NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

THESIS

THE ROLE OF STEREOSCOPIC DEPTH CUE AND IMMERSION IN MAINTENANCE TASKS

by

Douglas Yamashita de Moura

September 2018

Thesis Advisor: Amela Sadagic
Second Readers: Rolf E. Johnson
Michael J. Guerrero

Research for this thesis was performed at the MOVES Institute.

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE			*Form Approved OMB No. 0704-0188*

1. AGENCY USE ONLY *(Leave blank)*	**2. REPORT DATE** September 2018	**3. REPORT TYPE AND DATES COVERED** Master's thesis

4. TITLE AND SUBTITLE THE ROLE OF STEREOSCOPIC DEPTH CUE AND IMMERSION IN MAINTENANCE TASKS	**5. FUNDING NUMBERS**
6. AUTHOR(S) Douglas Yamashita de Moura	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Postgraduate School Monterey, CA 93943-5000	**8. PERFORMING ORGANIZATION REPORT NUMBER**
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) N/A	**10. SPONSORING / MONITORING AGENCY REPORT NUMBER**

11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release. Distribution is unlimited.	**12b. DISTRIBUTION CODE** A

13. ABSTRACT (maximum 200 words)

Maintenance operations play a critical role in both civilian and military domains, and they can influence the state of their operational readiness. Thus, having access to superior solutions that can be used to train maintenance personnel is essential. Virtual reality (VR) technology, with its capability to simulate 3D objects with high fidelity, is a good candidate for maintenance training solutions. The main component of a large majority of maintenance tasks includes assembly and disassembly of physical setups. These tasks involve judgment of distance, depth, sizing and fit. Various factors may influence the operator's performance in a VR system by affecting perception of the components and, consequently, task execution. Two such factors are stereoscopic depth cue and immersion. This study uses assembly tasks as a context for exploring operator performance while manipulating virtual objects positioned within arm's reach. A user study collected a comprehensive data set over four distinct experimental conditions: immersive stereoscopic, immersive non-stereoscopic, non-immersive stereoscopic, and non-immersive non-stereoscopic. Data analysis suggests that the immersive stereoscopic condition was superior when compared to others; most people in that condition finished their assembly tasks, and they did it in shortest time. No significant simulator sickness issues were recorded in any condition.

14. SUBJECT TERMS maintenance, stereoscopic depth cue, immersion, assembly, virtual reality	**15. NUMBER OF PAGES** 175
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	**18. SECURITY CLASSIFICATION OF THIS PAGE**	**19. SECURITY CLASSIFICATION OF ABSTRACT**	**20. LIMITATION OF ABSTRACT**
Unclassified	Unclassified	Unclassified	UU

NSN 7540-01-280-5500 Standard Form 298 (Rev. 2-89) Prescribed by ANSI Std. 239-18

THIS PAGE INTENTIONALLY LEFT BLANK

Approved for public release. Distribution is unlimited.

THE ROLE OF STEREOSCOPIC DEPTH CUE AND IMMERSION IN MAINTENANCE TASKS

Douglas Yamashita de Moura
Captain, Brazilian Air Force
BCE, Technological Institute of Aeronautics, 2007

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN MODELING, VIRTUAL ENVIRONMENTS, AND SIMULATION

from the

NAVAL POSTGRADUATE SCHOOL
September 2018

Approved by: Amela Sadagic
Advisor

Rolf E. Johnson
Second Reader

Michael J. Guerrero
Second Reader

Peter J. Denning
Chair, Department of Computer Science

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

Maintenance operations play a critical role in both civilian and military domains, and they can influence the state of their operational readiness. Thus, having access to superior solutions that can be used to train maintenance personnel is essential. Virtual reality (VR) technology, with its capability to simulate 3D objects with high fidelity, is a good candidate for maintenance training solutions. The main component of a large majority of maintenance tasks includes assembly and disassembly of physical setups. These tasks involve judgment of distance, depth, sizing and fit. Various factors may influence the operator's performance in a VR system by affecting perception of the components and, consequently, task execution. Two such factors are stereoscopic depth cue and immersion. This study uses assembly tasks as a context for exploring operator performance while manipulating virtual objects positioned within arm's reach. A user study collected a comprehensive data set over four distinct experimental conditions: immersive stereoscopic, immersive non-stereoscopic, non-immersive stereoscopic, and non-immersive non-stereoscopic. Data analysis suggests that the immersive stereoscopic condition was superior when compared to others; most people in that condition finished their assembly tasks, and they did it in shortest time. No significant simulator sickness issues were recorded in any condition.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF FIGURES

LIST OF TABLES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF ACRONYMS AND ABBREVIATIONS

2D	two-dimensional
3D	three-dimensional
3DTV	3D-ready television
6DOF	six degrees of freedom
AR	augmented reality
CAD	computer auxiliary design
CAVE	Cave Automatic Virtual Environment
CITI	Collaborative Institutional Training Initiative
COTS	commercial-of-the-shelf
DoD	Department of Defense
FOV	field-of-view
HMD	head-mounted display
INS	Immersive Non-Stereoscopic
IR	infrared
IRB	Institutional Review Board
IS	Immersive Stereoscopic
LED	light-emitting diode
NINS	Non-Immersive Non-Stereoscopic
NIS	Non-Immersive Stereoscopic
NPS	Naval Postgraduate School
PC	personal computer
SDK	software development kit
SSQ	Simulator Sickness Questionnaire
STL	Standard Triangle Language or Standard Tessellation Language
SUS	System Usability Scale
TPF	Proposal Approval Form for Thesis & Reports
TV	television
VE	virtual environment
VR	virtual reality
VRTK	Virtual Reality Toolkit

THIS PAGE INTENTIONALLY LEFT BLANK

ACKNOWLEDGMENTS

First and foremost, I would like to thank my family for their love and support through the years. Marina and Carlos, I would not be here if you had not supported my dreams and put yourselves second to your children. Paula, thank you for our trips and for supporting me during my surgery recovery. Hugo and Raquel, thank you for our conversations and motivation.

I would like to thank my thesis advisor, Dr. Amela Sadagic, for her willingness to make the project better and for her hard work and patience in supporting me throughout the thesis process. Thank you for the conversations, ideas, insights, and motivation.

I would like to thank the Brazilian Air Force for the support and opportunity of undertaking this master's program and living in this amazing city of Monterey.

I would like to thank the MOVES FutureTech for the assistance with the development of the research application software. In particular, I would like to acknowledge my thesis second readers, Erik Johnson and Michael Guerrero, for their patience and help solving problems, and Ryan Lee, for the great work with the 3D models.

I would like to thank my cohorts for their support, help, and conversations throughout our time in MOVES. Beyond academic knowledge, I learned so much from all of you, and sharing experiences made this an invaluable period.

Finally, I would like to thank my friends for all the conversations, barbecues, game nights, and great moments we spent together. You made this a pleasant, worthwhile journey.

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION

A. RESEARCH DOMAIN

Over the last five years, we have witnessed emergence and maturation of various low cost technologies that are used to visualize, interact with and manipulate virtual three-dimensional (3D) objects. Collectively, they have been developed and improved to the point that they show clear promise of their large-scale adoption. Benefits from masses of users adopting an innovation mean that the same individuals have conditions to receive training, improve their performance with low cost solutions which means reduction of investment in training domain, manufacturing costs and, consequently, end product price, all resulting in large scale saving of material and human resources.

Those technologies include but are not limited to virtual reality (VR) headsets, non-immersive stereoscopic displays, tracking systems, hand controllers, and haptic devices. Considering that, "generally, complex data can be interpreted more effectively when displayed in three dimensions" (Reichelt et al. 2010, p. 1), the ability to reproduce 3D objects and their physics in high fidelity have been making those devices an interesting opportunity for training of tasks that involve judgment of distances, depth, sizing, and fit. The demand that those tasks pose to the users is to master navigation and manipulation of those objects at distances that are characterized as 'within arms' reach'; such tasks are typically performed by doctors, dentists, sculptors, craftsmen, and maintenance personnel.

Different factors may influence user's performance when manipulating 3D objects in a virtual environment (VE); those factors can improve human perception of the components and environment and, consequently facilitate execution of the tasks. Two significant factors are stereoscopic depth cue and immersion (Wichansky 1991; Pausch, Proffitt, and Williams 1997).

Immersion is defined as an extent to which simulated information surrounds and envelops a user (Slater and Wilbur 1997), while presence is defined as a sense of being present in an artificial environment that is different from immediate (physical) environment (Slater and Wilbur 1997; Draper, Kaber, and Usher 1998; Bowman and Mcmahan 2007;

Slater et al. 2009); this is done by presenting a human with any subset of stimuli that mimics interactions in real time—visual (images), sounds, haptics, and others. Visual stimulus is often associated with information presented inside VR headsets, also called head-mounted displays (HMDs). Those displays show stereoscopic images calculated by an image generator, usually a desktop computer or a video-game console, and presented to the user's eyes. Interaction between the user and virtual environment in typically conducted with the help of tracking system (tracks user's head, hands or even full body) and input controllers operated by user's hands. Examples of VR headsets that provide virtual reality experiences are HTC VIVE, Oculus Rift, GearVR and PlayStation VR.

Stereoscopic depth cue is a phenomenon closely related to the way human see the world. This type of binocular cue consists of the illusion of three-dimensional depth from given two-dimensional images (Howard and Rogers 1996), each viewed by one eye. They are largely used on VR headsets (spatial separation of images form left and right eye) and on desktop-style stereoscopic displays, such as zSpace (use of polarized glasses) and Cave Automatic Virtual Environment – CAVE (use of active shutter glasses).

B. RESEARCH PROBLEM AND MOTIVATION

The importance of maintenance tasks has always been significant in both civilian and military domain, especially in all Department of Defense (DoD) services. Maintenance tasks can be very complex; they require the knowledge of specific techniques and procedures that involve use of cognitive memory and fine motor ability (Gutiérrez et al. 2010).

Main components of maintenance tasks include assembly and disassembly steps; they, in turn, involve judgement of distance, depth, sizing, and fit. Different factors may influence the operator's performance in assembly and disassembly tasks done using VR technology. They include but are not limited to variety of depth cues (both monocular and binocular), immersion, level of realism, presence of haptic cues, and interactive techniques made available to the user. They can affect operator's perception of the components in the environment and, consequently, influence the execution of tasks. This study uses assembly

tasks as a context for exploring operator's performance while manipulating the objects made of multiples components positioned within arms' reach in a VE.

Two main parameters most closely related to maintenance task in VR system are stereoscopic depth cue and immersion. We use stereoscopic depth cue images as a factor in this study due to its potential to improve ratings of subjective image quality and increase in judgement of depth (Reinhart, Beaton, and Snyder 1990; Ware and Franck 1996). Immersion is used due to its potential benefits of spatial understanding, decrease in information clutter, and increase in peripheral awareness (Bowman and Mcmahan 2007). Both immersion and stereoscopic depth cues are typical for VR headsets, and with emergence of low-cost headsets, we wanted to examine their potential in being advantageous in maintenance domain.

The purpose of this thesis is to study the influence of stereoscopic depth cue and immersion on performance of human operator in tasks that heavily involve judging of distances, depth, sizing, and fit. VR environments and VR technology will be used to provide fully immersive / non-immersive, and stereoscopic / non-stereoscopic depth cues user experiences.

C. RESEARCH QUESTIONS

The following research questions provided much of the motivation for this study and are the focal points in this thesis:

- Can stereoscopic depth cue improve human performance in assembly tasks?
- Can immersion improve human performance in assembly tasks?
- What is the type of VR display solution that provides best results for assembly tasks?

D. SCOPE

The scope of this thesis is to design and execute an experiment that requires subjects to perform two maintenance procedure tasks, and compare their performance in four

different study conditions. The maintenance tasks will be performed using following low-cost commercial-of-the-shelf (COTS) display solutions:

1. Immersive stereoscopic VR display: Oculus Rift
2. Non-immersive stereoscopic VR display: 3D-ready TV (3DTV) with shutter glasses
3. Non-immersive non-stereoscopic VR display: 3DTV used in monoscopic mode
4. Immersive non-stereoscopic VR display: Oculus Rift with same image presented to both eyes

An additional comparison to a virtual maintenance task will be done using 3D printed object that is a replica of virtual model used in this study.

E. APPROACH

In order to properly address the research questions, the methodology used for this study consisted of several steps.

First, a literature review was done on domains closely related to this thesis. This included studies focused on maintenance tasks, depth perception, stereoscopic platforms, immersive and non-immersive environments, and haptic feedback.

Our second step included task analysis; this work was performed to identify the main elements of maintenance task and their order, sensory stimuli typical for this task, conditions under this task is performed, profile of a typical user, skills required for the users, and the standards of human performance.

After the task analysis, a user study was designed; this included identifying target audience (subjects), experimental conditions, procedures that subjects will go through, the apparatus that will allow us to support subjects' activity and collect data, type of data and methods of data collection, and the metrics to evaluate users' performance. IRB documentation was also developed and approval to execute study was received prior to execution of the study.

A virtual reality application was developed using Unity game engine; this software provided participants with capability of interacting with virtual environment and assembling virtual objects. Before the main study was executed, we conducted a pilot test to identify elements of application that could be improved and to test user study design. The main study was then executed and a comprehensive set of data was collected.

Finally, after the data collection, qualitative and quantitative data analysis were performed and, based on the results of the analysis, conclusions were stated.

F. THESIS STRUCTURE

The remainder of this thesis is structured as follows:

Chapter II reviews previous studies in domains related to this thesis: depth perception, stereoscopic platforms, fully and non-fully immersive environments, VR use in maintenance tasks and haptic feedback.

Chapter III provides a task analysis on basic maintenance tasks, including assembly and disassembly procedures.

Chapter IV details the user study, describing the study design, participants, methodology and apparatus used, IRB process and metrics of performance.

Chapter V describes the system development for each of the platforms used on this study.

Chapter VI details the work done on the pilot test, lessons learned and modifications applied to the final experiment.

Chapter VII presents the analysis of the data collected in the experiments and the results of this analysis.

Chapter VIII presents the conclusions of the study and details the future work.

THIS PAGE INTENTIONALLY LEFT BLANK

II. BACKGROUND

This chapter reviews previous studies in domains related to this thesis: depth perception, stereoscopic platforms, fully and non-fully immersive environments, and use of VR technology in support of maintenance tasks.

A. DEPTH PERCEPTION

Depth perception can be characterized as the visual ability to perceive the world, i.e. see things, in three dimensions and to judge how far the object is. It is typically influenced by different monocular and binocular depth cues that, together, cognitively lead to a specific depth perception (Kleiber and Winkelholz 2008).

Cues for depth perception can be classified in oculomotor, which includes the known ocular near triad of oculomotor responses (this includes accommodation, convergence, and myosis, also known as pupillary constriction cues), and in visual, which includes binocular and monocular cues (Reichelt et al. 2010). Monocular depth cues can be classified as static or pictorial depth cues, such as interposition, linear perspective, and light and shadow distribution, and motion-based cues, such as motion parallax, kinetic depth effect, and dynamic occlusion.

Reinhart, Beaton, and Snyder (1990) analyzed the impact of monocular and binocular depth cues on objective task performance and subjective image quality. Their results indicated that stereoscopic depth cues strongly improved ratings of subjective image quality and that stereoscopic images may provide subjectively more compelling depth information than images containing only monocular cues.

B. STEREOSCOPIC VR PLATFORMS

Ponce and Born (2008) defined stereopsis as “the use of differences in the images projected onto the retinas of the two eyes — so-called ‘binocular disparity’ — to reconstruct the third visual dimension of depth” (p. R845). This human ability to process both images is widely taken advantage of by VR platforms. Two main types of stereoscopic VR platforms are head mounted displays (also known as fully-immersive VR displays) and

non-HMD type of 3D displays that use either polarized (passive stereo) or shutter glasses active stereo) to separate images and 'deliver' them to appropriate eye.

Stereoscopic head mounted displays present slightly different images (one for left eye and one for right eye) presented on the screens that are placed in front of each eye (Costello 1997). Besides the ability to induce the sense of depth, this type of platform has the advantage of providing the user with a 360^{o} field of regard, i.e., independent of the direction the user is looking, a visual image will be computed in real time and delivered to the displays for human use. Contemporary examples of stereoscopic VR HMDs are Oculus Rift and HTC VIVE.

Stereoscopic 3D displays (non-HMD VR display solutions) need to present two slightly different images on the same screen. There are two major methods for doing this: images can be displayed onto the screen simultaneously or they can be displayed alternatively i.e. they are separated over the time dimension (temporal separation). For the first method, users wear glasses with polarizing filters, which allow only light polarized in the same direction to pass and reach the one of the user's eye. The second method uses alternated images that are synchronized with the shutter glasses, which become opaque if the image is not intended for that specific eye, or transparent otherwise. Examples of 3D displays with either polarized or shutter glasses are 3D TV, zSpace, CAVE or any size of display solution that use passive stereo like in movie theaters.

C. IMMERSIVE AND NON-IMMERSIVE ENVIRONMENTS

Slater and Wilbur (1997) defined immersion as a characteristic of a technology that describes the extent to which simulated information surrounds and envelops a user. Their definition is based on the following dimensions of immersion: inclusive (extent to which physical reality is inhibited i.e. removed from human observation), extensive (range of sensory modalities accommodated), surrounding (extent to which VR is panoramic rather than limited to a narrow field of view), vivid illusion (resolution, fidelity, and variety of energy simulated within a particular modality), and body matching and tracking. Potential benefits of immersion include better spatial understanding, decrease in information clutter, and increase in peripheral awareness (Bowman and Mcmahan 2007).

Costello (1997) ranked immersive and non-immersive environments in three different categories, according to the sense of immersion, or degree of presence they provide: non-immersive, semi-immersive, and fully immersive. Non-immersive and semi-immersive systems use the least immersive implementation of VR technologies and, generally, do not require the highest level of graphics performance. Additionally, they use one or more screen monitor or projector systems that can be combined with shutter glasses, and their field of regard (total area that can be captured by a movable sensor) is limited. On the other hand, immersive systems use the most immersive implementation of VR technologies and, generally, require the highest level of graphics performance. Usually they consist of HMDs or CAVEs, and provide the highest sense of presence and the largest field of regard.

For the purpose of this study, we consider both non-immersive and semi-immersive environments simply as "non-immersive", and fully immersive environment as "immersive".

D. USE OF VR TECHNOLOGY IN SUPPORT OF MAINTENANCE TASKS

Pausch, Proffitt, and Williams (1997) showed on their study that VR can improve users' performance on a search task when compared to stationary monitor and hand-input device, being significantly better if they already had searched the virtual environment. Thus, researchers started to analyze the use of VR in different domains, such as virtual maintenance.

Duan et al. (2012) defined virtual maintenance as "a process of maintenance evaluation or maintenance audit by applying the virtual reality technology in virtual environment created by computer auxiliary design (CAD)" (p. 1396), and described some of its main training advantages: capability of exhibiting a realistic training environment without costing any actual equipment, enhancement of trainees' science and technology qualities, simulation of the actual training field as accurate as possible, and capability of providing different virtual training set ups. Additionally, virtual maintenance has become an "important force in improving product development efficiency, reducing costs,

promoting information integration and standardization system construction" (Cheng et al. 2011, p. 3546).

In a recent study, Murcia-Lopez and Steed (2018) analyzed the effectiveness of virtual training and physical training for learning transfer of a bimanual assembly task, in which participants assembled three versions of 3D burr puzzles. Results on their study showed that the performance of virtually trained participants was promising.

E. SUMMARY

This chapter detailed previous works in main domains related to this study, including depth perception, stereoscopic platforms, immersive and non-immersive environments, and use of VR technology in support of maintenance tasks.

III. TASK ANALYSIS

This chapter provides a task analysis on basic maintenance tasks, including assembly and disassembly procedures. The identification of tasks actions was done by observing how maintenance personnel performs their regular tasks.

A. IDENTIFICATION OF TASK ACTIONS

Basic maintenance tasks consist of assembly and disassembly steps, in which the operator needs to judge distance, depth, sizing, and fit of objects. Additionally, those steps demand the operator to master navigate and manipulate objects at distances that are characterized as 'within arms' reach'. In this study, our main task is to assembly objects, for which we identified the following main task actions:

- Grab: user grabs the object with their hand.
- Move: when grabbing an object, user manipulates and repositions it, whether by moving across the space or rotating it.
- Release: user releases the object from their hand.
- Connect: user moves the object to a specific position—close enough or touching another object—that makes the object to connect to the other one.

B. USER SKILLS

Skills required for maintenance personnel vary based on the type of job. However, there are a few skills that are common for basic maintenance tasks:

- Spatial perception: the user's ability to be aware of their relationships with themselves and with the surrounding environment.
- Spatial orientation: the user's ability to maintain body orientation and posture in relation to the surrounding environment.

- Time management: users must be able to manage their time and execute tasks in the allotted time.
- Problem-solving ability: when facing a problem, users are responsible for diagnosing its root cause and applying a solution.
- Instruction/reading comprehension: the user's ability to read instruction and execute them as expected.

C. ENVIRONMENT

The environment in which maintenance personnel works varies according to the type of job. Nonetheless, for tasks that mainly involve manipulation of objects within arms' reach, the usual environment consists of a chair and a table, in which the equipment and tools stays on top of.

