Investigation of Cognitive Gains in Object Assembly by using Augmented Reality

Author:
Abhiraj DESHPANDE

Supervisor:
Dr. Inki KIM

Department of Systems and Information Engineering
UNIVERSITY OF VIRGINIA

A thesis submitted to the University of Virginiain accordance with the requirements of the degree of MASTER OF SCIENCE in Systems and Information Engineering.

Charlottesville, Virginia

AUGUST 2017

APPROVAL SHEET

This Thesis
is submitted in partial fulfillment of the requirements
for the degree of
Master of Science

Author Signature: _____

This Thesis has been read and approved by the examining committee:

Advisor: Dr. Inki Kim

Committee Member: Dr. Gregory Gerling

Committee Member: Dr. Stephanie Guerlain

Committee Member: _____

Committee Member: _____

Committee Member: _____

Accepted for the School of Engineering and Applied Science:

Craig H. Benson, School of Engineering and Applied Science

August 2017

ABSTRACT

There has been a growing demand for Ready-To-Assemble (RTA) products due to the savings in their transport and labor cost. RTA furniture market is forecasted to increase at a Compound Annual Growth Rate (CAGR) of 4.75% during 2014-2019. Despite the rise in RTA's popularity, non-intuitive assembly instructions continue to plague widespread adoption of RTAs. The lack of cognitive support in these "static manuals", increases the mental workload during the installation. Evidence warns that perceived difficulties during RTA assembly may lead to emotional frustration or even physical damages to the user and the product. Yet, no technologies have been developed to assist the RTA assembly efficiently. From the standpoint of cognitive processing, people perceive it difficult to assemble RTA products. The main reasons being the lack of spatial problem-solving abilities or continual creation of mental representations of the product-in-assembly. This study hypothesizes that with additional spatial information along with mental representation, a cognitively heavy RTA assembly can relieve the mental workload of users as well as decrease the installation time and errors. Augmented Reality (AR) enabled Microsoft HoloLens™ headset was used to enhance the physical parts with computer-generated virtual objects. This is a comparative study in which 4 tasks – simple without AR, simple with AR, complex without AR, and complex with AR – are used to measure the effectiveness of the AR application on 14 subjects. AR has demonstrated its potential to improve the spatial awareness thereby reducing the demand for the cognitive resources. However, no application has been developed to test AR on the scale of RTA furniture. The application would guide the entire installation process by showing which parts to pick up and how the parts are connected. Overall,

the ability to overlay and register information on the workspace in spatially meaningful ways is likely to trigger diverse types of AR applications for advanced aid and instructional tools. However, it is crucial to understand specific cognitive gains that can be earned from the use of AR applications. This knowledge will help evaluate the current tools and justify design improvements. Results of this study showed that AR application reduced the installation completion time and errors, but there was an increase in the cognitive load. Also, the spatial problem-solving ability of the person was improved. Task complexity was found to be an important characteristic that impacted the applicability of the AR application. Finally, this study explained the factors that can improve the design of an AR app and gave insights into the advantages and disadvantages of Microsoft HoloLens™.

ACKNOWLEDGMENTS

I would like to thank my advisor, Dr. Inki Kim of the Department of Systems and Information Engineering (SIE) at the University of Virginia for the continuous support during my research and study. His active participation, encouragement motivated me to come up with a high-quality research. He made sure that I always took steps in the right direction. I am fortunate to have an advisor and mentor like him in my graduate life. Also, I would like to thank each of my thesis committee members, Dr. Gregory Gerling and Dr. Stephanie Guerlain, for their time and valuable feedback.

I would also like to thank all the students of Human-Centered Systems and Innovation Laboratory, and the staff of SIE. I take this opportunity to thank Debo and Adwait, for their continual support, valuable advice, reviewing this document, and helping me out with my thesis proposal and defense presentations. Also, Pallavi for giving me company during writing this document, helping me with the formatting of plots for this document and for all the support and food - thank you very much. Thanks in particular to Brian for his help, advice, and his company in the lab. I would also thank Snehal who has always been there for me. I thank many others who have helped me in some or the other way to complete my thesis.

Finally, I would express my heartfelt gratitude to Mom and Dad, without their support and love, it would have been impossible for me to achieve anything in life. I can't thank you both enough for all your efforts and sacrifices that have helped me progress in life.

TABLE OF CONTENTS

INTRODUCTION

1.1 Overview

The choice of dorm rooms and first-time apartment renters - Ready-to-Assemble (RTA) furniture - is seeing a revival in the U.S. per a 2015 study by Technavio [8]. According to research conducted by PBM Strategic Insights, 83% of Millennials have purchased RTA furniture, 67% of Millennials purchased entertainment furniture, while 64% bought bedroom furniture, and 62% purchased home office furniture [8]. E-commerce is viewed as the biggest source of growth in the US furniture and home furnishings industry [19]. Though the RTA market is increasing, there are certain problems related to the assembly processes that need to be addressed. There is evidence suggesting that practical difficulties faced during installation of self assembly products can lead to frustration, damage of the product and injury [66]. This thesis was developed to investigate the cognition gains due to an Augmented Reality solution to the object assembly problems, in particular RTA furniture assembly. Although there has been research in cognition in object assembly, problems in object assembly haven't been assessed to a great extent using AR. The remainder of this chapter focuses on the trends of this new technology,

the furniture market, overview of cognition and object assembly, and finally the goal, objectives, and contributions of the conducted research.

1.2 Motivation

Smart-phones have proved beneficial to connect people, perform daily tasks and provide information with a tap on the screen. Despite the advantages, it's excessive usage has had ill-effects on users. There have been reports on the symptoms of musculoskeletal disorders associated with the amount of text message and the time for daily usage of smart-phone during the study of characteristics of smart-phone usage and posture of users [40]. Smart-glasses not only help in relieving the "text neck" [15] but also overcome the inherent short-comings of a smart-phone. Display size and non-hands-free are the two biggest limitations of using a smart-phone. Smart glasses allow one to look straight out at the world without having to rely on the use of one's hands. This provides a completely different interaction situation and radically new applications [38]. By wearing a smart-glass, the environment becomes a display area on which virtually generated information can be displayed. This is due to the combination of smart-glasses and Augmented Reality (AR). The variety of smart-glasses can be seen in the figure 1.1. With the rise of smart-glasses, the mobile app-based world is most likely to become a smart-glass app-based world and thereby opportunities to redefine the applications and customizing options are likely to increase even more.

FIGURE 1.1. Smart-glasses in AWE2016 [9]

Schweizer [68] explains different applications of smart-glasses - production, educational, health, commerce, etc. and expects large investment in the research and development of smart-glasses due to the applications in military and entertainment industries. Researchers have defined Augmented Reality (AR) in a way that requires the use of Head-Mounted Displays (HMDs) [27]. Some of the application areas of AR include Archaeology, Architecture/Civil Construction, Art, Commerce, Construction, Education, Emergency management/search and rescue, Entertainment, Everyday, Gaming, Industrial Design, Spatial Immersion and Interaction, Medical, Military, Navigation, Office Workplace, Sports and Entertainment/Television, Task Support, Tourism and Sightseeing and Translation.

Due to the hands-free operation and vast display size, the application areas of AR are diverse. The ability to overlay and register information on the workspace in a spatially meaningful way allows AR to be a more effective instructional medium [73]. AR has helped to reduce head and neck movements during a repair and allowed mechanics to locate individual tasks more quickly [49]. There is a commercial AR product called 'WorkLink' [24] that helps in the remote maintenance and repairs of products. Due to AR's success in maintenance and repair operations as an instructional medium, this technology may be used in installation or assembly too. These operations require both hands to be free to perform the tasks. Thus, smart-glasses and AR can help due to the advantages mentioned above. Having said that, there isn't a product targeting installation or assembly of products. The main reason could be that smart-glasses haven't reached the intended population yet. So the companies can't ship the AR installation manual application with the product.

One such application in this domain is assembly of Ready-to-Assemble (RTA) products. RTA furniture is comprised of pre-finished or unfinished, non-upholstered furniture that is purchased in component form and then assembled by the consumer [13]. RTA products require little to extensive customer assembly while fully assembled products are ready to use right out of the box [12]. When defining the assembly of consumer products, we are considering products that

require one-off assembly at home, not a batch assembly on production lines [28]. RTA products have become common these days as they offer good savings and are cost-efficient by reducing transport and labor costs [71].

The US Furniture Market Report by Fung Business Intelligence Centre (FBIC) published on 9th February 2016 [19] mentions that the US furniture market will grow at a 2.9% compound annual rate from $96.4 billion (in 2014) to $111 billion (in 2019). Also, in 2014, sales at Overstock.com, Amazon and Wayfair, the three largest retailers of furniture online, grew by 14.8%, 19.5% and 44.0%, respectively. So e-commerce is making a big impact on the growing furniture industry with the number of furniture stores decreasing and increase in the online sale of furniture. Technavios's analysis forecasts the RTA furniture market in the US to grow at a Compound Annual Growth Rate (CAGR) of 4.75% during the period 2014-2019 [14]. E-commerce is viewed as the biggest source of growth in the US furniture and home furnishings industry [19]. In 2015, retail e-commerce revenue from furniture and home furnishing sales amounted to 26.7 billion U.S. dollars. E-tailers are witnessing double-digit revenue growth and a host of new startups are offering convenient and personalized options that cater to the evolving tastes of the US consumer [19]. As shown in figure 1.2, the U.S. furniture and home furnishings e-commerce revenue is set to increase by nearly 42% by 2019 than what it is in 2016.

Though the RTA market is increasing, there are certain problems related to the assembly processes that need to be addressed. There is evidence suggesting that practical difficulties faced during installation of self assembly products can lead to frustration, damage of the product and injury [66]. One of the difficulties faced includes, removal of all parts of the product from the tightly packaged carton for identification prior to assembly [31]. All the RTA products come with an instruction manual to guide the installation. These manuals are illustrations only or have text as well. Typical RTA furniture takes a significant amount of time to assemble, often leading a consumer to pay someone else to assemble the RTA furniture who has a greater amount of skill and/or more time to perform the work [50]. The problems experienced by the customers in the
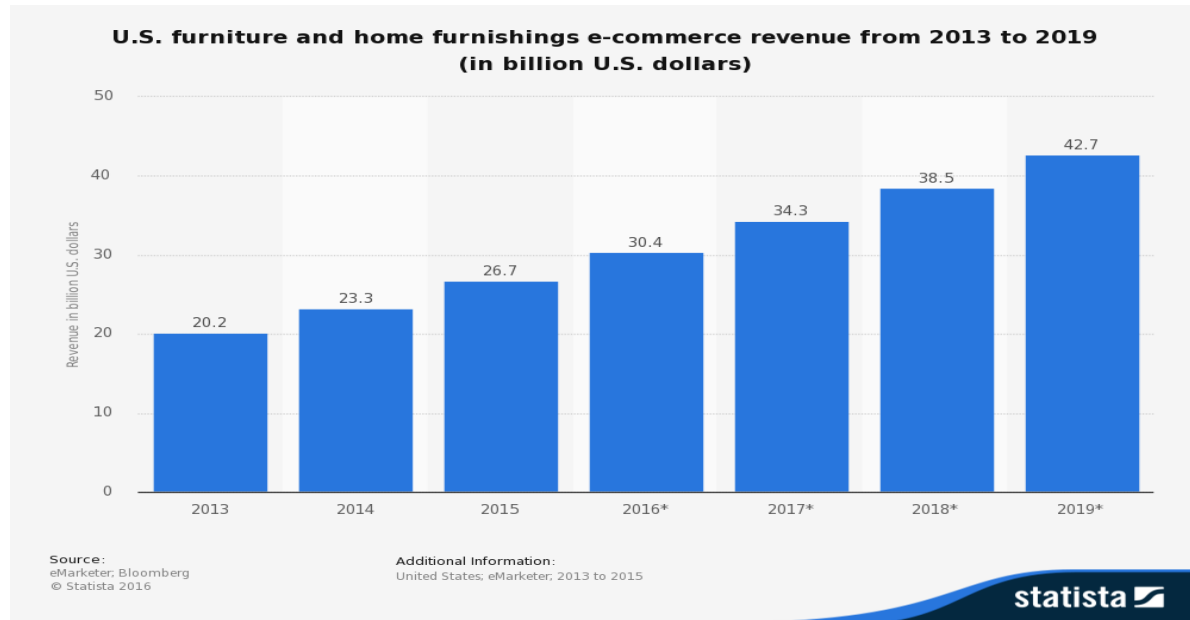
FIGURE 1.2. U.S. furniture and home furnishings e-commerce revenue from 2013 to 2019 (in billion U.S. dollars) [19].

assembly processes are mainly related to: grouping of parts, missing and/or defective parts and unclear instructions [71]. Results of a survey done by RatedPeople.com, demonstrated that 80% of the respondents would be willing to pay more for pre-assembled products. In fact, 25% would pay as much as 10% more to have the products pre-assembled in order to reduce frustration with the assembly process [37]. The complexity of an assembly can be very high. Typically, RTA products consist of upto several hundred parts [31]. It is more likely that these assemblies are one-time. The assembly would take place without any prior training and with the help of an instruction manual, even if the assembler is experienced [28].

There exists evidence regarding the problems faced by the consumers and statistics related to injuries caused during it. In a survey of UK consumers by the Office for National Statistics relating to just one type of self-assembly product, Richardson [65] reports that 52% of adults (approximately 23–24 million) stated - they had assembled self-assembly furniture in the previous

two years. Within these respondents, 67% reported some type of difficulties during the self-assembly process. Issues included accidents, such as consumers damaging the item that they were assembling (13% of the respondents) or causing minor injuries to themselves (7.8% of the respondents). When the large proportion of adults undertaking self-assembly tasks is considered, these percentages project to approximately 3 million UK adults who are likely to damage the item being assembled. More seriously, a potential of 100,000 people had required medical attention for injuries sustained during assembly [65]. These findings make it clear that consumers indeed face problems with self-assembly consumer products, resulting in issues such as damage to the item or personal injury [28]. Thus, there exists a need to simplify and ease out the assembly process of RTA products.

1.3 Overview of Cognition and Object Assembly

There is a tendency to think object assembly as a simple task that can be performed using a manual. But, it is a cognition-heavy task. Although there has been research in cognition in object assembly, problems in object assembly haven't been assessed to a great extent using AR. In order to understand more about object assembly and cognition, let us consider the following use case.

Howard Wattson, aged 57 yrs, buys a unit for his home music system from Wal-Mart online. The unit is shipped in a flat-box and delivered to his home. He opens the box and has all the parts of the unit spread across his living room. He has to install it by himself by following a 10-page installation manual. He tries to understand the labeling of the component and the structure of the unit to install it properly. Doing all of this by matching it with the instructions and illustrations given in the manual. He refers to the image of the final product given in the manual as well as imagines it as he installs the unit. He struggles comprehending the information from manual placed on his left side, the parts in front of him and finally the unit, still incomplete, on his right. After 2-3 hrs, he finally installs the unit.

As we can see that Howard is trying to understand how the individual parts of the unit

map spatially in the final product. It is like a puzzle to him and he has to solve the problem in hand by arranging the parts. His spatial problem solving ability is being tested. During this process, Howard refers to the physical and mental representations. Object assembly for Howard seems to be a heavy cognitive task and requires him to process all of the information present before him to perform the next step. The problems faced by Howard - *spatial problem solving*, *mental representations* and *cognitive load* - are discussed in the literature of cognition and object assembly. There has been little analysis of the role of human cognition in assembly performance [69]. Apart from cognition, design for assembly, predictive models for assembly time and assembly object characteristics are the other areas of object assembly [66]. These areas involve the user as well as the designer.

Let us understand more about the three problems that Howard faced:

Spatial problem solving ability:

According to Lohman [55], the spatial ability is composed of three major factors - spatial visualization, spatial relations and spatial orientation. Spatial visualization is the ability to mentally transform a complex stimuli, like a 3-D object in space. Spatial relation is the ability to rapidly recognize the identity of a simple item, like a 2-D object under rotation in a task. Spatial orientation is the ability to image how stimuli will appear from another perspective [55] [35]. Spatial visualization is the strongest and the most consistent factor supported in the literature. Hegarty and Waller defined spatial visualization as "the ability to mentally manipulate, rotate, twist, or invert objects without reference to one's self" [47].

Creation of mental representations:

Psychology dictionary [22] defines 'mental representations' as an entirely hypothetical being or entity which, philosophers believe stand for - perception, thought, memory and other mental processes during cognitive operations and functions during experiments. Representations are defined by dictionaries as something exhibited to the mind; a likeness, portrait, image or description; a sign or symbol, picture, plastic art or statue; and type or example. As stated by Paivio [61],

these definitions indicate that representations are physical or mental, that they are symbolic and that they vary in abstractness.

Physical representations are of two types - *picture-like* and *language-like* [61]. Picture-like representations are described as having analog, iconic, continuous and referentially isomorphic properties, whereas language-like representations are characterized as being non-analog, non-iconic, digital or discrete, referentially arbitrary and propositional [61]. Palmer's (1978) analysis mentions that representations depend on processing operations such that operations determine the relations that hold among the (represented) object elements [62]. Despite having or creating any representations, there is a need to process information about it. Thus having to create a mental representation and then process it continuously is a different task in itself.

Contextual human information-processing and Cognitive load:

Cognitive load can be defined as a multidimensional construct representing the load that performing a particular task imposes on the learner's cognitive system [60]. According to the general model presented by Paas and van Merriënboer [60], the construct has a causal dimension reflecting the interaction between task and learner characteristics and an assessment dimension reflecting the measurable concepts of mental load, mental effort, and performance.

In an experiment that was conducted, the process of how students' learn while problem solving, was examined. There exists a continual process as shown in the figure 1.3. People in general, are influenced by the context of the activities and the person's response affects the immediate physical and social environment [59].

From the figure 1.3, we can say that humans understand the context and then process information to respond again. This cycle continues till the goal is achieved. Thus, it can be inferred that context influences the cognitive processing required and processing affects the context. In order to reduce the cognitive processing, it is essential that the context presented to the humans is accurate. The lesser the cognitive processing required, lesser is the cognitive load.
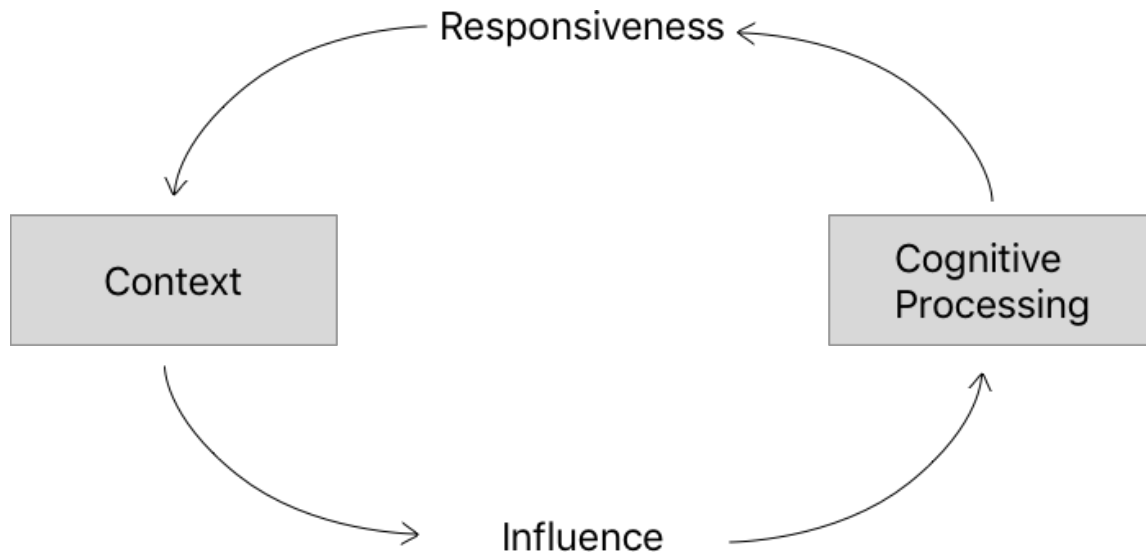
FIGURE 1.3. Interaction between context and cognitive processing.

