the participants. The in-view mode had pictures of every piece needed for a step displayed in the order they would need to be assembled (the pictures of the pieces that would go on top of another appeared above the picture of the piece that would be the base)—see Figure 3. The last picture for each step was a picture of what the assembled object should look like at the end of that step. The next step showed the pieces needed for that step and a picture of progress that should have been made when the new pieces were added to the structure from the last step. Figures 2 & 3 below show what each mode looks like. To help with observation, the users were encouraged to think out loud while working through the tasks.

Figure 2

In-Situ House Manual Mode

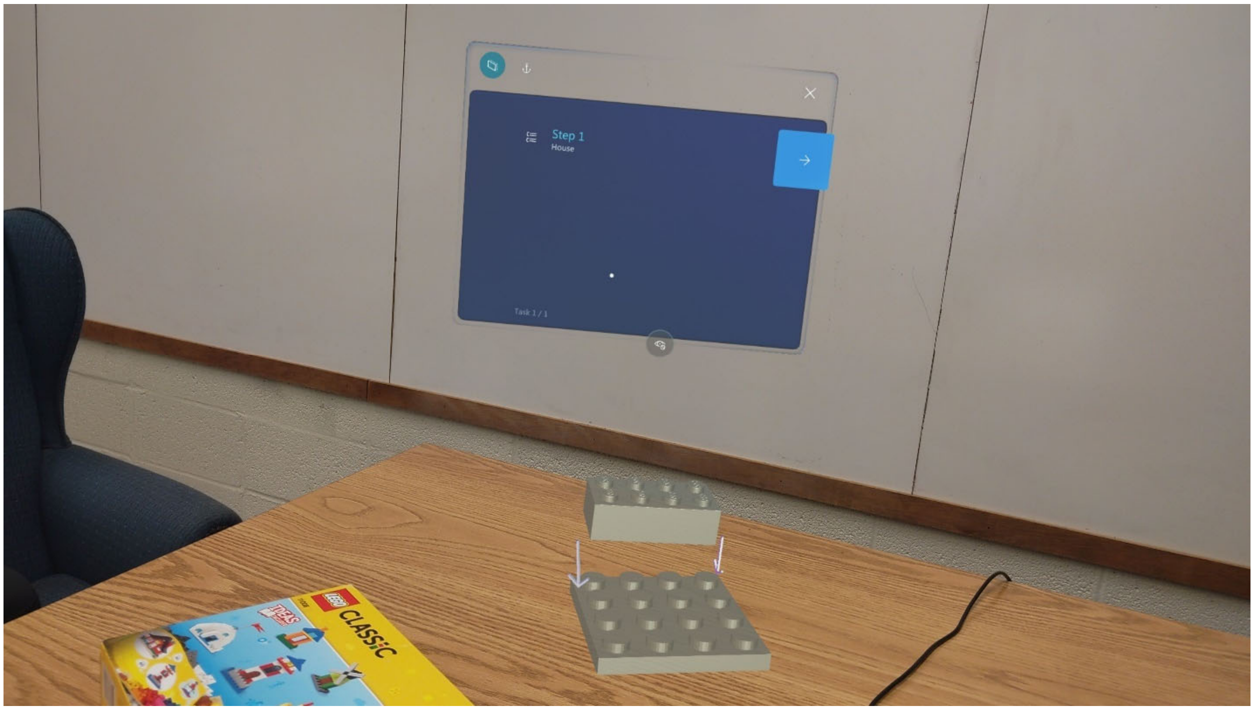


Figure 3

In-View Bridge Manual Mode

The Task

Two items needed to be assembled in both modes creating four tasks that resulted in the test sessions averaging an hour and half—although some participants completed all tasks in just under an hour. A house or a bridge were the goal of each assembly task (Figure 4) with the bridge providing for an increase in difficulty to the house assembly since it called for more pieces—some of which needed to be pieced together to simulate a whole (for instance, the base of the bridge columns (Figure 5)). The needed pieces were identified prior to testing and placed in a Ziploc bag with a few additional pieces so the participants still had to pay attention to identify what pieces were needed. The assembled piece was also taken apart after each task was completed so the pool of pieces was always the same for each task.