D. TYPE OF SENSORY INFORMATION USED

Sensory information consists of the all the information that the human sensory systems collects from the surrounding environment, and make available to the brain for further processing. A set of sensory modalities consist of visual (sight), tactile (touch), auditory (hearing), vestibular, proprioception, olfactory (smell), and gustation (taste). The sensory modality inputs that are most closely related to maintenance tasks are vision, touch (somatosensation), and hearing (audition). The sense that humans rely heavily while operating in physical environment is the sight: maintenance worker use it to visually identify the components and tools they will use to perform the task, to identify the size of each component, their depth (position in space) and its shape. Besides sight, they largely use touch—in particular haptic, that is related to the perception and manipulation of objects through touch (e.g. feeling when parts collide and when they connect)—and hearing—the ability to perceive sounds (e.g. hearing when two parts collide or connect).

E. TOOLS AND MANUALS

Most of maintenance tasks require the use of specific tools in order to be performed. Depending on the task complexity, or even on the tool complexity, manuals are necessary

to correctly guide the user in the procedures. Common tools used in maintenance operations are hand tools; depending on the type of job and system that needs to be worked on, they can include wrench, screwdriver, hammer, sleeve, pliers, scissors, utility knife or similar. Manuals, also known as user guides, are technical communication documents intended to give assistance to users using a tool or system.

F. EXAMPLE MAINTENANCE TASK

1. Preventive

Preventive maintenance focuses on maintaining a level of service on an equipment. Programmed inspections or interventions are performed with pre-defined frequency over the time to lessen the likelihood of an equipment failing. The approaches to different types of preventive maintenance are based on time, failure finding, risk, condition, and prediction. Examples of preventive maintenance tasks are: daily lubrication of an engine, control console lamps testing, tightening of connections, and calibration of system pressure.

2. Corrective

Corrective maintenance focuses on fixing failures found in a system, i.e., it is performed when a failure has occurred and the system functionality needs to be reinstated. The approaches to different types of corrective maintenance are based on deferrable and emergency. Examples of corrective maintenance task are: replacement of a defective component, overhaul of an engine, welding a circuit board, and repairing a tire.

G. SUMMARY

This chapter detailed the identification of basic maintenance tasks actions of assembly procedures used on this study.

THIS PAGE INTENTIONALLY LEFT BLANK

IV. USER STUDY

This chapter details the elements of user study that was designed and executed as a part of this thesis research. Sections in this chapter review study research goals and expectations, subject pool, methodology and apparatus, study procedure, IRB process, and metrics of user performance used in the study.

A. RESEARCH GOALS AND EXPECTATIONS FOR THE STUDY

The goal of this research was to ascertain the effectiveness of stereoscopic depth cue and immersion in maintenance related tasks, i.e., determine if and how the performance in assembly tasks is affected when human operator uses immersive or non-immersive environments, and stereoscopic or non-stereoscopic images on VR displays.

The expectations for the study were to acquire guidance and indication of the type of low-cost commercial of the shelf VR display solution that assured best results for assembly task.

B. STUDY DESIGN

Study has been designed to have two independent variables with two different levels each. It was classified as a 2x2 between-group design, resulting in four distinct study conditions. The independent variables are Stereoscopic Depth Cue and Immersion. The levels for the first variable were “Immersive (I)” and “Non-Immersive (NI)”, and for the second variable, the levels were “Stereoscopic (S)” and “Non-Stereoscopic (NS)”. The combination of these two variables provided four study conditions, as shown in Table 1.

Table 1. Study conditions of the 2x2 between-group design

		STEREOSCOPIC DEPTH CUE	
		Stereocopic (S)	**Non-Stereoscopic (NS)**
IMMERSION	**Immersive (I)**	Immersive Stereoscopic (IS)	Immersive Non-Stereoscopic (INS)
	Non-Immersive (NS)	Non-Immersive Stereoscopic (NIS)	Non-Immersive Non-Stereoscopic (NINS)

Participants in immersive conditions (IS and INS) used the following equipment: Oculus Rift headset, Oculus Rift sensors for tracking, and Oculus Touch controllers, as shown in Figure 1. Participants in non-immersive conditions (NIS and NINS) used the following equipment: 3D-ready TV with shutter glasses, TrackIR head tracker, and Oculus Touch controllers, as shown in Figure 2.

Figure 1. Example of participant in immersive conditions (IS and INS)

Figure 2. Example of participant in non-immersive conditions (NIS and NINS)

C. HYPOTHESIS

The hypothesis for this study were:

- Null hypothesis (H_o): There is no difference in user performance and user satisfaction between the four experimental conditions.
- Alternative hypothesis (H_a): There is difference in user performance between experimental conditions, and Immersive Stereoscopic (IS) condition results in best user performance and user satisfaction.

D. PARTICIPANTS

Participants in this study included general public (adults over 18 years old) with no specific skill set required. They were recruited at the Naval Postgraduate School (NPS) via one bulk email, and in person by the researcher. To minimize coercion and undue influence during the recruitment process, participants voluntarily signed up for the experiment using a free online scheduling service that allowed them to select date and time of their choice. The researcher did not have any command, academic, or employment influence over participants and explained the voluntary nature of the experiment in person and via email

during recruitment process. The researcher also explained the voluntary nature of the study to participants during the informed consent process before any tasks begun.

E. PROCEDURE

Each participant was assigned to one of the four treatment conditions. The list of steps executed for each subject, including the approximated time for each step (in parenthesis), consisted of the following:

1. Participant arrived to the research location, and was given the opportunity to provide informed consent (5 minutes). If conditions of the study were accepted by participant, the experimenter continued with the remainder of the session and participant's involvement in it.

2. Participant completed the baseline Simulator Sickness Questionnaire (SSQ) (Kennedy et al. 1993) (2 minutes).

3. Participant received initial instruction about the system (2 minutes).

4. Participant got familiarized with VR device and interface, and did a training session (5 minutes).

5. Participant completed the second Simulator Sickness Questionnaire (SSQ) (2 minutes).

6. Participant received instructions about the main experimental tasks.

7. Participant executed task #1 in virtual environment (10 minutes).

8. Participant completed the third Simulator Sickness Questionnaire (SSQ) (2 minutes).

9. Participant completed the post-task #1 questionnaire (3 minutes).

10. Participant executed task #2 in virtual environment (10 minutes).

11. Participant completed the fourth Simulator Sickness Questionnaire (SSQ) (2 minutes).

12. Participant completed the post-task #2 questionnaire (3 minutes).

13. Participant received instructions about task #3 (1 minute).

14. Participant executed task #3 in physical environment (5 minutes).

15. Participant completed the post-task #3 questionnaire (3 minutes).

16. Participant completed System Usability Scale (SUS) questionnaire (Brooke 1995) and demographic survey (5 minutes).

17. Participant received a short debrief/explanation of the study and were permitted to ask questions (5 minutes).

Copy of IRB application form that details the procedure used in this study is listed in Appendix A. Copy of checklist (step-by-step procedure) that was used by the experimenter is listed in Appendix B.

F. APPARATUS

The main elements of study hardware and software setups are:

1. Hardware

1. Oculus Rift bundle (headset, touch controllers, and two tracking sensors)

2. 3DTV bundle: Sony Bravia 3D TV with shutter glasses

3. Desktop computer: Alienware PC with NVIDIA GeForce GTX 1080 graphics card

4. TrackClip Pro Bundle (TrackIR 5, and TrackClip Pro)

5. Video camera: Sony HDR-XR520

6. Video camera stand: Bogen Manfrotto 3063

7. 3D printer with polymer filament: MakerBot Replicator+

2. Software

1. Scatterfix 2018, in-house developed software. Source: (Yamashita de Moura, Sadagic, Heine, Johnson, & Lee, personal communication, July 22, 2018)[1]
2. Unity 2017.4.2.f2
3. Oculus Rift Software
4. Oculus Utilities for Unity
5. Oculus Avatar Software Development Kit (SDK)
6. Oculus Integration Asset
7. Virtual Reality Toolkit (VRTK) SDK Manager
8. TrackIR Software
9. TrackIR Enhanced SDK
10. MakerBot Print Software

Details about the selection of the main elements of the VR setup—Oculus Rift, 3DTV, and TrackIR—that were used for this study are described in Chapter V.

G. LIMITATIONS IN THE APPARATUS

Several issues related to both hardware and software setup limited the creation of an ideal system environment for the study. Given those physical limitations, we created the best possible environment that could be provided with selected equipment.

[1] Sadagic conceptualized the original application for zSpace, and the development team (Heine, Johnson, & Lee) programmed the application and created the 3D objects. The thesis author, based on the original software, created the application for Oculus Rift and 3DTV.

1. Oculus Rift

Oculus Rift presents a predefined setting for the field-of-view (FOV) when on the monoscopic mode, and its value could not be modified. Thus, the FOVs of IS (110°) and INS (96°) conditions were different.

2. 3D TV and Shutter Glasses

It has been noticed that the brightness and contrast of the TV image with the 3D mode turned on (NIS condition) were different from the ones with the 3D mode turned off (NINS condition). In NIS mode, the TV image appeared to us as slightly darker and with less contrast.

3. TrackIR

TrackIR has a FOV of 51.7º—this is the area with optimal tracking performance, which limits the participant's head movements. If the participant's head is out of the FOV, TrackIR automatically resets the current head position to the initial position.

4. Unity

In order to provide 3D video signals for the 3DTV for the NIS condition, we created a side-by-side image in Unity. We used the images captured from two distinct cameras (1920x1080), horizontally separated by 3 cm, to create rendered textures, which were horizontally scaled to 50% and applied to (i.e., the images were projected to) two planes of 960x1080 pixels each located side-by-side, creating a combined plane of 1920x1080. A third camera captured the combined plane (with views from both left and right eye cameras) and its signal was sent to the 3DTV, which processed the side-by-side 3D signal and turned it into 3D images synchronized with the shutter glasses. Thus, there was a loss of 50% on the horizontal resolution when creating the 3D video signal. It was also noticed that the processing of the video signal was sometimes slow, resulting in flickering and 'stuttering image' effects on the 3DTV.

H. IRB PROCESS

The study involved human subjects research, and it required creation of official IRB package that was submitted to the NPS Institutional Review Board. The approval process lasted around three weeks. The IRB package consisted of the following documents:

1. Initial Review Application
2. Scientific Review Form
3. Conflict of Interest Disclosure Form
4. Informed Consent Form
5. Collaborative Institutional Training Initiative (CITI) Program certificates
6. Recruitment Email
7. Recruitment Flyer
8. Post-task questionnaires
9. Demographics and SUS questionnaires
10. Simulator Sickness Questionnaire (SSQ)
11. Proposal Approval Form for Thesis & Reports (TPF)
12. Approved Thesis Proposal

I. METRICS OF USER PERFORMANCE

Objective and subjective data sets for each participant were collected during the study.

1. Objective Data Set

Objective data recorded from each participant includes the following elements:

- Time on each task

- Success of assembly
- Time stamped events of correct assembly
- Time stamped events of grabbing action
- Time stamped events of releasing action
- Time stamped events of collision between two objects (two parts of one large object that needed to be assembled)
- Time stamped events of entering the snap-drop zone

2. Subjective Data Set

Subjective data recorded from each participant includes the following elements:

a. Simulator Sickness Questionnaire (SSQ)

Participants filled the SSQ (Appendix C) at the beginning of the study – this allowed us to form a baseline SSQ data set, and after each use of the VR application, i.e., after the training session, Task 1, and Task 2. The SSQ consists of 16 symptoms of simulator sickness on a four-point scale (0-3), which are weighted and summed together to obtain a single score, resulting in an overall simulator sickness score for a given simulator.

b. Post-task Questionnaires

Participants filled post-task questionnaires after each of the three tasks—Tasks 1, 2 and 3 (Appendixes D, E, and F). The post-task questionnaires included, but it was not limited to, participants' information about their success in completing the task within the allotted time (including the reasons for not completing it if that happened), their experience with selecting, manipulating, and assembling the objects, presence—their sense of being in the virtual room, and their rating of difficulty level with different elements of assembly task.

c. Demographic Survey

Participants filled the demographic survey (Appendix G) after they completed the third post-task questionnaire. The demographic survey included, but it was not limited to information about year of birth, sex, use of regular glasses, occupation, video game experience, and prior use of VR headsets.

d. System Usability Scale (SUS)

Participants filled out a modified SUS (Appendix H) at the end of the study. The SUS questionnaire was created by John Brooke, and is typically described as a "quick and dirty" reliable tool for measuring the usability of a system.

J. SUMMARY

This chapter discussed the design of user study. Research goals and expectations were presented, as well as the study design, participants, methodology used on the study, and procedure that was followed for each participant. Also, it described the tasks each participant executed, the apparatus used on the study and their limitations, the IRB process, and the metrics of user performance recorded during the experiment.

V. SYSTEM DEVELOPMENT

This chapter details a design and development of the system architecture and software application that were used in support of this study.

A. HARDWARE ENVIRONMENT

1. Immersive Display Solution

Two alternatives were considered for immersive display solution: Oculus Rift and HTC VIVE, as shown in Figure 3, which were two VR major makers of head-mounted displays at the time when this research was done.

Figure 3. HTC VIVE (left) and Oculus Rift (right) bundles

Both Oculus Rift and HTC VIVE provide immersive VR experience (they are visual displays), support six degrees of freedom (6DOF) navigation, they have established community of developers, have development packages for Unity, and their costs are affordable. Also, both have very similar technical specifications, as shown in Table 2, and similar minimum system requirements, as shown in Table 3.

Table 2. Oculus Rift and HTC VIVE specs comparison. Adapted from Eva (2017).

	Oculus Rift	**HTC VIVE**
Display	OLED	OLED
Resolution	2160 x 1200	2160 x 1200
Refresh Rate	90 Hz	90 Hz
Field of View	110º	110º
Tracking Area	5 ft x 11 ft	15 ft x 15 ft
Built-In Audio Output	Yes	Yes
Built-In Audio Input	Yes	Yes
Store Platform	Oculus Home	SteamVR & VIVEPORT
Connections	HDMI, USB 2.0 & USB 3.0	HDMI, USB 2.0 & USB3.0
Sensors Camera	Accelerometer, magnetometer, gyroscope and Constellation tracking	Accelerometer, gyroscope, front facing camera and Lighthouse system
Controller Options	Oculus Touch, XBox One Controller	HTC VIVE Controller or any compatible PC gamepad

Table 3. Oculus Rift and HTC VIVE minimum requirements comparison. Adapted from Eva (2017).

	Oculus Rift	**HTC VIVE**
Graphics Card	NVIDIA GeForce GTX 960 / AMD Radeon RX 470 or greater	NVIDIA GeForce GTX 960 / AMD Radeon RX 470 or greater
Processor	AMD FX4350 / Intel Core i3-6100	Equivalent to Intel Core i5-4590 or greater
Memory	8GB+ RAM	4GB+ RAM
Video Output	Compatible HDMI 1.3	Compatible HDMI 1.3
Required Free Ports	2x USB 3.0	1x USB 2.0
Minimum OS Compatibility	Windows 7 SP1	Windows 7 SP1

The key differences between those two VR systems are consisted in the tracking area, range of sensors, and form of hand controller. Our major concern was related to two factors: size of tracking area and suitability of hand controllers for assembly task. In the study the participant is required to stay seated and move the arms to manipulate parts of the object that is been assembled. Although HTC Vive has a larger tracking area, the area covered by Oculus sensors is also sufficiently large for the purpose of this study. Regarding the tracking sensors, both provide a high-quality tracking of user's head in the space. The size and weight of hand controllers, and the way they are held by the user, are different though, as shown in Figure 4. HTC Vive controller weights 190g each while Oculus Touch controller weights 150g each. Also, HTC Vive controller is bigger than Oculus Touch controller due to the size of tracking elements at its front end, limiting the user when they want to bring them very close to each other. This limitation inhibits the user to correctly replicate the proprioceptive feedback one would feel when assembling two parts together, as shown in Figure 4 (top).

The decision to choose Oculus Rift was based on the size of the controller and the way user holds them while interacting with virtual world. As the objects used for the assembly tasks on this study were relative small, most of them measuring a few inches, smaller controllers could provide more natural user experience when manipulating and connecting the parts to each other. The way controllers are held by the users (the grip) is more similar to the way that humans would hold pieces of object that is been assembled, and in the end this ultimately led to decision to choose Oculus Touch and consequently Oculus Rift HMD.

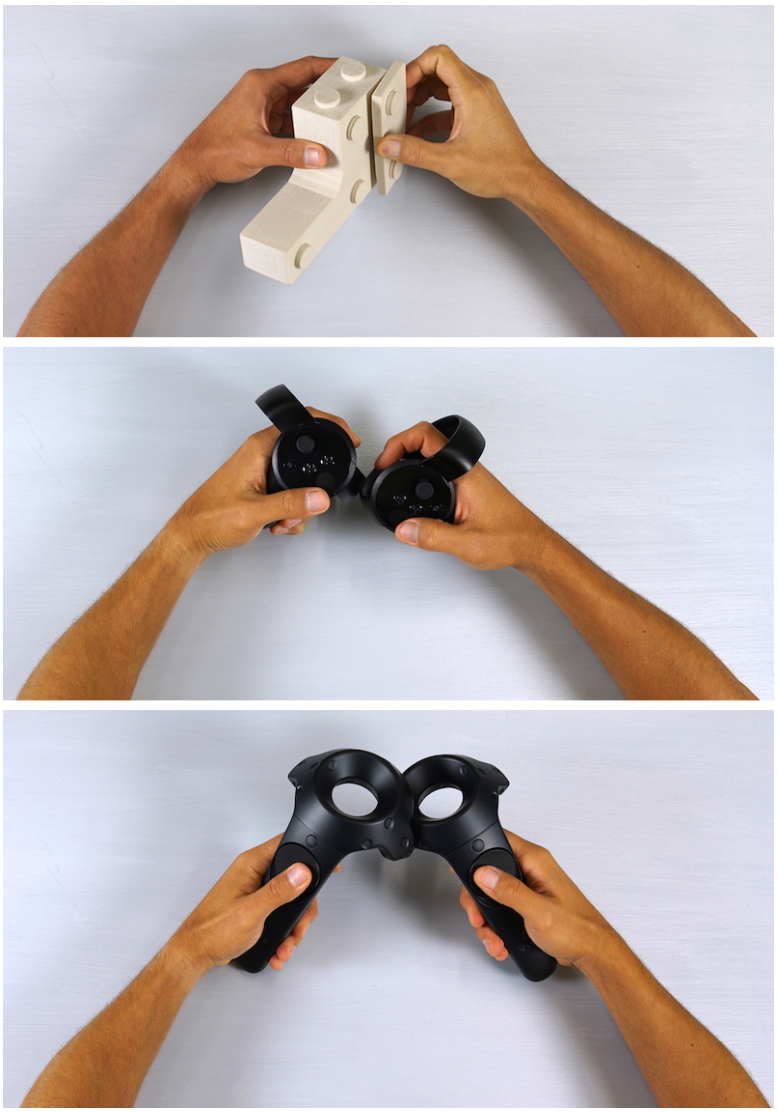


Figure 4. Position of operator's hands while holding physical parts (top), Oculus Touch controllers (middle) and HTC VIVE controllers (right)

2. Non-Immersive Display Solution

The initial choice for non-immersive display solution was zSpace display platform; this display combines both stereoscopic capability and head tracking in the same device that looks like a tablet. With the use of a stylus, the system allows manipulation of virtual objects, which can be seen as three-dimensional objects with the aid of polarized glasses. Although zSpace seemed as a good candidate, two main issues prevented its use: number

of input devices that can be used simultaneously and the type of basic interaction technique supported in each platform.

zSpace uses only one stylus at the time to conduct selection and manipulations of 3D objects, as shown in Figure 5. This is a concern because the performance of an operator who is assembling objects could be dramatically impaired if he would use only one input device and not two (similar to using only one hand and not two to assemble the objects). More natural mode of interaction would be if the operator would use two input devices at the same time, such as Oculus Touch controllers.

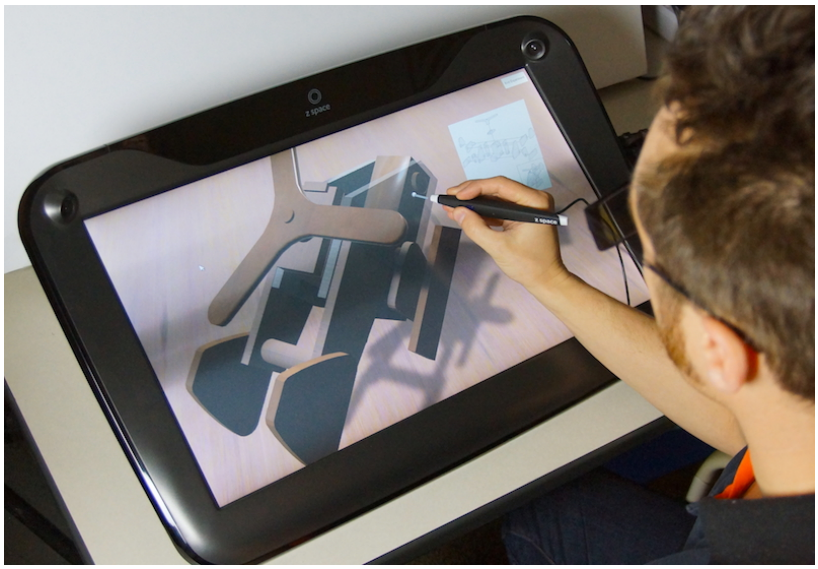


Figure 5. Non-immersive zSpace display with head tracking and unimanual stylus interaction

The second issue relates to the type of interaction technique that allows operator to interact with virtual objects. The zSpace stylus uses an infrared (IR) light-emitting diode (LED) that, when pointed to the 3D object, works like a lever, keeping a distance between the stylus and the object. The opposite was the case for Oculus Rift controllers: the operator's interaction with the virtual objects is similar to manipulation of objects with their own hands i.e. direct manipulation.