1.4   Problem Statement

In object assembly, the spatial ability plays an important role in solving the problem of object assembly successfully. Visualizing the part of the product in relation to the entire product and then orienting it correctly to match the visualization is necessary. This process has to be done quickly and error free for any object assembly to be efficient. The physical representation, in the form of a manual or an illustration, that is provided has to be re-created mentally by the person. The user has to refer to it continuously and at times re-create it multiple times. This does increase the cognitive processing and the utilization of resources every time every time. As seen earlier that the context plays an important role in the information processing and impacts the cognitive load, it is essential to understand the context of an operation done during the assembly process.

In the use case discussed earlier, let us consider that Howard has an AR application provided by the supplier. The process after opening the box would be completely different. In this scenario, the application guides him in installing the table. Smart-glasses show him the manual and

detect the parts of the unit. The smart-glasses display the immediate steps needed to be done by Howard to install the unit that he is already seeing virtually through his glasses. The app shows him the order of the components to be used along with its spatial orientation. The augmented world does not change the real-world, instead adds additional information to provide spatial awareness to help in the problem solving process. There is always an illustration that he is able to see of how the final product will look. This virtually generated representation will substitute the mental representations that he needed to create earlier. These representations will remain in the world as augmented objects. Thus, this scenario was a well-guided process instead of the struggle without this application. Computer generated information and its representation in the real-world reduces the amount of processing required by Howard. The information that needs to be processed by his mind would already be processed and displayed virtually. Thus, AR has the potential to become a viable solution to tackle RTA product assembly problems.

Figures 1.4 and 1.5 are examples of how the application was proposed. Figure 1.4 displays the skeletal representation of the unit that is being assembled. Figure 1.5 shows an example of how the user will be guided in the installation.

FIGURE 1.4. App View - Skeletal representation of the product.

FIGURE 1.5. App View - One of the steps during the assembly of the product.

1.5   Goal, Objectives and Hypotheses

The goal was to find a technical solution to reduce the cognition load on humans in object assembly task. The objectives in achieving the goal were to develop an Augmented Reality application that acts as an instruction manual and perform human-subject experiments using the application. The app would address the problems and present solutions to ease the object assembly process.

The hypotheses of the thesis were as follows:

- H_1: Augmented Reality application reduces object assembly time as compared to the time without the application.

- H_2: Augmented Reality application reduces errors during object assembly as compared to the errors without the application.

- H_3: Augmented Reality application improves spatial understanding of the assembly.

- H_4: Augmented Reality application reduces cognitive load during the assembly.

1.6 Contributions

The research provides contributions to the fields of Human Factors and cognition in Augmented Reality. The thesis provides a solution towards the spatial solving ability and help reduce time and errors during the installation of RTA furniture. The thesis also aimed at reducing the cognitive load required to complete a task of object assembly. This research thus projects AR as a technology solution that helps tackle human cognition problems in object assembly. Also, the application developed is an exemplary application that can be referred to while creating applications for more complicated object assembly tasks.

1.7 Document Outline

Chapter 2 gives an overview of the RTA products and object assembly, and discusses cognition in them. Also, it gives an overview of Augmented Reality (AR), it's commercial applications, and the cognitive aspect of AR. Chapter 3 includes the details of the AR app that was built, and the experimental setup of the human-subject experiments. Analysis of the data collected during the experiment is explained in chapter 4. Chapter 5 presents the conclusions, limitations, and discusses the future work.

The sections that follow contain information reported in the literature. The sections discuss RTA products and object assembly, and Augmented Reality. The cognition aspects are also mentioned in the individual sections.

2.1 Ready-to-Assemble Products and Object Assembly

2.1.1 Overview

RTA furniture is comprised of pre-finished or unfinished, non-upholstered furniture that is purchased in component form and then assembled by the consumer [13]. Due to lack of research, there is only anecdotal evidence [56] that people often find assembling self-assembly products complicated. Spatial problem solving ability impacts the assembly [63]. People construct mental representations of the assembly or the final product in order to comprehend and manipulate spatial information [36], thus demand cognitive resources in order to complete the assembly. Cognitive load theory [72] deals with the interaction of cognitive structures and information and its implications. It is based upon the limitations of the human information-processing system,

such as working memory capacity. The amount of processing being performed is referred to as Cognitive load, and can be caused due to inherent complexity of the assembly [63].

2.1.2 Cognition in Object Assembly

Oxford dictionary [10] defines a problem as a proposition in which something has to be constructed. Since in object assembly, we assemble parts and construct an object, we may say that the task of object assembly is a problem that needs to be solved by assembly. Problem solving is generally regarded as the most important cognitive activity in everyday and professional contexts [52].

A cognitive perspective of assembly is taken by Helander and Willen [48], who provide an informative overview in Table 2.1 of Design for Human Assembly (DHA) and suggest Human Factors principles based on the aspects of human information processing, including perception, decision making and action [34] [28].

Table 2.1: Human Factors Principles in DHA [28].

Why	What	How
Minimize perception time	Visible parts	Nothing hidden
	Visual discrimination	Size, color
	Tactile discrimination	Texture, size
Minimize decision time	Ease the formation of a mental model	Visible parts
	Reduce choice reaction time	Minimize number of parts
	Spatial compatibility	Collocation of associated items
	Visual, auditory, and tactile feedback	Assembly looks different, auditory and tactile snaps
Minimize manipulation time	Ease of manipulation	Fixture to hold parts, parts that are easy to grip and don't tangle, fasteners that are easy to use
	Physical affordances and constraints	Self-locating parts, increase tolerances
	Design for transfer of learning	New product similar to old

The figure 2.1 is an instruction manual for assembling a floor lamp. Consider a person trying

Installing the floor lamp:
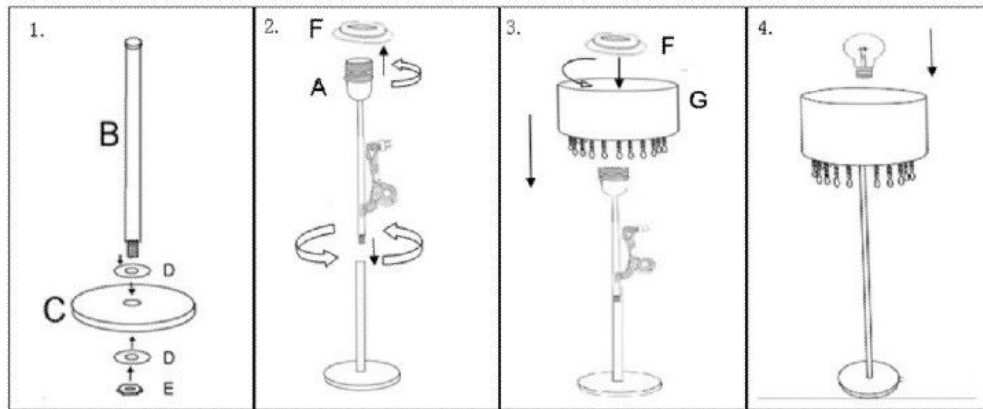
◆ **Technical Data:**

    230V~，50Hz，bulb:E27 ,1 x Max 40 Watt   Class Ⅱ，   IP20

FIGURE 2.1. An instruction manual for a floor lamp assembly. Source: http://www.alibaba.com

to perform this assembly. The person is presented with numerous parts when fixed together form a lamp. The person uses the manual and follows all the instructions step-by-step to finally construct the floor lamp. We may say that the instruction manual is a step-by-step solution to solve the problem of 'How to assemble a floor lamp'.

Jonassen [52] has categorized problems into 11 types and provided 7 characteristics for each of those problem types. Although object assembly satisfies 4 of the 7 characteristics for it to be classified as a 'Design Problem', the classification isn't explicit and detailed enough. Greeno [43] classified problems based on general psychological skills and knowledge needed to solve them: the problems of *arrangement*, *inducing structure*, and *transformation*. This classification serves as a criterion to determine the nature of a problem by examining underlying cognitive skills related to arrangement, inducing structure, transformation, or a combination of the three. Table 2.2 explains this classification with respect to object assembly. In a generic way, if we classify object assembly according to Greeno, it may fall under the combination of Arrangement and

Transformation.

Table 2.2: Greeno's problem type classification according to a generic object-assembly

General problem-solving process includes	Arrangement	Inducing Structure	Transformation	A generic object-assembly task
Problem Definition	A set of objects to be arranged to get a correct solution	Understanding the patterns to get a correct solution	Perform set of operations on the initial state to reach the goal state	Arrange a set of objects of a product to install the final desired product
Goal / Final state	Is not provided	Is not provided	Is provided	Is provided
Initial information	A set of objects	Analogical information	Initial state of the product is given	A set of objects that may or may not have a relevant grouping
Process / Operations	The objects are arranged to reach the solution	Identifying the fixed patterns and relationship	Perform operations that changes current state of the solution	Arrangement of objects by joining, rotating and placing them in a particular order that matches the final state
Evaluation criteria	Can have one or more correct solutions	The analogy/ pattern holds correct for other examples	The goal state is achieved	The solution matches the final state

Problem solving is "any goal-directed sequence of cognitive operations" [26]. Those operations have two critical attributes. First, problem solving requires the mental representation of the situation in the world. That is, human problem solvers construct a mental representation (or mental model) of the problem, known as the problem space [58]. Thus, a certain assembly task too would require an essential set of cognitive skills. An arrangement and transformation problem requires the following skills according to Greeno [43]:

1. Ability to generate possibilities:

   To solve a given problem, one needs to perform operations that lead to a solution. In this process, the problem solver is generating partial solutions and discarding those that appear unpromising. Thus, this procedure of 'generating possibilities' requires the problem solver to recognize and analyze the operations and the partial solutions. During this process, a lot of information needs to be processed by the problem solver. The amount of this human-information processing is referred to as Cognitive load [63]. Increase in amount of this processing would increase the Cognitive load so the problem solver needs to be fluent in generating possibilities in order to restrict the Cognitive load. Consider an example of a Lego block toy. The installer needs to recognize where a particular block can fit and discard the possibilities where it can't.

2. Ability to retrieve solution patterns:

   There is a contextual influence of the activities and person's response affects the immediate physical and social environment [59]. Problem solver would try to find a context that is accurate to reduce the cognitive load. While processing information about the problem, the problem solver may try to create a context that resembles one of the familiar matches. This context may be a partial solution or a similar series of operations that are performed.

3. Knowledge of constraints for efficient search:

   Read [64] explains 'knowledge of principles' in a puzzle example as knowing the frequency with which the letters occur together. This is similar to eliminating possibilities based on certain principles. Considering with respect to object assembly, most of the assembly products have screws that join two parts of the product. Understanding the dimensions of the screws and the holes in which it needs to be fixed should help guide the assembly. Unlikely combination can thus be avoided. Understanding the unlikely combination will thus constrain the search of the solution. As the number of combinations reduces, lesser information processing would be needed, reducing the cognitive load.

4. Means-end heuristics:

   Transformation problems provide a goal state rather than requiring problem solvers to produce it. Means-end analysis is a strategy that can be used to solve transformation problems by eliminating differences between the initial and goal states [64]. This analysis relies on the comparison of the current problem state with the goal state. Transformation problems rely on the means-end analysis according to Greeno [43] as they have a given goal state. During assembly of products, the assembler always knows what the final product is going to look like. By comparing the final product image/illustration, one can select the operations and take decisions to achieve the goal.

2.1.3 Related Works

Tang et. al, 2003 [73] have done a similar study to provide evidence to support the proposition that AR systems improve task performance and can relieve mental workload on assembly tasks. The 3 hypotheses were reduction of time, errors, and cognitive load and improve accuracy. They also focused on head and eye movement, attention switching, spatial cognition and mental transformation. Their results concluded that information overlay did not emerge as an advantage as expected. Also, visual interference was pointed out by the participants. Difficulty to view the overlaid information was a concern as well. Despite all the concerns, AR helped in improving performance due to hands-free operation. Participants made significantly less dependent errors while using AR as they were able to determine position and orientation easily. Mental workload was measured using the NASA Task Load Index (NASA TLX) and the results supported the hypothesis.

2.2 Augmented Reality

2.2.1 Overview

Augmented Reality (AR) is a real-time direct or indirect view of a physical real-world environment that has been enhanced/augmented by adding virtual computer generated information [32]. As seen in the figure 2.2, the real-world comprises of just a text 'DUCK'. But when watched through AR glasses as shown in figure 2.3, the text is shown with a virtually generated duck.

FIGURE 2.2. Physical real-world environment.
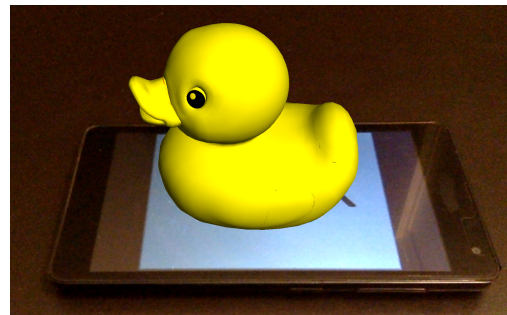
FIGURE 2.3. World augmented with virtually generated objects.

AR allows the user to support the real-world with an additional set of information by using sound, video, graphics or some other form of data. This modification of reality results in the technology functioning in order to enhance the user's current perception of reality [42]. AR overlays information needed by the user while Virtual Reality (VR) immerses user in a virtually created who may or may not be part of real world or would have opted for virtuality. AR still supports interaction with the information portrayed to the user and creates an environment in which there is information displayed artificially on the real-world objects [5][57][51]. AR supplements reality, rather than completely replacing it. Ideally, it would appear to the user that the virtual and real objects coexisted in the same space [27].

AR has emerged in 1950s when Morton Heilig, a cinematographer, thought of cinema is an activity that would have the ability to draw the viewer into the onscreen activity by taking in

all the senses in an effective manner [32]. The applications of AR have grown in recent years. The variety of areas of application include medical, entertainment, commerce, education, gaming, navigation, sports, task support, etc. Since AR is an additional layer of information that is presented to the user, most of the applications use an additional device to render this information. These devices comprise of Head-up Displays (HUD), wearable glasses, projectors and mobile devices. With the ever increasing market of smart-phones and increase in the processing power they possess, the AR applications are centered around the mobile apps (including tablets and phablets). As stated by Livingston, there is a paucity of investigations regarding human factors in AR systems compared to the enabling technologies [54].

FIGURE 2.4. AR Ecosystem [9].

AR is a technology that consists of multiple systems that work together and form a system-of-systems. This ecosystem is shown in figure 2.4. The color coding helps us to understand what the different layers are. With the advancements done in individual layers, AR is improving and advancing. Advancements are done in the hardware to reduce the weight of the glasses, to

increase its battery life and reduce over-heating. ARToolkit, Wikitude and EasyAR SDK are few of the SDKs that are being improved so that more and more developers use them to c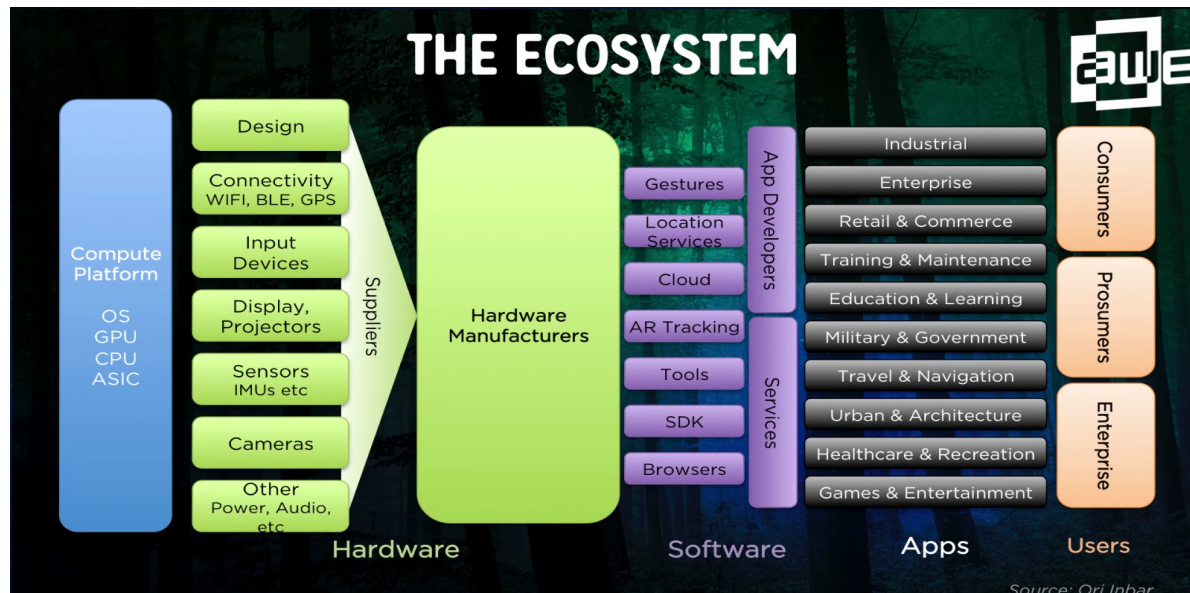reate AR apps for multiple areas like retail and commerce, education, entertainment, navigation, health-care, etc. With the use and the management of the depth sensors, AR applications' possibilities just got pushed to the absolute limits. This sensor will help in the recreation of the surrounding reality of the users with precision and in real-time. All this information will help in augmenting the real-world with the virtually generated objects in a more natural and precise manner [25].

2.2.2 Commercial AR Systems

There exists a current commercial application 'WorkLink' [24] that helps in remote assistance of products using AR. This application includes step-by-step 3D animated instructions enhanced with intuitive layers of text, images and video that empower users with expert knowledge right at the time and place they need it most, projected onto the task at hand as shown in figure 2.5. It uses converted CAD data, instruction templates and simple tools to create complex animations, where authors can publish new content quickly, with little to no training. WorkLink captures analytics and meta-data (such as how long it takes to perform steps and procedures), images, annotations, checklist verifications and measurements so organizations can ensure accountability, effectiveness and efficiency on every task. The proposed application would be designed to address the core installation process in a much more easier and accessible platform instead of remote support.

Augmented Furniture [2] is a commercial app that allows augmenting your house with furniture in real-time. ViewAR [21] is another such application. Atheer has developed AR suite, glasses and OS that can be used by clinicians for monitoring vitals, checking reports and manage the medicine dosage of the patient [23]. Aerospace giant Boeing recently completed a successful Google Glass pilot, reducing production time of aircraft 'wire harnesses' and reducing related error rates in half. The company also says it is 'very close' to launching a live smart-glass system

FIGURE 2.5. WorkLink illustration [24].

[67]. This was reported on July 13th, 2016.

2.2.3 Cognition in AR

To develop highly effective AR applications, there is a need to incorporate user-centered perspective and understand the impact of characteristics of AR systems on human activities [77]. An important category of human factors in AR is the Cognitive issues that relate to the users' cognitive process for understanding an AR environment during interaction [53]. Cognitive functioning may be inhibited by the synthetic scene if there isn't a balance between the visual complexity of real world background and the virtual information. [70] There are three main types of co-ordinate systems for mobile wearable AR. They are *head-stabilized* where "information is fixed to the user's viewpoint", *body-stabilized* where "information is fixed to the user's body position", and *world-stabilized* where "information is fixed to real world locations" [30]. The information layout would impact the effectiveness of understanding the information at a glance [29]. The modal display of information impacts users' interaction during navigation in defined space. Users move their attention between information display and physical surroundings to cognitively map their present location and spatial relationships with targeting plots [75].

EXPERIMENT

3.1   Overview

A comparative study was conducted in order to explore physical and cognitive functions that occur during furniture assembly, and then to examine the effectiveness of AR as an aid to support the task. The study was conducted in 2 phases - Non-AR and AR. Since, it was a comparative study, the subjects were first tested with the Non-AR task and then installed the other task using the AR application. To further test the effectiveness of AR, 2 different tasks were chosen - simple and complex. The task assignment is explained in the following section. For generalization of the findings, two furniture with different levels of difficulties were selected. Considering RTA furniture are typically assembled once, no group was allocated to assemble the same type of furniture twice. This also removed the learning bias of the task. The effectiveness was measured in terms of time of the installations and errors occurred during installation. A detailed explanation of the setup of the 2 phases is explained in the subsequent sections followed by the details of the AR app.

3.2 Experimental Setup

3.2.1 Task

Table 3.1 illustrates the subject bifurcation for the tasks. As shown in table 3.1, subject A will perform simple Non-AR task and complex AR task. While subject B will perform complex Non-AR task and simple AR task. Figure 3.1 explains how subjects are randomized for the experiment.