Figure 4

House and Bridge Goal

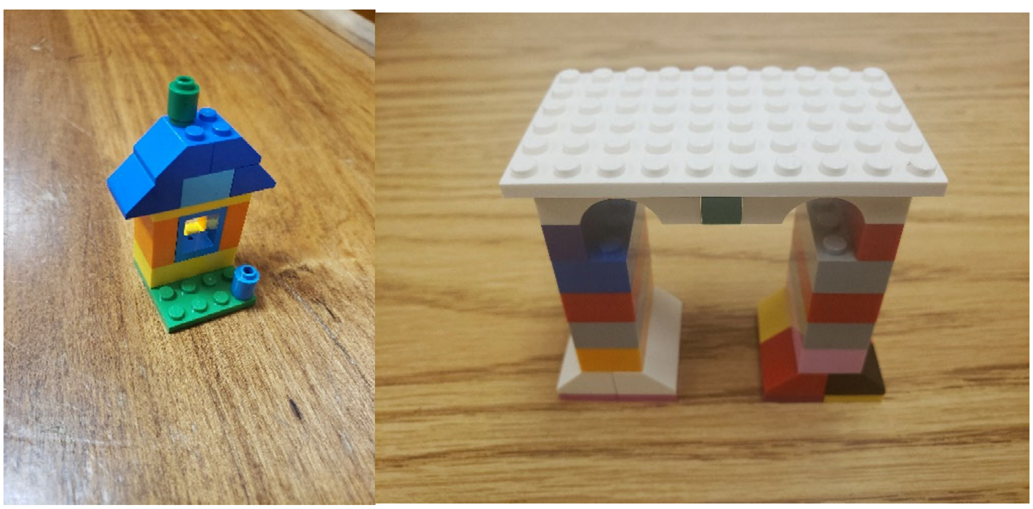
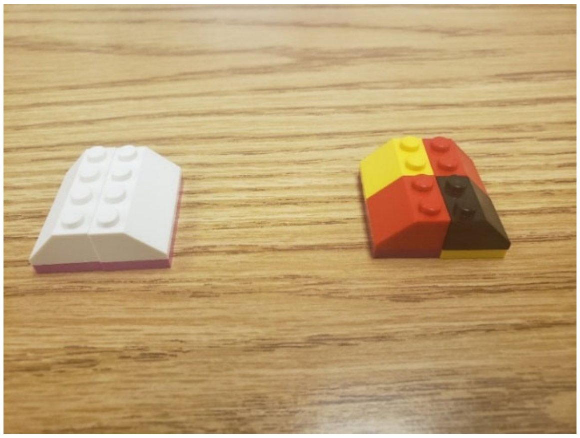


Figure 5

Base of the Bridge Columns



The house pieces were initially separated from the bridge pieces, but after feedback from Participant 1, all the pieces were left together for the duration of the test period.

The test sequence included:

1. Pre-test questionnaire
2. Demo video showing how to navigate the mode to be tested
3. An offering to play around with the HMD (which all participants turned down)
4. A task
5. Completion of the NASA TLX questionnaire to gauge cognitive load for that task
6. An offering to watch the demo video of the other mode if the next task was a different mode (if not, offering a rewatch of the demo video for the previous task)
7. Repeat 4 – 6 until the final task
8. Post-test questionnaire
9. Open discussion of the experience

Interestingly, none of the participants took the offers to play with the HMD or rewatch the demo video for each mode.

Earlier plans included using a Quick Response (QR) code placed on a physical object to provide a shortcut to access the appropriate manual for a task. QR codes are easily generated, and like barcodes, information embedded in QR codes are accessed by scanning the code with a reader. This feature in Dynamics 365 Guides proved more complex than time allowed to troubleshoot, hence the hologram anchor feature was used instead. This alternate feature allowed participants to choose the location of the holograms for the in-situ mode. Although they were also required to anchor the guide in the in-view mode, it did not have an impact since there were no holograms for that mode. Due to the time needed to acquire the needed skills with the program of choice for delivering the manual, the initial plan to use holograms that are projected directly over the physical pieces telling the user where to place them was adjusted to having the holograms assembling the pieces step-by-step (Figure 2).

Chapter 4: Results

A total of 11 participants took part in the testing phase of the study and their ages ranged from 18 to 40. Out of these 11, seven tested the original prototypes and four participants tested the updated in-situ prototypes that incorporated the suggested changes from the initial participants. Out of the seven that tested the original in-situ bridge prototype, four of them successfully completed the task, while three successfully completed the in-situ house. The in-view prototypes remained the same for all 11 participants. Out of the 11 that tested out the in-view bridge, eight were able to successfully complete the task and 9 successfully completed the in-view house task. The table below shows the time it took the first seven participants to report completing each task. The bolded entries in red signify those that did not complete the task successfully.

Table 1

Time to Completion for Each Task for the first seven participants

Participant	**Time to Completion (secs)**			
	In-Situ Bridge	**In-Situ House**	**In-View Bridge**	**In-View House**
1	458	600	518	240
2	**1062**	607	**628**	258
3	567	**469**	353	**172**
4	**507**	**371**	**628**	180
5	384	172	893	227
6	663	**234**	248	200
7	**1058**	**761**	357	180
Average	**671.2857143**	**459.1428571**	**517.8571429**	**208.1428571**

Table 2

Time to Completion for Each Task for the last four participants

Participant	**Time to Completion (In-Situ Iteration 2) (secs)**			
	In-Situ Bridge (v2)	**In-Situ House (v2)**	**In-View Bridge**	**In-View House**
8	356	**384**	460	127
9	878	343	748	240

10	**2580**	**285**	**388**	**351**
11	663	**383**	606	144
Average	**1119.25**	**348.75**	**550.5**	**215.5**

Table 3

Participant 2 Time of Completion for In-Situ v2 Tasks

Participant	**Time to Completion (in-situ Iteration 2) (secs)**	
	In-Situ Bridge	**In-Situ House**
2	627	309

Table 2 shows the time to completion for participants eight to eleven. The entries in red signify those that did not complete the task successfully. From the earlier participants, Participant 2 returned four days later to test the updated version of the in-situ modes. Their time to completion can be seen in Table 3 (this time around they were able to complete both tasks successfully in significantly less time). Table 4 shows the number of errors recorded at the end of each task for each guide mode.

Table 4

Number of Errors for Each Task.