Thus, in order to support two input devices and have type of interaction that closely resembles to the way humans do assembly task in real world, zSpace display could not be used and we had to identify another (different) non-immersive display with stereoscopic capability. The solution was found in 3D-ready TV (3D TV) with shutter glasses, and that display was integrated with Oculus Touch hand controllers (Oculus tracking was used to track both hand controllers). A 3D TV has two distinct modes for image rendering: stereoscopic mode, which creates 3D images with the aid of shutter glasses (active stereo), and monoscopic mode, which is the default mode with the same image for both eyes.

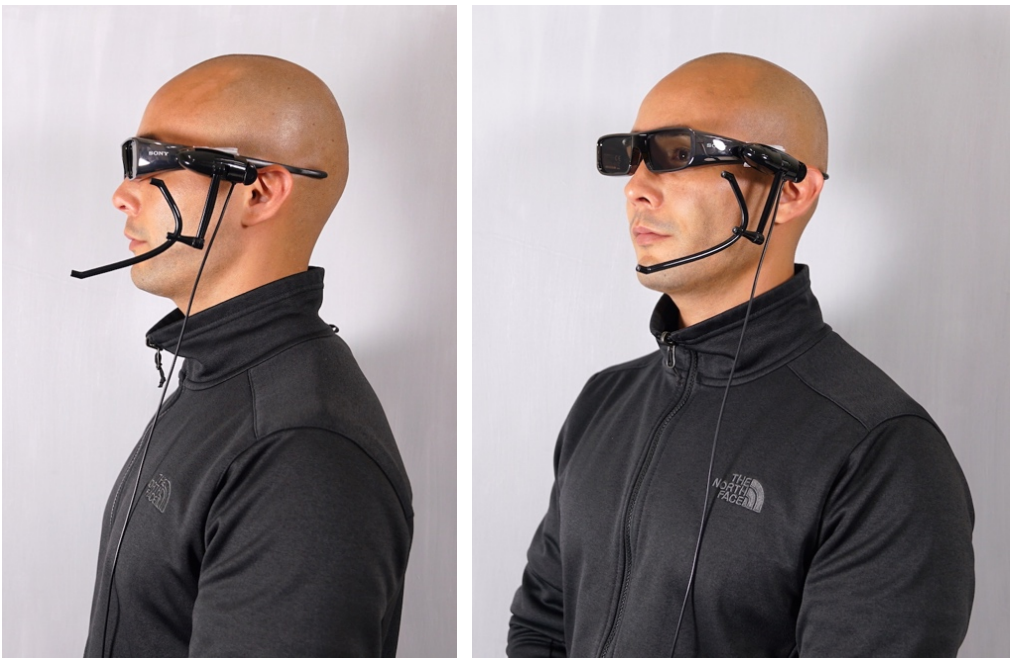
The 3DTV chosen for this study was the Sony Bravia XBR-52LX900 and the active shutter glasses were the Sony TDG-BR100 3D Glasses. This 3DTV has three different 3D formats: over-under, side-by-side, and simulated 3D. The first two formats require a video signal input with two different images, one for each eye, and they are horizontally arranged for the over-under format, and vertically arranged for the side-by-side format. On the other hand, the third format uses a regular video signal and converts regular two-dimensional (2D) images in simulated 3D images. We decided to use the side-by-side format. Thus, the video signal input to the 3DTV needed to have a resolution of 1920x1280 pixels, consisting of two images of 960x1280 pixels. It is important to emphasize that the 3DTV scales both images to 1920x1280 pixels, so these images should be previously scaled down to half of their width.

3. Head Tracking

In order to enable the same mode of navigation for all four study conditions, it was necessary to integrate head tracking for the NIS and NINS conditions. The decision was to use lightweight TrackIR 5 with TrackClip Pro, as shown in Figure 6, which provides a FOV of 51.7° and six degrees of freedom (6DOF) technology. The sensor form factor and its light weight allowed us to mount it on shutter glasses, as shown in Figure 7 (shutter glasses and TrackClip Pro sensor weight an approximately total of 100 grams); Figure 8 shows an example of a participant wearing shutter glasses with tracked sensor mounted on it.

Figure 6. TrackClip Pro and TrackIR 5 devices. Source: Natural Point (n.d.).

Figure 7. Tracking sensor mounted on shutter glasses



User's full profile on the left, and user's half profile on the right.

Figure 8. User wearing shutter glasses with tracked sensor mounted on it

4. Computer Systems

Oculus Rift requires a high performance computer system with special graphics card to function correctly and to be able to generate stereoscopic images with satisfactory frame rate. Alienware PC with the following specification was used on this study:

- Processor: Intel® Core i7-5820K CPU @ 3.30Gz
- Installed Memory (RAM): 16.0 GB
- System Type: 64-bit Operating System, x64-based processor
- Operating System: Windows 10 Home
- Video board: NVIDIA GeForce GTX 1080

5. 3D Printing

The third task the participants were asked to execute in this study was to assembly a physical version of the second task's virtual object. Thus, this researcher used MakerBot

Replicator+, as shown in Figure 9, to 3D print fifteen individual parts that make up chosen object

Figure 9. MakerBot Replicator+ 3D Printer. Source: Makerbot Industries, LLC (n.d.).

B. SOFTWARE ENVIRONMENT

1. Unity Editor

Unity is a game development platform used to build high-quality 3D and 2D games, which can be deployed in mobile, desktop, VR/AR, and consoles. This system was chosen for the application development because of three main reasons:

1. Unity was already well stablished in the game development area and had a large developer community, which was very helpful when developing a software.

2. Unity provided ways to integrate different VR platforms using the same application, an important requirement for this study (study conditions include both immersive and non-immersive VR platforms).

3. The software in which the new application was based was developed in Unity.

A few years ago, MOVES Institute developed a software for zSpace named Scatterfix, in which the user could assembly and disassembly objects using the zSpace stylus. We used some of its core classes—mainly related to objects relationships and behaviors—and developed new application that supported user study.

2. Oculus Rift Software

Oculus Rift Software required to use Oculus Rift bundle. It was downloaded and installed on the PC, and an Oculus account was also created. The software provided the means to setup the Oculus Rift bundle, and download and install apps.

3. TrackIR

TrackIR 5.4 required to use TrackClip Pro bundle. It was downloaded and installed on the PC; it allowed the user to setup the bundle and change the settings of the sensor.

4. MakerBot Print

MakerBot Print software was used to load the digital objects, on the Standard Triangle Language (STL, also known as Standard Tessellation Language) format, and send them to the MakerBot Replicator+ 3D Printer.

5. System Development Kits (SDKs)

System Development Kits are packages that help the user quickly and easily develop applications for a specific development environment; they also facilitate the integration of COTS devices to other platforms or software. We used SDKs to integrate Oculus Rift and Track IR 5 to Unity.

a. Oculus Integration Asset

Oculus Integration is an asset designed to provide advanced support to Oculus Rift, Oculus Touch, and Gear VR for rendering, audio, social, and avatars (“Oculus Integration - Asset Store” 2018). We used it mainly to create the virtual hands on the VE, as shown in Figure 10.

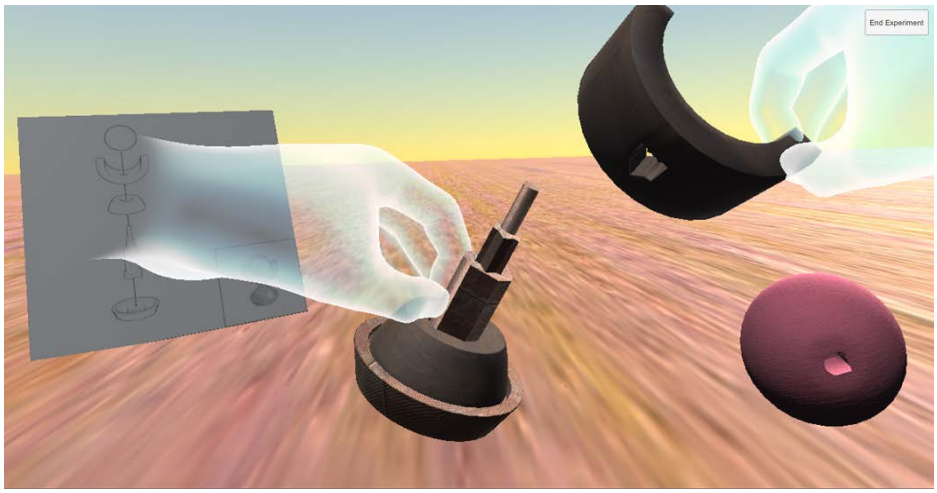


Figure 10. Avatar of virtual hands on the VE

b. Oculus Utilities for Unity

The Oculus Utilities for Unity is a basic package designed to help developers with the essential scripts, prefabs, and other resources to supplement Unity's built-in support, including "an interface for controlling VR camera behavior, a first-person control prefab, a unified input API for controllers, advanced rendering features, object-grabbing and haptics scripts for Touch, debugging tools, and more" ("Oculus Utilities for Unity | Developer Center | Oculus" 2018). This researcher used it mainly to integrate Oculus Rift cameras (HMD right-eye and left-eye cameras) to the application.

c. Virtual Reality Toolkit (VRTK) SDK Manager

VRTK is a package that includes useful scripts and concepts to help developers to build VR applications in Unity. Among others, the solution for interactions like touching, grabbing, and using objects, was an important aid on the development of this study's application.

d. TrackIR Enhanced SDK

In order to integrate TrackIR to Unity, this researcher used the TrackIR Enhanced SDK, gently provided by the TrackIR manufacturer, NaturalPoint Inc.

C. SYSTEM ARCHITECTURE

This study used two distinct system architectures: one for the immersive conditions, (conditions IS and INS), and one for the non-immersive conditions (conditions NIS and NINS). The first system architecture includes Oculus Rift bundle (with Oculus Touch controller) and PC, as shown in Figure 11.

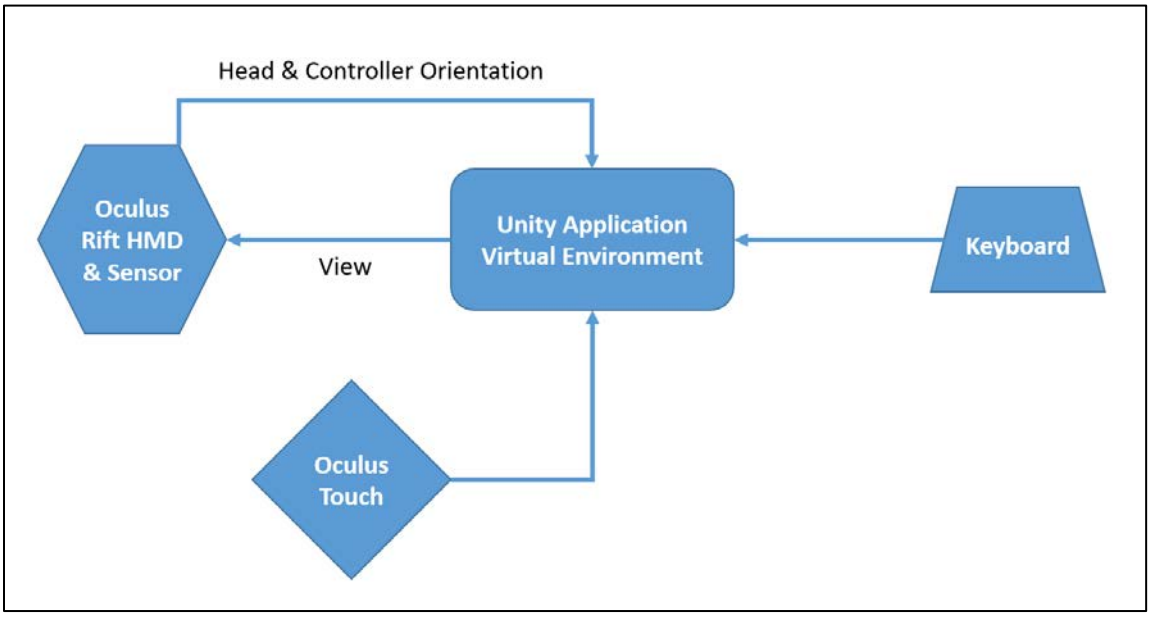


Figure 11. System architecture for immersive conditions

The second architecture includes Oculus Touch controller, PC, 3DTV, shutter glasses, and TrackIR, as shown in Figure 12.

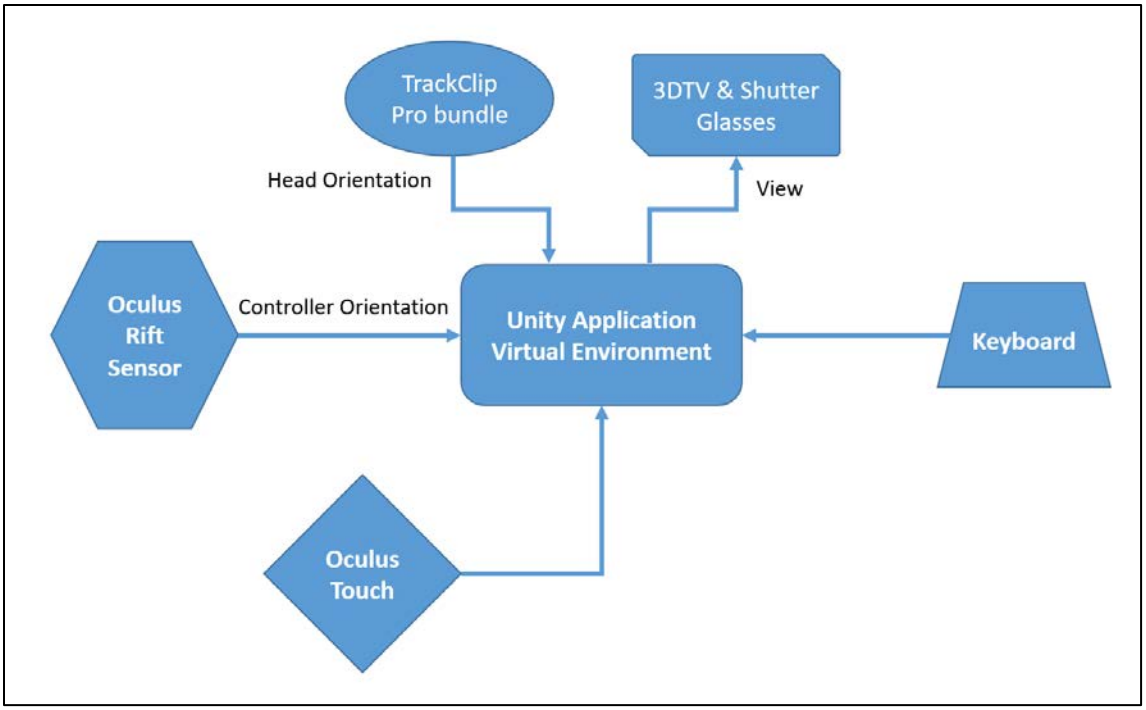


Figure 12. System architecture for non-immersive conditions

D. USER TASKS

Three user tasks were designed and used in this study:

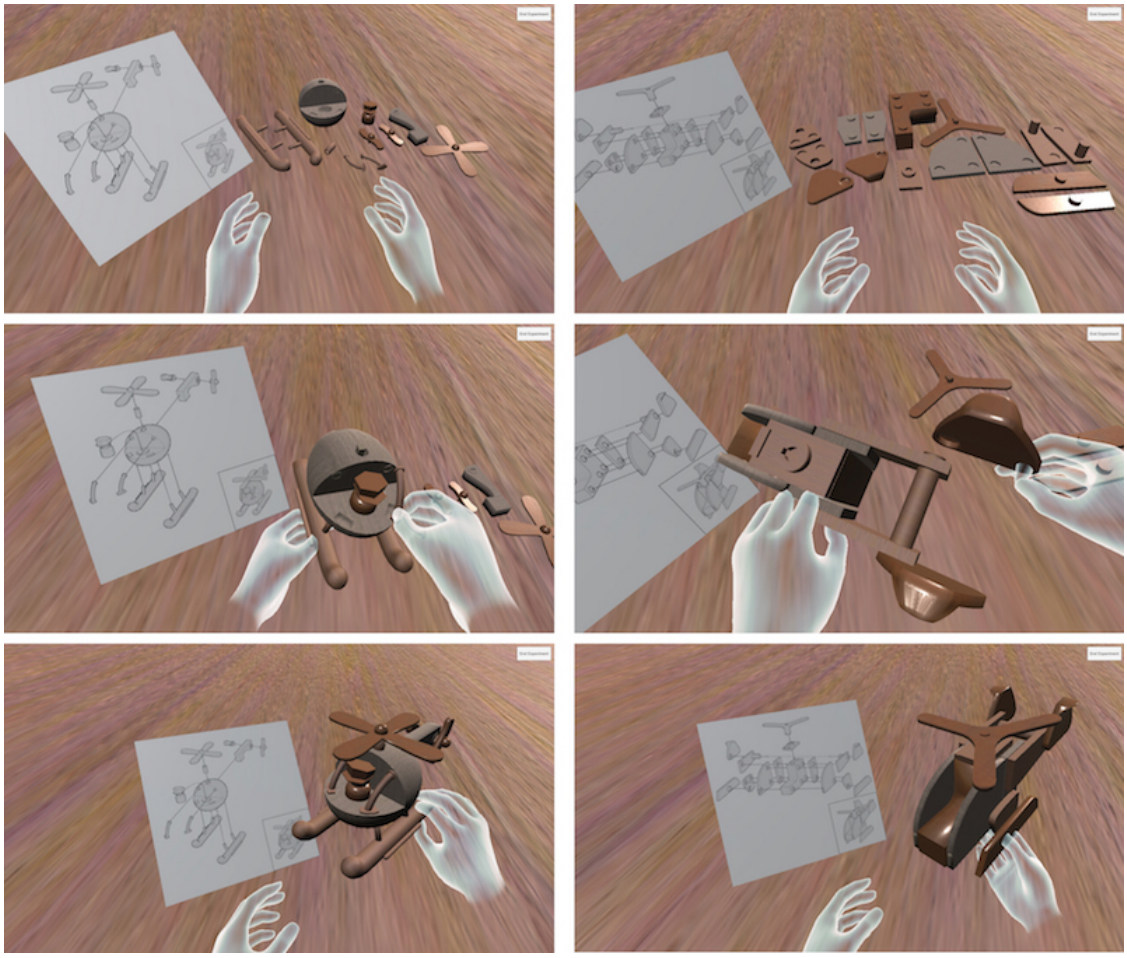
1. Task 1: Assembly of a virtual toy helicopter, as shown in the images on the left of Figure 13.

2. Task 2: Assembly of a more complex toy helicopter, as shown in the images on the right of Figure 13.

3. Task 3: Assembly of a physical object that was a 3D-printed version of the object from Task 2, as shown in Figure 14.

Task 3 served two main purposes: (1) it was an opportunity for participant self-discovery—to see the same parts in their physical form and comment on potential issues that prevented them from completing Task 2, and (2) to compare participant's body posture and assembly strategies in real life with body postures and assembly strategies used during virtual assembly tasks.

An additional task—a training task, as shown in Figure 15—was designed to allow participants to get familiarized with the virtual environment and with the devices used on the experiment. Participants got used to holding hand controllers, wearing headset or shutter glasses (depending on condition that they were in). The learned interaction techniques that would be used in the experimental tasks, how to manipulate the assembly diagram, and they got familiar with visual system feedback.



Stages of Tasks 1 and 2 assemblies: start stage (top), intermediate stage (middle), and end stage (bottom).