Table 3.1: Task bifurcation

Task Type	**Non-AR**	**AR**
Simple	A	B
Complex	B	A
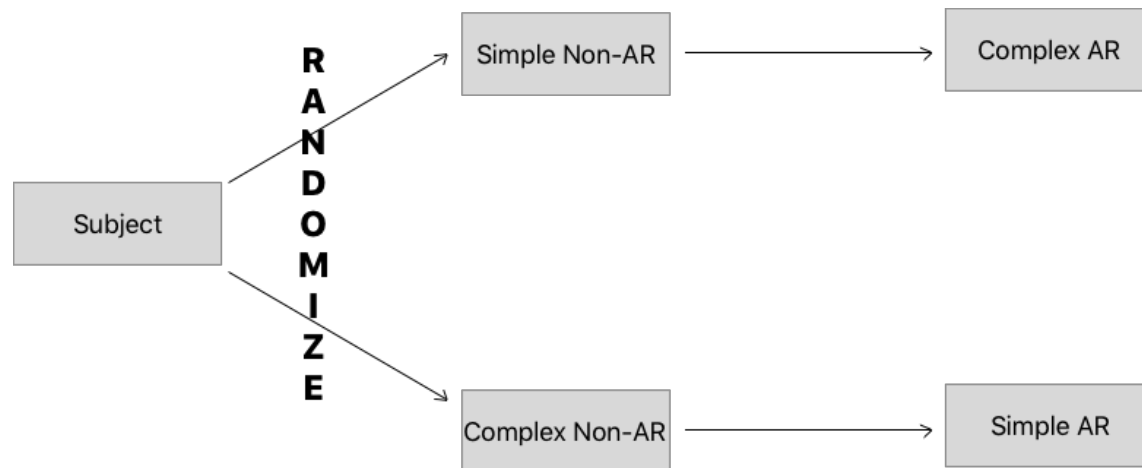
FIGURE 3.1. Task Allocation.

The tasks were categorized into simple and complex based on factors explained in table 3.2. The factors were chosen depending on the skills that are required for problem-solving as explained in the section 2.1.2. The simple task (figure 3.2) included a repeated set of steps and took less time. The repetitive pattern within the furniture, and the time taken for installation

were the important factors considered in the classification. The complex task (figure 3.3) took approximately double the time of the simple task. Though there was an inherent pattern, but the transformations required made the task complex. Since, the spatial problem solving ability of the subject was to be tested, a few simple steps like, fixing the dowels and 2 fixed screws was already done. Also, the focus was on the spatial understanding of the subject and not the subject's skill of using a screwdriver during installation. The initial configuration of the installation procedure was kept consistent across the task sections.

Table 3.2: Task type factors

Factors	**Simple**	**Complex**
Time of Installation	10 minutes	20 minutes
Symmetry of parts	Yes	Yes
Does symmetry help?	Yes	No
Arrangement of parts	Similar	Different
Transformation of parts	Similar to each other	Similar but multiple transformations
Patterns present?	One pattern	Multiple patterns
Screws needed?	No	Yes
Number of parts	25	15

FIGURE 3.2. Simple installation task.

FIGURE 3.3. Complex installation task.

3.2.2 Equipment

Throughout the research process of data collection and analysis, a range of hardware and software were used. The furniture was a Corner 5-Shelf Open Bookcase (Beech) and a Mainstays Home 8-Shelf Bookcase (Multiple Finishes). These were bought online via Wal-Mart. Each of these furnitures came with an installation manual of its own, either booklet or printed on the packaging of the furniture. A Philips screwdriver was used along with an electric drill. The data during the first phase - Non-AR - was collected using Eye Tracking Glasses (ETG, SensoMotoric Instruments, Inc., Boston, MA 02110 USA). The video data was collected by cameras (DVR4-4350™ Digital Video Recorder, Swann Communications U.S.A. Inc., Santa Fe Springs, CA 90670 USA) throughout the experiment. The camera was mounted on a tripod. Video data was also collected for 2 subjects using Microsoft Kinect™ for Windows. A computer (OptiPlex 7440 All-in-One, Dell Inc., Round Rock, TX) installed with Windows 10 Operating System was used for development and analysis purpose. In the AR phase of the experiment, an AR enabled headset (HoloLens™, Microsoft Corporation, Redmond, WA) was used. In order to capture the video from the HoloLens™, its built-in interface was accessed using a web browser (Chrome, Google, Mountain View, CA).

Data collected using the ETG was analyzed using SMI BeGaze™ software. The data exported from this software was stored into a Comma Separated Value (CSV) format using Microft Excel

and then analyzed using R Studio and Spyder. The 3D models were created using a free and open 3D creation software (Blender, Blender Foundation, Amsterdam, The Netherlands). The application was developed using a game development engine (Unity 3D 5.4, Unity Technologies, San Francisco, CA) and a toolkit (HoloToolkit, Microsoft Corporation, Redmond, WA) that allowed easier integration of HoloLens™ in the Unity 3D engine. The videos captured were converted into the appropriate format for analysis using a video converter (DVD Video Converter, WondexFox). The video analysis was done using a video editing software (After Effects, Adobe Systems, San Jose, CA).

3.2.3 Subjects

The experiments conducted involved human subjects so a new IRB protocol was filed. A Informed Consent form and an Audio Visual (AV) Consent form was sent for approval to IRB along with the Protocol form, Debriefing statement and the Study Protocol and Instruments form. The Informed Consent form has the details about the experiment such as, the purpose of the study, what is the person supposed to do, risks associated, time, confidentiality and participation. The AV Consent form were required as the subjects were audio video recorded during the experiment. This form gives option to the subjects to select if they want their data to be used in presentations or only for analysis. The Debriefing Statement contained information about AR technology and how the subject's contribution will help the subject recruiters in their research. This statement was read after the subject finished the experiment. All of these forms were approved by IRB and the protocol number '#2017-0034-00' titled 'Investigation of Cognitive Gains in Object Assembly using Augmented Reality' was accepted for a period from February 9, 2017 to February 8, 2018.

Subjects were recruited after the protocol was accepted. A total of 14 people - 12 male and 2 female - were recruited to perform the experiment. Thus, there were 7 instances in each of the task-type and complexity section. The subjects were undergraduates and graduates at the University of Virginia. The average age of participants was 24.5 years (SD = 3.1).

3.2.4 Design

As explained in the section 3.2.1, there are 2 independent variables with 2 levels each. Complexity with 'Simple' and 'Complex' as 2 levels, and Technology with 'Non-AR' and 'AR' as 2 levels. In order to negate the learning bias, 'between subjects' subject design was chosen. The dependent measures include time of assembly, errors occurred, cognitive load and spatial understanding. The allocation of the task to a subject was completely randomized. The experiment was designed for 2 phases, according to the independent measures. The phases of the experiment are explained in the following sections.

3.2.4.1 Phase One - Non-AR

In the first phase of experiments, the subjects were asked to install the furniture using the manual provided with the furniture. The manual for the simple task was printed on the box of the furniture. So, an image of it was printed and given to the subjects (refer figure 3.4). The complex task had a printed manual (refer figure 3.5) that had instructions in 2 languages, English and Spanish. All the subjects used the English instructions.
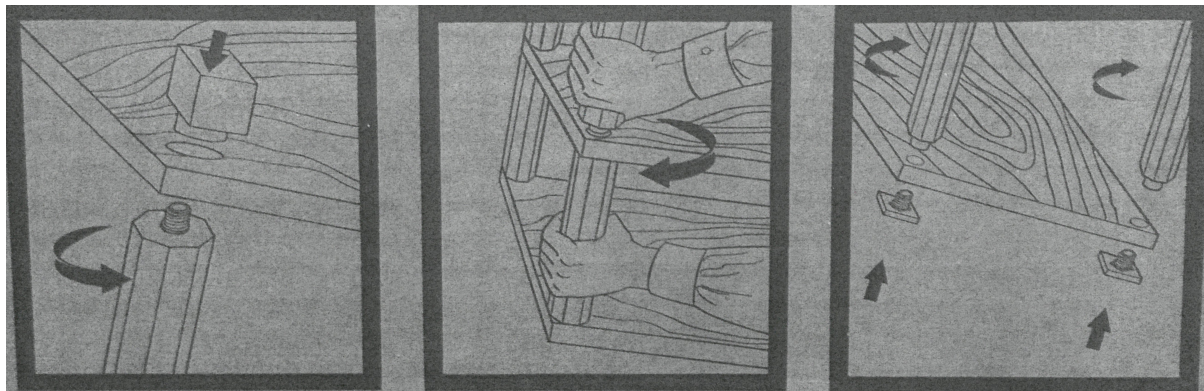
FIGURE 3.4. Simple task manual.

During the phase one of experiment, the subjects had to wear SMI's Eye-Tracking Glasses (ETG). The subjects gaze during the entire installation was captured through the ETG. In order
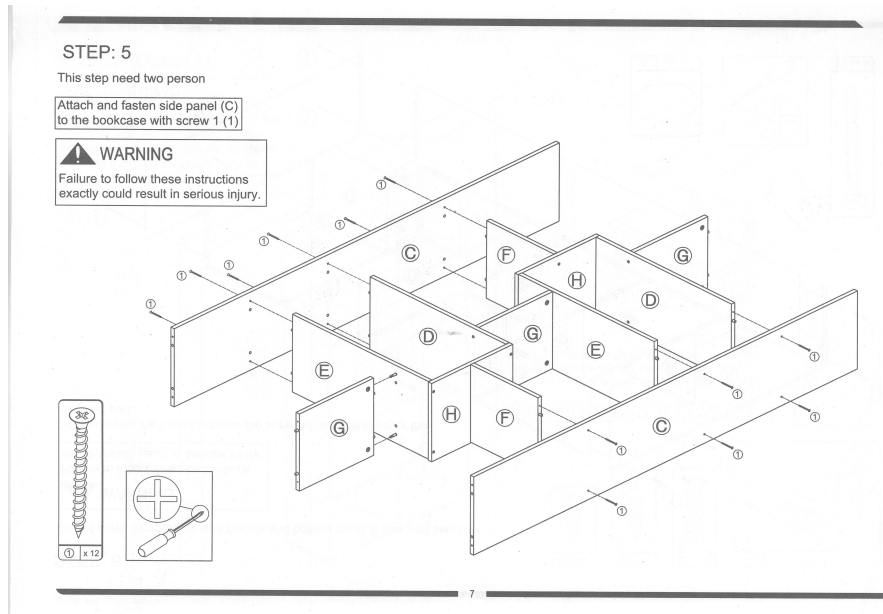
FIGURE 3.5. Complex task manual.

to analyze the user's actions throughout the installation a taxonomy of his/her intent was created. This taxonomy was built on a consensus among the four analysts who worked with the data. This was done:

1. To minimize the chance to classify a sub-task into multiple categories by using elementary actions.

2. To clarify the start and end point of each action where each point can be objectively determined based on observation, rather than conjecture.

The taxonomy helped in the analysis of the collected data through ETG and understand which areas of installation does a subject spend the most time on, and the number of transitions between the different intents of the subject during the installation. The analysis of the data collected is given the chapter 4. The taxonomy of the intents of the user is as follows:

1. Viewing Surrounding (VS): When the subject is not looking at furniture or manual.

29

2. Verbal Questioning (VQ): When the subject asks questions.

3. Viewing Instruction (VI): When the subject is looking/reading the instruction manual.

4. Positioning of the Part (PP): When the subject grabs a part and this action continues until the part stabilizes on a space.

5. Examining Options (EO): When the subject looks at multiple parts and then makes the most observable action.

6. Examining the Part (EP): When the subject examines the part to figure out the orientation of the part while positioning the part (PP).

7. Joining (J): When the subject takes action to combine parts together.

8. Redoing (R): When the subject redoes a previously performed activity.

9. Reviewing (RV): When the subject views the outcome of previous actions.

10. Positioning Oneself (PO): When the subject shifts himself from one place to another.

11. Slip (SL): When the subject does an unintentional motion. It is usually one-time and considered as an error.

ETG records the first person view video. This video was used for the intent analysis. On completion of the installation, the subject was asked to fill out a questionnaire (Appendix A) and NASA TLX form (Appendix A). The questionnaire was created to understand if the rationale behind deciding the complexity of the task was justified. Also, it collected data about the subject's age, gender, experience with RTA furniture, subjective reasoning behind the ease of assembly, remarks about the AR app (if used) and feedback about the overall conduct of experiment. One of the important reasons behind creating this questionnaire was to understand the AR experience of the subjects. The NASA TLX is a multi-dimensional rating scale in which information about the magnitude and sources of six workload-related factors are combined to derive a sensitive

and reliable estimate of workload [45]. These ratings would correspond to the cognitive load during the object assembly task. The factors would help understand the effort, mental demand and frustration while using a paper manual vs using the app. This could help evaluate if the technology has advanced and can be accepted by people to help in once-a-while RTA assembly task.

3.2.4.2 Phase Two - AR

The phase two of the experiment includes subjects using Augmented Reality technology through Microsoft HoloLens™. This was the main phase of the thesis experiment wherein, the subjects installed the furniture using a Microsoft HoloLens™. There was a comparative study that was done with the data collected in the phase one of the experiment. The main purpose of this phase was to test the AR application. In this step, the subjects that installed the simple task in the earlier step, installed the complex task and vice-versa. The application aimed at reducing the installation time, errors made and the cognitive load of the subject. These were the hypotheses that were tested in this phase.

For this phase, applications were developed for each of the two RTA furniture installation and deployed on HoloLens™. No additional manual was given to the subject. The instructions were provided within the app. The app is explained in the section 3.4 in detail. HoloLens™was calibrated for all the subjects before starting the experiment. This calibration is done by the built-in app provided in the HoloLens™. The initial configuration of the furniture part was kept consistent with phase one. The installation process was video recorded as well as video from the HoloLens™was also captured. The analysis of the data collected is given in chapter 4. On completion of the installation, the subject was asked to fill out the same questionnaire (Appendix A) and NASA TLX form (Appendix A) as in phase one.

3.3 Mircrosoft HoloLens™

Microsoft HoloLens™as shown in figure 3.6 is the world's first fully untethered holographic computer. HoloLens™facilitates personal computing through holographic experiences to empower you in new ways. It blends cutting-edge optics and sensors to deliver 3D holograms pinned to the real world around you. HoloLens™has a Intel 32 bit architecture with TPM 2.0 support and a custom-built Microsoft Holographic Processing Unit (HPU 1.0) with 64GB flash memory and 2GB RAM. It has 4 environment understanding cameras and a depth camera. It has a camera to capture photo and HD video, 4 microphones, and 1 ambient light sensor. For human understanding, it has spatial sound, gaze tracking, gesture input and voice support. HoloLens™doesn't track the actual eye gaze of the person. This term is explained in the next section. HoloLens™supports a Wi-Fi 802.11ac, Bluetooth 4.1 LE and Micro USB 2.0 for connectivity. It has built-in speakers. The battery life of HoloLens™is 2-3 hours of active use. All of the above specifications are as listed on the Microsoft website [4].

FIGURE 3.6. Microsoft HoloLens™.

3.4   AR App

The goal of the AR application was to assist the user in the installation of the RTA furniture and reduce the time of installation, errors during installation and reduce the cognitive load. The app aimed to help in the spatial-problem solving ability and provide contextual information to reduce the information processing; thus allowing the user to make a decision quickly and accurately. The app was deployed on Microsoft HoloLens™ which enabled viewing of 3D holograms placed in the environment. In this app, 3D holograms of the furniture to be installed was shown. The user was guided through a step-by-step installation of the RTA furniture. In every step, parts of furniture or the assembly done in the previous step were shown. These 3D holographic parts animate,i.e., transform and rotate from their original position to the final position in the assembly. Microsoft HoloLens™ applications require the Universal Windows Platform. Application deployed on it is a Universal Windows Application. In order to make use of the Windows Holographic Application Program Interface (API), Microsoft recommends the applications to be built using Unity3D game engine[17]. Unity 3D is a cross-platform game engine developed by Unity Technologies, which is primarily used to develop video games and simulations for PC, consoles, mobile devices and websites [18]. Developers may also use DirectX and other Windows APIs to build their own engines. Microsoft has an online Academy [1] wherein there are tutorials to understand how to create application for Microsoft HoloLens™.

The HoloLens™ application development was divided into following three steps. These three steps are explained in detail in the next sub sections.

1. Furniture model creation

2. Unity application

3. Deployment on Microsoft HoloLens™

3.4.1 Furniture model creation

In order to display a virtual 3D object in real world, the object has to be modelled initially and imported into the scene, in this case, it was imported into Unity3D. An AR application may or may not contain a 3D object. So, this step can be skipped if the application does not have a 3D virtual object. In many applications only textual information is displayed. In those cases, there isn't a need to create any model in a 3D modelling software. 3D text is already present in Unity3D as a built-in component. For this application, the software that was used for this 3D modelling was Blender. A screen shot of the Blender application's interface is shown in the figure 3.7. The dimensions of the furniture parts were measured and created to scale in Blender. The texture/color of the objects was also maintained so that the virtual objects resembled the real objects. The 3D models for the simple and the complex task are shown in the figures 3.8 and 3.9 respectively.

FIGURE 3.7. Blender software interface.
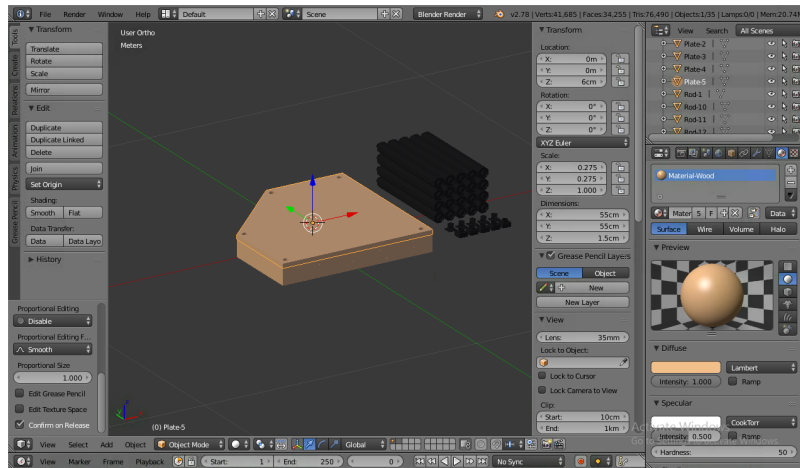
In the complex installation, the parts were addressed with alphabets, A to H. So, these letters were also inscribed in the 3D model. This was done for the user to identify parts distinctly. On the actual furniture parts the labels were too small to be seen from a distance. By embedding letter, the user doesn't have to depend on the features of the individual part like dimensions,
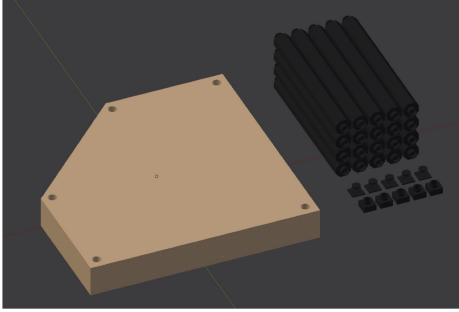
FIGURE 3.8. 3D Model of simple task.
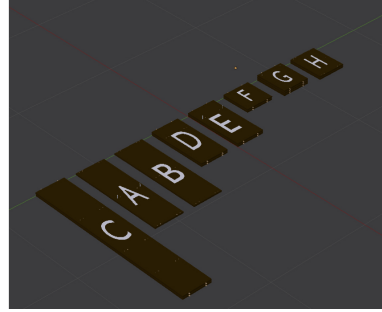
FIGURE 3.9. 3D Model of complex task.

number of holes, etc. For the simple task, the letters weren't embedded as the position of the similar parts didn't matter. All the flat plates, rods, screws and bolts were exactly identical to the other parts of their category. In the simple task, the rods were also modelled exactly similar to the actual part. The dimensions of the screw part of the rod were measured as well. All the measurements were then translated into the Blender software while creating the 3D models.