Participant	**In-Situ Bridge**	**In-Situ House**	**In-View Bridge**	**In-View House**
1	0	0	0	0
2	1	0	2	0
3	0	2	0	1
4	1	1	1	0
5	0	0	0	0
6	0	2	0	0
7	1	2	0	0
8	0	1	0	0
9	0	0	0	0
10	1	1	1	1
11	0	1	0	0
Total	**4**	**10**	**4**	**2**
	1	3	Number of Errors (New In-Situ)	

The in-situ house model recorded the highest number of errors with seven for the original prototype and three for the updated prototype. The in-view house model on the other hand recorded the lowest number of errors with just two errors being recorded throughout the tests. The in-situ and in-view bridge models recorded an equal number of four errors, with the original in-situ bridge model responsible for three out of the four errors. The in-view house took the shortest average time to completion while the in-situ bridge had the longest average time to completion. There were seven successful completions of the in-situ bridge task, four for the in-situ house, eight for the in-view bridge and nine for the in-view house.

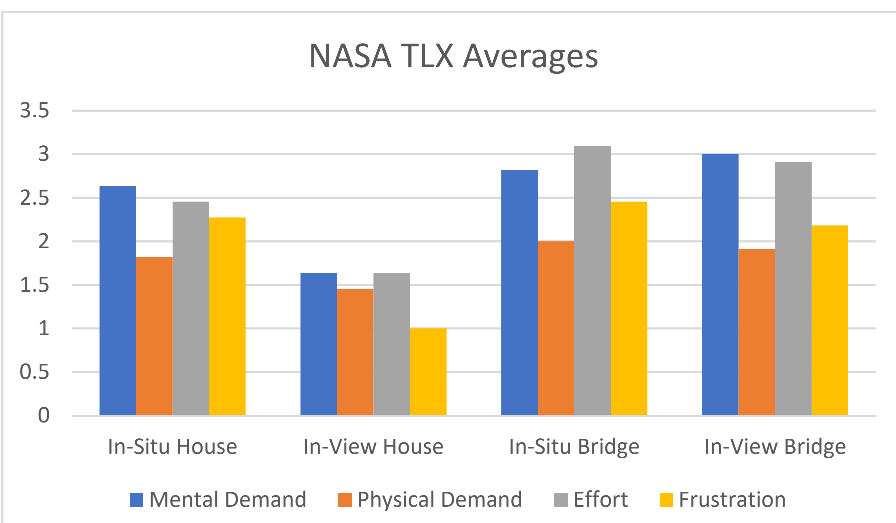
Chapter 5: Discussion

As previously mentioned, it was important to not increase cognitive load in this study. The goal is to contribute to Industry 4.0 without adding unnecessary stress to the intended users, especially since it is possible that this solution will be used in environments that require a significant allocation of working memory to environmental awareness. Cognitive load, however, can be tricky to measure, thus, it is often measured subjectively by asking participants how they felt performing tasks. For this study, the NASA TLX questionnaire was used to attempt to gain an understanding of how each model impacted the user. The NASA TLX questionnaire asks the user to rate their experience with a task based on five categories (Mental Demand, Physical Demand, Temporal Demand, Performance, Effort, and Frustration) across 21 notches that result to an equivalent of a 7-point Likert scale, with 1 being the lowest and 7 the highest. The averages of these ratings were then taken and compared to identify which mode may have been the most mentally tasking. It should be noted that the ratings for performance and temporal demand were exempted, as they didn't seem relevant to the information being sought after.

As shown in Figure 6 below, the in-view house model had the lowest average overall for cognitive load. It was rated as very low in mental and physical demand, effort required, and frustration experienced. In-situ bridge on the other hand, had the highest averages except for mental demand. In-view bridge had the highest average for mental demand.

Figure 6

NASA TLX Averages



While it was expected that the bridge tasks would be perceived as more complex than the house tasks, it was not expected that the level of frustration would not be proportional to the level of mental demand. It was observed that for all four guides, the level of frustration was reported as lower than the level of perceived mental demand. It is inferred that this had to do with the idea of building upon schemas to reduce the amount of cognitive load introduced to the user. By showing a depiction of what each finished step should look like, the guides were both accounting for the change in environment as suggested by (Blattgerste et al., 2017), and building on the schema accomplished in the previous step as suggested by (Paas et al., 2003). Only with in-situ bridge did the amount of effort exerted surpass the mental demand reported. From observations, this can be attributed to participants moving around to see all sides of the 3D models, and the increased complexity of the task compared to the house.

The fact that most of the participants had no prior experience with AR may have also contributed to the amount of effort required for each task, as not only were they processing the information needed to be successful, but they were also simultaneously familiarizing themselves with the concept of working in the real world while being fully immersed in AR via an HMD. This characteristic makes them the perfect candidates, as it is still expected that the majority of workers in the targeted industries will also be novices to AR technology. Something that was anticipated and reflected in Figure 6 is the amount of effort increasing as the ratings for mental demand increased. As Paas et al. (2003) stated, “increases in effort or motivation can increase the cognitive resources devoted to a task.” Thus, it was not surprising to see an increase in reported applied effort directly proportional to an increase in reported mental demand. The dip in the chart for physical demand was also not surprising since the tasks involved building structures with Legos and most participants completed the tasks while sitting.

In-View House vs In-Situ House (Time and Error)

In comparing the manuals for building the house, it was expected that the 3D models would make it easier to locate the needed parts and their respective positions, thus reducing confusion and the number of errors at the end of each task. While this was mostly true for identifying parts, assembling instructions proved confusing. From observation, I recommend requiring users to have a "play session" with the manual to build a dummy object before proceeding with the main task. While this was presented as an option, all participants declined the option to play with a demo guide that would familiarize them with using an AR guide to build real life Legos. It was later observed that a good number of participants were initially confused by the 3D models, what pieces to use, and where the pieces should go even after watching a demo video. I believe that if they had used the demo manual to build a demo structure (which was an abbreviated version of the actual manual that was used for the test), some of the confusion faced by the users in relation to the arrows used, where pieces should go, and their positioning (offset or aligned) could have been avoided or reduced.