Figure 13. Models and stages of Task 1 (left) and Task 2 (right)

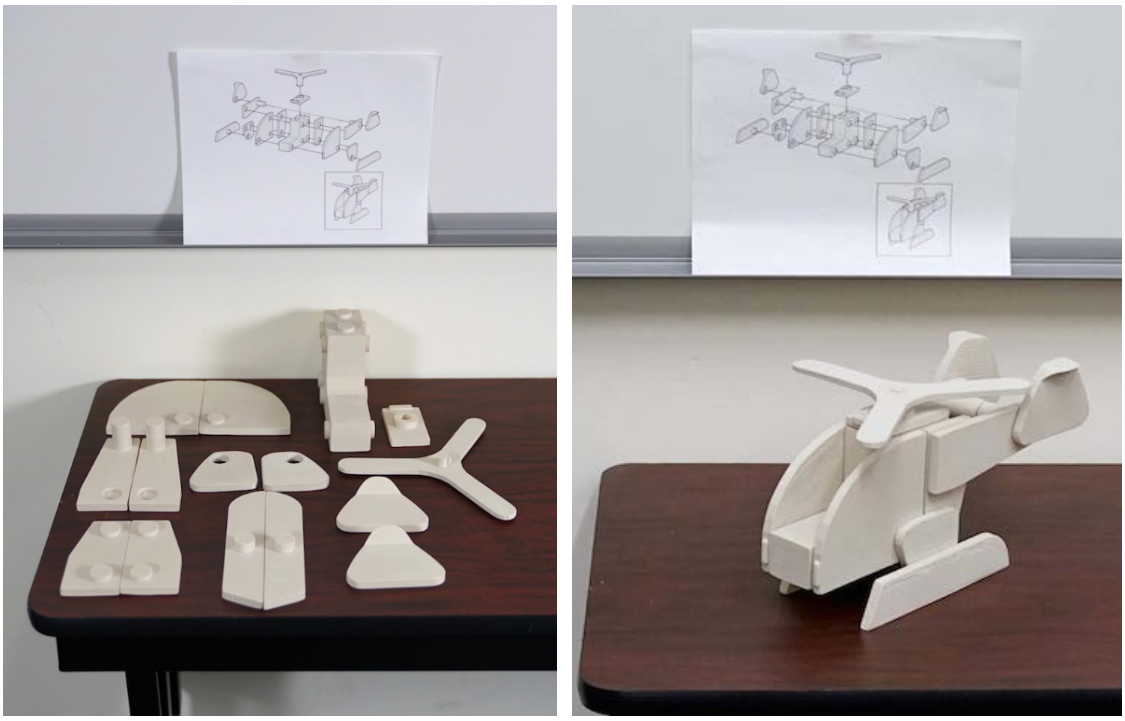


Figure 14. Initial setup layout and model assembled of Task 3

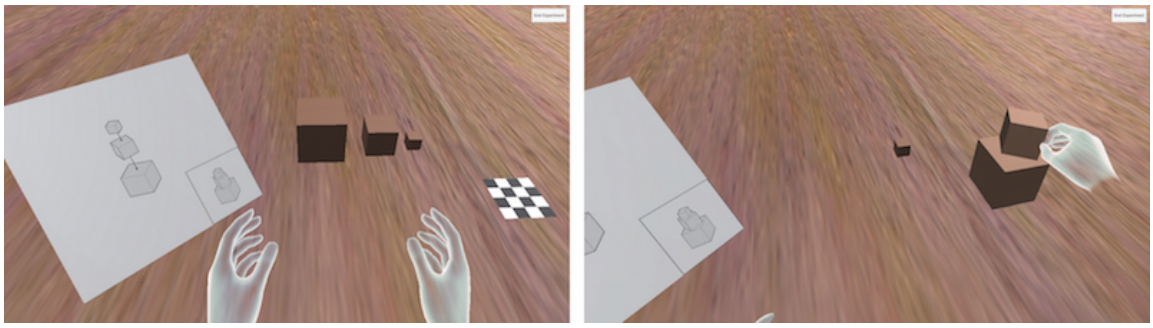


Figure 15. Start (left) and end (right) stages of the training task

In order to guide the participants on how to assemble the objects, we created a diagram for each of the tasks that showed how all parts should be connected, and how the object should look like at the end of the assembly, as shown in Figure 16. Participants could move and rotate the diagram and position it wherever it suited them.

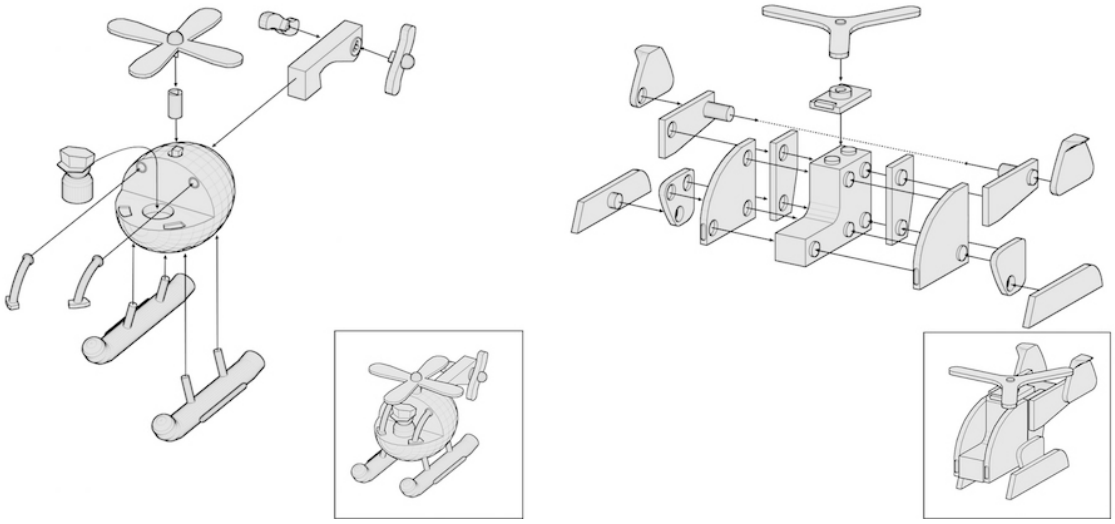


Figure 16. Assembly diagrams for Task 1 (left) and Task 2 (right)

E. SUMMARY

This chapter detailed the development of the system used to support user study, and elaborated the reasons that led us to choose specific hardware and software environment.

VI. PILOT TESTING

This chapter details the work done on refining the experiment through a pilot testing. Also, it describes the lessons learned and the modifications on the system and user study design that were done before we submitted it to IRB Committee.

A. USABILITY AND PHYSICAL FACTORS

A pilot testing was performed in order to identify factors that could affect the system usability, participant's physical conditions, and experiment pace. Our preliminary test participants were students, faculty and interns at NPS, and they went through partial and complete user study sessions. At the end, they provided feedback about the experiment set up, such as chair height and order of events; physical factors, such as arm's position while manipulating objects and tiredness level at the end of the session; and difficulty of the assembly tasks. The following avenues are the main feedbacks and lessons learned:

- **Experiment set up:** We defined the location of the apparatus used on the study, including chairs, table, 3DTV, PC, and video camera locations, as shown in Figure 17 and Figure 18.

- **Participant's position:** We adjusted the participant position both in physical and virtual environment so they could execute similar movements in all experimental conditions.

- **Heights of chair and virtual table:** We adjusted the chair and virtual table heights so the participants could keep their arms in a comfortable position, which minimized the tiredness effect on their performance.

- **Task difficulty:** We increased the distance and rotation angle tolerances of the snap drop zone (region in which the parts snap together if released) due to the difficulty presented by some of the participants when trying to connect the parts.

- **Task time limit:** We adjusted the time limit for the three tasks to ten minutes, so a larger number of participants could complete the tasks.
- **3D TV brightness and contrast:** We changed the brightness and contrast parameters values in order to make the image seen by the participants in all conditions as similar as possible.
- **Electronic questionnaires:** We fixed the input format so the participants could fill in the questionnaires in a faster and easier way.
- **3D objects:** We modified some parts of the 3D objects to make them more distinguishable from the other parts, and to highlight their correct orientation. Ambiguities in shape were rectified, and 3D models were changed accordingly.
- **3D printed object setup:** We defined a standard initial set up for the 3D printed parts.

Figure 17. Experiment setup

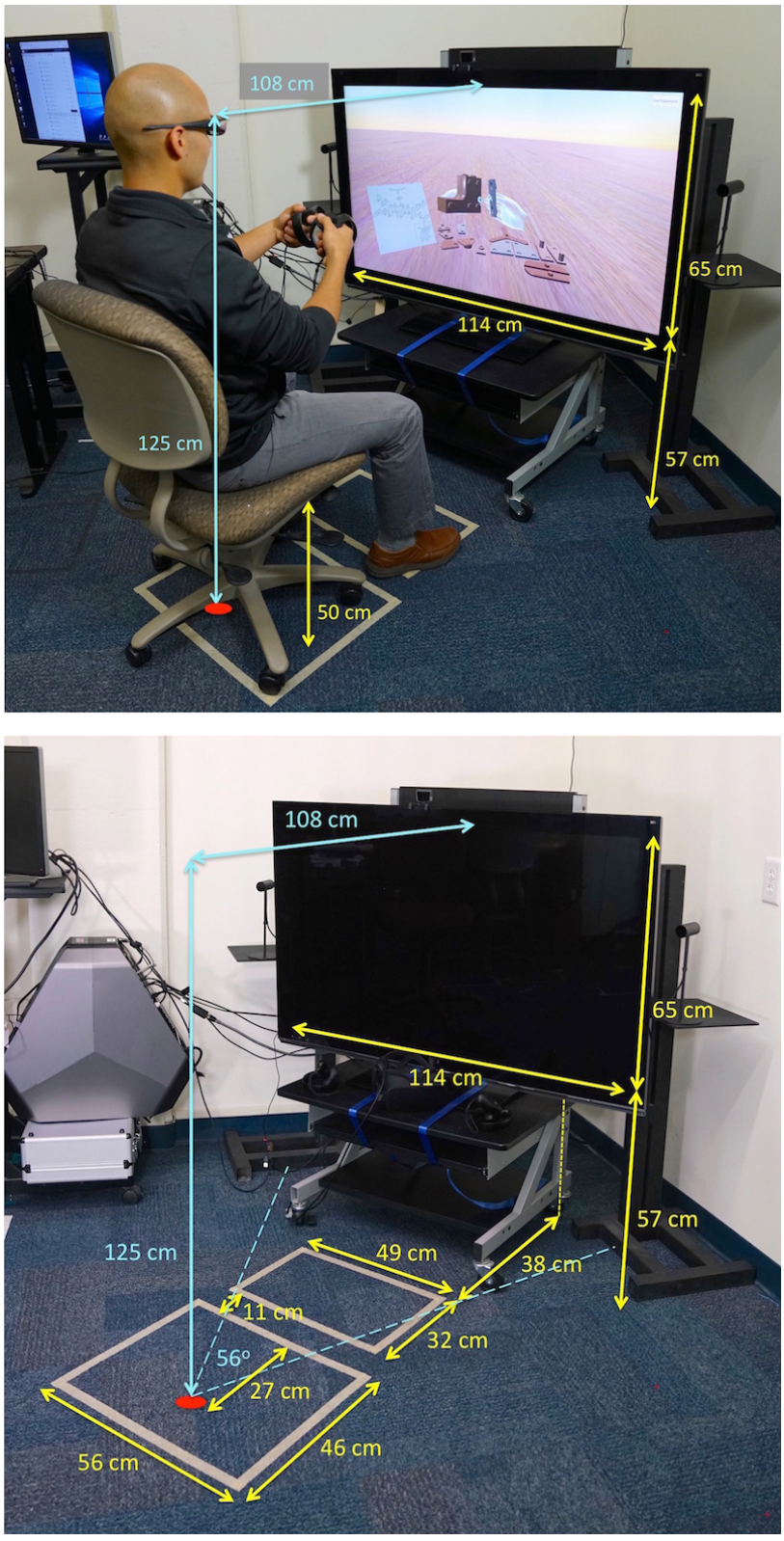


Figure 18. Main visualization station setup with dimensions

B. 3D PRINTING

We made 3D printed version of the toy helicopter that was used in Task 2, and noticed that some parts were easily detaching from each other during the assembly of its parts. This prompted us to slightly modify our virtual model and print a modified version. Modification on the connections included increasing the diameter size of the circular joints, adding bumps to the joints, and increasing their depth. We tested different types of connections and the one that worked best was the connection with two bumps around the circular joint, as shown in Figure 19.

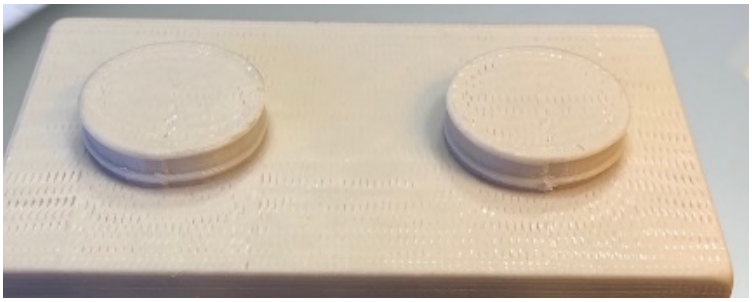


Figure 19. Connections on 3D printed parts with bumps added for easier assembly

C. SUMMARY

This chapter detailed the work done on the pilot testing, lessons learned, and modifications applied to the final experiment.

VII. RESULTS AND DISCUSSIONS

This chapter presents the results of user study. Sections in this chapter analyses the user performance, and discusses the results. The data sets that were analyzed include objective data set (system logs) and subjective data set (questionnaires).

A. RESULTS

This section presents a summary of the analysis of objective and subjective data set that were collected during the user study. The data set included participant's performance on each of the three tasks, post-task questionnaires, demographic survey and system usability questionnaires.

1. Demographic Survey Questionnaire

A total of 68 participants (24 female, 44 male; average age 37.28 years, standard deviation (SD) = 9.95) completed this study. Table 4 shows the distributions of participants per experimental condition, including their average age, and number of participants that successfully completed the tasks. Seven participants wore regular glasses and all of them kept the glasses during the sessions. When asked about what hand they used to manipulate the computer mouse, 65 reported they used the right hand and 3 reported they use either hand. Additionally, half of the participants used to play video games, and 29 of 68 (42.65%) have used VR HMD before the experiment. Regarding the problems faced during the tasks, a few participants reported that the image contrast was low on the NIS condition, making it difficult to identify objects' features.

Table 4. Distribution of participants per experimental condition

			IS		INS		NIS		NINS	
			Female	Male	Female	Male	Female	Male	Female	Male
Number of participants			12	5	7	10	1	16	4	13
Age (years)	**Average**		40.25	37.40	40.43	41.30	29.00	34.81	35.50	33.92
	SD		11.71	7.20	8.47	14.69	0.00	6.98	3.04	7.29
	Overall average		40.94		39.41		34.47		34.29	
	Overall SD		10.67		12.52		6.90		6.58	
Completed	**Task 1**	**#**	11	5	7	8	1	8	2	11
		%	92	100	100	80	100	50	50	85
	Task 2	**#**	9	5	3	6	1	8	3	11
		%	75	100	43	60	100	50	75	85

2. Objective Data Set

a. Type of Error

Unsuccessful assembly completions during tasks were due to one reason: the time allocated for object assembly has run out, i.e., participants did not complete the object assembly within the given maximum time (610s). Completion time for participants that did not assemble the objects were corrected by assigning them the upper time limit (610s). It is important to highlight that we considered the total time limit as 610 seconds instead of 600 seconds due to imprecision on starting and stopping the time watch.

b. Number of Successful Assemblies

The number of successful and unsuccessful assemblies in Tasks 1, 2 and 3 is shown in Figure 20. Figure 21 shows the number of successful and unsuccessful assemblies in the VE (Task 1 and Task 2) and overall (all three tasks)

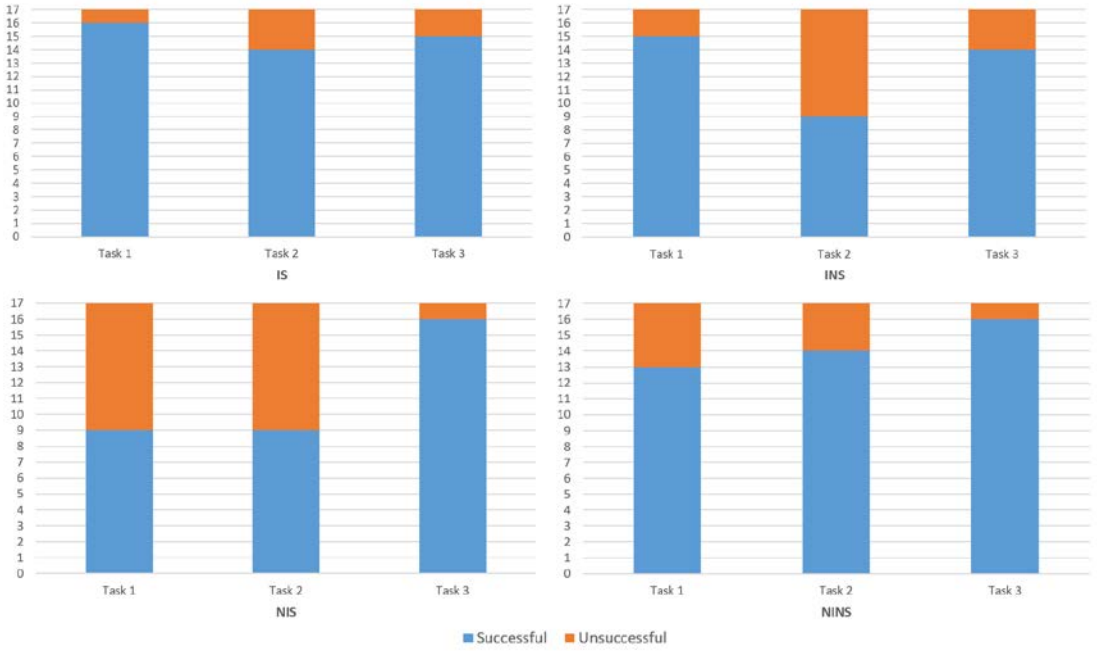


Figure 20. Number of successful (blue) and unsuccessful (orange) assemblies in Tasks 1, 2 and 3 for each experimental condition

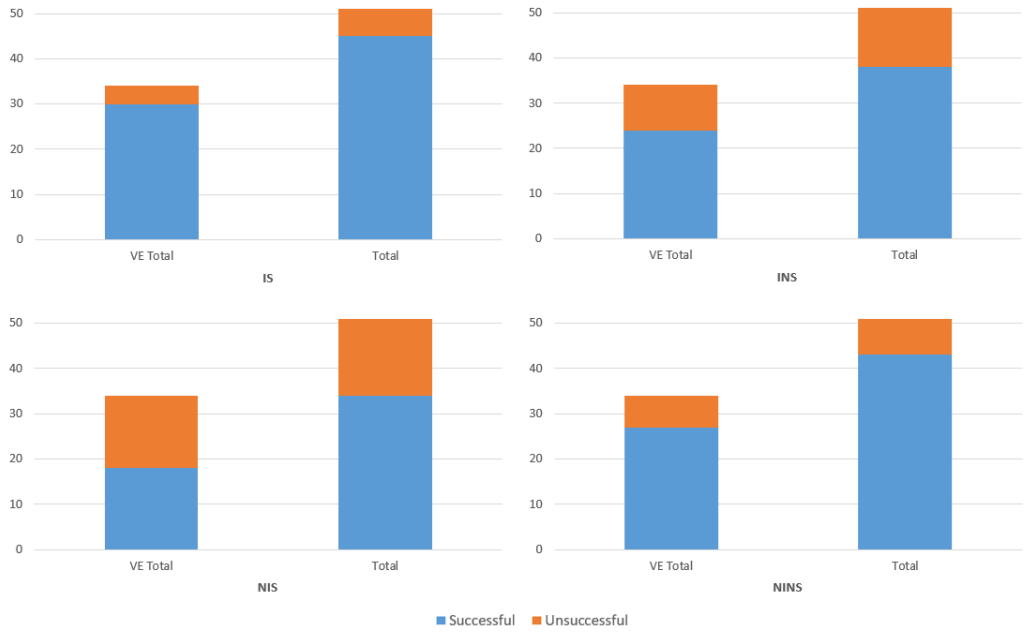


Figure 21. Number of successful (blue) and unsuccessful (orange) assemblies for each experimental condition in the VE (VE Total) and overall (Total)

IS condition resulted with the largest number of successful assemblies for Task 1, Task 2 (tied with NINS condition), VE Total and Overall Total, as show in Table 5, in which maximum values are highlighted in red. For the third task, NIS and NINS conditions resulted with the largest number of successful assemblies.

Table 5. Number of successful assemblies for each task, VE Total and Overall Total

	IS	INS	NIS	NINS
Task 1	**16**	15	9	13
Task 2	**14**	9	9	**14**
Task 3	15	14	**16**	**16**
VE Total	**30**	24	18	27
Overall Total	**45**	38	34	43

c. Assembly Times

Boxplots with assembly times for each task are shown in Figure 22, and boxplots with the assembly times for VE Total and Overall Total are shown in Figure 23. We performed a non-parametric statistical analysis for assembly times—Task 1, Task 2, VE Total and Overall Total times—because the collected data was not normally distributed as shown by a Shapiro-Wilk test in Appendix I.

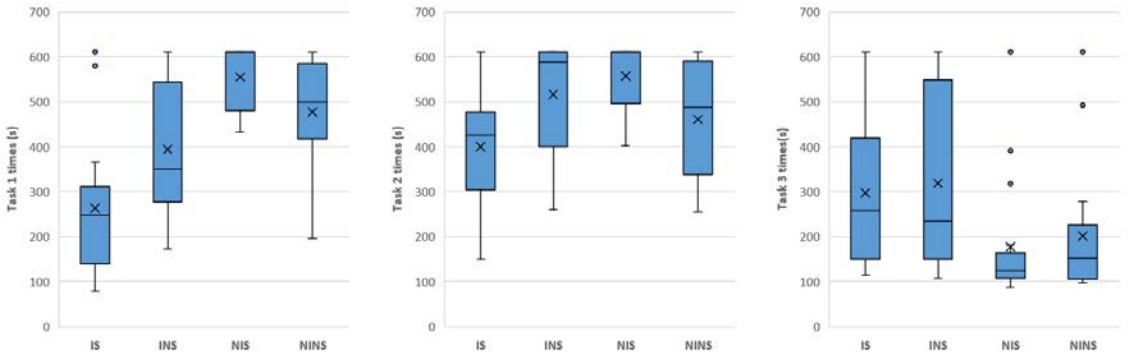


Figure 22. Boxplot containing assembly times for each task

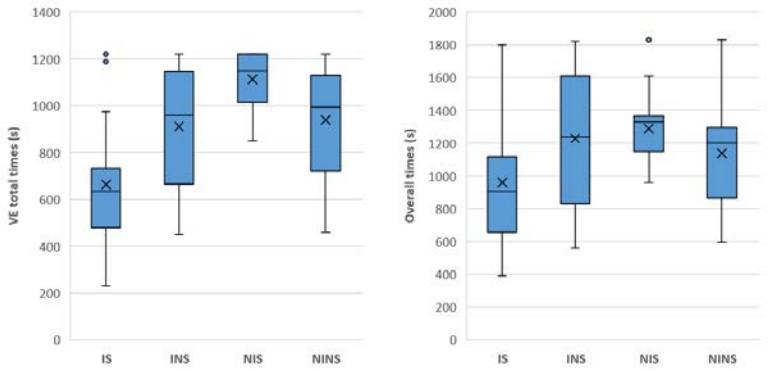


Figure 23. Boxplot containing VE Total and Overall Total assembly times

We performed a Kruskall-Wallis H test to check if there was an overall statistically significant difference in assembly times between the four experimental conditions for each of the three tasks, VE Total and Overall Total times. Additionally, we performed pairwise comparisons using Dunn's method for joint ranking with Bonferroni adjustment.