3.4.2 Unity3D application

3.4.2.1 Environment Setup

Microsoft has an extensive documentation on how to build applications for Microsoft HoloLens™ using Unity3D. Thus, promoting it as its default game development engine. The tutorials that are posted on their Academy site are also based on Unity3D. Also, Microsoft has created a HoloToolkit that can be directly imported as a package into Unity3D. Unity3D is an excellent development engine for applications to be deployed on Microsoft HoloLens™. Unity3D's ability to export the application for different platforms makes it a truly cross-platform tool and helps developers to create one application and build for multiple platforms a the same time. Its seamless integration and support from the widely used modelling tools makes Unity3D a popular choice. Unity3D provides developers tools to create content for applications that are deployed on Microsoft HoloLens™. In Unity3D, the 'scene' contains the objects of the game. It is an environment in

which objects that are shown in the game or application are placed. Figure 3.10 shows different steps that need to be done to make the scene HoloLens™ ready.
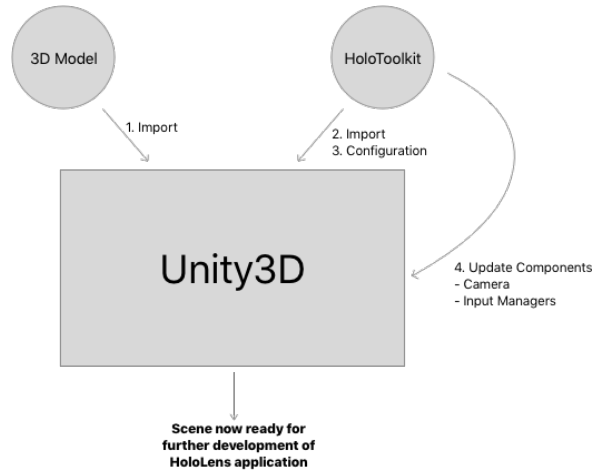
FIGURE 3.10. Steps taken to make the scene, HoloLens™ ready.

Step 1 is to import the 3D model into Unity3D that has been made in some modelling tool. Once imported, Unity3D treats it as an object in the scene. Step 2 is to import the HoloToolkit components into the project. Step 3 is to configure the project and the open scene for HoloLens™. This makes the project HoloLens™ ready. Step 4 is where the components are updated. Further details are explained in this section. This step enables components like camera, gaze and makes the scene input-ready. After these steps are completed, the scene is ready for further development. Step 3 and 4 need to be repeated for all the scenes in the project. Let us look at the above mentioned individual steps in detail.

The 3D model created in Blender can directly be imported as an asset into Unity. All of the materials that are used in Blender are also imported into Unity. Individual parts of the model can be separated in the Unity scene and can be treated as separate objects. This is useful if only a part of the entire model has to be transformed. A screen shot of Unity's interface is shown in the figure 3.11. Successful addition of the model into the scene completes the step 1.

The HoloToolkit is a collection of scripts and components intended to accelerate development
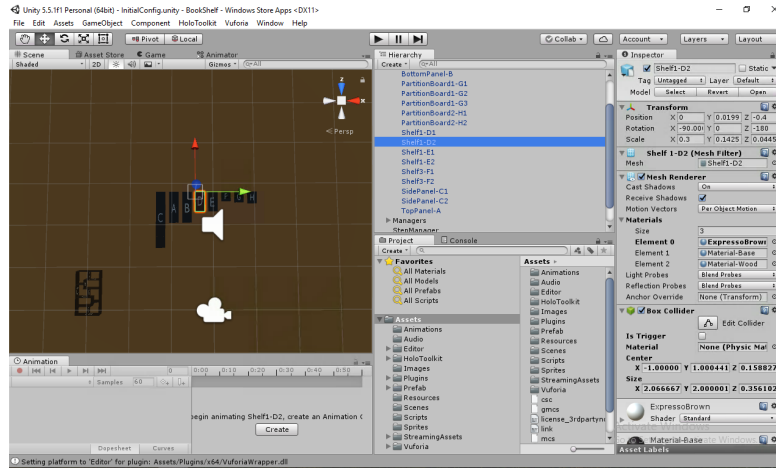
FIGURE 3.11. Unity software interface.

of holographic applications targeting Windows Holographic [7]. This toolkit contains features like camera, gestures, text, build setting, etc. that can be dragged and dropped into the Unity scene. In order to use HoloToolkit, it has to be imported into the Unity application's asset. HoloToolkit is imported in the project as an asset. Step 2 is completed when the asset folder contains HoloToolkit folder and the menu bar has a HoloToolkit menu. The next step - Step 3 - after importing the HoloToolkit into the application is to configure the application for HoloLens™. This can be done through the 'HoloToolkit' option added in the menu bar. The configuration includes Apply HoloLens Scene Settings (to the open scene), Apply HoloLens Project Settings and Apply HoloLens Capability Settings. These update the necessary settings that need to be done in the Unity environment.

Step 4 is to configure the essential objects like camera and input manager in the scene. This is done by removing the default camera in the scene and using the HoloLens camera. For this camera, the objects that are placed in the scene are rendered as holograms and the rest of the world is rendered black. Microsoft HoloLens™does not render anything on the area that is rendered as black in the scene. The application when run on a HoloLens simulator on Visual Studio is shown as in figure 3.12. The black area that is shown is the real-world and nothing is

rendered in this space. Microsoft HoloLens™uses the user's gaze as the cursor to navigate in the scene. Gaze is a like a ray that has been cast perpendicular to the HoloLens™™and is not the user's eye gaze. It is actually the head position. In figure 3.12, the gaze is shown in the form of a purple ring. This means that the HoloLens is ready to accept gestures. If the HoloLens was not ready to accept the gestures then the ring would be in the form of a dot. This gaze is the direction of the camera in the scene.
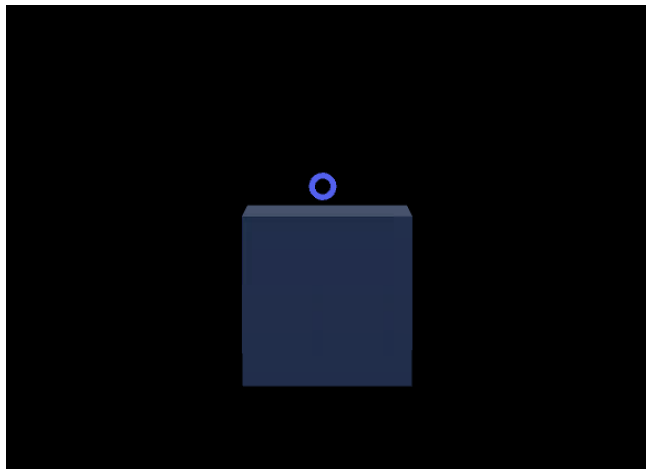
FIGURE 3.12. Scene through the HoloLens Emulator.

Part of the step 4 is to include Input Manager that includes interaction sources, gestures and event system. This components recognizes user gestures like Bloom, Air Tap, and Tap and Hold [20]. There is also a Keyword Manager that was used to accept voice commands from the user. This manager listens to the user's voice and then based on the keywords entered in the manager takes the appropriate action. For e.g., the word 'start' can be mapped to an action that has to be performed when the user says the word 'start'. The Keyword Manager has 2 recognizer start values - 'Auto Start' or 'Manual'. For the RTA application, the recognizer start value was set to 'Manual'. This allowed the voice commands only to be active when the user looked at a button, and this allowed the user to perform an action via voice command instead of Air Tap gesture.

3.4.2.2 Designing the application

The depth in 3D space corresponds to the Z-axis in the coordinate system of the HoloLens camera while the X-axis goes horizontal to the camera and Y-axis goes vertical. Figure 3.13 shows the optimal distance for placing holograms from the user. This guide was used to place the holograms in the scene. This completes the configuration of the scene and now scene is HoloLens™ ready.
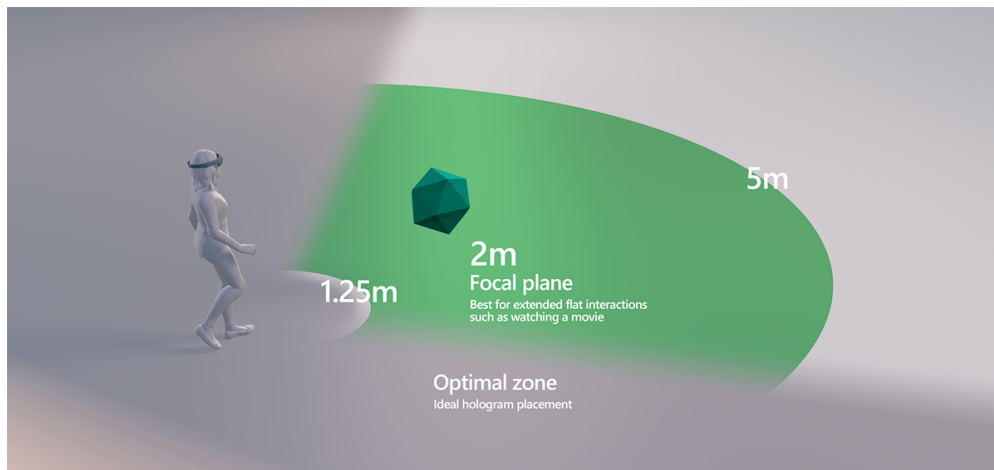
FIGURE 3.13. Optimal distance for placing holograms from the user. [3]

The actual design and development of the application starts from this point on-wards. The installation was divided into multiple steps depending on the manuals given with the furniture as well as the videos recorded during the phase one. Each of the steps was encoded in a scene. There was a initial configuration scene that explained how to use the application and showed the initial configuration of the parts of the furniture. The rest of the scenes included the individual steps. In these steps, only the parts of the furniture involved in the particular step were shown. This reduced the clutter in the surrounding and did not increase the complexity of the layout. These scenes included the animations that showed how the parts are finally oriented. In every step, holograms of the parts involved are projected as shown in figure 3.14. The placement of the hologram was mapped to the initial configuration.

Audio instructions were used instead of text. The subject could hear the instructions again by
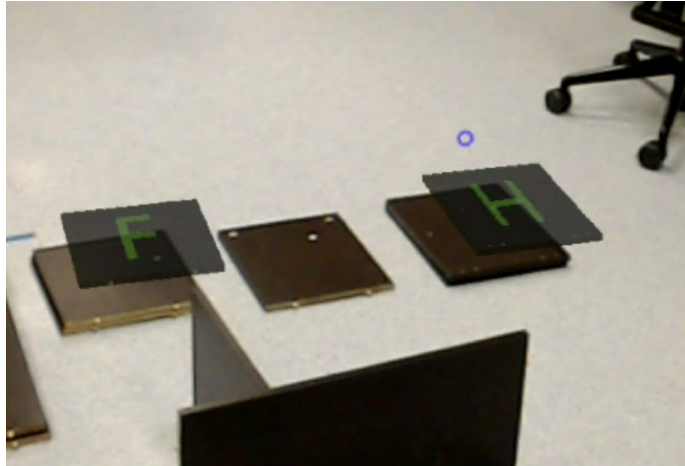
FIGURE 3.14. Placement of holograms of the parts involved in a step of installation.

click or by voice. The user could navigate from one step to the other step. Figure 3.15 shows the navigation panel shown to the user to navigate through the steps of the installation. This panel would let the user know which step was currently performed. Also, this panel allowed the user to hear the instructions again for the particular step by saying 'Repeat' by looking at the panel. This panel accepted voice commands. Whenever the user's gaze would be on the panel, it shows the voice commands below that can be used instead of Air Tap. This representation of a 'microphone' icon with the word besides it, is consistent with the other applications of Microsoft HoloLens™.

When the subject moves the gaze of the HoloLens™onto the part, it would animate to its final orientation. The animation was restricted to linear transformation along the X, Y or Z axes one at a time, and the rotation of the part at the same place. The transformation and rotation was decoupled to maintain simplicity of steps. These are explained in figures 3.16 and 3.17. The subject was able to view the animation again by clicking on the redo option, as shown in figure 3.18, or by voice commands. The start time of the animation as well as the duration of the animation was chosen so that user does not have to move his head quickly.

At all times during the installation, there is a hologram of the complete installation shown on the left side of the subject. The subject may drag the hologram closer to himself. On clicking
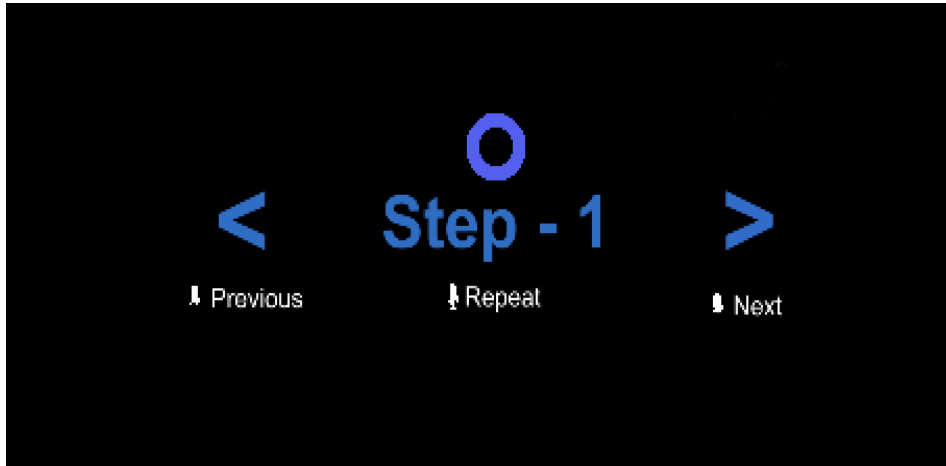
FIGURE 3.15. Holographic panel to navigate through the installation steps.
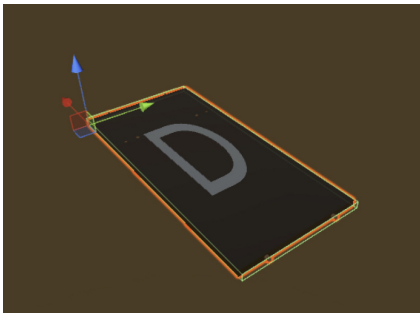
FIGURE 3.16. Transformation animations.
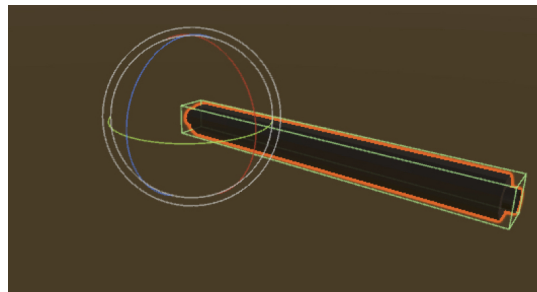
FIGURE 3.17. Rotation animations.

the hologram, the model bursts into a skeletal view of the product. This helps in analyzing the current state and the final state of installation. Figure 3.19 shows the final installation of the simple task and figure 3.20 shows the skeletal view of the installation when clicked on the hologram.

The scripts for the Unity application were written in C#. These scripts were mainly used to navigate through the scenes, access objects, control animations. These scripts allowed certain objects to be hide and shown on a particular action of the user. These also helped to start and stop the animations.
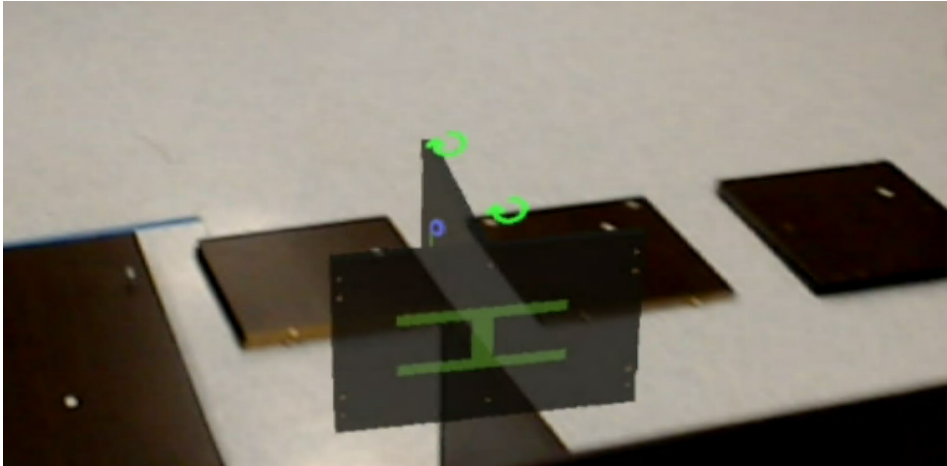
FIGURE 3.18. Redo option shown in the application.
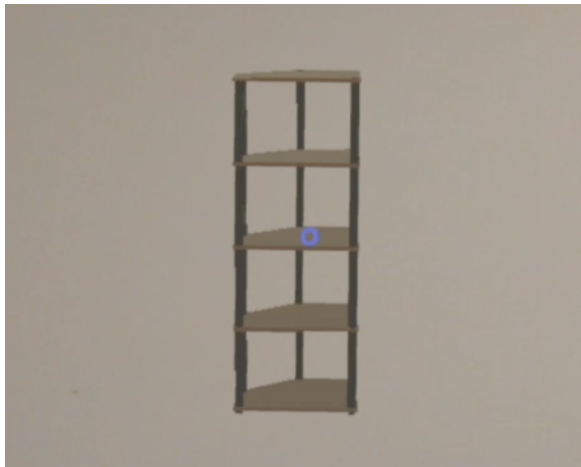
FIGURE 3.19. Hologram of the simple task.
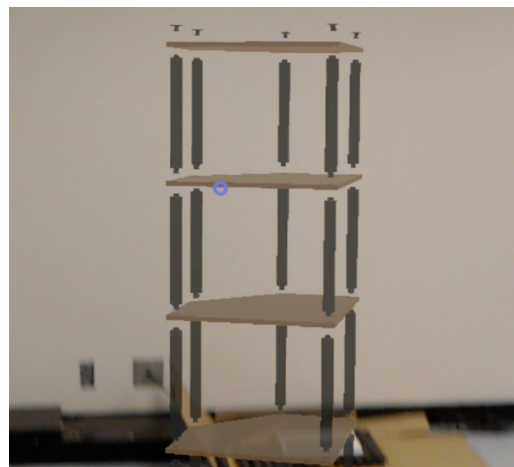
FIGURE 3.20. Skeletal view of the simple task.

3.4.3 Deployment on Microsoft HoloLens™

The HoloToolkit provides an interface as shown in figure 3.21, integrated into Unity, to build the application for Microsoft HoloLens™. The build creates a Universal Windows app project for Visual Studio that can be deployed on the Microsoft HoloLens™. This deployment happens remotely by entering the IP address of the Microsoft HoloLens™as it is considered a remote

machine during deployment. The app while deployed can be debugged using Visual Studio. During deployment, the app is installed on the Microsoft HoloLens™and can be accessed later on as an app without going through the Visual Studio solution. The app needs to be deployed again when any changes are done.
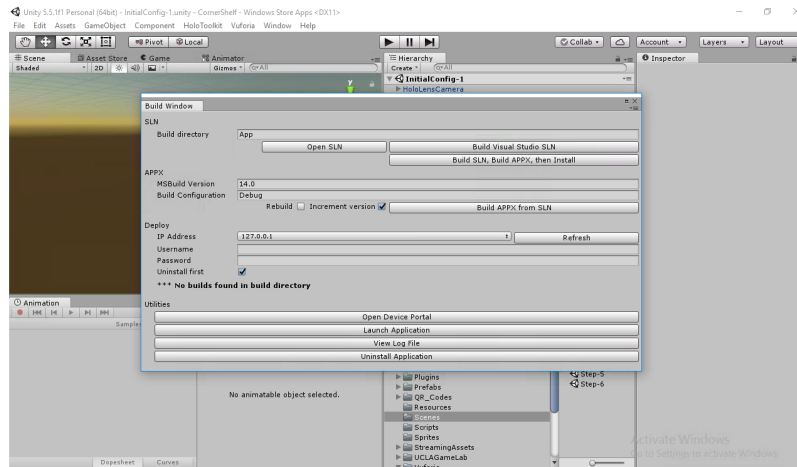
FIGURE 3.21. HoloToolkit User Interface in Unity.

ANALYSIS

4.1   Overview

In this section, we would analyze the data collected during the 2 phases of the experiment. The analysis is focused on testing the four hypotheses stated in the section 1.5; mainly completion time of installation, errors, spatial-problem solving ability and mental workload with and without using the AR application. To the test the hypotheses, the analysis section is divided into three parts - Behavioral Analysis, Cognitive Analysis and Questionnaire Analysis. Each of these sections uses the data collected in both the phases of the experiment and present comparative results. The analysis was mainly performed on the video data that was collected via the ETG, the video camera recorder and Microsoft HoloLens™, the questionnaires answered and the NASA TLX form filled by the participants after the experiments after both the phases.