The higher error rate that was recorded with the in-situ guide can be attributed to this lack of practice; although both guides introduced a learning curve of receiving instructions through an HMD, users could apply their experience with current manuals that have pictures on paper pages to interpreting the in-view guides. Hence, there was an added advantage/familiarity with the in-view than the in-situ house. As expected, the in-situ had a longer time of completion than the in-view, because it was anticipated that the time taken to familiarize oneself with this new mode would have accounted for more time unlike the in-view that was expected to be more familiar. This finding supports what (Blattgerste et al., 2017) found when they said the paper instruction was quicker to complete than other modes. It was expected that since the in-view mode resembled the paper manual most, the in-view mode would have been quicker to complete than the in-situ mode. While this study showed that in-situ took more time and had more errors, it is anticipated that applying the above recommendations will reduce the number of errors made and meet the original expectations of less errors and more time.

In-View Bridge vs In-Situ Bridge (Time and Error)

With the in-situ bridge, the longer time spent was also expected, as in-situ mode took more time on average than the in-view mode. Given that the bridge was a more complex assembly task, it was expected that there would also be more errors with the bridge tasks than the house tasks. What was unexpected, however, was that the in-situ bridge would yield the same number of errors as the in-view bridge, and less errors than the in-situ house (Table 4). This finding supports feedback from some participants that suggested using in-situ for more complex tasks and in-view for simpler tasks.

With the in-view bridge, users were initially confused as to what to do with the pieces. It was observed multiple times that participants would pick the pieces and lay them across the worktable without being sure of when to start assembling. They usually would advance to step two or three before understanding that the final picture on each step was an image of what their Legos should look like at the end of that step. At that point, they would revert to step one to start building the Legos as expected. Some participants voiced confusion as to not being sure what they were supposed to do initially. With the in-situ mode, the confusion usually arose from the pieces not being flushed on all sides of the hologram. With the first prototype, users were not sure if the Legos were supposed to be stacked flush on top of one another because the holograms were leaning forward. With the second prototype, if there was a slight offset with the holograms when viewed from behind or from a different angle, users took this literally and would adjust their physical build to try and match the offset. Typically, when they saw the final picture at the last step, they adjusted the physical Legos to match what it should originally have looked like. Although, in some instances, participants were able to use context clues to figure out the pieces needed to be flushed as opposed to staggered.

The in-situ bridge was however credited multiple times for being clearer at step five where they had to build the bridge top. This step required that Lego pieces be put on the reverse side of the 6 x 10 plate that was used as the bridge top (Figure 7). Participant feedback suggested that the in-view mode didn’t make it clear because everything was white. However, with the in-situ mode, they were able to see what piece they needed because of the contrasting colors. The arrows used also indicated that they had to flip it over. Some participants counted the indented circles and recognized that since the circles were indented, it was the reverse side of the Lego piece and not the top side. This was one of the instances cited by participants as a good instance to use the in-situ model for a complex task such as building the bridge.

Figure 7a

Step 5 from In-View Bridge Manual

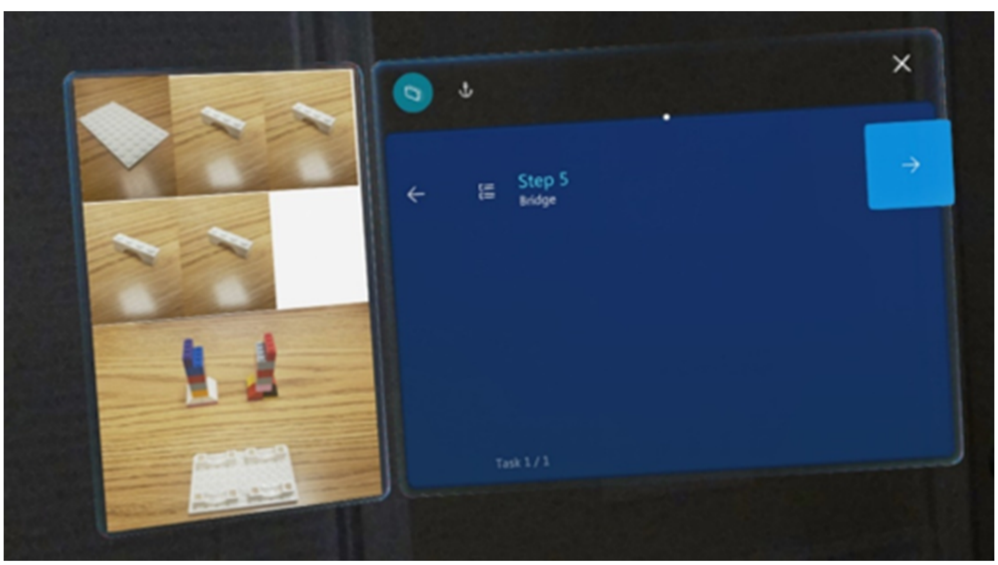


Figure 8b

Step 5 from In-Situ Bridge Manual

The errors that were occurring with the bridge models were similar across both modes, and they suggest that a combination of both in-situ and in-view would be helpful for future guides for more complex tasks. That is having both 3D holograms (which would help identify the pieces that are needed for that task) and the pictures that are displayed in in-view as a backup to better illustrate what the assembly should look like. Although a participant suggested using words as well, words were not considered because the goal is a generic manual that can be used by any and everybody irrespective of language. Therefore, having words would automatically create a language barrier that non-English speakers would encounter with using this tool. It is also advised that contrasting colors are used to depict different assembly pieces.