(1) Task 1

The means and standard deviations of assembly times for the first task are shown in Figure 24.

Means and Std Deviations

Level	Number	Mean	Std Dev	Std Err Mean	Lower 95%	Upper 95%
IS	17	262,588	147,555	35,787	186,72	338,45
INS	17	393,882	148,010	35,898	317,78	469,98
NIS	17	554,824	68,434	16,598	519,64	590,01
NINS	17	477,235	115,026	27,898	418,09	536,38

Figure 24. Means and standard deviations of assembly times for Task 1

The Kruskall-Wallis H test for the first task showed that there was an overall statistically significant difference in assembly times between the four experimental conditions, $\chi^2(3) = 28.1916$, $p < 0.0001$, with a mean rank score of 16.5882 for IS, 30.3824 for INS, 50.9706 for NIS, and 40.0588 for NINS, as shown in Figure 25.

Wilcoxon / Kruskal-Wallis Tests (Rank Sums)

Level	Count	Score Sum	Expected Score	Score Mean	(Mean-Mean0)/Std0
IS	17	282,000	586,500	16,5882	-4,334
INS	17	516,500	586,500	30,3824	-0,991
NIS	17	866,500	586,500	50,9706	3,985
NINS	17	681,000	586,500	40,0588	1,340

1-Way Test, ChiSquare Approximation

ChiSquare	DF	Prob>ChiSq
28,1916	3	<,0001*

Figure 25. Kruskall-Wallis H test of assembly times for Task 1

The pairwise comparison using Dunn's method showed that there was statistically significant difference in assembly times for the first task. There was a statistically significant difference between NIS (mean rank = 50.9706) and IS (mean rank = 16.5882) ($p < 0.0001$), NINS (mean rank = 40.0588) and IS (mean rank = 16.5882) ($p = 0.0031$), and NIS (mean rank = 50.9706) and INS (mean rank = 30.3824) ($p = 0.0139$), as shown in Figure 26.

Nonparametric Comparisons For All Pairs Using Dunn Method For Joint Ranking

Level	- Level	Score Mean Difference	Std Err Dif	Z	p-Value
NIS	IS	34,3235	6,738043	5,09399	<,0001*
NINS	IS	23,4118	6,738043	3,47456	0,0031*
NIS	INS	20,5294	6,738043	3,04679	0,0139*
INS	IS	13,7353	6,738043	2,03847	0,2490
NINS	INS	9,6176	6,738043	1,42736	0,9208
NINS	NIS	-10,8529	6,738043	-1,61070	0,6435

Figure 26. Pairwise comparisons using Dunn's method of assembly times for Task 1

(2) Task 2

The means and standard deviations of assembly times for the second task are shown in Figure 27.

Means and Std Deviations

Level	Number	Mean	Std Dev	Std Err Mean	Lower 95%	Upper 95%
IS	17	400,294	139,435	33,818	328,60	471,99
INS	17	516,000	133,518	32,383	447,35	584,65
NIS	17	557,000	79,451	19,270	516,15	597,85
NINS	17	461,000	127,459	30,913	395,47	526,53

Figure 27. Means and standard deviations of assembly times for Task 2

The Kruskall-Wallis H test for the second task showed that there was an overall statistically significant difference in assembly times between the four experimental conditions, $\chi^2(3) = 11.8441$, $p = 0.0079$, with a mean rank score of 23.5294 for IS, 39.6176 for INS, 44.3824 for NIS, and 30.4706 for NINS, as shown in Figure 28.

Wilcoxon / Kruskal-Wallis Tests (Rank Sums)

Level	Count	Score Sum	Expected Score	Score Mean	(Mean-Mean0)/Std0
IS	17	400,000	586,500	23,5294	-2,694
INS	17	673,500	586,500	39,6176	1,253
NIS	17	754,500	586,500	44,3824	2,426
NINS	17	518,000	586,500	30,4706	-0,985

1-Way Test, ChiSquare Approximation

ChiSquare	DF	Prob>ChiSq
11,8441	3	0,0079*

Figure 28. Kruskall-Wallis H test of assembly times for Task 2

The pairwise comparison using Dunn's method showed that there was statistically significant difference in assembly times for the second task. There was a statistically significant difference between NIS (mean rank = 44.3824) and IS (mean rank = 23.5294) ($p = 0.0103$), as shown in Figure 29.

Nonparametric Comparisons For All Pairs Using Dunn Method For Joint Ranking

Level	- Level	Score Mean Difference	Std Err Dif	Z	p-Value
NIS	IS	20,7941	6,631595	3,13561	0,0103*
INS	IS	16,0294	6,631595	2,41713	0,0939
NINS	IS	6,8824	6,631595	1,03781	1,0000
NIS	INS	4,7059	6,631595	0,70962	1,0000
NINS	INS	-9,0882	6,631595	-1,37044	1,0000
NINS	NIS	-13,8529	6,631595	-2,08893	0,2203

Figure 29. Pairwise comparisons using Dunn's method of assembly times for Task 2

(3) Task 3

The means and standard deviations of assembly times for the third task are shown in Figure 30.

Means and Std Deviations

Level	Number	Mean	Std Dev	Std Err Mean	Lower 95%	Upper 95%
IS	17	297,529	163,546	39,666	213,44	381,62
INS	17	318,647	196,007	47,539	217,87	419,42
NIS	17	177,706	138,073	33,488	106,72	248,70
NINS	17	201,059	143,662	34,843	127,19	274,92

Figure 30. Means and standard deviations of assembly times for Task 3

The Kruskall-Wallis H test for the third task showed that there was an overall statistically significant difference in assembly times between the four experimental conditions, $\chi^2(3) = 13.3850$, $p = 0.0039$, with a mean rank score of 43.0000 for IS, 43.0882 for INS, 23.1765 for NIS, and 28.7353 for NINS, as shown in Figure 31.

Wilcoxon / Kruskal-Wallis Tests (Rank Sums)

Level	Count	Score Sum	Expected Score	Score Mean	(Mean-Mean0)/Std0
IS	17	731,000	586,500	43,0000	2,041
INS	17	732,500	586,500	43,0882	2,062
NIS	17	394,000	586,500	23,1765	-2,721
NINS	17	488,500	586,500	28,7353	-1,382

1-Way Test, ChiSquare Approximation

ChiSquare	DF	Prob>ChiSq
13,3850	3	0,0039*

Figure 31. Kruskall-Wallis H test of assembly times for Task 3

The pairwise comparison using Dunn's method showed that there was statistically significant difference in assembly times for the second task. There was a statistically significant difference between NIS (mean rank = 23.1765) and IS (mean rank = 43.0000) (p = 0.0213), and NIS (mean rank = 23.1765) and INS (mean rank = 43.0882) (p = 0.0204), as shown in Figure 32.

Nonparametric Comparisons For All Pairs Using Dunn Method For Joint Ranking

Level	- Level	Score Mean Difference	Std Err Dif	Z	p-Value
NINS	NIS	5,5000	6,777992	0,81145	1,0000
INS	IS	0,0294	6,777992	0,00434	1,0000
NINS	IS	-14,2059	6,777992	-2,09588	0,2166
NINS	INS	-14,2941	6,777992	-2,10890	0,2097
NIS	IS	-19,7647	6,777992	-2,91601	0,0213*
NIS	INS	-19,8529	6,777992	-2,92903	0,0204*

Figure 32. Pairwise comparisons using Dunn's method of assembly times for Task 3

(4) VE Total

The means and standard deviations for the VE Total assembly times are shown in Figure 33.

Means and Std Deviations

Level	Number	Mean	Std Dev	Std Err Mean	Lower 95%	Upper 95%
IS	17	662,88	271,140	65,761	523,5	802,3
INS	17	909,88	261,112	63,329	775,6	1044,1
NIS	17	1111,82	126,923	30,783	1046,6	1177,1
NINS	17	938,24	228,799	55,492	820,6	1055,9

Figure 33. Means and standard deviations of the VE Total assembly times for all experimental conditions

The Kruskall-Wallis H test for the VE Total assembly times showed that there was an overall statistically significant difference in assembly times between the four experimental conditions, χ^2 (3) = 21.3136, $p < 0.0001$, with a mean rank score of 18.5294 for IS, 33.9412 for INS, 49.6765 for NIS, and 35.8529 for NINS, as shown in Figure 34.

Wilcoxon / Kruskal-Wallis Tests (Rank Sums)

Level	Count	Score Sum	Expected Score	Score Mean	(Mean-Mean0)/Std0
IS	17	315,000	586,500	18,5294	-3,849
INS	17	577,000	586,500	33,9412	-0,128
NIS	17	844,500	586,500	49,6765	3,657
NINS	17	609,500	586,500	35,8529	0,320

1-Way Test, ChiSquare Approximation

ChiSquare	DF	Prob>ChiSq
21,3136	3	<,0001*

Figure 34. Kruskall-Wallis H test of assembly times for VE tasks

The pairwise comparison using Dunn's method showed that there was statistically significant difference in the VE Total assembly times. There was a statistically significant difference between NIS (mean rank = 49.6765) and IS (mean rank = 18.5294) ($p < 0.0001$), as shown in Figure 35.

Nonparametric Comparisons For All Pairs Using Dunn Method For Joint Ranking

Level	- Level	Score Mean Difference	Std Err Dif	Z	p-Value
NIS	IS	31,0882	6,763729	4,59632	<,0001*
NINS	IS	17,2647	6,763729	2,55254	0,0642
NIS	INS	15,6765	6,763729	2,31773	0,1228
INS	IS	15,3529	6,763729	2,26989	0,1393
NINS	INS	1,8529	6,763729	0,27395	1,0000
NINS	NIS	-13,7647	6,763729	-2,03508	0,2511

Figure 35. Pairwise comparisons using Dunn's method of assembly times for VE tasks

(5) Overall Total

The means and standard deviations for the Overall Total assembly times are shown in Figure 36.

Means and Std Deviations

Level	Number	Mean	Std Dev	Std Err Mean	Lower 95%	Upper 95%
IS	17	960,41	398,264	96,59	755,6	1165,2
INS	17	1228,53	416,748	101,08	1014,3	1442,8
NIS	17	1289,53	215,731	52,32	1178,6	1400,4
NINS	17	1139,29	311,029	75,44	979,4	1299,2

Figure 36. Means and standard deviations of the Overall Total assembly times for all experimental conditions

The Kruskall-Wallis H test for the Overall Total assembly times showed that there was an overall statistically significant difference in assembly times between the four experimental conditions, $\chi^2(3) = 9.1096$, $p < 0.0279$, with a mean rank score of 23.6471 for IS, 38.3529 for INS, 43.1176 for NIS, and 32.8824 for NINS, as shown in Figure 37.

Wilcoxon / Kruskal-Wallis Tests (Rank Sums)

Level	Count	Score Sum	Expected Score	Score Mean	(Mean-Mean0)/Std0
IS	17	402,000	586,500	23,6471	-2,606
INS	17	652,000	586,500	38,3529	0,921
NIS	17	733,000	586,500	43,1176	2,068
NINS	17	559,000	586,500	32,8824	-0,382

1-Way Test, ChiSquare Approximation

ChiSquare	DF	Prob>ChiSq
9,1096	3	0,0279*

Figure 37. Kruskall-Wallis H test for of assembly times for all tasks

The pairwise comparison using Dunn's method showed that there was statistically significant difference in the VE Total assembly times. There was a statistically significant difference between NIS (mean rank = 38.3529) and IS (mean rank = 23.6471) (p = 0.0252), as shown in Figure 38.

Nonparametric Comparisons For All Pairs Using Dunn Method For Joint Ranking

Level	- Level	Score Mean Difference	Std Err Dif	Z	p-Value
NIS	IS	19,4118	6,782201	2,86216	0,0252*
INS	IS	14,6471	6,782201	2,15963	0,1848
NINS	IS	9,1765	6,782201	1,35302	1,0000
NIS	INS	4,7059	6,782201	0,69386	1,0000
NINS	INS	-5,4118	6,782201	-0,79794	1,0000
NINS	NIS	-10,1765	6,782201	-1,50047	0,8010

Figure 38. Pairwise comparisons using Dunn's method f of assembly times for all tasks

d. Correct Connections/Parts

Boxplots with number of correct connections/parts for each task are shown in Figure 39, and boxplots with VE Total and Overall Total number of correct connections/parts are shown in Figure 40. We performed a non-parametric statistical analysis for number of correct connections/parts—Task 1, Task 2, VE Total and Overall

Total—because the collected data was not normally distributed as shown by a Shapiro-Wilk test in Appendix J.

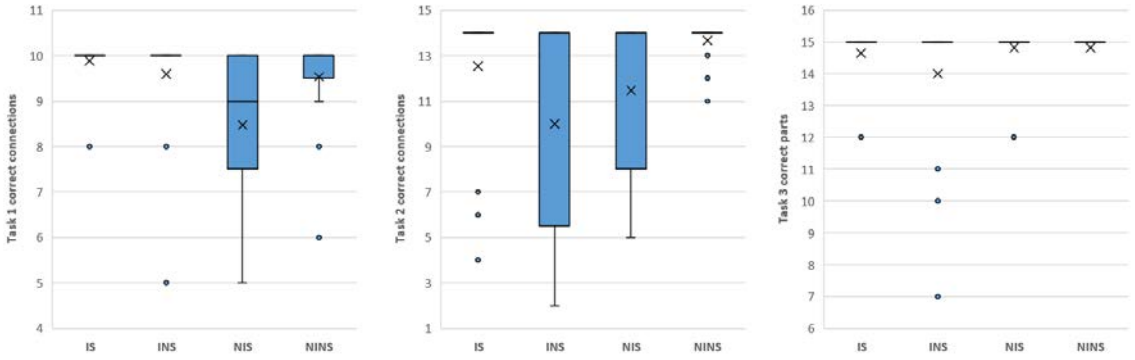


Figure 39. Boxplot containing number of correct connections/parts for each task

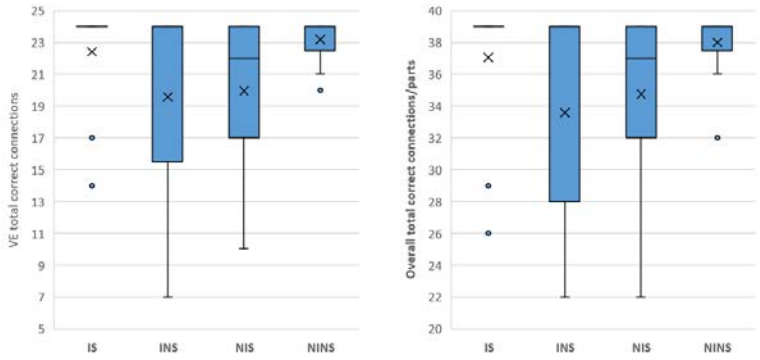


Figure 40. Boxplot containing VE Total and Overall Total number of correct connections/parts

We performed a Kruskall-Wallis H test to check if there was an overall statistically significant difference in number of correct connections/parts between the four experimental conditions for each of the three tasks, VE Total, and Overall Total. Additionally, we performed pairwise comparisons using Dunn's method for joint ranking with Bonferroni adjustment.

(1) Task 1

The means and standard deviations of number of correct connections for the first task are shown in Figure 41.

Means and Std Deviations

Level	Number	Mean	Std Dev	Std Err Mean	Lower 95%	Upper 95%
IS	17	9,88235	0,48507	0,11765	9,6330	10,132
INS	17	9,58824	1,27764	0,30987	8,9313	10,245
NIS	17	8,47059	1,80685	0,43823	7,5416	9,400
NINS	17	9,52941	1,06757	0,25892	8,9805	10,078

Figure 41. Means and standard deviations of number of correct connections for Task 1

The Kruskall-Wallis H test for the first task showed that there was an overall statistically significant difference in number of correct connections between the four experimental conditions, $\chi^2(3) = 12.5779$, $p < 0.0056$, with a mean rank score of 40.5882 for IS, 38.2059 for INS, 24.0882 for NIS, and 35.1176 for NINS, as shown in Figure 42.

Wilcoxon / Kruskal-Wallis Tests (Rank Sums)

Level	Count	Score Sum	Expected Score	Score Mean	(Mean-Mean0)/Std0
IS	17	690,000	586,500	40,5882	1,964
INS	17	649,500	586,500	38,2059	1,192
NIS	17	409,500	586,500	24,0882	-3,366
NINS	17	597,000	586,500	35,1176	0,191

1-Way Test, ChiSquare Approximation

ChiSquare	DF	Prob>ChiSq
12,5779	3	0,0056*

Figure 42. Kruskall-Wallis H test of correct connections for Task 1

The pairwise comparison using Dunn's method showed that there was statistically significant difference in number of correct connections for the first task. There was a statistically significant difference between NIS (mean rank = 24.0882) and INS (mean rank

= 38.2059) (p = 0.0315), and NIS (mean rank = 24.0882) and IS (mean rank = 40.5882) (p = 0.0066), as shown in Figure 43.

Nonparametric Comparisons For All Pairs Using Dunn Method For Joint Ranking

Level	- Level	Score Mean Difference	Std Err Dif	Z	p-Value
NINS	NIS	10,9706	5,037437	2,17781	0,1765
INS	IS	-2,3235	5,037437	-0,46125	1,0000
NINS	INS	-3,0294	5,037437	-0,60138	1,0000
NINS	IS	-5,4118	5,037437	-1,07431	1,0000
NIS	INS	-14,0588	5,037437	-2,79087	0,0315*
NIS	IS	-16,4412	5,037437	-3,26380	0,0066*

Figure 43. Pairwise comparisons using Dunn's method of correct connections for Task 1

(2) Task 2

The means and standard deviations of number of correct connections for the second task are shown in Figure 44.

Means and Std Deviations

Level	Number	Mean	Std Dev	Std Err Mean	Lower 95%	Upper 95%
IS	17	12,5294	3,31884	0,8049	10,823	14,236
INS	17	10,0000	4,63681	1,1246	7,616	12,384
NIS	17	11,4706	3,29996	0,8004	9,774	13,167
NINS	17	13,6471	0,86177	0,2090	13,204	14,090

Figure 44. Means and standard deviations of number of correct connections for Task 2

The Kruskall-Wallis H test for the second task showed that there was no overall statistically significant difference in number of correct connections between the four experimental conditions, χ^2 (3) = 7.7643, p = 0.0511, with a mean rank score of 38.8235 for IS, 27.6176 for INS, 30.5882 for NIS, and 40.9706 for NINS, as shown in Figure 45.

Wilcoxon / Kruskal-Wallis Tests (Rank Sums)

Level	Count	Score Sum	Expected Score	Score Mean	(Mean-Mean0)/Std0
IS	17	660,000	586,500	38,8235	1,245
INS	17	469,500	586,500	27,6176	-1,986
NIS	17	520,000	586,500	30,5882	-1,125
NINS	17	696,500	586,500	40,9706	1,867

1-Way Test, ChiSquare Approximation

ChiSquare	DF	Prob>ChiSq
7,7643	3	0,0511

Figure 45. Kruskall-Wallis H test of correct connections for Task 2

(3) Task 3

The means and standard deviations of number of correct parts for the third task are shown in Figure 46.

Means and Std Deviations

Level	Number	Mean	Std Dev	Std Err Mean	Lower 95%	Upper 95%
IS	17	14,6471	0,99632	0,24164	14,135	15,159
INS	17	14,0000	2,34521	0,56880	12,794	15,206
NIS	17	14,8235	0,72761	0,17647	14,449	15,198
NINS	17	14,8235	0,72761	0,17647	14,449	15,198

Figure 46. Means and standard deviations of number of correct parts for Task 3

The Kruskall-Wallis H test for the third task showed that there was no overall statistically significant difference in number of correct parts between the four experimental conditions, $\chi^2(3) = 2.0786$, $p = 0.05563$, with a mean rank score of 34.1765 for IS, 31.6471 for INS, 36.0882 for NIS, and 36.0882 for NINS, as shown in Figure 47.

Wilcoxon / Kruskal-Wallis Tests (Rank Sums)

Level	Count	Score Sum	Expected Score	Score Mean	(Mean-Mean0)/Std0
IS	17	581,000	586,500	34,1765	-0,134
INS	17	538,000	586,500	31,6471	-1,289
NIS	17	613,500	586,500	36,0882	0,712
NINS	17	613,500	586,500	36,0882	0,712

1-Way Test, ChiSquare Approximation

ChiSquare	DF	Prob>ChiSq
2,0786	3	0,5563

Figure 47. Kruskall-Wallis H test of correct parts for Task 3

(4) VE Total

The means and standard deviations of the VE Total number of correct connections are shown in Figure 48.