The experiment was designed to measure and analyze the dependent measures - the time of assembly, errors occurred, cognitive load and spatial understanding. There were 14 subjects in all that included 12 males and 2 females. The average age of the sample was at 24.5 years (SD = 3.1). The subjects were named S01 to S16. Subjects S02 and S07 dropped after the first phase hence

the data for those subjects was excluded. Table 4.1 shows the task division between the subjects.

Table 4.1: Task allocation of subjects

Task Type	Simple	Complex
Non AR	S01, S03-S06, S08, S16	S09-S15
AR	S09-S15	S01, S03-S06, S08, S16

This allocation meant that the sample size for statistical analysis was 14. Due to the small sample size, a power analysis was performed to calculate the power for the specified sample size. Power analysis allows us to determine the probability of detecting an effect of a given size with a given level of confidence, under sample size constraints [11]. If the probability is unacceptably small, the experiment would need alteration or should be abandoned. For this analysis, the completion time using AR was chosen. The power was calculated to be 0.82, which meant there was 82% chance of seeing a significant effect by conducting the experiment. Before the selection of the test to report the significance of the statistical values, the data were analyzed for a few assumptions. The data consisted of independent samples since the comparison was between installation done different sets of people. Since the sample size is not large, the normality of the data was tested. Figures are attached in the Appendix A. The distribution adhered to the assumption of normality condition as it was close to normal. Therefore, 2 Sample t-test was chosen [16]. To cross check the findings assuming the data was not normally distributed, Mann-Whitney U test was also performed [6]. The other assumptions for the Mann-Whitney U test include the dependent variable be measured on an ordinal scale and the independent variable be two independent groups [6], which were also satisfied.

4.2 Behavioral Analysis

One of the most important issues of object assembly is the spatial-problem solving ability of the person. The other concerns discussed in the previous chapters of this document include the necessity to create mental representations as well as contextual human-information processing. All of these are associated with the user's actions in the entire duration of the assembly. Thus, in this section, the subjects' actions during the installation were analyzed with the help of the video data collected. This behavioral analysis is targeted towards testing the first three hypotheses reported in the section 1.5 which are completion time of assembly, errors, and spatial problem-solving ability. Completion time and errors are absolute measures which were calculated directly through the video. Spatial problem-solving ability was tied to intents of the user mentioned in the section 3.2.4.1.

Various studies have been conducted in order to measure the spatial abilities of a person [74] [46] [33], but most have association with mathematics. These metrics were not used in this experiment. Ekstrom et. al measured spatial ability a standardized paper-folding and card-rotation test [39]. A similar approach was used in this experiment. Spatial problem-solving ability was measured by the duration taken by the subject to place the part in its correct position. This duration was calculated by Positioning Part (PP) intent of the user. Creation of mental representations was mapped to the Reviewing (RV) intent of the user. Contextual human-information processing was calculated by the duration of Viewing Instructions (VI) of the user. Each of the metrics was considered in terms of percentage of the entire installation to give a better estimate. Table A.1 shows the above mentioned details while figure 4.1 shows the aggregate assembly completion time across tasks.

In the non-AR phase, on an average, the subjects took 8.14 minutes (SD = 1.21) to complete the simple task while they took on an average 14.89 minutes (SD = 1.43) to complete the simple task in the AR phase. The complex task took 43.19 minutes (SD = 15.33) to complete using the manual, while the subjects finished the complex task in 40.80 minutes (SD = 2) using the AR
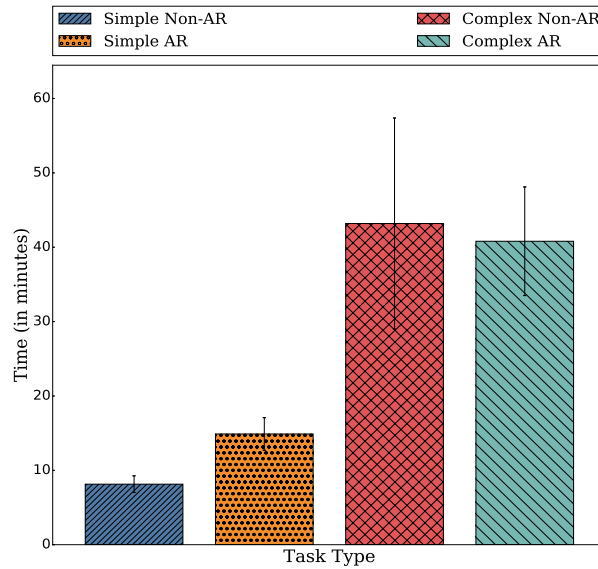
FIGURE 4.1. Assembly completion time across tasks.

application. Thus, subjects took almost 83% more time to complete the simple task and took about 6% less time to complete a complex task.

Table 4.2: Mann Whitney U test for Completion Time of Task

Task	*U-value*	*Z-score*	*p-value*
Simple	0	-3.07	0.002
Complex	21	-0.38	0.70

Table 4.3: 2 Sample t-test for Completion Time of Task

Task	*t-value*	*p-value*
Simple	-6.71	0.01
Complex	0.37	0.36

Thus, from tables 4.2 and 4.3, the p-values of simple task are less than 0.05 which means that the timings are significantly different. Thus, for the simple task, the AR application increased the time significantly. For $p < 0.05$, we can reject the hypothesis H_1 that *Augmented Reality application reduces object assembly time as compared to the time without the application*. Although the time taken was less for a complex task completed using the AR application, it wasn't statistically significant.
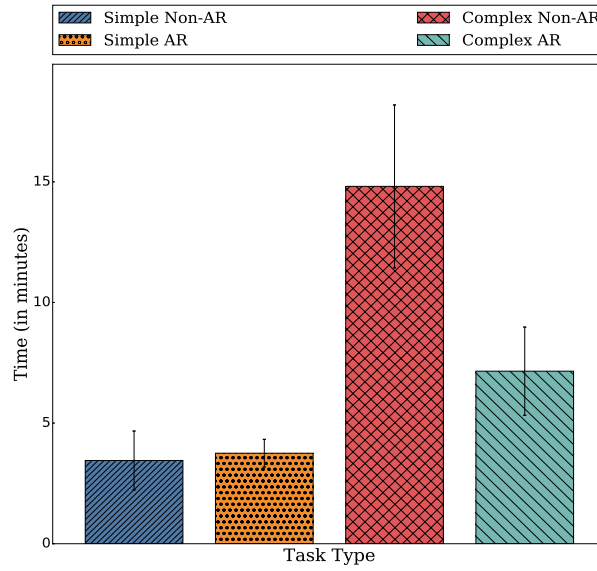
FIGURE 4.2. PP duration across tasks.

As shown in figure 4.2, while performing the simple non-AR task, subjects spent 3.45 minutes (SD = 1.32) (42%) of their total installation time on placing the part in the final location. While this duration was 14.81 minutes (SD = 3.41) (34%) for the complex non-AR task. When the subjects completed the task using the Microsoft HoloLens™, they spent 3.75 minutes (SD = 0.62) (25%) placing the part. It was seen that the time taken to place the part was almost the same in this case. Thus, having 3D holographic instructions didn't affect the subjects when it came to placing the parts in the simple task. There was a big difference when it came to the complex AR task. Subjects spent 7.15 minutes (SD = 1.97) (18%) of their time, positioning the part. This

duration is almost 52% less than the time taken during the non-AR phase. It can be stated that

AR app helped the subjects understand the position of the part and allowed them to make the

correct decision quickly.

Table 4.4: Mann Whitney U test for Positioning of Part Time

Task	U-value	Z-score	p-value
Simple	21	-0.38	0.70
Complex	1	2.94	0.003

Table 4.5: 2 Sample t-test for Positioning of Part Time

Task	t-value	p-value
Simple	-0.55	0.3
Complex	4.89	0.001

Thus, from tables 4.4 and 4.5 for $p < 0.05$, we can accept the hypothesis H_3 that *Augmented*

Reality application improves spatial understanding of the assembly for the complex task and

reject H_3 for the simple task.

The duration of the subject viewing the instructions (VI) was calculated. Figure 4.3 shows

that the subjects see instructions for 0.58 minutes ($SD = 0.22$) (7.14%) of the total installation

time during the simple non-AR task while spent 6.08 minutes ($SD = 1.63$) (14.09%) of the total

installation time. Since the Microsoft HoloLens™does not track the gaze of the subject's eyes, it

was not possible to divide the VI intent with the RV intent. The percentages, therefore, increased

to 3.24 minutes ($SD = 1.5$) (21.73%) and 7.49 minutes ($SD = 1.16$) (18.35%) for simple AR and

complex AR respectively.

The complex task was divided into eight steps that comprised of different operations to be

performed on various parts or the assembly done in the previous step. The transformations of
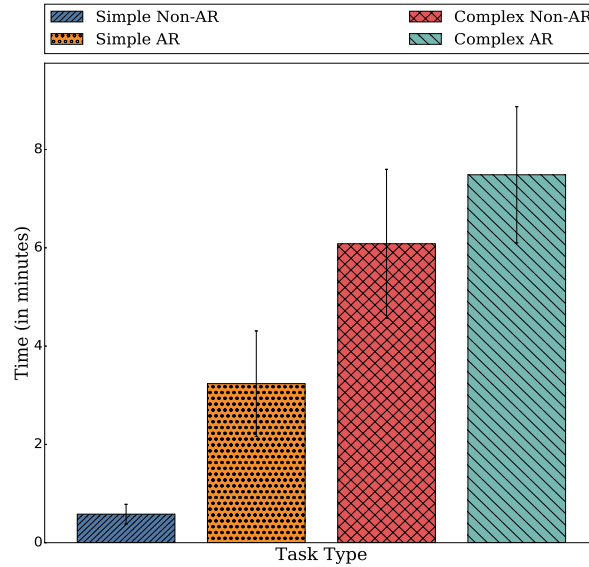
FIGURE 4.3. VI duration across tasks.

the components were critical as they had a unique position in the assembly. It was observed that subjects looked at the manual on an average 86 times (SD = 15.37) during the entire installation done with the help of a paper manual. This count reduced to an average value of 27 (SD = 5.79) for the AR app. The reduced count showed that the subjects didn't have to change their focus from the installation at hand to the manual over and over again. Subjects looked at the manual given on an average 10 times (SD = 1.90). On the contrary, while using the AR app to assemble the simple task, the subject was forced to listen to the instructions to complete the task. Aforementioned increased the overall time of the installation. Despite this, the number of times the subject looked at the instructions remained almost similar at an average value of 11 times (SD = 2.31). Figure 4.5 shows these counts. Figure 4.4 shows the transformations mapped to the 'Positioning of Parts (PP)' intent were 74 (SD = 18.62) for simple task and 23 (SD = 6.82) for complex task on an average during the non-AR phase of the experiment. The same intent was done 32 (SD = 10.34) times for simple task and 31 (SD = 1.73) times for complex task on an average when installation was done with the AR app. Thus, there wasn't much of a difference

between them.
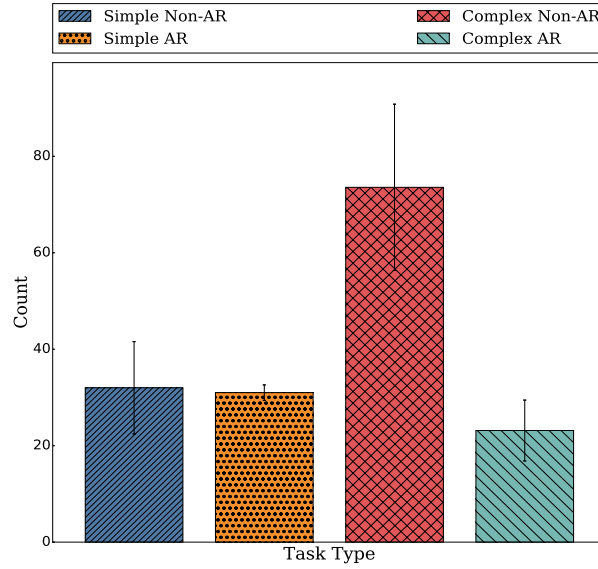
FIGURE 4.4. PP Count across tasks.

Table 4.6: Mann Whitney U test for Count of Positioning of Part

Task	U-value	Z-score	p-value
Simple	21	-0.38	0.85
Complex	0	3.07	0.002

Table 4.7: 2 Sample t-test for Count of Positioning of Part

Task	t-value	p-value
Simple	0.25	0.41
Complex	6.73	0.001

Thus, from tables 4.6 and 4.7 for *p < 0.05*, we can state that Augmented Reality application

helped in reducing the transformations needed for the complex task and not for the simple task.
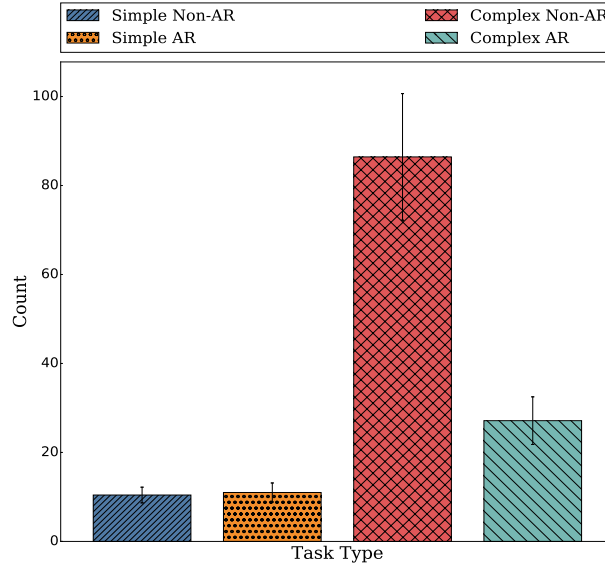
FIGURE 4.5. VI Count across tasks.

Table 4.8: Mann Whitney U test for Count of Viewing Instructions

Task	U-value	Z-score	p-value
Simple	22.5	-0.19	0.70
Complex	0	3.07	0.002

Table 4.9: 2 Sample t-test for Count of Viewing Instructions

Task	t-value	p-value
Simple	-0.51	0.69
Complex	9.55	0.001

Thus, from tables 4.8 and 4.9 for $p < 0.05$, we can state that Augmented Reality application

helped in reducing the mental representation effort needed for the complex task and not for the simple task.

Errors during the installation were defined as the operations made by the subject that resulted in redoing the operation on the same part. For, e.g., a subject placed a part in an incorrect orientation or an incorrect position and had to remove it from that location to put it in a different orientation or somewhere else. Tables A.2 and A.3 show the error count of the individual subjects in the phase one and phase two of the experiments while the figure 4.6 shows the error count across tasks. It can be seen that three subjects didn't make any errors during the simple non-AR task and the error count was at most 2 for rest of the subjects. While just one subject completed the complex non-AR task error-free and one subject made one error, the rest of the subjects made almost six errors on average. In the AR phase of experiments, the error count was reduced considerably. Three subjects didn't make any errors; three subjects made one error each, and one subject made seven errors while performing the simple task. There were no errors made by five subjects during the complex installation task, and two subjects made 2 and 12 errors.
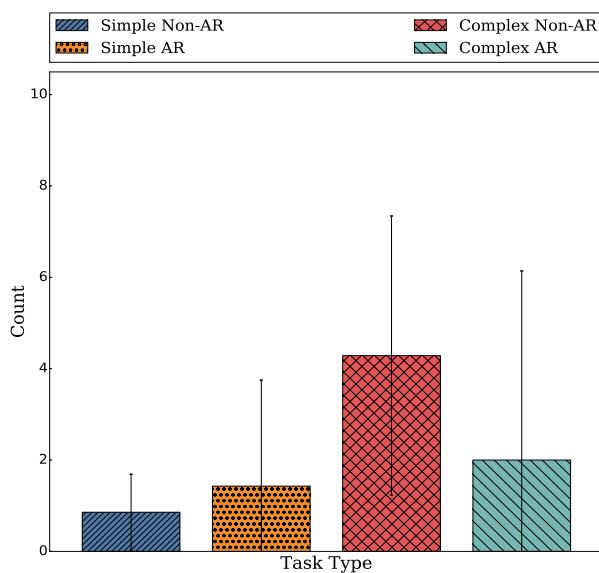
FIGURE 4.6. Errors made by subjects.

Table 4.10: Mann Whitney U test for Error Count

Task	U-value	Z-score	p-value
Simple	23.5	0.06	0.95
Complex	11.5	1.6	0.11

Table 4.11: 2 Sample t-test for Error Count

Task	t-value	p-value
Simple	-0.57	0.71
Complex	1.09	0.15

Thus, from tables 4.10 and 4.11 for *p < 0.05*, we can reject the hypothesis H_2 that *Augmented Reality application reduces errors during object assembly as compared to the errors without the application*. Although the count of the errors was reduced more for the complex task, the results aren't significant statistically.

4.3   Cognitive Analysis

Evaluation of mental workload is a key point to establish how efficient and viable a technology can be, especially when it is used as an instructional medium. To check the workload level of the current method of installation - using a paper manual - as well as to investigate if the AR application helps in reducing the current levels, Cognitive Analysis was done. This analysis tested the hypothesis H_4 that *Augmented Reality application reduces cognitive load during the assembly*.

The cognitive load of the task was measured using the NASA TLX ratings. After nearly 30 years of use, NASA-TLX has achieved a certain venerability and is being used as a benchmark against which the efficacy of other measures, theories, or models are judged [44]. Most of the

studies using NASA TLX addressed some sort of question about interface design or evaluation:
Visual and auditory displays, vocal and manual input devices, virtual or augmented vision [44].
The six factors included in the NASA TLX form are Mental Demand, Physical Demand, Temporal
Demand, Performance, Effort, and Frustration. Coincidentally, these dimensions also correspond
to various theories that equate workload with the magnitude of the demands imposed on the
operator, physical, mental, and emotional responses to those demands or the operator's ability to
meet those demands [44]. Each of these factors was given a weight of 1, and the total load was
calculated. Figure 4.7 shows the comparison between the average ratings for each of the factors
for all the tasks.
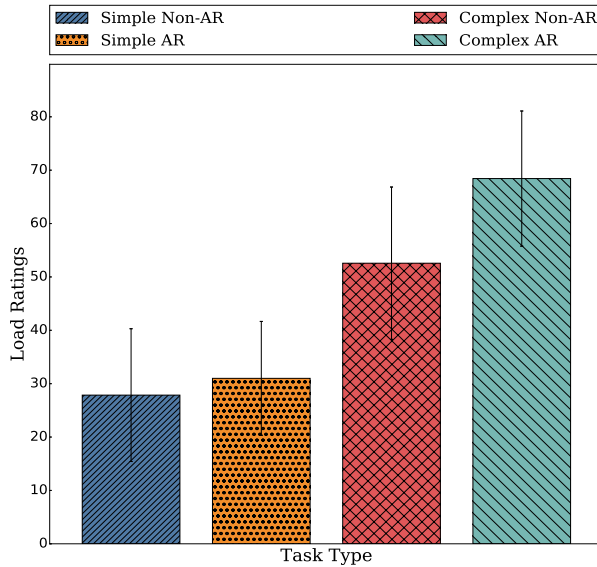
FIGURE 4.7. NASA TLX Ratings across tasks.

The average NASA TLX rating for the simple non-AR task was 28 (SD = 13.43) while that for
simple AR task was at 31 (SD = 11.52). Also, there was an increase in the NASA TLX ratings
for the AR task. The complex non-AR task rating was 53 (SD = 15.41) while that of the complex
AR task was 68 (SD = 13.67). For the simple task, the subjects marked the mental demand and

physical demand to have increased by a rating of 2, while the performance for all the subjects was 'Perfect' when it came to using the AR app. The other ratings remained more or less constant. The complex task was where the cognitive load increased by 30% than what it was for the simple task. The factors contributing to this rise were the mental demand, effort, and frustration as they increased by 4, 3.5 and 4.7 respectively. Also, the subjects marked their performance to have gone down by 4 points by using the AR application. The reasons for this increase in the ratings is explained in the next section where the responses of the subjects to the questionnaire are analyzed.

Table 4.12: Mann Whitney U test for NASA TLX Ratings

Task	U-value	Z-score	p-value
Simple	19	-0.64	0.52
Complex	10.5	-1.73	0.09

Table 4.13: 2 Sample t-test for NASA TLX Ratings

Task	t-value	p-value
Simple	-0.47	0.68
Complex	-2.04	0.97

Thus, from tables 4.12 and 4.13 for *p < 0.05*, we can reject the hypothesis H_4 that *Augmented Reality application reduces cognitive load during the assembly*.