House vs Bridge

The house and bridge models were created to provide two separate levels of complexity for the assembly tasks; with the house being the simple level and the bridge being the more complex level. The bridge also incorporated more (and smaller) pieces than the house did in a bid to require participants to pay more attention to the task, and to handle different sizes of tools/objects as expected in the real world when assembling things like a car. You would need to handle screws and bigger portions like a car door for example. The House on the other hand, had mostly block pieces and about two smaller pieces. One of the house pieces did not have an exact 3D model replica. However, it was anticipated that participants will be able to see the available 3D model and use context clues to decipher what the appropriate piece should be. This was the piece that ended up being the window for the house. Through observations it was noticed that this piece did cause some confusion for participants as they dwelled on identifying the accurate piece for this section of the house, and some were unsure if that piece should be used at all. While some were able to use context clues as expected, others were not quite sure, and they voiced how the piece they were using was not identical to the 3D hologram. They proceeded with their choice (the right one) because they felt that was the piece that most likely fit. This suggests that for industry 4.0, companies need to be very specific with their 3D models; they should not depend on generic models, rather they would need to create custom 3D models for every piece that would be needed for each assembly task. They also need to ensure that the virtual pieces are assembled flush (as needed), and as close to reality as possible.

Most of the errors made for the bridge involved assembling the top portion of the bridge and the base for the legs of the bridge (Figure 5). Observations from watching participants assemble the base of the legs suggest that in industry 4.0, manufacturers need to ensure that pieces are flush at all angles. The confusion arose when the holograms were not 100% flush against each other for the tiny pieces that represented the base of the bridge. However, that was not an issue with the in-view guide for the most part. With the in-view mode, the issue was identifying the pieces and where they should go. Hence, a combination of in-view and in-situ will be needed for complex tasks to help users identify the needed pieces through in-situ (and loosely where they should go), and then in-view will give the clearer picture of what the actual product should look like once that assembly step has been completed. This approach should help mitigate confusions caused by offsets that may be seen from different angles of the hologram.

The other portion of the bridge assembly with an issue was assembling the top. Due to some limitations of Microsoft Dynamics 365 Guides, arrows were not used in later steps. This led to some confusion as to where some of the smaller pieces needed to go to create the top of the bridge (the ones that needed to go on the flip side). This is another scenario where a combination of both in-situ and in-view would be needed, as in-situ was very helpful in letting participants know they had to flip the 6x10 Lego plate upside down, while in-view was helpful in letting them know where to place the smaller pieces that needed to go underneath the bridge top. All in all, the expectation of the house being the simpler assembly task was met by this study.

Iteration 1 vs 2 (In-Situ Guides)

There were two iterations of the in-situ manual mode. The first mode involved 3D models that were stacked, but upon deployment was confusing to participants. Each Lego in the 3D model assembly leaned forward, hence it wasn't clear that they needed to be flush with each other when assembled. The original idea behind the first iteration was to leave gaps between each

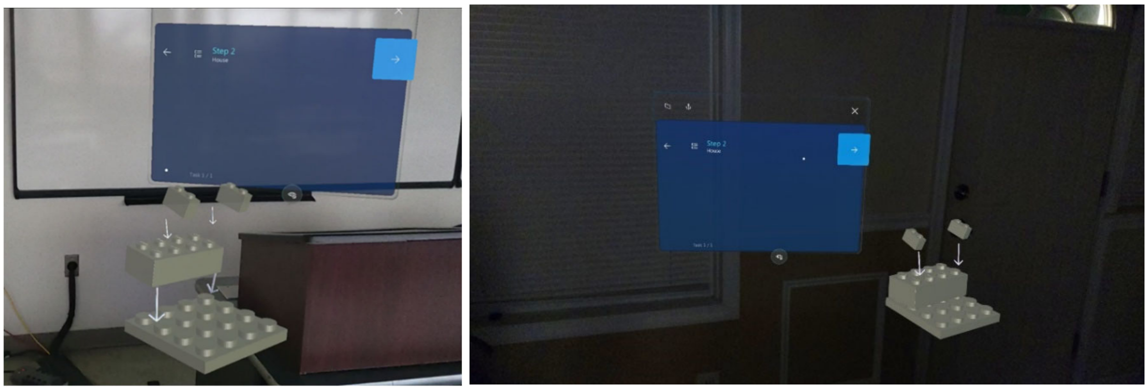
Lego piece to show which model went over the other. However, this intention was not perceived, as the participants interpreted the depiction to mean the spaces were required. Therefore, instead of assembling blocks snugly, they assembled them offset against each other to replicate the space that was seen in the manual. This resulted in having to modify the in-situ manual, such that the spaces and what can be referred to as a leaning tower effect for the assembled pieces were removed.

In the second iteration, the spaces between each block were removed to give a clearer depiction that the blocks needed to be stacked on top of each other with the new block flush with the previous one. The goal of the second iteration was to make it less confusing for the participants. It was observed that while the second iteration got rid of some of the confusion that could be seen in the first iteration, it introduced a new behavior of participants moving around the 3D model to look at it from a 360-degree angle versus just the angle that faces them. Therefore, if there was any slight offset at the other angles (for instance at what would have been the back of the model), they still interpreted that to mean that an offset was required versus focusing on just the front view that depicted what the object should look like once assembled physically.