Means and Std Deviations

Level	Number	Mean	Std Dev	Std Err Mean	Lower 95%	Upper 95%
IS	17	22,4118	3,58920	0,8705	20,566	24,257
INS	17	19,5882	5,42055	1,3147	16,801	22,375
NIS	17	19,9412	4,22005	1,0235	17,771	22,111
NINS	17	23,1765	1,33395	0,3235	22,491	23,862

Figure 48. Means and standard deviations of correct connections for VE tasks

The Kruskall-Wallis H test for the VE Total showed that there was an overall statistically significant difference in number of correct connections between the four experimental conditions, χ^2 (3) = 9.4311, p = 0.0241, with a mean rank score of 41.8529 for IS, 30.0882 for INS, 26.0294 for NIS, and 40.0294 for NINS, as shown in Figure 49.

Wilcoxon / Kruskal-Wallis Tests (Rank Sums)

Level	Count	Score Sum	Expected Score	Score Mean	(Mean-Mean0)/Std0
IS	17	711,500	586,500	41,8529	1,958
INS	17	511,500	586,500	30,0882	-1,172
NIS	17	442,500	586,500	26,0294	-2,257
NINS	17	680,500	586,500	40,0294	1,471

1-Way Test, ChiSquare Approximation

ChiSquare	DF	Prob>ChiSq
9,4311	3	0,0241*

Figure 49. Kruskall-Wallis H test of correct connections for VE tasks

However, the pairwise comparison using Dunn's method showed that there was no statistically significant difference ($p < 0.05$) in VE Total number of correct connections. The difference that was closest to be statistically significant was between NIS (mean rank = 26.0294) and IS (mean rank = 41.8529) ($p = 0.0590$), as shown in Figure 50.

Nonparametric Comparisons For All Pairs Using Dunn Method For Joint Ranking

Level	- Level	Score Mean Difference	Std Err Dif	Z	p-Value
NINS	NIS	13,9412	6,106747	2,28291	0,1346
NINS	INS	9,8824	6,106747	1,61827	0,6336
NINS	IS	-1,7647	6,106747	-0,28898	1,0000
NIS	INS	-4,0000	6,106747	-0,65501	1,0000
INS	IS	-11,7059	6,106747	-1,91688	0,3315
NIS	IS	-15,7647	6,106747	-2,58152	0,0590

Figure 50. Pairwise comparisons using Dunn's method of correct connections for VE tasks

(5) Overall Total

The means and standard deviations for the Overall Total number of correct connections/parts are shown in Figure 51.

Means and Std Deviations

Level	Number	Mean	Std Dev	Std Err Mean	Lower 95%	Upper 95%
IS	17	37,0588	4,36564	1,0588	34,814	39,303
INS	17	33,5882	6,40370	1,5531	30,296	36,881
NIS	17	34,7647	4,69746	1,1393	32,349	37,180
NINS	17	38,0000	1,87083	0,4537	37,038	38,962

Figure 51. Means and standard deviations of the Overall Total number of correct connections/parts

The Kruskall-Wallis H test for the first task showed that there was an overall statistically significant difference in number of correct connections between the four experimental conditions, χ^2 (3) = 8.6899, p = 0.0337, with a mean rank score of 41.6765 for IS, 29.6765 for INS, 26.8235 for NIS, and 39.8235 for NINS, as shown in Figure 52.

Wilcoxon / Kruskal-Wallis Tests (Rank Sums)

Level	Count	Score Sum	Expected Score	Score Mean	(Mean-Mean0)/Std0
IS	17	708,500	586,500	41,6765	1,911
INS	17	504,500	586,500	29,6765	-1,282
NIS	17	456,000	586,500	26,8235	-2,045
NINS	17	677,000	586,500	39,8235	1,416

1-Way Test, ChiSquare Approximation

ChiSquare	DF	Prob>ChiSq
8,6899	3	0,0337*

Figure 52. Kruskall-Wallis H test of number of correct connections/parts for all tasks

However, the pairwise comparison using Dunn's method showed that there was no statistically significant difference (p < 0.05) in Overall Total number of correct connections/parts. The difference that was closest to be statistically significant was between NIS (mean rank = 26.8235) and IS (mean rank = 41.6765) (p = 0.0925), as shown in Figure 53.

Nonparametric Comparisons For All Pairs Using Dunn Method For Joint Ranking

Level	- Level	Score Mean Difference	Std Err Dif	Z	p-Value
NINS	NIS	12,9412	6,106819	2,11914	0,2045
NINS	INS	10,0882	6,106819	1,65196	0,5913
NINS	IS	-1,7941	6,106819	-0,29379	1,0000
NIS	INS	-2,7941	6,106819	-0,45754	1,0000
INS	IS	-11,9412	6,106819	-1,95538	0,3032
NIS	IS	-14,7941	6,106819	-2,42256	0,0925

Figure 53. Pairwise comparisons using Dunn's method of number of correct connections/parts for all tasks

e. Collisions

Boxplots with number of collisions for tasks 1, 2, and VE Total are shown in Figure 54. We performed a non-parametric statistical analysis of number of collisions—Task 1, Task 2, and VE Total—because the collected data was not normally distributed as shown by a Shapiro-Wilk test in Appendix K.

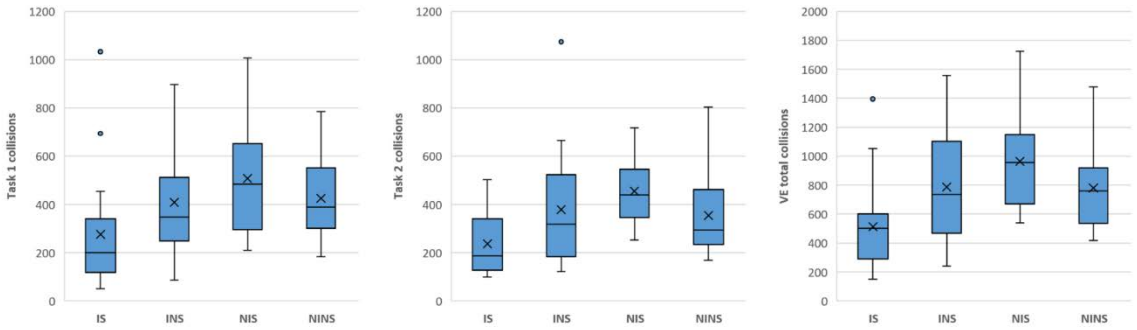


Figure 54. Boxplot containing number of collisions for each task

We performed a Kruskall-Wallis H test to check if there was an overall statistically significant difference in number of collisions between the four experimental conditions for tasks 1, 2, and VE Total. Additionally, we performed pairwise comparisons using Dunn's method for joint ranking with Bonferroni adjustment.

(1) Task 1

The means and standard deviations of number of collisions for the first task are shown in Figure 55.

Means and Std Deviations

Level	Number	Mean	Std Dev	Std Err Mean	Lower 95%	Upper 95%
IS	17	275,882	251,315	60,953	146,67	405,10
INS	17	408,529	227,913	55,277	291,35	525,71
NIS	17	508,000	211,709	51,347	399,15	616,85
NINS	17	425,588	170,189	41,277	338,09	513,09

Figure 55. Means and standard deviations of number of collisions for Task 1

The Kruskall-Wallis H test for the first task showed that there was an overall statistically significant difference in number of collisions between the four experimental conditions, $\chi^2(3) = 13.0145$, $p = 0.0046$, with a mean rank score of 20.4706 for IS, 35.2059 for INS, 43.7353 for NIS, and 38.5882 for NINS, as shown in Figure 56.

Wilcoxon / Kruskal-Wallis Tests (Rank Sums)

Level	Count	Score Sum	Expected Score	Score Mean	(Mean-Mean0)/Std0
IS	17	348,000	586,500	20,4706	-3,371
INS	17	598,500	586,500	35,2059	0,163
NIS	17	743,500	586,500	43,7353	2,217
NINS	17	656,000	586,500	38,5882	0,977

1-Way Test, ChiSquare Approximation

ChiSquare	DF	Prob>ChiSq
13,0145	3	0,0046*

Figure 56. Kruskall-Wallis H test of number of collisions for Task 1

The pairwise comparison using Dunn's method showed that there was statistically significant difference in number of collisions for the first task. There was a statistically significant difference between NIS (mean rank = 43.7353) and IS (mean rank = 20.4706)

(p = 0.0037), and NINS (mean rank = 38.5882) and IS (mean rank = 20.4706) (p = 0.0465), as shown in Figure 57.

Nonparametric Comparisons For All Pairs Using Dunn Method For Joint Ranking

Level	- Level	Score Mean Difference	Std Err Dif	Z	p-Value
NIS	IS	23,2059	6,782265	3,42155	0,0037*
NINS	IS	18,0588	6,782265	2,66265	0,0465*
INS	IS	14,6765	6,782265	2,16395	0,1828
NIS	INS	8,4706	6,782265	1,24893	1,0000
NINS	INS	3,3235	6,782265	0,49003	1,0000
NINS	NIS	-5,0882	6,782265	-0,75023	1,0000

Figure 57. Pairwise comparisons using Dunn's method of number of collisions for Task 1

(2) Task 2

The means and standard deviations of number of collisions for the second task are shown in Figure 58.

Means and Std Deviations

Level	Number	Mean	Std Dev	Std Err Mean	Lower 95%	Upper 95%
IS	17	236,765	121,532	29,476	174,28	299,25
INS	17	378,706	245,620	59,572	252,42	504,99
NIS	17	454,824	129,183	31,332	388,40	521,24
NINS	17	354,647	175,900	42,662	264,21	445,09

Figure 58. Means and standard deviations of number of collisions for Task 2

The Kruskall-Wallis H test for the second task showed that there was an overall statistically significant difference in number of collisions between the four experimental conditions, $\chi^2(3) = 16.0585$, p = 0.0011, with a mean rank score of 21.0000 for IS, 34.5294 for INS, 48.1765 for NIS, and 34.2941 for NINS, as shown in Figure 59.

Wilcoxon / Kruskal-Wallis Tests (Rank Sums)

Level	Count	Score Sum	Expected Score	Score Mean	(Mean-Mean0)/Std0
IS	17	357,000	586,500	21,0000	-3,243
INS	17	587,000	586,500	34,5294	0,000
NIS	17	819,000	586,500	48,1765	3,286
NINS	17	583,000	586,500	34,2941	-0,042

1-Way Test, ChiSquare Approximation

ChiSquare	DF	Prob>ChiSq
16,0585	3	0,0011*

Figure 59. Kruskall-Wallis H test of number of collisions for Task 2

The pairwise comparison using Dunn's method showed that there was statistically significant difference in number of collisions for the second task. There was a statistically significant difference between NIS (mean rank = 48.1765) and IS (mean rank = 21.0000) (p = 0.0004), as shown in Figure 60.

Nonparametric Comparisons For All Pairs Using Dunn Method For Joint Ranking

Level	- Level	Score Mean Difference	Std Err Dif	Z	p-Value
NIS	IS	27,1176	6,782265	3,99832	0,0004*
NIS	INS	13,5882	6,782265	2,00350	0,2707
INS	IS	13,4706	6,782265	1,98615	0,2821
NINS	IS	13,2353	6,782265	1,95146	0,3060
NINS	INS	-0,1765	6,782265	-0,02602	1,0000
NINS	NIS	-13,8235	6,782265	-2,03819	0,2492

Figure 60. Pairwise comparisons using Dunn's method of number of collisions for Task 2

(3) VE Total

The means and standard deviations for the VE Total number of collisions are shown in Figure 61.

Means and Std Deviations

Level	Number	Mean	Std Dev	Std Err Mean	Lower 95%	Upper 95%
IS	17	512,647	313,738	76,09	351,34	674,0
INS	17	787,235	414,958	100,64	573,88	1000,6
NIS	17	962,824	308,557	74,84	804,18	1121,5
NINS	17	780,235	312,349	75,76	619,64	940,8

Figure 61. Means and standard deviations of the VE Total number of collisions for all experimental conditions

The Kruskall-Wallis H test for the VE Total number of collisions showed that there was an overall statistically significant difference in number of collisions between the four experimental conditions, χ^2 (3) = 15.8991, p = 0.0012, with a mean rank score of 19.8824 for IS, 35.4706 for INS, 46.7059 for NIS, and 35.9412 for NINS, as shown in Figure 62.

Wilcoxon / Kruskal-Wallis Tests (Rank Sums)

Level	Count	Score Sum	Expected Score	Score Mean	(Mean-Mean0)/Std0
IS	17	338,000	586,500	19,8824	-3,512
INS	17	603,000	586,500	35,4706	0,227
NIS	17	794,000	586,500	46,7059	2,932
NINS	17	611,000	586,500	35,9412	0,340

1-Way Test, ChiSquare Approximation

ChiSquare	DF	Prob>ChiSq
15,8991	3	0,0012*

Figure 62. Kruskall-Wallis H test of number of collisions for VE tasks

The pairwise comparison using Dunn's method showed that there was statistically significant difference in the VE Total number of collisions. There was a statistically significant difference between NIS (mean rank = 46.7059) and IS (mean rank = 19.8824) (p = 0.0005), as shown in Figure 63.

Nonparametric Comparisons For All Pairs Using Dunn Method For Joint Ranking

Level	- Level	Score Mean Difference	Std Err Dif	Z	p-Value
NIS	IS	26,7647	6,782330	3,94624	0,0005*
NINS	IS	16,0000	6,782330	2,35907	0,1099
INS	IS	15,5294	6,782330	2,28969	0,1322
NIS	INS	11,1765	6,782330	1,64788	0,5963
NINS	INS	0,4118	6,782330	0,06071	1,0000
NINS	NIS	-10,7059	6,782330	-1,57850	0,6867

Figure 63. Pairwise comparisons using Dunn's method of number of collisions for VE tasks

3. Subjective Data Set

a. Simulator Sickness Questionnaire (SSQ)

Participants filled a total of 272 reports of SSQ (four reports per participant). For the most part, participants reported no symptoms of any kind, and only one participant reported having severe symptom. The most common symptoms were reported with 'slight' level rating: fatigue (23.16%), eye strain (23.16%), difficulty focusing (11.40%), sweating (10.66%), and blurred vision (10.29%). The most common reports of moderate level symptoms were: eye strain (2.21%), general discomfort (1.84%), and sweating (1.10%). Of the baseline SSQs, 26 of 68 participants (38.23%) reported having some slight symptoms, and 2 of 68 (2.94%) reported having moderate symptom. The most common baseline reports of symptoms with slight level were: fatigue (23.53%), eye strain (10.29%), difficulty focusing (8.82%), and sweating (8.82%). The baseline symptoms reported with moderate level were: general discomfort (2.94%), and sweating (1.47%). The results of collected data are shown in Table 6.

Table 6. SSQ overall results

	None		Slight		Moderate		Severe	
Symptom	**Number**	**%**	**Number**	**%**	**Number**	**%**	**Number**	**%**
1. General discomfort	243	89.34	24	8.82	5	1.84	0	0.00
2. Fatigue	209	76.84	63	23.16	0	0.00	0	0.00
3. Headache	261	95.96	11	4.04	0	0.00	0	0.00
4. Eye strain	203	74.63	63	23.16	6	2.21	0	0.00
5. Difficulty focusing	241	88.60	31	11.40	0	0.00	0	0.00
6. Salivation increasing	268	98.53	4	1.47	0	0.00	0	0.00
7. Sweating	240	88.24	29	10.66	3	1.10	0	0.00
8. Nausea	266	97.79	5	1.84	1	0.37	0	0.00
9. Difficulty concentrating	264	97.06	8	2.94	0	0.00	0	0.00
10. "Fullness of the Head"	261	95.96	11	4.04	0	0.00	0	0.00
11. Blurred vision	243	89.34	28	10.29	1	0.37	0	0.00
12. Dizziness with eyes open	265	97.43	7	2.57	0	0.00	0	0.00
13. Dizziness with eyes closed	261	95.96	10	3.68	1	0.37	0	0.00
14. *Vertigo	268	98.53	4	1.47	0	0.00	0	0.00
15. **Stomach awareness	254	93.38	18	6.62	0	0.00	0	0.00
16. Burping	259	95.22	11	4.04	0	0.00	2	0.74

* Vertigo is experienced as loss of orientation with respect to vertical upright.

** Stomach awareness is usually used to indicate a feeling of discomfort which is just short of nausea.

The order of questionnaires filled by the participants was:

- 1st (Baseline): At the very beginning of the entire session (before any exposure to VE)
- 2nd: After training period (length of exposure to immersive VE: 10 min)
- 3rd: After first experimental session (length of exposure to immersive VE: 10 min)
- 4th: After second experimental session (length of exposure to immersive VE: 10 min)

Table 7 shows how the participants reported symptoms for each of the four questionnaires, in numbers, and Table 8 details their percentage. It is noticeable that most symptoms were reported on the third and fourth questionnaires, immediately after the first and second experimental sessions, respectively.

Additionally, during the user study, only one participant reported severe symptom related to simulator sickness for burping. It is important to make a remark that the experimenter did not notice this symptom on the participant. Only eight participants reported symptoms of simulator sickness with moderate level after they started the VE sessions (five reported eye strain, two presented sweating, one presented nausea, one presented blurred vision, and one presented dizziness with eyes closed); two participants reported moderate symptoms before they started the VE sessions. Forty-nine participants reported slight symptoms related to simulator sickness.

The most common simulator sickness symptom reported in the IS condition was eye strain, which affected eight participants (47.05%) after they started the VE sessions (three participants reported moderate symptom, and five reported symptoms at the slights level). For the INS condition, the most common symptoms were general discomfort, eye strain, difficulty focusing, and sweating (two participants reported moderate symptoms and one participant presented severe symptom). For the NIS condition, the most common symptoms were eye strain, difficulty focusing, blurred vision and stomach awareness (no participant reported moderate symptom). And, for the NINS condition, the most common symptoms were general discomfort, eye strain, difficulty focusing, and blurred vision (three participants reported moderate symptoms).

Table 7. SSQ reported symptoms according to number of participants

Symptom	Session	Levels															
		None (number)				Slight (number)				Moderate (number)				Severe (number)			
		IS	INS	NIS	NINS	IS	INS	NIS	NINS	IS	INS	NIS	NINS	IS	INS	NIS	NINS
1. General discomfort	1st (Baseline)	15	16	16	16	0	1	1	1	2	0	0	0	0	0	0	0
	2nd	16	16	16	16	0	1	1	1	1	0	0	0	0	0	0	0
	3rd	16	13	16	14	0	4	1	3	1	0	0	0	0	0	0	0
	4th	16	13	16	12	0	4	1	5	1	0	0	0	0	0	0	0
2. Fatigue	1st (Baseline)	9	12	15	16	8	5	2	1	0	0	0	0	0	0	0	0
	2nd	10	13	15	16	7	4	2	1	0	0	0	0	0	0	0	0
	3rd	12	12	15	14	5	5	2	3	0	0	0	0	0	0	0	0
	4th	11	12	14	13	6	5	3	4	0	0	0	0	0	0	0	0
3. Headache	1st (Baseline)	17	16	17	17	0	1	0	0	0	0	0	0	0	0	0	0
	2nd	17	16	16	17	0	1	1	0	0	0	0	0	0	0	0	0
	3rd	17	16	16	16	0	1	1	1	0	0	0	0	0	0	0	0
	4th	16	15	16	16	1	2	1	1	0	0	0	0	0	0	0	0
4. Eye strain	1st (Baseline)	12	16	17	16	5	1	0	1	0	0	0	0	0	0	0	0
	2nd	12	16	16	17	5	1	1	0	0	0	0	0	0	0	0	0
	3rd	10	12	9	10	6	5	8	6	1	0	0	1	0	0	0	0
	4th	10	11	9	10	5	5	8	6	2	1	0	1	0	0	0	0
5. Difficulty focusing	1st (Baseline)	15	14	17	16	2	3	0	1	0	0	0	0	0	0	0	0
	2nd	16	16	16	16	1	1	1	1	0	0	0	0	0	0	0	0
	3rd	15	14	16	15	2	3	1	2	0	0	0	0	0	0	0	0
	4th	14	13	14	14	3	4	3	3	0	0	0	0	0	0	0	0
6. Salivation increasing	1st (Baseline)	17	17	17	16	0	0	0	1	0	0	0	0	0	0	0	0
	2nd	17	17	17	16	0	0	0	1	0	0	0	0	0	0	0	0
	3rd	17	17	17	17	0	0	0	0	0	0	0	0	0	0	0	0
	4th	17	16	16	17	0	1	1	0	0	0	0	0	0	0	0	0
7. Sweating	1st (Baseline)	15	14	17	15	2	2	0	2	0	1	0	0	0	0	0	0
	2nd	15	16	17	17	2	1	0	0	0	0	0	0	0	0	0	0
	3rd	13	14	16	14	4	3	1	2	0	0	0	1	0	0	0	0
	4th	14	13	16	14	3	3	1	3	0	1	0	0	0	0	0	0
8. Nausea	1st (Baseline)	17	17	17	17	0	0	0	0	0	0	0	0	0	0	0	0
	2nd	17	17	17	17	0	0	0	0	0	0	0	0	0	0	0	0
	3rd	17	15	17	16	0	2	0	1	0	0	0	0	0	0	0	0
	4th	17	15	17	16	0	2	0	0	0	0	0	1	0	0	0	0
9. Difficulty concentrating	1st (Baseline)	16	16	17	16	1	1	0	1	0	0	0	0	0	0	0	0
	2nd	17	16	17	16	0	1	0	1	0	0	0	0	0	0	0	0
	3rd	17	17	17	15	0	0	0	2	0	0	0	0	0	0	0	0
	4th	17	17	17	16	0	0	0	1	0	0	0	0	0	0	0	0
10. "Fullness of head"	1st (Baseline)	16	15	17	17	1	2	0	0	0	0	0	0	0	0	0	0
	2nd	17	17	17	17	0	0	0	0	0	0	0	0	0	0	0	0
	3rd	16	17	16	16	1	0	1	1	0	0	0	0	0	0	0	0
	4th	16	16	16	15	1	1	1	2	0	0	0	0	0	0	0	0
11. Blurred vision	1st (Baseline)	16	16	17	17	1	1	0	0	0	0	0	0	0	0	0	0
	2nd	16	17	16	17	1	0	1	0	0	0	0	0	0	0	0	0
	3rd	14	15	15	15	3	2	2	2	0	0	0	0	0	0	0	0
	4th	11	14	14	13	6	3	3	3	0	0	0	1	0	0	0	0
12. Dizziness with eyes open	1st (Baseline)	17	17	17	17	0	0	0	0	0	0	0	0	0	0	0	0
	2nd	16	17	17	17	1	0	0	0	0	0	0	0	0	0	0	0
	3rd	16	17	17	16	1	0	0	1	0	0	0	0	0	0	0	0
	4th	15	17	16	16	2	0	1	1	0	0	0	0	0	0	0	0
13. Dizziness with eyes closed	1st (Baseline)	17	17	17	17	0	0	0	0	0	0	0	0	0	0	0	0
	2nd	15	17	16	17	2	0	1	0	0	0	0	0	0	0	0	0
	3rd	15	17	16	17	2	0	1	0	0	0	0	0	0	0	0	0
	4th	15	16	16	16	1	1	1	1	1	0	0	0	0	0	0	0
14. *Vertigo	1st (Baseline)	17	17	17	17	0	0	0	0	0	0	0	0	0	0	0	0
	2nd	17	17	17	17	0	0	0	0	0	0	0	0	0	0	0	0
	3rd	16	17	17	17	1	0	0	0	0	0	0	0	0	0	0	0
	4th	15	16	17	17	2	1	0	0	0	0	0	0	0	0	0	0
15. **Stomach awareness	1st (Baseline)	15	16	17	17	2	1	0	0	0	0	0	0	0	0	0	0
	2nd	16	16	17	17	1	1	0	0	0	0	0	0	0	0	0	0
	3rd	15	14	16	17	2	3	1	0	0	0	0	0	0	0	0	0
	4th	16	14	14	17	1	3	3	0	0	0	0	0	0	0	0	0
16. Burping	1st (Baseline)	17	15	17	17	0	2	0	0	0	0	0	0	0	0	0	0
	2nd	16	14	17	17	1	2	0	0	0	0	0	0	0	1	0	0
	3rd	17	15	15	17	0	1	2	0	0	0	0	0	0	1	0	0
	4th	17	16	16	16	0	1	1	1	0	0	0	0	0	0	0	0