4.4   Questionnaire Analysis

A set of questions was asked after the subjects performed the experiment in both the phases. The questionnaire was prepared to get a qualitative feedback and insights regarding the task and subjective responses regarding the assembly task and the AR app. The questionnaire was kept

same for both the non-AR and AR phases. Figures 4.8, 4.9 and 4.10 are some of the responses displayed graphically and the data from it are referred to in the following sections. The following sections discuss the questionnaire responses about the individual sections within the experiments.

4.4.1 Simple - Non-AR task

The difficulty of the task was rated at 1.5/5 (SD = 0.53), suggesting that the task was easy. Out of the seven subjects, four subjects agreed that the support materials given were helpful for the installation. Whereas, three subjects gave a 2.33/5 rating for the support material to be effective. In all, six subjects hardly referred to the manual according to them. The majority of participants felt that they were able to next steps and the repeated procedures were most helpful in executing this task successfully. Most of the suggestions were related to designing the manual. Those were:

- Manual: One of the participants suggested that it should be made clear in the manual which was the top or bottom screw. This suggestion was a good idea as one of the subjects did assemble the furniture upside down and didn't realize it. One of the subjects also suggested that a more detailed manual would have made the assembly easier.
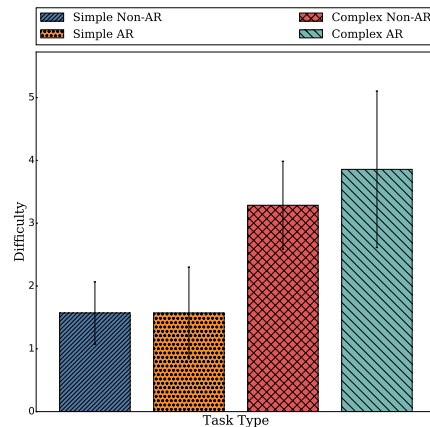
FIGURE 4.8. Task Rating given by subjects across tasks.

4.4.2   Complex - Non-AR task

The difficulty of the task was rated at 3.29/5 (SD = 0.76), suggesting that the task was difficult. Only one subject rated the task 2/5 while the other six subjects were divided equally between 3/5 and 4/5. Five subjects rated themselves 4/5 and above when asked about how frequently they referred to the manual, and almost all subjects felt that the printed support material was helpful. Aforementioned suggests that the task heavily relied on the manual that came along with the furniture. There weren't many suggestions for improving the assembly process when it came to this task. Suggestions included:

- Equipment: One of the subjects suggested that an electric screw driver should be given for the installation as there were a lot of screws that needed to be fastened.

- Parts: Color coded parts would have been easier to identify.
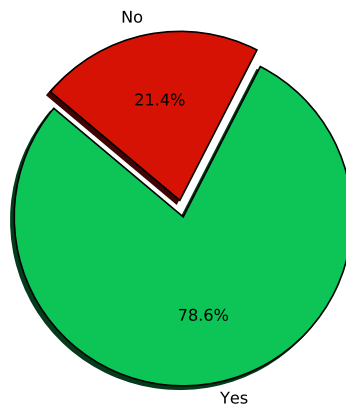
FIGURE 4.9. Subjects who had assembled RTA furniture before the experiment.

4.4.3 Simple - AR task

The difficulty of the task was rated at 1.6/5 (SD = 0.79), which was similar to the non-AR task. All of the subjects found the instructions that were given in the app extremely helpful. Subjects found the audio instructions and the animations helpful. Suggestions and feedback given by the subjects include:

- Voice Commands: Some subjects suggested that the voice commands should have been better.

- App: One of the subjects suggested that the app should have sensed the completion of the step and moved to the next step on its own.

- Hologram Alignment: Aligning the holograms with the actual parts was recommended by a subject.

- Part Identification: Since the parts of the furniture were black, the subjects found it difficult to locate the hologram as Microsoft HoloLens™ doesn't render black and treats it as real-world.

- HoloLens™ FOV: Overall, using the HoloLens™ made the installation easier, but the subjects did find a learning curve to it and said the head movement to view the animation was a little difficult as the view through the HoloLens™ was small.

4.4.4 Complex - AR task

The difficulty of the task was rated at 3.9/5 (SD = 1.35), which is greater than the non-AR task. Out of the seven subjects, four subjects found the instructions that were given in the app helpful and rated it higher than 4/5. One subject strongly disagreed that the instructions were helpful while one other rated 1/5 for the instructions to be helpful. There was a mixed response to
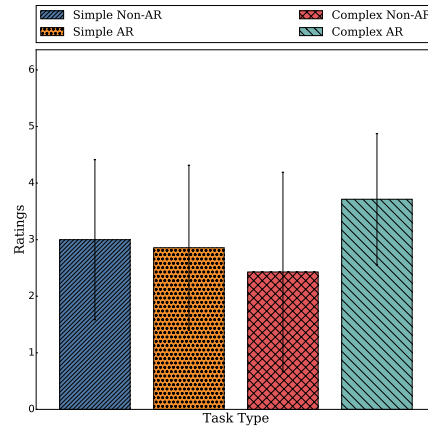
FIGURE 4.10. Ratings of subjects when asked if the prior assembly experience helped.

identification of the next steps in the installation. Suggestions and feedback given by the subjects include:

- Screw Placement: The main improvement suggested for the app was to show the placement of the screws instead of just the audio instructions.

- HoloLens™ FOV: Subjects found the head movement and small FOV difficult to manage.

- App: A few subjects suggested in detailing the installation steps even further. There was also a suggestion to give feedback by detecting if the performed action was correct.

- Hologram Overlay: Overlaying the holograms over the furniture was found to be a little confusing as the holograms did not overlay completely and were fixed relative to the origin (HoloLens™). This made the task a little more cumbersome.

All of these responses suggested that using the HoloLens™ for the first time made the task more difficult and certainly contributed towards increasing the NASA TLX ratings. Also, the adaptability to Microsoft HoloLens™ was a big concern seen throughout experiment. As shown in the figure 4.11, several subjects were seen struggling to get used to the FOV. The subjects

adapted to the Microsoft HoloLens™ as the installation progressed suggesting that there is a learning curve to the Microsoft HoloLens™ usage.
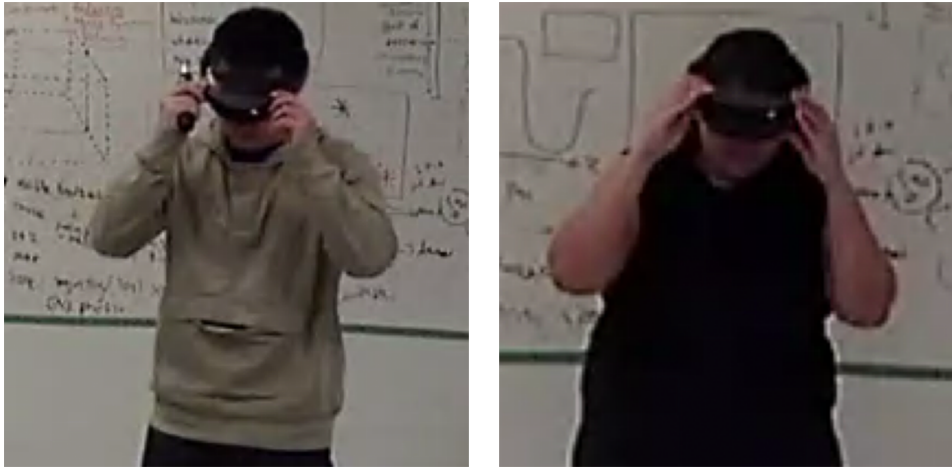
FIGURE 4.11. Subjects struggling with Microsoft HoloLens™FOV.

CONCLUSIONS

5.1 Overview

The goal of this thesis was to find a technical solution to reduce the cognition load on humans in the object assembly task. The motivation behind it was the lack of spatial problem-solving ability and cognitive load on humans [63]. Also, due to lack of research, there is only anecdotal evidence [56] that people often find assembling self-assembly products complicated [63]. With the help of a comparative study, this thesis showed that Augmented Reality technology could be used for the assembly of a complex ready-to-assemble furniture. This chapter contains the outline of the findings from the analysis performed and discusses the future topics. This chapter also presents another important aspect of this thesis was the use of Microsoft HoloLens™.

5.2 Discussion

5.2.1 Complexity of the task

In the previous chapter, detailed statistical analysis along-with the analysis of the other qualitative data was done. The hypotheses were also tested. There were some expected and some unexpected findings of the conducted experiment. Each of the three analysis done was crucial and gave insights into the task of object assembly as well as use of Augmented Reality to complete the task. One of the important aspect in determining the applicability of AR in performing the task that emerged from this analysis was the task complexity. The complex nature of the task depends on the two types of inputs: required acts and information cues [76], which were evident in this experiment.

The Behavioral Analysis helped in analyzing the spatial problem-solving ability of the person by calculating the time taken to position a part correctly. The AR app assisted in improving the spatial problem-solving ability for the complex task and helped in reducing the time to place the part correctly, while for the simple task the duration increased. The length of the entire assembly also decreased for the complex task, and increased for the simple task. Thus, it can be concluded that the complexity of the task plays a major role in the applicability of AR technology. A task that is simple to perform and requires fewer instructions might be completed quickly if the subject didn't have to be instructed through an AR app. Whereas, a complex and lengthy task can be performed in a faster way through instructions via an AR app. There was a considerable decrease in the number of errors that were made using the AR application than using the paper manual. Just two subjects made errors in the complex AR experiment as opposed to five subjects making five errors on an average. The number of errors made in the simple task was almost equal.

The Cognitive Analysis was performed on the NASA Task Load Index ratings given by the subjects after they performed the experiment. The analysis proved that the cognitive load increased across both the tasks when it came to performing the task using HoloLens™. There was

an increase in the NASA TLX ratings for the simple task while the scores increased by a greater percentage for the complex task. The main factors were mental demand, effort, and frustration that contributed to a higher rating for the AR phase.

The Questionnaire Analysis gave insights on why the cognitive load increased for the AR phase despite it helping reduce the duration and errors made by the subjects. Although people found the AR application to be handy, there were issues related to the particular aspects of it like the showing the placement of screws, efficient voice commands, and aligning the holograms with the real-world.

The analysis shows us that the augmented reality application proved to be more efficient for the complex task. It helped reduce time as well as the errors during the installation. The complexity of the tasks is explained in chapter 3, wherein factors such as time, symmetry, arrangement, transformations, patterns, the number of parts, etc. contribute to the complexity. Also, the number of steps given in the manual can be considered as an estimate of the complexity of the task. A simple task would comprise of repetitive small steps whereas a complex task would be a combination of multiple non-repetitive steps. In this thesis experiment, the simple task required only three different operations to be performed multiple times. Hence, once the subject understood the operations, there wasn't need to view instructions again. Whereas, the complex task consisted of multiple steps with different transformations in each step. The transformation count, as well as the time taken to do them, was reduced considerably using the AR app as compared to the paper manual for the complex task, while it was almost similar to the simple task.

With all these observations, it can be deduced that the complexity of the task mattered when it came to the effectiveness of the AR app. When the required actions were less and, lesser information cues were needed to identify the part, and its orientation, the complexity of the task was simple enough for the subject to complete it without additional assistance. These align with the complexity construct reported by Wood [76].

5.2.2   Cognition and Performance

There are multiple definitions of workload. These typically include the amount of "work" that is "loaded" on an individual, level of effort exerted, meeting of task requirements and physiological or psychological consequences. Thus, one individual's "workload" rating may reflect her assessment of task difficulty while others' might reflect the level of effort exerted [45]. So, it is subjective yet critical to understand the reported NASA TLX ratings reported in the experiment and find the relationship with other aspects of the experiment.

The unexpected results of this experiment were related to the relation of NASA TLX ratings and the performance (in terms of completion time and errors). It was observed that the cognitive load increased when it came to using the AR app despite reducing time and errors. Further analysis of this inverse correlation was done, and adaptability to Microsoft HoloLens™ was the causal factor. The increased effort to view things through the small FOV resulted in the increase of frustration and effectively the mental load among the subjects.

5.2.3   Design of the Augmented Reality App

The AR app was designed to map the mental model of the subject to the virtual objects shown in front as well as its resemblance to the real objects. The virtual objects were created to scale which allowed the subject to see the virtual installation exactly similar to the actual installation. The AR app showed the transformation of the parts from their initial configuration to the final position. Thus, the subjects understood the action that was needed to be done by them. While referring to the manual, the subjects had to process the information shown and map it spatially. But through the AR app, they were able to see the spatial transformation in front of them virtually.

There can be other ways to design the app. It would be interesting to find out the effectiveness of the AR app when the virtual objects are not to scale and are not overlaid on the real objects. The current app didn't include the positioning of the screws in the installation as it would have created more virtual clutter for the subject. There could also be a feedback loop that checks with

the user if the current step has been completed. There are multiple applications related to object assembly that have been created and deployed on Microsoft HoloLens™. Most of the applications are confined to one area of operation. Evans et al. evaluated HoloLens™by developing an assembly application [41]. This application was also table-top and restricted to a certain area. Although they didn't test the effectiveness of the application but presented user interface guidelines and best practices that can be followed while designing an AR application.

Helander and Willen [48] have suggested Human Factors principles for Design for Human Assembly (DHA) that target at minimizing perception time, decision time and manipulation time (refer table 2.1). For the thesis application, visual discrimination was considered to reduce perception time, which was done by embedding letters on the holographic parts. Decision time was reduced by placing holograms of parts on the actual parts and by showing transformation and rotation one at a time. One of the ways suggested was by easing the formation of the mental model, which was done by displaying the hologram of the final product at all times. Also, holograms of only the parts used in the ongoing step were shown. All of the above reduced the decision-making time as well as reduced the errors.

5.2.4   Augmented Reality Technology

The advantages of AR as an instructional medium can be seen in this experiment as it helped in the showing the complex transformations that may not be obvious on a 2D medium like paper, identification of parts and interactivity with the virtual objects being some of the critical points. There are still areas in which AR as technology needs to be improved and made robust. Image recognition is a significant factor for the effectiveness of AR. In this experiment, the image recognition wasn't used as it was found inconsistent during the early stages of development. Therefore, the virtual objects were placed in an absolute position and required the subject to adjust the installation according to the set locations. The other aspect of the AR technology that is an issue is the dependency on a device. Most of the devices do not have the capability

of spatially mapping the environment which is essential to position objects in the real world. Microsoft HoloLens™ is the only AR head-mounted display (HMD) that is capable of spatially mapping the environment.

Overall, this experiment helped in understanding the effectiveness of augmented reality as a technology in an object assembly task. Although it proved useful in a complex object assembly task, there are still areas in which work needs to be done. AR helped in showing spatial transformations of parts effectively. The cognitive load increased while using the AR app but improved the performance. Since the addition of the cognitive load was a result of Microsoft HoloLens™, further study can be done by using another device that overcomes the shortcomings of it. The experiment also creates a foundation for some further studies that can be done which are discussed in the next section.

5.3   Future Work

The most important aspect of this study would be to incorporate image recognition into the application and provide object tracking. Recognition and tracking would enable the user to move freely around the installation area. The study could also be done on a different headset with a larger FOV. Currently, Microsoft HoloLens™ makes the user move his/her head as the gaze is tied to the head movement. But a headset with the user's gaze tracker would allow a better usage of the FOV that is available.

AR applications can also be used in a different area other than RTA furniture assembly. The results show that the scale of the installation matters in the applicability of the technology. Thus, it is necessary to investigate even further by taking different use cases, such as installation wherein the area of operation is small, but the parts involved are small.

The effectiveness of interactive holograms can be tested by allowing the user to scale and move the holograms at run-time. These interactions can be improved by going beyond tap/click gesture, hold and drag gesture, and voice commands. More natural interactions like moving the

hologram just by sliding it aside with hand, grabbing a hologram as one would grab in real, etc. need to be incorporated. These will make the augmented instructions much more realistic and user-friendly.

Similar applications could be developed on the mobile platform to target a larger audience. Such applications may have a few shortcomings like there would be a need to hold the mobile device while performing the installation, the device may not be able to show a 1:1 scaled model due to display size restrictions, and interactions with holograms will be a challenge. The missing depth sensors on most of the mobiles will make spatial mapping challenging.

Despite the numerous usability issues of Microsoft HoloLens™, and concerns regarding the overlay and alignment of holograms, there is still potential for AR to provide a feasible solution to object assembly and other problems. Its capacity to improve spatial problem-solving ability and to reduce the mental representation effort is shown in this study. As the complexity of the task increases, AR shows its power to make life easier for us. As this technology matures in the next few years, the impact will be much more, and our life will undoubtedly be augmented.

APPENDIX A

This section contains the following additional documents that were used:

- Data Tables

- Normality Plots

- Questionnaire

- NASA TLX form

Table A.1: Duration for PP, RV and VI intents.

Subject	Positioning Part (PP)		Reviewing (RV)		Viewing Instructions (VI)	
	Time (in minutes)	% of total time (in minutes)	Time (in minutes)	% of total time (in minutes)	Time (in minutes)	% of total time (in minutes)
S01	2.37	30.28	0.75	9.64	0.44	5.64
S03	4.02	57.41	0.04	0.57	0.35	5.01
S04	1.95	25.19	0.20	2.58	0.58	7.43
S05	2.44	28.99	0.26	3.04	1.02	12.09
S06	3.52	51.57	0.00	0.00	0.64	9.38
S08	4.11	46.98	0.24	2.72	0.47	5.32
S16	5.74	55.18	0.04	0.41	0.57	5.52
S09	15.93	21.22	4.08	5.44	8.33	11.76
S10	14.77	34.43	1.84	4.29	7.83	18.26
S11	17.67	48.07	2.03	5.53	4.38	11.93
S12	19.50	40.55	3.13	6.50	5.84	12.15
S13	15.83	43.40	4.40	12.06	5.35	14.68
S14	10.60	31.53	1.87	5.56	5.56	16.55
S15	9.38	31.92	2.10	7.14	4.78	16.26

Table A.2: Errors made by subjects during the non-AR phase of experiment.

Task Type	Subject	Time (in minutes)	Error Count
Simple	S01	7.83	1
	S03	7.00	2
	S04	7.75	0
	S05	8.42	0
	S06	6.83	2
	S08	8.75	0
	S16	10.40	1
Complex	S09	65.08	5
	S10	42.90	4
	S11	36.75	0
	S12	48.10	10
	S13	32.48	6
	S14	33.62	4
	S15	27.38	1

Table A.3: Errors made by subjects during the AR phase of experiment.

Task Type	Subject	Time (in minutes)	Error Count
Complex	S01	40.25	0
	S03	25.65	2
	S04	39.52	0
	S05	38.10	0
	S06	45.77	12
	S08	47.32	0
	S16	49.02	0
Simplex	S09	19.38	1
	S10	14.05	1
	S11	13.12	0
	S12	12.43	0
	S13	15.97	7
	S14	13.55	1
	S15	15.73	0

Table A.4: NASA TLX ratings for Simple Non-AR task

Scale	S01	S03	S04	S05	S06	S08	S16
Mental Demand	3	3	1	1	5	4	14
Physical Demand	7	4	4	9	6	6	8
Temporal Demand	2	3	4	3	9	9	11
Performance	1	2	1	3	3	1	5
Effort	5	3	4	5	4	7	10
Frustration	3	3	1	1	8	3	6
Total	21	18	15	22	35	30	54

Table A.5: NASA TLX ratings for Complex Non-AR task

Scale	S09	S10	S11	S12	S13	S14	S15
Mental Demand	8	14	9	8	11	5	13
Physical Demand	10	10	13	8	9	13	16
Temporal Demand	9	17	8	10	7	10	10
Performance	5	1	3	1	1	1	4
Effort	11	15	14	5	5	13	16
Frustration	2	13	4	3	3	15	15
Total	45	70	51	35	36	57	74

Table A.6: NASA TLX ratings for Complex AR task

Scale	S01	S03	S04	S05	S06	S08	S16
Mental Demand	15	13	11	10	15	14	19
Physical Demand	13	7	11	16	11	11	17
Temporal Demand	10	3	1	10	16	7	13
Performance	5	1	9	4	10	6	9
Effort	16	13	15	14	15	16	15
Frustration	15	11	14	7	17	11	13
Total	74	48	61	61	84	65	86

Table A.7: NASA TLX ratings for Simple AR task

Scale	S09	S10	S11	S12	S13	S14	S15
Mental Demand	5	8	6	3	10	9	4
Physical Demand	7	8	8	4	15	13	4
Temporal Demand	9	3	3	3	9	6	10
Performance	1	1	1	1	1	1	1
Effort	2	8	3	2	8	10	4
Frustration	1	13	3	1	3	1	4
Total	25	41	24	14	46	40	27

Table A.8: Average NASA TLX ratings across tasks.