As stated before, it wasn't expected that in-situ mode would have more errors than in-view mode. However, it should still be noted that once the adjustments were made for both the bridge and the house in-situ guides, the second iteration saw fewer errors than the first. Out of the four errors that were recorded for both iterations of the in-situ bridge guide, only one of them was attributed to the second prototype. And out of the 10 errors that were reported for the in-situ house manual, only three of those were accorded to the new prototype. When comparing the performances of Participant 2 who tested both in-situ versions four days apart, no errors were made with the second prototype for building both the bridge and the house unlike their previous attempts that had an error recorded for the in-situ bridge build. When asked if retained memory might have contributed to the fact that no errors were made, the participant responded that they couldn't remember the steps that were taken four days prior. They had anticipated that new tasks would be assigned that would require them to assemble different structures. This shows that with improvement, in-situ guides can help reduce the number of errors recorded in assembly tasks, especially for locating the right pieces.

Figure 8

Step 2 from Iteration 1 (Left) and Iteration 2 (Right)



Arrows & Angles

I expected the 3D models to make it clearer what pieces were needed at each stage and where those pieces should go. I also expected the in-situ guides to have longer completion times than the in-view guides to account for the time the participants would need to familiarize themselves with that mode of instruction. What I didn't account for however, is that the 3D models will inspire them to move around and view the holograms from a 360^0 perspective. This finding supports (Baird & Barfield, 1999)'s discovery that floating instructions were distracting to users. With the mindset that physical manuals are viewed from one direction, the initial prototype expected the guides to be read from the front. Thus, the right alignments could be seen from the front, while less attention was paid to what the back of the hologram looked like.

I quickly learned however, that not only did they want to see all directions, the angle at which they anchored their holograms also impacted what they saw. Some used context clues to make up for the confusion, while others followed what they saw literally. The arrows that were applied to show directions also became a point of confusion for those that were impacted by the placement of their hologram anchors. To them it was unclear what part of the pieces needed to be aligned even with the arrows present. The initial prototype also had the pieces hovering over one another in an effort to show how they should be lined up, this unfortunately was another source of confusion, as the participants would see that and misalign the Lego pieces, offsetting them when they needed to be flush. It was these observations and feedback that led to the iteration that produced the second prototype. The second prototype had the same contents except the 3D models were flushed against each other as they would be when building the Legos physically (Figure 8).

Recommendations

Based on what we've covered thus far, and the responses provided by participants in the post-test questionnaire, in-view is suggested to be used for simple tasks. It got a vote of six out of 11, and it had the least number of errors at the end of the house tasks. There was no significant difference in number of errors (difference of 1) for the bridge guides, however, seven participants did prefer the in-view mode for constructing the bridge. It is also recommended that pictures are provided for every step in the in-situ guides instead of only at the end. This was also supported by (Blattgerste et al., 2017), as they suggest using a combination of in-situ for picking the pieces and pictorial which in this case would be in-view for assembly instructions. This was also backed up by some of the observations made during the testing phase and the comments made by participants about expecting a picture at the end of each step for the In-situ mode; just like a paper manual would.

It is also advised that when a new technology is being introduced to users (in this case AR), the researcher should enforce completing a demo session before starting the main task to reduce confusion. Although this increases the time cost for each session, I expect that taking the extra time to familiarize the users with the basic skills needed to complete the upcoming task will be beneficial once the actual task begins. In addition, generating custom 3D models for each piece needed in in-situ guides, and ensuring the assembled 3D models are accurately depicted on all sides (not just from one angle) will help reduce confusion and time/effort needed to understand the instructions. Finally, ensure there is adequate space in the work area for users to move around and view the 3D models from all angles.

Chapter 6: Conclusion

The research question was aimed at identifying the impact that an AR assembly manual's presentation style may have on user experience and time to task completion. Upon comparing both guides, it became evident that participants leaned towards the in-view guides more than they did towards the in-situ. The total number of errors for each mode also showed that the in-view resulted in less errors than the in-situ guides. It is therefore suggested that the in-view guide be used for simpler tasks, while the in-situ guide can be used for more complex tasks. Furthermore, for the more complex tasks, it is also encouraged that a combination of both the in-situ and the in-view approach is used in tandem. That is, pictures are used to depict the steps, while the in-situ approach of using 3D models or holograms that are projected onto the workspace are used to encourage accurate identification of the required pieces. The features from the in-view mode can also be used as a reassurance of what the finished product should look like. This approach will be building on the user's existing experience with using paper manuals with pictures (building schemas), such that they are not jarred by the newness of adjusting to both being fully immersed and learning how to interpret the 3D depictions. A full in-situ guide can be introduced later in a person's career once they're confident and comfortable with using the hybrid guides. The progression suggested would be in-view for simple tasks, hybrid for more complex tasks, and in-situ for experts.

Future work

For future iterations of this study, I expect to explore launching the guides by scanning a QR code strategically placed in the work area instead of having to anchor the hologram. This should anchor the hologram at a uniform angle that serves the user best. Enabling a geotag feature that restricts access to a guide to a specific location is another feature that could be useful for privacy and reducing information overload in scenarios where there are many guides to choose from, but only a select number are needed by individuals in a certain department.