Table 8. SSQ reported symptoms according to percentage of participants

Symptom	Session	Levels															
		None (%)				Slight (%)				Moderate (%)				Severe (%)			
		IS	INS	NIS	NINS	IS	INS	NIS	NINS	IS	INS	NIS	NINS	IS	INS	NIS	NINS
1. General discomfort	1st (Baseline)	88.24	94.12	94.12	94.12	0.00	5.88	5.88	5.88	11.76	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	2nd	94.12	94.12	94.12	94.12	0.00	5.88	5.88	5.88	5.88	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	3rd	94.12	76.47	94.12	82.35	0.00	23.53	5.88	17.65	5.88	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	4th	94.12	76.47	94.12	70.59	0.00	23.53	5.88	29.41	5.88	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2. Fatigue	1st (Baseline)	52.94	70.59	88.24	94.12	47.06	29.41	11.76	5.88	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	2nd	58.82	76.47	88.24	94.12	41.18	23.53	11.76	5.88	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	3rd	70.59	70.59	88.24	82.35	29.41	29.41	11.76	17.65	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	4th	64.71	70.59	82.35	76.47	35.29	29.41	17.65	23.53	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
3. Headache	1st (Baseline)	100.00	94.12	100.00	100.00	0.00	5.88	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	2nd	100.00	94.12	94.12	100.00	0.00	5.88	5.88	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	3rd	100.00	94.12	94.12	94.12	0.00	5.88	5.88	5.88	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	4th	94.12	88.24	94.12	94.12	5.88	11.76	5.88	5.88	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
4. Eye strain	1st (Baseline)	70.59	94.12	100.00	94.12	29.41	5.88	0.00	5.88	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	2nd	70.59	94.12	94.12	100.00	29.41	5.88	5.88	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	3rd	58.82	70.59	52.94	58.82	35.29	29.41	47.06	35.29	5.88	0.00	0.00	5.88	0.00	0.00	0.00	0.00
	4th	58.82	64.71	52.94	58.82	29.41	29.41	47.06	35.29	11.76	5.88	0.00	5.88	0.00	0.00	0.00	0.00
5. Difficulty focusing	1st (Baseline)	88.24	82.35	100.00	94.12	11.76	17.65	0.00	5.88	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	2nd	94.12	94.12	94.12	94.12	5.88	5.88	5.88	5.88	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	3rd	88.24	82.35	94.12	88.24	11.76	17.65	5.88	11.76	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	4th	82.35	76.47	82.35	82.35	17.65	23.53	17.65	17.65	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
6. Salivation increasing	1st (Baseline)	100.00	100.00	100.00	94.12	0.00	0.00	0.00	5.88	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	2nd	100.00	100.00	100.00	94.12	0.00	0.00	0.00	5.88	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	3rd	100.00	100.00	100.00	100.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	4th	100.00	94.12	94.12	100.00	0.00	5.88	5.88	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
7. Sweating	1st (Baseline)	88.24	82.35	100.00	88.24	11.76	11.76	0.00	11.76	0.00	5.88	0.00	0.00	0.00	0.00	0.00	0.00
	2nd	88.24	94.12	100.00	100.00	11.76	5.88	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	3rd	76.47	82.35	94.12	82.35	23.53	17.65	5.88	11.76	0.00	0.00	0.00	5.88	0.00	0.00	0.00	0.00
	4th	82.35	76.47	94.12	82.35	17.65	17.65	5.88	17.65	0.00	5.88	0.00	0.00	0.00	0.00	0.00	0.00
8. Nausea	1st (Baseline)	100.00	100.00	100.00	100.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	2nd	100.00	100.00	100.00	100.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	3rd	100.00	88.24	100.00	94.12	0.00	11.76	0.00	5.88	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	4th	100.00	88.24	100.00	94.12	0.00	11.76	0.00	0.00	0.00	0.00	0.00	5.88	0.00	0.00	0.00	0.00
9. Difficulty concentrating	1st (Baseline)	94.12	94.12	100.00	94.12	5.88	5.88	0.00	5.88	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	2nd	100.00	94.12	100.00	94.12	0.00	5.88	0.00	5.88	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	3rd	100.00	100.00	100.00	88.24	0.00	0.00	0.00	11.76	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	4th	100.00	100.00	100.00	94.12	0.00	0.00	0.00	5.88	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
10. "Fullness of head"	1st (Baseline)	94.12	88.24	100.00	100.00	5.88	11.76	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	2nd	100.00	100.00	100.00	100.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	3rd	94.12	100.00	94.12	94.12	5.88	0.00	5.88	5.88	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	4th	94.12	94.12	94.12	88.24	5.88	5.88	5.88	11.76	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
11. Blurred vision	1st (Baseline)	94.12	94.12	100.00	100.00	5.88	5.88	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	2nd	94.12	100.00	94.12	100.00	5.88	0.00	5.88	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	3rd	82.35	88.24	88.24	88.24	17.65	11.76	11.76	11.76	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	4th	64.71	82.35	82.35	76.47	35.29	17.65	17.65	17.65	0.00	0.00	0.00	5.88	0.00	0.00	0.00	0.00
12. Dizziness with eyes open	1st (Baseline)	100.00	100.00	100.00	100.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	2nd	94.12	100.00	100.00	100.00	5.88	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	3rd	94.12	100.00	100.00	94.12	5.88	0.00	0.00	5.88	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	4th	88.24	100.00	94.12	94.12	11.76	0.00	5.88	5.88	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
13. Dizziness with eyes closed	1st (Baseline)	100.00	100.00	100.00	100.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	2nd	88.24	100.00	94.12	100.00	11.76	0.00	5.88	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	3rd	88.24	100.00	94.12	100.00	11.76	0.00	5.88	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	4th	88.24	94.12	94.12	94.12	5.88	5.88	5.88	5.88	5.88	0.00	0.00	0.00	0.00	0.00	0.00	0.00
14. *Vertigo	1st (Baseline)	100.00	100.00	100.00	100.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	2nd	100.00	100.00	100.00	100.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	3rd	94.12	100.00	100.00	100.00	5.88	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	4th	88.24	94.12	100.00	100.00	11.76	5.88	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
15. **Stomach awareness	1st (Baseline)	88.24	94.12	100.00	100.00	11.76	5.88	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	2nd	94.12	94.12	100.00	100.00	5.88	5.88	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	3rd	88.24	82.35	94.12	100.00	11.76	17.65	5.88	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	4th	94.12	82.35	82.35	100.00	5.88	17.65	17.65	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
16. Burping	1st (Baseline)	100.00	88.24	100.00	100.00	0.00	11.76	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	2nd	94.12	82.35	100.00	100.00	5.88	11.76	0.00	0.00	0.00	0.00	0.00	0.00	0.00	5.88	0.00	0.00
	3rd	100.00	88.24	88.24	100.00	0.00	5.88	11.76	0.00	0.00	0.00	0.00	0.00	0.00	5.88	0.00	0.00
	4th	100.00	94.12	94.12	94.12	0.00	5.88	5.88	5.88	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

b. Post-task Questionnaires

Participants filled post-task questionnaires about their experience in each of the tasks using a 7-point scale varying from “1 – It was very difficult” to “7 – It was very easy”. Table 9 shows the overall results for the self-reported data on post-task questionnaires of Task 1 and Task 2, and highest ratings are highlighted in red.

The results show that participants in IS condition reported the most favorable responses for both tasks, which were perceived as the easiest, level of realism as the highest, and difficulty while interacting as the lowest. Also, the sense of being in the room with parts on the table, and of being able to imagine themselves as interacting with the parts, were qualified as the highest in IS condition. In case of Task 1, closest second was INS condition, and in Task 2, it was NINS condition.

Table 9. Self-reported data on post-task questionnaires of Task 1 and Task 2

			IS		INS		NIS		NINS	
Task	**Question**	**Description**	**Average**	**SD**	**Average**	**SD**	**Average**	**SD**	**Average**	**SD**
1	**3**	Selecting	5.12	1.45	**5.35**	1.75	4.94	1.39	4.24	1.77
		Manipulating	**5.41**	1.54	5.29	1.49	4.88	1.45	4.82	1.69
		Assembling	**4.65**	1.57	4.18	1.58	3.00	0.91	3.76	1.63
	4	Overall experience	**4.94**	1.59	4.65	1.57	3.71	0.96	3.94	1.76
	5	Performance rating	**5.65**	1.57	5.47	1.42	3.88	1.53	4.71	1.67
	6	Presence	**5.82**	1.04	5.47	1.58	4.59	1.37	4.47	1.54
	7	Post-task interaction	**6.59**	0.60	6.29	0.82	5.82	1.10	5.53	1.42
	8	Visual representation	**6.47**	0.78	6.18	0.98	5.88	1.28	6.24	0.81
		Interaction (one part)	**5.94**	0.94	5.82	0.98	5.24	1.44	5.06	1.66
		Interaction (between parts)	**5.82**	1.25	5.35	1.13	4.71	1.74	4.18	1.69
		Hands and arms movements	**6.53**	0.70	6.41	0.84	6.06	1.26	6.12	1.18
		Head movement	**6.88**	0.32	6.71	0.75	6.12	1.45	6.00	1.14
	9	Task rating	**5.53**	1.42	4.59	1.42	4.00	1.37	3.94	1.73
	10	Distinguish depth	**6.12**	0.76	4.76	1.73	3.94	1.66	4.59	1.65
2	**3**	Selecting	**5.29**	1.74	5.06	1.55	4.76	1.48	4.94	1.35
		Manipulating	**5.41**	1.54	4.94	1.47	5.18	1.29	**5.41**	1.50
		Assembling	4.47	1.50	3.47	1.72	3.71	1.46	**4.71**	1.64
	4	Overall experience	**4.47**	1.79	3.65	1.57	3.76	1.48	**4.47**	1.68
	5	Performance rating	5.35	1.68	4.47	2.15	4.41	1.48	**5.76**	0.94
	6	Presence	**5.94**	1.06	5.12	1.78	4.88	1.25	5.00	1.53
	7	Post-task interaction	**6.35**	0.76	5.76	1.55	5.94	0.70	5.88	0.96
	8	Visual representation	6.24	1.06	5.88	1.13	6.06	1.58	**6.35**	0.68
		Interaction (one part)	**5.88**	1.37	5.41	1.42	5.47	1.25	5.06	1.55
		Interaction (between parts)	**5.59**	1.33	4.71	1.52	5.18	1.05	4.88	1.37
		Hands and arms movements	**6.29**	1.02	6.35	1.19	6.00	0.70	5.88	1.23
		Head movement	**6.71**	0.57	6.41	1.19	6.35	0.49	6.06	1.06
	9	Task rating	4.59	1.78	3.41	1.57	3.59	1.28	**4.65**	1.71
	10	Distinguish depth	**5.82**	1.29	4.59	1.78	4.41	0.99	5.00	1.81
	11	Task 2 compared to Task 1	2.88	1.18	2.47	0.98	3.47	0.90	4.59	1.54

One of the questions was how would they rate the assembly task. Boxplots with participants' ratings for Task 1 and Task 2, and average for VE Total, are shown in Figure 64. We performed a non-parametric statistical analysis for tasks ratings because the

collected data was not normally distributed as shown by a Shapiro-Wilk test in Appendix L.

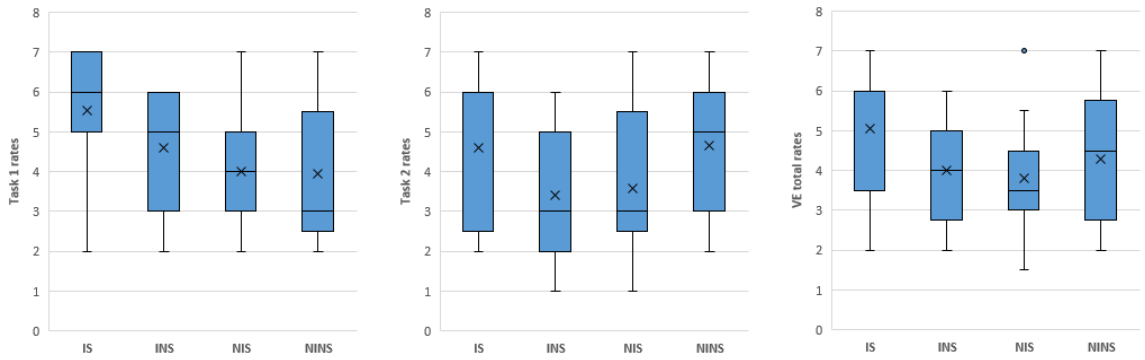


Figure 64. Boxplot containing participants' ratings for Task 1 and Task 2, and VE Total

We performed a Kruskall-Wallis H test to check if there was an overall statistically significant difference in task ratings between the four experimental conditions for each of the two tasks, and for VE Total. Additionally, we performed pairwise comparisons using Dunn's method for joint ranking with Bonferroni adjustment.

(1) Task 1

The means and standard deviations of participant's ratings for the first task are shown in Figure 65.

Means and Std Deviations

Level	Number	Mean	Std Dev	Std Err Mean	Lower 95%	Upper 95%
IS	17	5,52941	1,46277	0,35477	4,7773	6,2815
INS	17	4,58824	1,46026	0,35416	3,8374	5,3390
NIS	17	4,00000	1,41421	0,34300	3,2729	4,7271
NINS	17	3,94118	1,78433	0,43276	3,0238	4,8586

Figure 65. Means and standard deviations of participants' ratings for Task 1

The Kruskall-Wallis H test for the first task showed that there was an overall statistically significant difference in participants' ratings between the four experimental conditions, $\chi^2(3) = 10.5903$, $p = 0.0142$, with a mean rank score of 46.6765 for IS, 35.4412 for INS, 27.9118 for NIS, and 27.9706 for NINS, as shown in Figure 66.

Wilcoxon / Kruskal-Wallis Tests (Rank Sums)

Level	Count	Score Sum	Expected Score	Score Mean	(Mean-Mean0)/Std0
IS	17	793,500	586,500	46,6765	2,976
INS	17	602,500	586,500	35,4412	0,223
NIS	17	474,500	586,500	27,9118	-1,607
NINS	17	475,500	586,500	27,9706	-1,593

1-Way Test, ChiSquare Approximation

ChiSquare	DF	Prob>ChiSq
10,5903	3	0,0142*

Figure 66. Kruskall-Wallis H test for the first task rating

The pairwise comparison using Dunn's method showed that there was statistically significant difference in participants' ratings for the first task. There was a statistically significant difference between NINS (mean rank = 27.9706) and IS (mean rank = 46.6765) ($p = 0.0309$), and NIS (mean rank = 27.9118) and IS (mean rank = 46.6765) ($p = 0.0300$), as shown in Figure 67.

Nonparametric Comparisons For All Pairs Using Dunn Method For Joint Ranking

Level	- Level	Score Mean Difference	Std Err Dif	Z	p-Value
NINS	NIS	0,0000	6,664545	0,00000	1,0000
NINS	INS	-7,4118	6,664545	-1,11212	1,0000
NIS	INS	-7,4706	6,664545	-1,12095	1,0000
INS	IS	-11,1765	6,664545	-1,67700	0,5612
NINS	IS	-18,6471	6,664545	-2,79795	0,0309*
NIS	IS	-18,7059	6,664545	-2,80678	0,0300*

Figure 67. Pairwise comparisons using Dunn's method for the first task ratings

(2) Task 2

The means and standard deviations of participant's ratings for the second task are shown in Figure 68.

Means and Std Deviations

Level	Number	Mean	Std Dev	Std Err Mean	Lower 95%	Upper 95%
IS	17	4,58824	1,83912	0,44605	3,6426	5,5338
INS	17	3,41176	1,62245	0,39350	2,5776	4,2460
NIS	17	3,58824	1,80481	0,43773	2,6603	4,5162
NINS	17	4,64706	1,76569	0,42824	3,7392	5,5549

Figure 68. Means and standard deviations of participants' ratings for Task 2

The Kruskall-Wallis H test for the second task showed that there was no overall statistically significant difference in participants' ratings between the four experimental conditions, $\chi^2(3) = 6.2576$, $p = 0.0997$, with a mean rank score of 39.7941 for IS, 27.4118 for INS, 30.0294 for NIS, and 40.7647 for NINS, as shown in Figure 69.

Wilcoxon / Kruskal-Wallis Tests (Rank Sums)

Level	Count	Score Sum	Expected Score	Score Mean	(Mean-Mean0)/Std0
IS	17	676,500	586,500	39,7941	1,297
INS	17	466,000	586,500	27,4118	-1,739
NIS	17	510,500	586,500	30,0294	-1,094
NINS	17	693,000	586,500	40,7647	1,536

1-Way Test, ChiSquare Approximation

ChiSquare	DF	Prob>ChiSq
6,2576	3	0,0997

Figure 69. Kruskall-Wallis H test for the second task rating

(3) VE Total

The means and standard deviations of participant's ratings for VE Total (average rating of Task 1 and Task 2) are shown in Figure 70.

Means and Std Deviations

Level	Number	Mean	Std Dev	Std Err Mean	Lower 95%	Upper 95%
IS	9	5,33333	1,65831	0,55277	4,0586	6,6080
INS	11	4,27273	1,42063	0,42834	3,3183	5,2271
NIS	8	4,12500	1,35620	0,47949	2,9912	5,2588
NINS	9	4,44444	1,74005	0,58002	3,1069	5,7820

Figure 70. Means and standard deviations of participants' ratings for VE Total

The Kruskall-Wallis H test for the VE Total showed that there was no overall statistically significant difference in participants' ratings between the four experimental conditions, $\chi^2(3) = 3.8874$, $p = 0.2739$, with a mean rank score of 24.8889 for IS, 16.9545 for INS, 15.9375 for NIS, and 18.3333 for NINS, as shown in Figure 71.

Wilcoxon / Kruskal-Wallis Tests (Rank Sums)

Level	Count	Score Sum	Expected Score	Score Mean	(Mean-Mean0)/Std0
IS	9	224,000	171,000	24,8889	1,897
INS	11	186,500	209,000	16,9545	-0,746
NIS	8	127,500	152,000	15,9375	-0,904
NINS	9	165,000	171,000	18,3333	-0,199

1-Way Test, ChiSquare Approximation

ChiSquare	DF	Prob>ChiSq
3,8874	3	0,2739

Figure 71. Kruskall-Wallis H test for the VE Total rating

c. System Usability Scale (SUS) Questionnaires

We slightly adjusted SUS scale from a 5-point scale to a 7-point scale for uniformity purposes. We used the majority of SUS questions as a useful instrument rather than using it in its original form because some of the original questions were not applicable to our system, what made us disregard them in our data analysis.