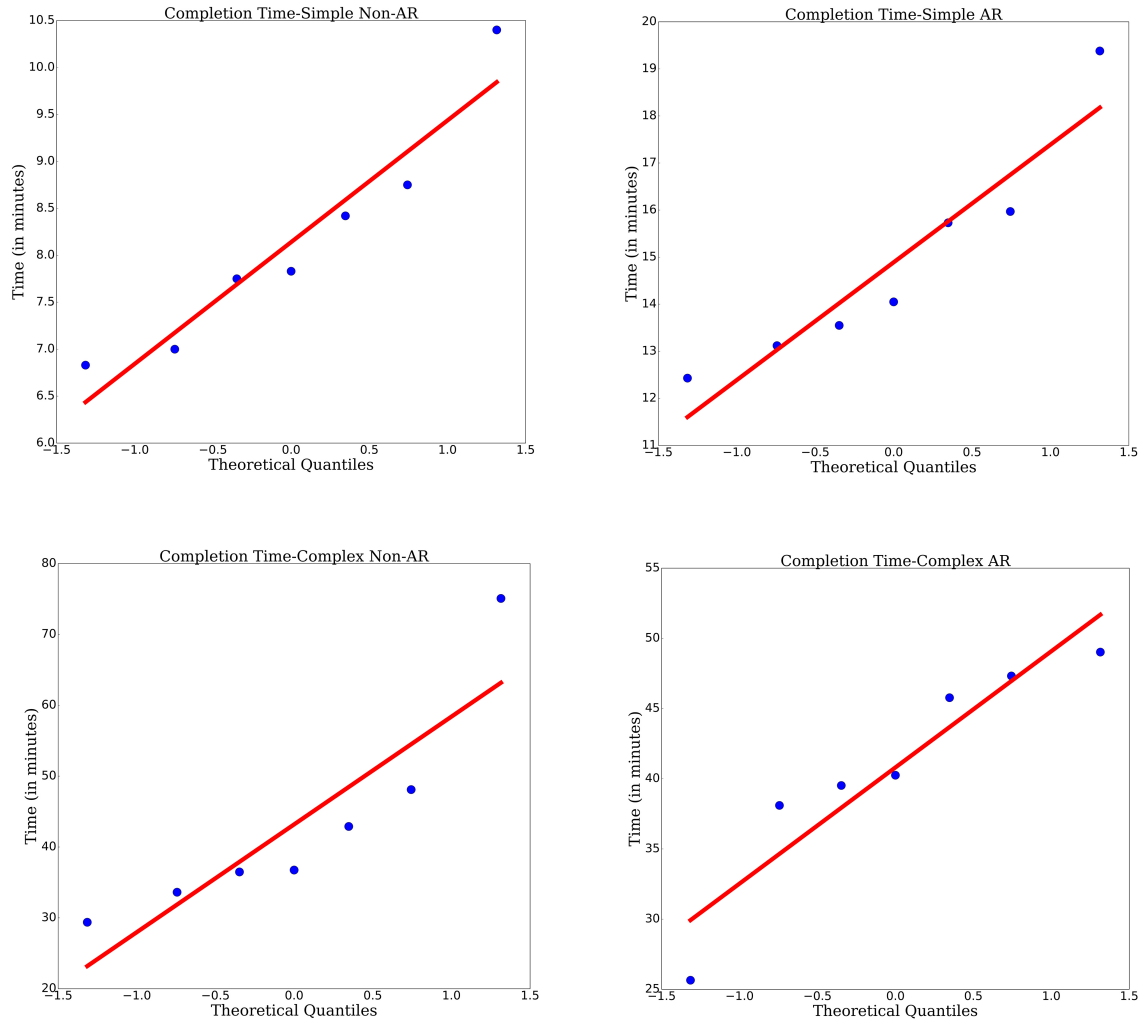| Scale	Simple Non-AR	Simple AR	Complex Non-AR	Complex AR
Mental Demand	4.43	6.43	9.71	13.86
Physical Demand	6.29	8.43	11.29	12.29
Temporal Demand	5.86	6.14	10.14	8.57
Performance	2.29	1.00	2.29	6.29
Effort	5.43	5.29	11.29	14.86
Frustration	3.57	3.71	7.86	12.57
Total	**27.86**	**31.00**	**52.57**	**68.43**

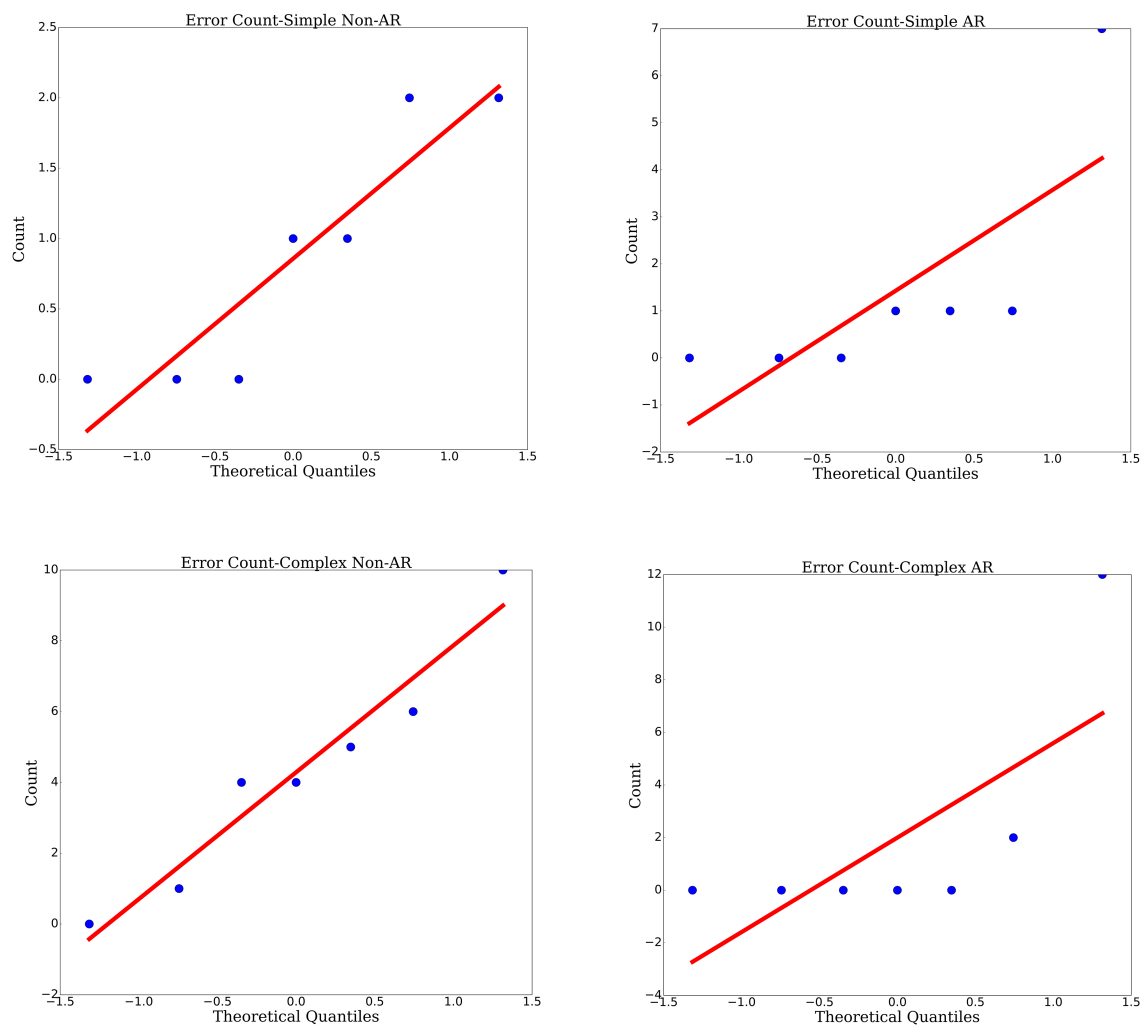FIGURE A.1. QQ Plot for Completion Time across tasks.
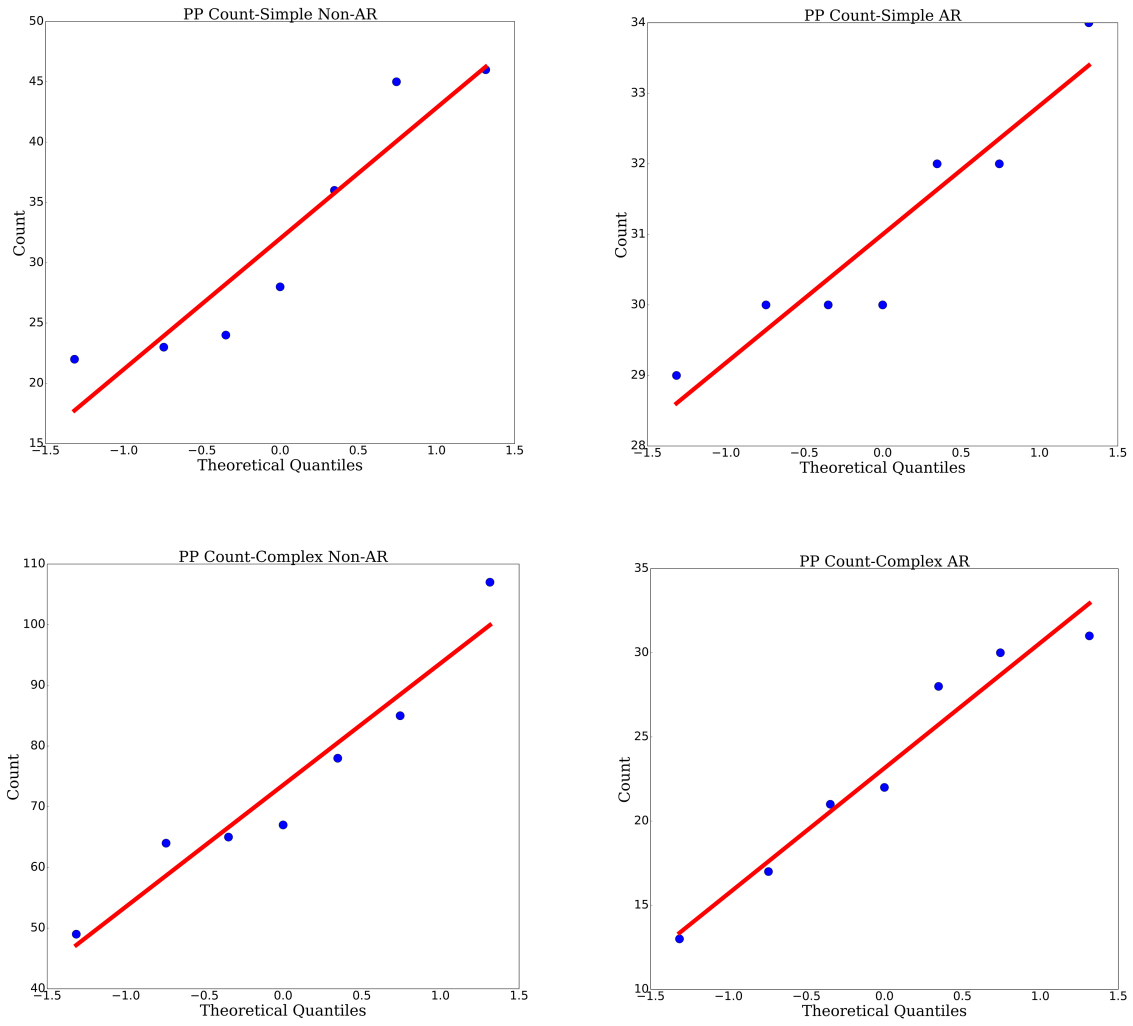
FIGURE A.2. QQ Plot for Error Count across tasks.

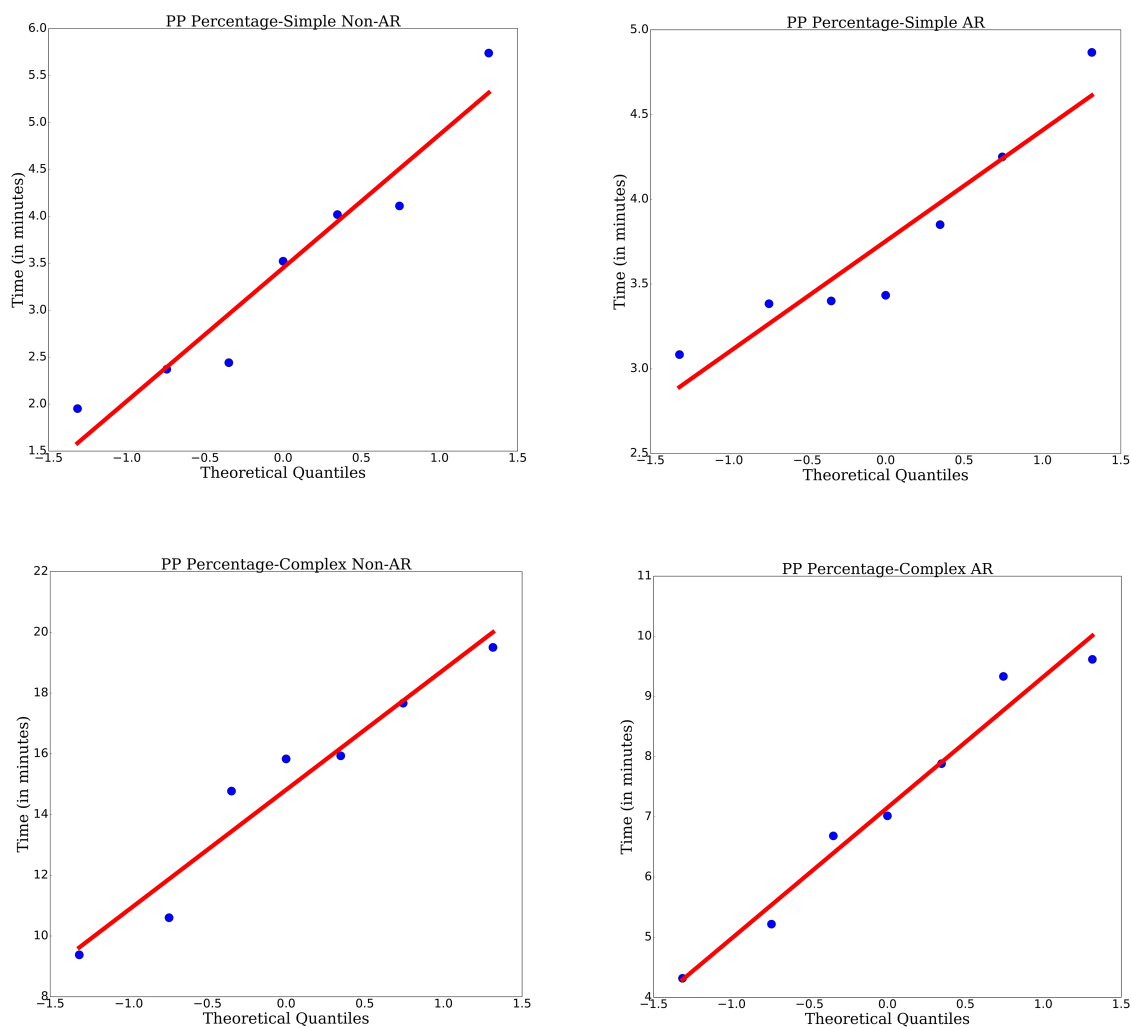FIGURE A.3. QQ Plot for PP Count across tasks.

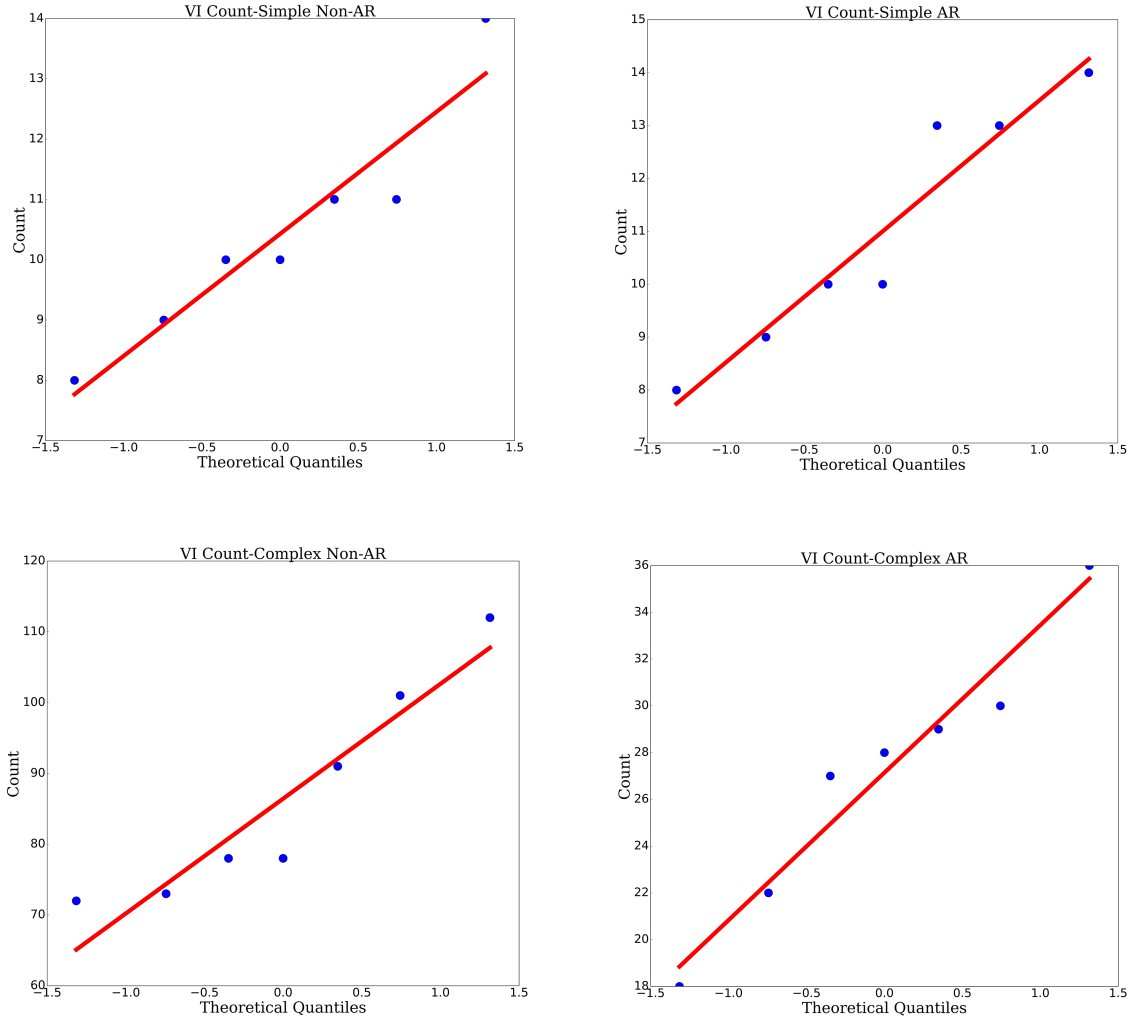FIGURE A.4. QQ Plot for PP Percentage across tasks.
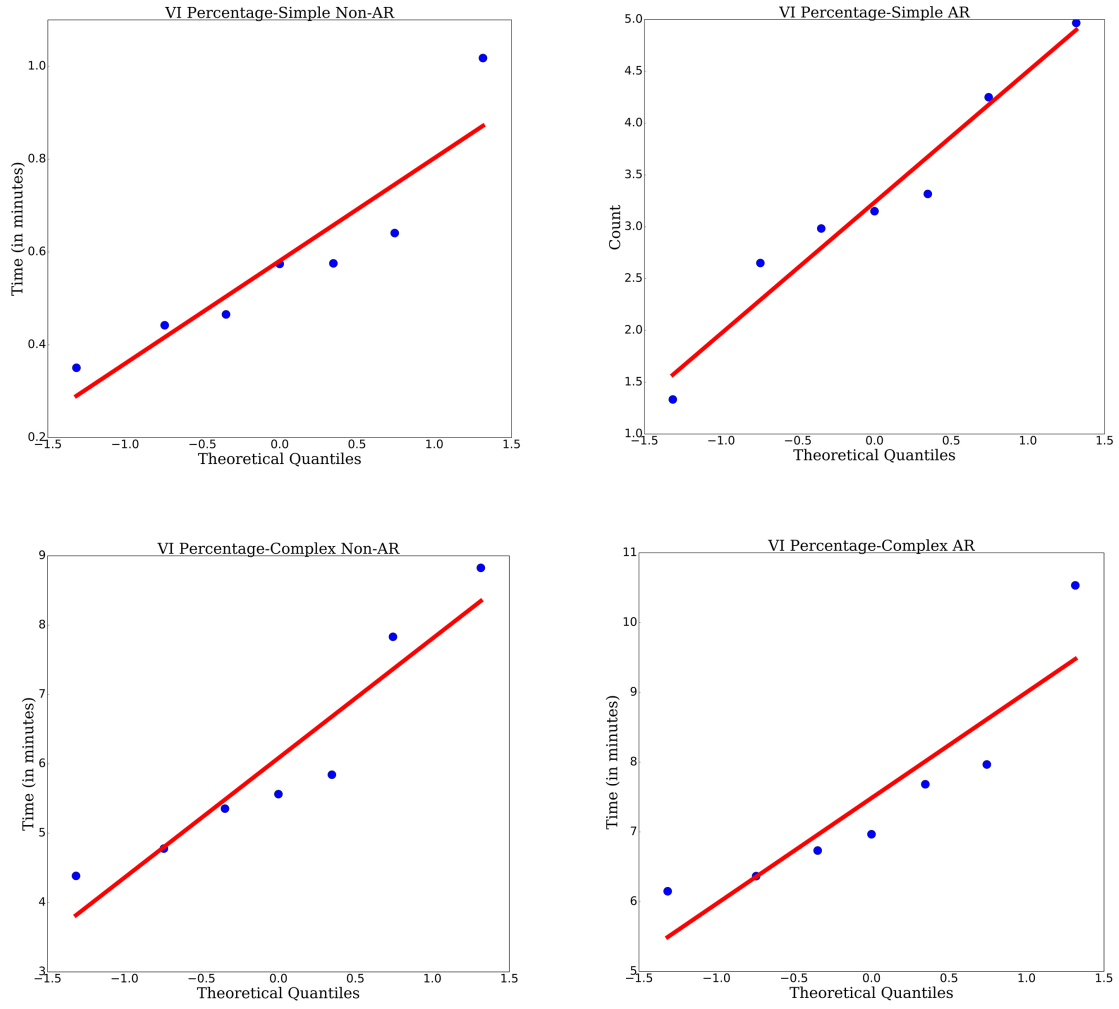
FIGURE A.5. QQ Plot for VI Count across tasks.