Limitations

One of the limitations experienced throughout this study was the skill level of the primary investigator. While I did learn as I progressed, it was quite evident that there were some features of the Dynamics 365 Guides that I was not familiar with and because of the time restriction of wrapping up the semester, I could not dwell on figuring out certain things like easier ways to better align the holograms, so that objects could be grouped and moved at once instead of being moved individually, which quickly led to hand fatigue. Eventually, the goal became to get the prototype to a point that was testable. However, it was apparent that the prototypes used were not at the optimal level, unlike If an expert had created them.

Dynamics 365 Guides also gave a maximum number of elements that could be imported into each step. Which is why at the latter steps of the in-situ guides, the arrows started depleting and eventually fizzled out as new pieces needed to be added to the assembly instructions. When participants were commenting on the later steps missing arrows, it was because there was no option to add arrows at that point. The maximum number of elements had been exhausted for that step. I had to get creative in how I presented the holograms and assembly instructions (trying my hardest to communicate positions and what pieces needed to be flush) to the participants without the use of arrows; hoping that it would be intuitive enough for them to follow through and build what was needed. It was evident especially with the in-situ bridge that participants would have preferred the consistency of having the arrows all the way to the end.

There was also the limitation of time, as the semester was progressing. And finally, the sample size and makeup of the sample size which consisted of 11 individuals recruited mainly through snowball sampling the primary investigator's contacts and supplementing with students at the University of Baltimore that were approached to partake in the study.

References

Association for Computing Machinery. (2018, July 17). *ACM Code of Ethics and Professional Conduct.* Retrieved from https://www.acm.org/binaries/content/assets/about/acm-code-of-ethics-booklet.pdf

Baird, K. M., & Barfield, W. (1999). Evaluating the effectiveness of augmented reality displays for a manual assembly task. *Virtual Reality*, 250-259.

Blattgerste, J., Renner, P., Strenge, B., & Pfeiffer, T. (2018). In-Situ Instructions Exceed Side-by-Side Instructions in Augmented Reality Assisted Assembly. *PETRA '18: Proceedings of the 11th PErvasive Technologies Related to Assistive Environments Conference* (pp. 133-140). New York, NY: Association for Computing Machinery.

Blattgerste, J., Strenge, B., Renner, P., Pfeiffer, T., & Essig, K. (2017). Comparing Conventional and Augmented Reality Instructions for Manual Assembly Tasks. *PETRA '17: Proceedings of the 10th International Conference on PErvasive Technologies Related to Assistive Environments* (pp. 75-82). New York, NY: Association for Computing Machinery.

Border States. (2021, April 14). *Exciting Industrial Applications for Augmented Reality (AR) in 2021 and Beyond*. Retrieved from Border States: https://solutions.borderstates.com/benefits-of-industrial-augmented-reality/

Carvalho, N. G. P., & Cazarini, E. W. (2020). Industry 4.0 - What Is It? In *Industry 4.0 - Current Status and Future Trends* (pp. 3–11). essay.

Dalenogare, L. S., Benitez, G. B., Ayala, N. F., & Frank, A. G. (2018). The expected contribution of Industry 4.0 Technologies for Industrial Performance. *International Journal of Production Economics*, *204*, 383–394. https://doi.org/10.1016/j.ijpe.2018.08.019

Deaton, J. P. (2009, May 11). *How automotive production lines work*. HowStuffWorks. Retrieved November 3, 2022, from https://auto.howstuffworks.com/under-the-hood/auto-manufacturing/automotive-production-line.htm

Essmiller, K., Asino, T. I., Ibukun, A., Alvarado-Albertorio, F., Chaivisit, S., Do, T., & Kim, Y. (2020, February 21). *Exploring mixed reality based on self-efficacy and motivation of users*. Research in Learning Technology. Retrieved November 10, 2022, from https://journal.alt.ac.uk/index.php/rlt/article/view/2331

Fast-Berglund, Å., Fässberg, T., Hellman, F., Davidsson, A., & Stahre, J. (2013). Relations between complexity, quality and cognitive automation in mixed-model assembly. *Journal of Manufacturing Systems*, *32*(3), 449–455. https://doi.org/10.1016/j.jmsy.2013.04.011

Fite-Georgel, P. (2011). Is there a reality in industrial augmented reality? *2011 10th IEEE International Symposium on Mixed and Augmented Reality*, 201–210. https://doi.org/10.1109/ismar.2011.6092387

Frank, A. G., Dalenogare, L. S., & Ayala, N. F. (2019). Industry 4.0 technologies: Implementation patterns in manufacturing companies. *International Journal of Production Economics*, *210*, 15–26. https://doi.org/10.1016/j.ijpe.2019.01.004

Funk, M., Bächler, A., Bächler, L., Kosch, T., Heidenreich, T., & Schmidt, A. (2017). Working with Augmented Reality?: A Long-Term Analysis of In-Situ Instructions at the Assembly Workplace. *Proceedings of the 10th International Conference on PErvasive Technologies Related to Assistive Environments*, 222–229. https://doi.org/10.1145/3056540.3056548

Funk, M., Kosch, T., & Schmidt, A. (2016). Interactive worker assistance. *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, 934–939. https://doi.org/10.1145/2971648.2971706

Gattullo, M., Scurati, G. W., Fiorentino, M., Uva, A. E., Ferrise, F., & Bordegoni, M. (2019). Towards augmented reality manuals for industry 4.0: A methodology. *Robotics and Computer-Integrated Manufacturing*, *56*, 276–286. https://doi.org/10.1016/j.rcim.2018.10.001

Gilchrist, A. (2016). Introducing industry 4.0. *Industry 4.0*, 195–215. https://doi.org/10.1007/978-1-4842-2047-4_13

Kirschner, P. A. (2002). Cognitive load theory: Implications of cognitive load theory on the design of learning. *Learning and Instruction*, *12*(1), 1–10. https://doi.org/10.1016/s0959-4752(01)00014-7

Lampen, E., Teuber, J., Gaisbauer, F., Bär, T., Pfeiffer, T., & Wachsmuth, S. (2019). Combining simulation and augmented reality methods for enhanced worker assistance in Manual Assembly. *Procedia CIRP*, *81*, 588–593. https://doi.org/10.1016/j.procir.2019.03.160

Medeiros, D., Sousa, M., Mendes, D., Raposo, A., & Jorge, J. (2016). Perceiving depth: optical versus video see-through. *Proceedings of the 22nd ACM Conference on Virtual Reality Software and Technology*, 237–240. https://doi.org/10.1145/2993369.2993388

Microsoft. (2022, May 11). *HoloLens 2 Pricing*. Retrieved from Microsoft: https://www.microsoft.com/en-us/hololens/buy

Microsoft. (2-22, May 11). *Dynamics 365 Guides*. Retrieved from Microsoft Dynamics 365: https://dynamics.microsoft.com/en-us/mixed-reality/guides/

NASA TLX Paper and Pencil Version. (2022, May 02). Retrieved from NASA-TLX: Task Load Index: https://humansystems.arc.nasa.gov/groups/tlx/downloads/TLXScale.pdf

NASA-TLX: Task Load Index. (2022, May 2). Retrieved from NASA-TLX: Task Load Index: https://humansystems.arc.nasa.gov/groups/tlx/

National Science Foundation - Where Discoveries Begin. Cyber-Physical Systems: Enabling a Smart and Connected World. (n.d.). Retrieved February 7, 2023, from https://www.nsf.gov/news/special_reports/cyber-physical/

Nijholt, A. (2021). Experiencing social augmented reality in public spaces. *Adjunct Proceedings of the 2021 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2021 ACM International Symposium on Wearable Computers*, 570–574. https://doi.org/10.1145/3460418.3480157

Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive load theory and instructional design: Recent developments. *Educational Psychologist*, *38*(1), 1–4. https://doi.org/10.1207/s15326985ep3801_1

PTC. (n.d.). *Augmented Reality*. Retrieved from ptc: https://www.ptc.com/en/technologies/augmented-reality

ThirdEye Gen. (2018, September 12). *Optical see through vs video see through on ThirdEye's X1 Smart Glasses*. Medium. Retrieved October 10, 2022, from https://medium.com/@thirdeyegen/optical-see-through-vs-video-see-through-on-thirdeyes-x1-smart-glasses-876dda5509d5#:~:text=One%20big%20advantage%20for%20video,mapped%20to%20the%20real%20world.

Uva, A. E., Gattullo, M., Manghisi, V. M., Spagnulo, D., Cascella, G. L., & Fiorentino, M. (2017). Evaluating the effectiveness of spatial augmented reality in Smart Manufacturing: A solution for manual working stations. *The International Journal of Advanced Manufacturing Technology*, *94*(1-4), 509–521. https://doi.org/10.1007/s00170-017-0846-4

Watson, G., Butterfield, J., Curran, R., & Craig, C. (2010). Do dynamic work instructions provide an advantage over static instructions in a small-scale assembly task? *Learning and Instruction*, *20*(1), 84–93. https://doi.org/10.1016/j.learninstruc.2009.05.001

What's the difference between human-centered design and design thinking? IDEO Design Thinking. (n.d.). Retrieved May 2022, from https://designthinking.ideo.com/faq/whats-the-difference-between-human-centered-design-and-design-thinking

Yauhen. (2015, October 7). *LEGO #21050-1 Architecture Studio "Create Your Own Architecrure"*. GrabCAD Community Library. Retrieved November 21, 2022, from https://grabcad.com/library/lego-21050-1-architecture-studio-create-your-own-architecrure-1/details?folder_id=1352364

Yauhen. (2018, May 8). *LEGO Set #4421-1 Box of Bricks +Extra Colours*. GrabCAD Community Library. Retrieved November 21, 2022, from https://grabcad.com/library/lego-set-4421-1-box-of-bricks-extra-colours-1/details?folder_id=4526134

Z, M., & S, A. (2017, July 18). *A Guide to Human-Centered Design Methodology and Process*. Retrieved from RubyGarage: https://rubygarage.org/blog/human-centered-design

Appendix: NASA TLX Questionnaire

Figure 8.6

NASA Task Load Index

Hart and Staveland's NASA Task Load Index (TLX) method assesses work load on five 7-point scales. Increments of high, medium and low estimates for each point result in 21 gradations on the scales.

Name	Task	Date

Mental Demand — How mentally demanding was the task?

Very Low — Very High

Physical Demand — How physically demanding was the task?

Very Low — Very High

Temporal Demand — How hurried or rushed was the pace of the task?

Very Low — Very High

Performance — How successful were you in accomplishing what you were asked to do?

Perfect — Failure

Effort — How hard did you have to work to accomplish your level of performance?

Very Low — Very High

Frustration — How insecure, discouraged, irritated, stressed, and annoyed wereyou?

Very Low — Very High