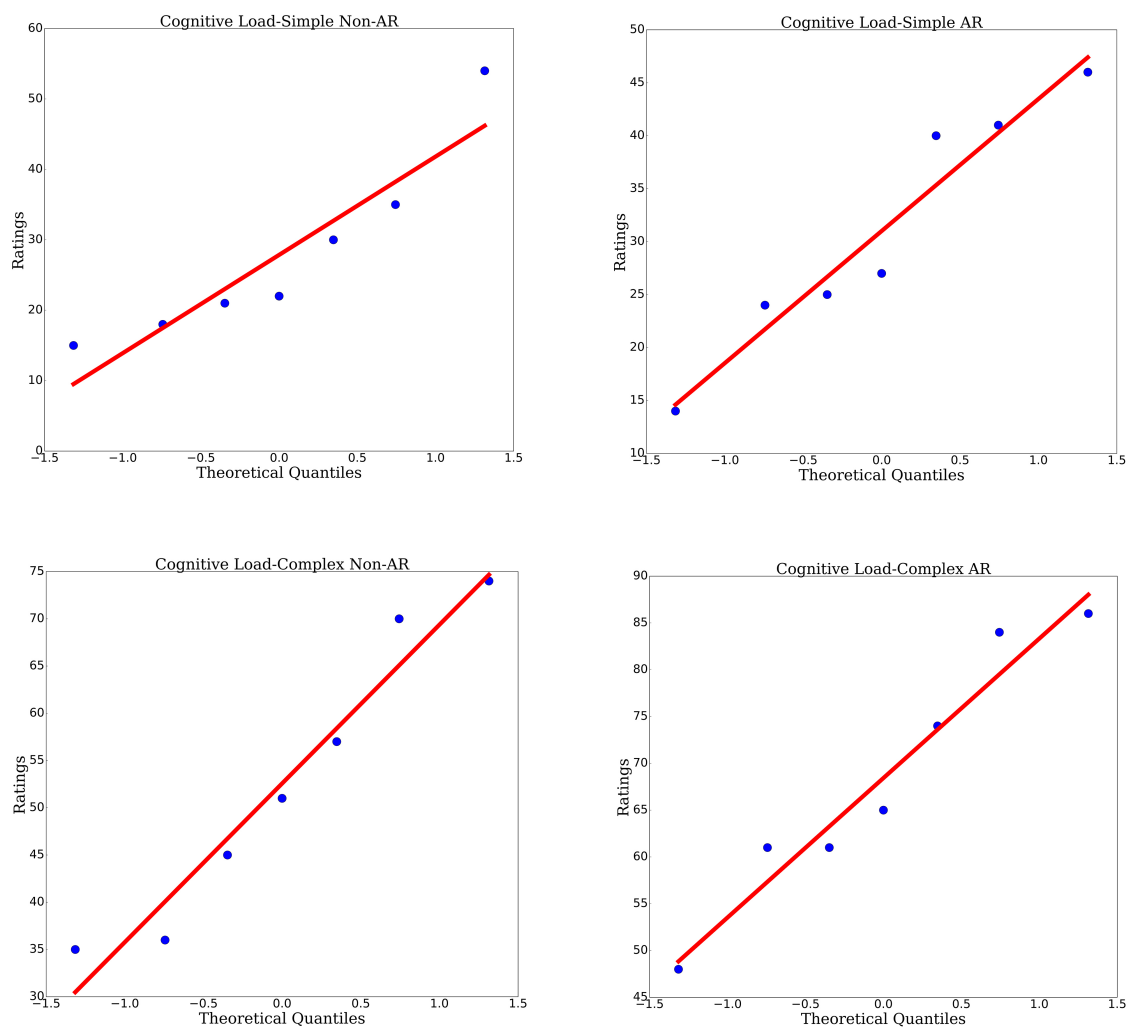FIGURE A.6. QQ Plot for VI Percentage across tasks.

FIGURE A.7. QQ Plot for Cognitive Load across tasks.

Post-Experiment Questionnaire

Subject Number: _____   **Gender**: [ ] Male   [ ] Female   **Age**: _____

Please answer the following questions:

1. Were the printed support materials helpful during assembly?
    Strongly Disagree   1   2   3   4   5   Strongly Agree

2. Did you refer to the manual frequently?
    Very Rarely   1   2   3   4   5   Very Frequently

3. Were you able to identify which part goes where?
    Strongly Disagree   1   2   3   4   5   Strongly Agree

4. Were you able to identify the next steps during assembly?
    Strongly Disagree   1   2   3   4   5   Strongly Agree

5. Were you able to identify repeated procedures during assembly?
    [ ] Yes [ ] No
    If yes,
    a. Did the repeated procedures help you with the assembly?
        Strongly Disagree   1   2   3   4   5   Strongly Agree

6. Have you assembled any ready-to-assemble furniture before?
    [ ] Yes [ ] No
    If yes,
    a. What did you assemble? (E.g., shelf, desk, chair, etc.)

        _____

    b. Did your prior experience help you with the assembly?
        Strongly Disagree   1   2   3   4   5   Strongly Agree

7. How will you evaluate the difficulty of the assembly you did just now?
    Easy   1   2   3   4   5   Difficult

    a. What steps of assembly made it easier/harder? (E.g., manual, placement of parts, size, etc.)

        _____

    b. What aid would have made the assembly easier?

        _____

8. Would you call a professional to assemble this?
    [ ] Yes [ ] No

9. How could the app have been improved?

    _____

10. Do you have any other comments, concerns, or suggestions?

    _____

FIGURE A.8. Post-experiment Questionnaire.

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 were you?

Very Low ———————————————— Very High

FIGURE A.9. NASA Task Load Index form.

83

BIBLIOGRAPHY

[1] *Academy*.

https://developer.microsoft.com/en-us/windows/mixed-reality/academy.

[2] *Augmented furniture - try furniture in real time, buy with confidence.*

[3] *Hologram stability*.

https://developer.microsoft.com/en-us/windows/mixed-reality/hologram_
stability.

[4] *Hololens hardware details*.

https://developer.microsoft.com/en-us/windows/mixed-reality/hololens_
hardware_details.

[5] *If you're not seeing data, you're not seeing | wired.*

[6] *Mann whitney u test*.

http://www.statisticshowto.com/mann-whitney-u-test/.

[7] *Microsoft/holotoolkit-unity: This is effectively part of the existing holotoolkit, but is the repo
that will contain all unity specific components.*

https://github.com/Microsoft/HoloToolkit-Unity.

[8] *Millennials and ready-to-assemble furniture | furniture today*.

http://www.furnituretoday.com/article/541341-millennials-and-ready-assemble-furniture.

[9] *Ori inbar (awe co-founder) opening keynote at awe 2016*.

[10] *Oxford dictionaries - dictionary, thesaurus, & grammar*.

[11] *Quick-r: Power analysis*.

http://www.statmethods.net/stats/power.html.

[12] *Ready-to-assemble furniture vs. fully assembled furniture | nbf blog*.

[13] *Report on the market for ready to assemble furniture | furniture world magazine*.

[14] *Rta (ready-to-assemble) furniture market us 2015-2019 | technavio - discover market opportunities*.

[15] *Text neck syndrome poster*.

[16] *Two-sample t-test: When to use it*.

http://www.statisticshowto.com/two-sample-t-test-difference-means/.

[17] *Unity development overview*.

https://developer.microsoft.com/en-us/windows/mixed-reality/unity_
development_overview.

[18] *Unity (game engine) - wikipedia*.

https://en.wikipedia.org/wiki/Unity_(game_engine).

[19] *Us furniture market report by fbic global retail tech feb 9 2016*.

[20] *Use gestures*.

https://support.microsoft.com/en-us/help/12644/hololens-use-gestures.

[21] *Viewar - augmented reality architecture & furniture specialized developers*.

[22] *What is mental representation? definition of mental representation (psychology dictionary)*.

[23] *With an eye on clinicians, atheer labs raises $8.8m for augmented reality glasses | mobi-healthnews*.

[24] *Worklink - scope ar*.

[25] *Three dimensions and more: augmented reality thrusts forward (metaio)*, 2014.

[26] J. R. ANDERSON, *Cognitive psychology and its implications .*, WH Freeman/Times Books/Henry Holt & Co, 1990.

[27] R. T. AZUMA, *A survey of augmented reality*, Presence: Teleoperators and virtual environments, 6 (1997), pp. 355–385.

[28] G. BAXTER, *Human factors and ergonomics in consumer product design: methods and techniques*, Ergonomics, 55 (2012), pp. 1124–1125.

[29] B. BELL, S. FEINER, AND T. HOLLERER, *Information at a glance [augmented reality user interfaces]*, Computer Graphics and Applications, IEEE, 22 (2002), pp. 6–9.

[30] M. BILLINGHURST, J. BOWSKILL, N. DYER, AND J. MORPHETT, *An evaluation of wearable information spaces*, in Virtual Reality Annual International Symposium, 1998. Proceedings., IEEE 1998, IEEE, 1998, pp. 20–27.

[31] T. BRAMORSKI, M. MADAN, J. MOTWANI, AND R. SUNDARRAJ, *Improving competitiveness of ready-to-assemble manufacturers through information technology*, Logistics Information Management, 13 (2000), pp. 201–209.

[32] J. CARMIGNIANI, B. FURHT, M. ANISETTI, P. CERAVOLO, E. DAMIANI, AND M. IVKOVIC, *Augmented reality technologies, systems and applications*, Multimedia Tools and Applications, 51 (2011), pp. 341–377.

[33] C. S. CARTER, M. A. LARUSSA, AND G. M. BODNER, *A study of two measures of spatial ability as predictors of success in different levels of general chemistry*, Journal of research in science teaching, 24 (1987), pp. 645–657.

[34] J. CASANOVA, *Re-thinking the image and assembly process of ready to assemble furniture*, Nov 2014.

[35] M. CHU AND S. KITA, *The nature of gestures' beneficial role in spatial problem solving.*, Journal of Experimental Psychology: General, 140 (2011), p. 102.

[36] L. A. COOPER, *The role of spatial representations in complex problem solving*, (1988).

[37] D. CRISPELL, *People patterns*, The Wall Street Journal. p. b1, (1992).

[38] B. L. DUE, *The future of smart glasses:: An essay about challenges and possibilities with smart glasses*, Working papers on interaction and communication, 1 (2014), pp. 1–21.

[39] R. B. EKSTROM, J. W. FRENCH, H. H. HARMAN, AND D. DERMEN, *Manual for kit of factor-referenced cognitive tests*, Princeton, NJ: Educational testing service, (1976).

[40] S.-H. EOM, S.-Y. CHOI, AND D.-H. PARK, *An empirical study on relationship between symptoms of musculoskeletal disorders and amount of smartphone usage*, Journal of the Korea Safety Management and Science, 15 (2013), pp. 113–120.

[41] G. EVANS, J. MILLER, M. I. PENA, A. MACALLISTER, AND E. WINER, *Evaluating the microsoft hololens through an augmented reality assembly application*, in SPIE Defense+ Security, International Society for Optics and Photonics, 2017, pp. 101970V–101970V.

[42] M. GRAHAM, M. ZOOK, AND A. BOULTON, *Augmented reality in urban places: contested content and the duplicity of code*, Transactions of the Institute of British Geographers, 38 (2013), pp. 464–479.

[43] J. G. GREENO, *Natures of problem-solving abilities*, Handbook of learning and cognitive processes, 5 (1978), pp. 239–270.

[44] S. G. HART, *Nasa-task load index (nasa-tlx); 20 years later*, in Proceedings of the human factors and ergonomics society annual meeting, vol. 50, Sage Publications Sage CA: Los Angeles, CA, 2006, pp. 904–908.

[45] S. G. HART AND L. E. STAVELAND, *Development of nasa-tlx (task load index): Results of empirical and theoretical research*, Advances in psychology, 52 (1988), pp. 139–183.

[46] M. HEGARTY AND D. WALLER, *A dissociation between mental rotation and perspective-taking spatial abilities*, Intelligence, 32 (2004), pp. 175–191.

[47] M. HEGARTY AND D. WALLER, *Individual differences in spatial abilities*, The Cambridge handbook of visuospatial thinking, (2005), pp. 121–169.

[48] M. HELANDER, B. WILLEN, W. KARWOWSKI, AND W. MARRAS, *Design for human assembly (dha)*, The Occupational Ergonomics Handbook, (1999), pp. 1849–1865.

[49] S. HENDERSON AND S. FEINER, *Exploring the benefits of augmented reality documentation for maintenance and repair*, IEEE transactions on visualization and computer graphics, 17 (2011), pp. 1355–1368.

[50] R. E. HENKEL, W. A. MATZKE, AND J. T. PHILLIPS, *Ready-to-assemble cabinet*, Sept. 2 1997.
US Patent 5,662,399.

[51] Z. HUANG, P. HUI, C. PEYLO, AND D. CHATZOPOULOS, *Mobile augmented reality survey: A bottom-up approach*, arXiv preprint arXiv:1309.4413, (2013).

[52] D. H. JONASSEN, *Toward a design theory of problem solving*, Educational technology research and development, 48 (2000), pp. 63–85.

[53] N. LI AND H. B.-L. DUH, *Cognitive issues in mobile augmented reality: an embodied perspective*, in Human Factors in Augmented Reality Environments, Springer, 2013, pp. 109–135.

[54] M. A. LIVINGSTON, *Issues in human factors evaluations of augmented reality systems*, in Human Factors in Augmented Reality Environments, Springer, 2013, pp. 3–9.

[55] D. F. LOHMAN, *Spatial ability: A review and reanalysis of the correlational literature.*, tech. rep., DTIC Document, 1979.

[56] J. MARGOLIS AND M. SEATON, *Working flat out*, The Sunday Times, (1996), p. 14.

[57] K. MAXWELL, *Augmented reality - definition of augmented reality, meaning of augmented reality, buzzword from macmillan dictionary.*

[58] A. NEWELL, H. A. SIMON, ET AL., *Human problem solving*, vol. 104, Prentice-Hall Englewood Cliffs, NJ, 1972.

[59] K. D. OWENS AND M. K. CLEMENTS, *Representations in spatial problem solving in the classroom*, The Journal of Mathematical Behavior, 17 (1998), pp. 197–218.

[60] F. G. PAAS AND J. J. VAN MERRIËNBOER, *Instructional control of cognitive load in the training of complex cognitive tasks*, Educational psychology review, 6 (1994), pp. 351–371.

[61] A. PAIVIO, *Mental representations: A dual coding approach*, Oxford University Press, 1990.

[62] S. PALMER, *Fundamental aspects of cognitive representation*, (1978).

[63] H. K. PILLAY, *Cognitive load and assembly tasks: Effect of instructional formats on learning assembly procedures*, Educational Psychology, 17 (1997), pp. 285–299.

[64] S. K. REED, *Cognition: Theories and applications*, CENGAGE learning, 2012.

[65] M. RICHARDSON, *Errors, accidents and self-assembly products*, CONTEMPORARY ER-GONOMICS, 2007 (2007), p. 307.

[66] M. RICHARDSON, G. JONES, M. TORRANCE, AND T. BAGULEY, *Identifying the task variables that predict object assembly difficulty*, Human Factors: The Journal of the Human Factors and Ergonomics Society, 48 (2006), pp. 511–525.

[67] A. SACCO, *Google glass takes flight at boeing*, 2016.

[68] H. SCHWEIZER, *Smart glasses: technology and applications*, 2014.

[69] V. L. SHALIN, G. V. PRABHU, AND M. G. HELANDER, *A cognitive perspective on manual assembly*, Ergonomics, 39 (1996), pp. 108–127.

[70] A. W. STEDMON, R. KALAWSKY, K. HILL, AND C. COOK, *Old theories, new technologies: cumulative clutter effects using augmented reality*, in Information Visualization, 1999. Proceedings. 1999 IEEE International Conference on, IEEE, 1999, pp. 132–137.

[71] R. SUNDARRAJ, M. S. MADAN, AND T. BRAMORSKI, *A customer-focus methodology for the manufacture of ready-to-assemble products*, International Journal of Operations & Production Management, 17 (1997), pp. 1081–1097.

[72] J. SWELLER, *Cognitive load during problem solving: Effects on learning*, Cognitive science, 12 (1988), pp. 257–285.

[73] A. TANG, C. OWEN, F. BIOCCA, AND W. MOU, *Comparative effectiveness of augmented reality in object assembly*, in Proceedings of the SIGCHI conference on Human factors in computing systems, ACM, 2003, pp. 73–80.

[74] D. VAN GARDEREN, *Spatial visualization, visual imagery, and mathematical problem solving of students with varying abilities*, Journal of learning disabilities, 39 (2006), pp. 496–506.

[75] E. VEAS, A. MULLONI, E. KRUIJFF, H. REGENBRECHT, AND D. SCHMALSTIEG, *Techniques for view transition in multi-camera outdoor environments*, in Proceedings of Graphics Interface 2010, Canadian Information Processing Society, 2010, pp. 193–200.

[76] R. E. WOOD, *Task complexity: Definition of the construct*, Organizational behavior and human decision processes, 37 (1986), pp. 60–82.

[77] F. ZHOU, H. B.-L. DUH, AND M. BILLINGHURST, *Trends in augmented reality tracking, interaction and display: A review of ten years of ismar*, in Proceedings of the 7th IEEE/ACM International Symposium on Mixed and Augmented Reality, IEEE Computer Society, 2008, pp. 193–202.