Participants reported different opinions about the system usability, using a 7-point scale grading, according to which experimental condition they were in. One factor that should be highlighted is the number of participants that responded "Strongly agree" to "I

though the system was easy to use": 9 of 17 (52.94%) in the IS condition, 5 of 17 (29.41%) in the INS condition, and only 3 of 17 (17.65%) for both NIS and NINS conditions. Another interesting observation is the number of participants that answered "Strongly agree" to "I felt very confident using the system": 9 of 17 (52.94%) in the IS condition, 4 of 17 (23.53%) in the INS condition, 5 of 17 (29.41%) for the NIS condition, and 7 of 17 (41.18%) in the NINS condition.

d. Behavioral

We recorded and monitored all sixty-eight experiment sessions for physical behavioral cues. The camera captured participants as they were seating while manipulating objects. The main observations are:

(1) Use of hands: The majority of participants used only one hand at a time to manipulate the objects.

(2) Movement of body and head: Participants in immersive conditions presented more movement of their body and head—leaning towards or to the sides—than participants in non-immersive conditions.

(3) Disorientation: No participant presented signs of disorientation, fell off the chair, or injured themselves during the sessions.

B. DISCUSSION

1. Task 1

Analysis of assembly times for the first task revealed a statistically significant difference between NIS and IS, NINS and IS, and NIS and INS conditions. Additionally, analysis of number of correct connections revealed a statistically significant difference between NIS and INS, and NIS and IS conditions. Also, analysis of number of collisions revealed a statistically significant difference between NINS and IS, and NIS and IS conditions. Furthermore, IS condition presented the largest numbers of successful assemblies, the highest ranking self-reported data, and the smallest number of collisions, indicating a better performance of the participants in this condition.

2. Task 2

Analysis of assembly times for the second task revealed a statistically significant difference between NIS and IS conditions. Additionally, analysis of number of collisions revealed a statistically significant difference between NIS and IS conditions. Also, there was no statistically significant difference between the conditions for number of correct connections. Although IS and NINS conditions presented the largest numbers of successful assemblies and the highest ranking self-reported data, IS condition presented the smallest number of collisions, suggesting a better performance of the participants in this condition.

3. VE Total

Analysis of VE Total (Task 1 and Task 2) assembly times revealed a statistically significant difference between NIS and IS conditions. Additionally, analysis of number of collisions revealed a statistically significant difference between NIS and IS conditions. We found no statistically significant differences in VE Total number of correct connections in VE Total ratings. Additionally, IS condition presented the largest numbers of successful assemblies and the smallest number of collisions, suggesting a slight better performance of the participants in this condition.

C. SUMMARY

This chapter presented the results of user study, detailing system performance and user performances, and analyzed and discussed about the objective and subjective data sets.

THIS PAGE INTENTIONALLY LEFT BLANK

VIII. CONCLUSIONS AND FUTURE WORK

This chapter highlights the main contributions of this study. It also discusses future work in the virtual reality and maintenance domains.

A. MAIN CONTRIBUTIONS

In this thesis we have presented a study that compares operator's performance on assembly task using virtual reality. The user study consisted of four distinct experimental conditions in which two factors were analyzed: stereoscopic depth cue and immersion.

Following a between-subjects experimental design, participants executed three assembly tasks: two in virtual environment and one in physical environment. We analyzed performance in terms of assembly times, number of correct connections, and number of successful assemblies. Additionally, we used subjective measurements—task ratings and system usability—to understand participants' perception and evaluation on the type of device used to perform the tasks. Our results show that the performance on the immersive stereoscopic condition was promising.

We found a statistically significant difference between the experimental conditions in assembly times (NIS and IS, NINS and IS, and NINS and INS) and number of correct connections (NIS and INS, and NIS and IS) in the first task. Additionally, immersive stereoscopic condition presented the largest number of successful assemblies and the highest participants' rating. When we analyzed data in the second task, we found a statistically significant difference between the experimental conditions in assembly times (NIS and IS). We found no statistically significant difference in number of correct connections. Additionally, immersive stereoscopic and non-immersive non-stereoscopic conditions presented the largest number of successful assemblies and the highest participants' ratings. Our results on virtual environment (Tasks 1 and Task 2) show a statistically significant difference the experimental conditions in assembly times (NIS and IS). We found no statistically significant difference in number of correct connections and participants' ratings. Additionally, immersive stereoscopic condition presented the largest number of successful assemblies.

Overall, the results suggest that immersive stereoscopic condition was superior when compared to other experimental conditions. The main reasons are: most people in that condition finished their assembly tasks, they did it in a shortest time, they correctly connected the largest number of parts, and they rated better this condition. Also, the results of this study provide an important input and guidance that people who work in training domain need to have before making their decision about acquisition of new solutions for training of assembly tasks.

Additionally, the work of this study contributes to several other domains that involve master navigation and manipulation of objects positioned within arms' reach, in which it is necessary to judge distance, depth, sizing, and fit. Besides maintenance personnel, a proper virtual training can benefit doctors, dentists, sculptors, and craftsmen, among others.

B. FUTURE WORK

As this study focused on the analysis of two specific visual factors (stereoscopic depth cue and immersion), the most obvious future work would be the inclusion of other factors that can affect operator's performance in assembly tasks using VR. Those factors can be visual (such as shades), auditory (such as sounds), and tactile (such as haptics feedback).

Also, it would be interesting to analyze operator's performance in more difficult levels. It can be done either using more complex objects, more parts, more difficult connections, or creating additional steps, such as disassembling.

Another critical extension of this work would be the creation of additional sessions with physical parts and measure human performance very closely, which would provide the researcher with ability of investigating a transfer of training. That way, besides the results for each of the sessions, it would provide guidance on the effectiveness of virtual training.

C. SUMMARY

This chapter detailed the main contributions of this thesis work and suggested future avenues of development and work in the virtual reality and maintenance domains.

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX A. IRB APPLICATION FORM

NPS Institutional Review Board
Initial Review Application
(new research protocol)

Purpose:
The Initial Review Application is used to submit new research projects involving the use of human subjects to the IRB for review and approval by the NPS President.

Form Instructions:
Only complete packages are reviewed by the HRPP Specialist. To ensure your application package is complete, refer to the initial review guidance document at the end of this application. Please note all IRB application packages must include a copy of each investigator's completion of CITI ethics training. Submit packages to IRB@nps.edu for review. An IRB administrator will contact you if additional information is required. For questions regarding this form or process, call the HRPP Specialist at 831-656-2998 or send an email to IRB@nps.edu.

MAC users, please use Adobe Reader for Macintosh. Do not use Apple Preview. Free Adobe reader can be found here.

Form Updated 3/24/17 (IRB Office Only) Protocol Number: _____

A. Protocol Basics

1. Title of the research. **User Study on The Role of Stereoscopic Depth Cue and Immersion in Maintenance Task**

2. Researchers. List all researchers. Attach CITI Training Completion Reports for each investigator. The IRB does not keep copies of CITI completion reports on file.

Principal Investigator			
Name	Title or Rank	Dept. or Outside Org. Name	Investigator Roles and Responsibilities in the Research
Amela Sadagic	Dr.	MOVES	Principal Investigator
Co-Investigators and Student Investigators			
Name	Title or Rank	Dept. or Outside Org. Name	Investigator Roles and Responsibilities in the Research
Douglas Yamashita de Moura	Capt	MOVES	Experimentor. Design and execute study. Analyze data.

3. Estimated completion date of the research. If student research, list student graduation date. September 2018

4. Is this research part of a sponsored project (e.g., reimbursable, RIP, NRP)?

☒ No

☐ Yes. List the job order number (JON):

5. Are you requesting an Exempt IRB review? If you believe the research meets an exempt review category please check "Yes" and provide the category number. A description of the categories is provided in the IRB Guidance Document found at the end of this application. If requesting an exempt review the following documents are not required:

- Scientific Review Form
- Recruitment Script
- Consent Form
- Consent Waiver Forms

If determined "not exempt" by the IRB, the above documents will be required.

☒ No.

☐ Yes. Exempt review category:

B. Research Summary

6. Summarize the objective(s) of the research including purpose, research question, hypothesis & background information, literature, etc.

The purpose of this study is to evaluate if performance in assembly tasks can be enhanced using different types of depth cues that are enabled in immersive / non-immersive and stereoscopic / non-stereoscopic VR devices.

The scope of this research is to design and execute an experiment that requires the subjects to perform assembly task and compare their performance in several different study setups (conditions). The study conditions will be: (1) Immersive Stereoscopic -IS, (2) Non-immersive Stereoscopic - NIS, (3) Immersive Non-stereoscopic - INS, and (4) Non-immersive Non-stereoscopic - NINS.

The objective of this research is to ascertain the effectiveness of stereoscopic depth cues and immersion in maintenance related tasks; assembly task is used as the main component of maintenance task that requires human spatial manipulation of each assembly part. The two main research questions are:

(1) Can stereoscopic depth cues improve performance in assembly tasks?
(2) Can immersion improve performance in assembly tasks?.

The null hypotheses are that there are no difference in performance either using stereoscopic depth cues or immersion in assembly tasks.

7. Describe the research study design.

Methodology:

1. Conduct literature review and task analysis.
2. Design and develop application
3. Design and execute user study: participants will complete two different levels of assembly tasks in virtual environment. Participants will be randomly assigned to one of the 4 study conditions: (1) Immersive Stereoscopic - IS, (2) Immersive Non-stereoscopic - INS, (3) Non-immersive Stereoscopic - NIS, or (4) Non-immersive Non-stereoscopic - NINS. A Simulation Sickness Questionnaire (SSQ) will be completed before, in-between and after the different levels. Additionally, participants will complete one assembly task in physical environment using 3D-printed parts of the same object of the second task. Also, after each of the three assembly tasks, participants will complete post-task questionnaires. The tasks consist of assembling objects made of multiple parts. The objects are very similar to wooden toddlers toys.
4. Analyze collected data set.
5. Make conclusions and recommendation for future work.

8. Describe in detail the tasks subjects will be asked to perform and the amount of time it will take to complete each task.

1. Participants will arrive to the research location, receive a short overview brief and given the opportunity to provide informed consent (5 minutes)
2. Participants will complete the baseline Simulator Sickness Questionnaire (SSQ) (3 minutes)
3. Participants will be familiarize with VR devices and interface and will do a training session (10 minutes)
4. Participants will complete the Simulator Sickness Questionnaire (SSQ) (2 minutes)
5. Participants will do task #1 in virtual environment (5 minutes)
6. Participants will complete the Simulator Sickness Questionnaire (SSQ) (2 minutes)
7. Participants will complete the post-task #1 questionnaire (3 minutes)

8. Participants will do task #2 in virtual environment (10 minutes)
9. Participants will complete a Simulator Sickness Questionnaire (SSQ) (2 minutes)
10. Participants will complete the post-task #2 questionnaire (3 minutes)
11. Participants will do task #3 in physical environment (2 minutes)
12. Participants will complete the post-task #3 questionnaire (3 minutes)
13. Participants will complete a demographic survey and System Usability Scale (SUS) questionnaire (5 minutes)
14. Participants will receive a short debrief/explanation of the study and be permitted to ask questions (5 minutes)

Estimated time requested of subjects is approximately 60 minutes.

9. Where will the research be performed?

Research will conducted on the NPS campus. Participants will conduct their part within a lab space located in Watkins 212A.

10. Are research subjects or research data located OCONUS?

☒ No

☐ Yes. Attach host country approval and ethics review.

11. Does the research involve the use of existing records?

Example of existing records are AARs, personnel records, medical records, databases, etc.

☒ No. Skip to question 12.

☐ Yes. Describe the records below. Include the data variables and the number of records to which you will have access.

11b. Are the records private (not available to the general public)?

☐ No

☐ Yes, attach proof of approval from the organization that owns the data stating you may access the data for your research.

11c. For what purpose will these records be used?

☐ To collect data that will be analyzed in the research.

☐ To identify potential subjects.

☐ Other. Describe below.

12. The following areas of research require approval outside of NPS. Check all that apply.

☐ Classified research

☐ Severe or unusual intrusions, either physical or physiological

☐ Potential or inherent controversial topics (those likely to attract media coverage or challenge by interest groups)

☒ Research involving Marine Corps population (requires USMC IRB administrative review and possibly USMC Survey Manager review)

C. Subject Population & Recruitment

13. Subject Populations. Check all that apply.

☐ Military Personnel (outside NPS)	☐ Government Contractors	☐ Foreign Nationals outside the U.S.
☐ DoD Personnel (outside NPS)	☐ Children, *under 18 years old*	☐ Pregnant Women or Fetuses
☒ NPS Students	☐ Elderly, *over 70 years old*	☐ Non-English Speakers
☒ NPS Civilian Employees	☒ General Public	

14. Describe subject inclusion and exclusion criteria.

All individuals at NPS to include members of the general public (DON contractors).

15. Provide the sample size (ex: 75) or range (ex: 75-100) and the rationale for why that number is chosen.

A sample size of 60 participants will allow the research team to meet the desired scope of the user study.
A calculation was done with values for Experimental Conditions of 4, Power of 0.80, Significance Level of 0.05, and Effect Size of 0.3, resulting in about 8 participants per condition (total of 32). The optimal number of 15 participants per condition (total of 60) is desired in order to minimize the Effect Size.

16. Will compensation be given to research subjects? Compensation may be monetary, raffles, meals/snacks, extra credit, etc. Reference DoDI 3216.02 for guidance on compensating research subjects.

☒ No

☐ Yes. Describe what the compensation consists of and the purpose for offering it.

17. Describe how potential subjects will be recruited to participate in the research.

Recruitment will be done via fliers posted about NPS and distributed via no more than three bulk emails via Ed Techs, or in person by experimenter.

18. How will you minimize coercion and undue influence during the recruitment process?

Participants will voluntarily contact experimenter to schedule a time of the potential participant's choosing.
Experimenter, who does not have any command, academic, or employment influence over participants will explain the voluntary nature of the experiment in person and via email during recruitment process. The experimenter will also explain the voluntary nature of the study to participants during the informed consent process before any tasks are begun.

D. Risk & Benefits

19. Does the research involve any of these possible risks or discomforts to subjects? Check all that apply.

- ☐ Use of deception
- ☐ Physiological risk
- ☐ Employment risk
- ☐ Legal risk
- ☐ Social or economical risk
- ☐ Physical risk
- ☐ Probing for personal or sensitive info
- ☐ Possible invasions of privacy of subjects or family
- ☐ Presentation of materials that might be considered sensitive, offensive, threatening or degrading
- ☐ Manipulation of physiological or social variables such as sensory deprivation, social isolation, psychological stresses

20. Describe any foreseeable risks or discomforts associated with the research.

Symptoms of cybersickness can occur when subjects use immersive virtual environment (symptoms that are similar to motion sickness symptoms.) The most common symptoms are general discomfort, headache, stomach awareness, nausea, vomiting, pallor, sweating, fatigue, drowsiness, and disorientation. Additionally, breach of confidentiality is a risk.

21. Explain what steps will be taken to minimize risks and discomforts (mentioned in Q19-20) and to protect subjects' welfare.

Steps to prevent virtual reality sickness will be to limit exposure times to the system All users be advised to take breaks if needed.
Experimenter will ensure participants are screened prior to their participation in the study.
Additionally, the subject will be continuously monitored for external signs of any discomfort. Signs of discomfort include nausea, dizziness, or verbal complaints of cybersickness by participants. If any symptoms are recognized we will halt the study.

22. Provide a description of the potential benefits of this research for individuals, subjects, society, military or DoD/DoN. Explain how risks are reasonable in relation to anticipated benefits.

We looked to explore how stereoscopic depth cues and immersion affect the user's performance when executing maintenance related tasks, most specifically the assembly task. This study research will provide us guidance and indication of the type of VR display solution that assures best results for assembly task.

E. Data Security & Monitoring

23. Will you record identifiers such as name, social security number, DoD ID #, address, telephone number or any combination of demographic data that could lead to the identification of a participant?

☐ No

☒ Yes. Explain below why it is necessary to collect these identifiers, state if you will use a coding system to protect against disclosure of identifiers and state when PII will be destroyed.

Participant name, email, and phone number is collected to schedule participation but is not recorded with participant data. Video recordings described in Q24 will involve collection of participant faces and possibly names if in uniform.

24. Will you audio or video record subjects?

☐ No

☒ Yes. Describe what will be recorded, why you are recording, if the recording will be transcribed and how you will safeguard the recording.

Participants will be recorded in order to evaluate their physical movements in the real world relative to their movements within the virtual environment. Videos will also demonstrate the degree of which the participants are able to familiarize themselves with virtual reality devices. Video recording will be done using a NPS camcorder and, at the end of each day, videos will be copied to NPS secure servers and deleted from the NPS camcorder. Recordings will kept on approved NPS secure servers at collection and at completion.

25. How will data and consent forms be kept confidential during collection, analysis, and long term storage after completion of the research? Please note electronic PII may only be stored on the NPS network.

Data will only be kept on approved NPS secure servers. Hard copies of informed consent forms and other survey data will be kept separately in a locked cabinet behind a locked door in PIs office. Hard copies of data will be associated only with the subject ID. A copy of the consent forms and de-identified data will be retained by the PI at completion.

26. When appropriate, a research plan is required to make adequate provisions for monitoring the data to ensure safety of subjects. Will you monitor data collection?

☐ No

☒ Yes. Describe the monitoring procedure below.

Participants will be monitored by the primary investigator while performing their roles within the application system. Before and after surveys will be conducted and additionally, thorough explanation, training, and familiarization will be provided before allowing participants into the virtual environment.

F. Consent Procedure (If requesting exempt review skip to Q29)

27(a). DoD regulations require that you obtain consent from subjects prior to data collection unless a waiver is approved by the IRB. Are you requesting a waiver of consent? *A waiver of consent is required if you do not intend to have subjects read and sign a consent form.*

☒ No

☐ Yes. *Complete Appendix A and skip to question 28.*

27(b). DoD regulations require subjects to sign the consent form unless a waiver is approved by the IRB. Are you requesting a waiver of signed consent? *A waiver of documented consent is required if you plan to provide a consent form to subjects, subjects will read and acknowledge it, but will not sign a consent form (e.g. online survey, phone interview, etc.)*

☒ No

☐ Yes. *Complete Appendix B.*

27(c). DOD regulations require that you provide a consent document to subjects (electronically or in hard copy) unless waived by the IRB. Are you requesting a waiver from the requirement to provide subjects with a consent form?

☒ No

☐ Yes. *Complete Appendix B.*

27(d). DOD regulations list 14 elements of informed consent that are required to be provided to subjects in the consent form script unless waived by the IRB. Are you requesting to exclude any of these elements? A waiver is required if the research involves deception.

☒ No

☐ Yes. *Complete Appendix A.*

28. Will you quote subjects? Please note: when quoting, signed informed consent is required.

☒ No

☐ Yes

29. Describe how you will obtain consent from subjects and how the potential for coercion or undue influence will be minimized. Note: if requesting a consent waiver please state that here.

As the participant arrives the first topic to be covered is the informed consent process; participants will sign the informed consent form and will be provided a blank copy so that they may retain a copy of the information for their records. During this process the experimenter will verbally explain the voluntary nature of the study and further that participants can discontinue participation at any time.
The experimenter has no command, academic or employment leverage over any potential participants.

Appendix A - Request Waiver of Documented Consent

1. Did you check "Yes" on question 27(b) or 27(c)?

☒ No, skip to Appendix B.

☐ Yes

2. Waiver request type.

☐ Waive the requirement to collect a signature on the consent form.

☐ Waive the requirement to provide subjects with a consent form.

3. Waiver applies to the following subject populations:
Note: Please state if the waiver request is for all subjects or certain subject populations. For example if your research involves only a online survey the waiver will request will befor the entire population. If your research involves interviews then the waiver request will only apply to subjects who participate over the phone.

Waiver Criteria
To be approved waiver criteria found in 4a or 4b must be affirmative.

4a. Does the research meet the following criteria?
- The research involves no more than minimal risk to subjects.
- Research involves no procedures for which written consent is normally required outside the research context.
- The information to be presented to subjects (which must be provided in written form as part of the IRB application), includes all required and any additional elements of informed consent.

☐ No, continue to 4b.

☐ Yes, skip to Appendix B.

4b. Does the research meet the following criteria?

- The only record linking the subject and the research is the consent document.
- Each subject will be asked whether he or she wants documentation linking the participant with the research, and the subjects' wishes will govern.
- The information to be presented orally to subjects (which must be provided in written form as part of the IRB application) includes all required and any additional elements of informed consent.

☐ Yes.

☐ No, research does not qualify for a waiver of consent.

Appendix B - Request Waiver of Consent or Elements of Consent

1. Did you check "Yes" on question 27(a) or 27(d)?

☒ No, skip to Principal Investigator Statement of Assurance.

☐ Yes

2. Are you requesting a waiver of consent?

☐ No, skip to Principal Investigator Statement of Assurance.

☐ Yes

3. Does the research involve experimental subjects?

Research involving a human being as an experimental subject: An activity, for research purposes, where there is an intervention or interaction with a living individual for the primary purpose of obtaining data regarding the effect of the intervention or interaction (DODI 3216.02).

☐ No

☐ Yes, waiver requires Secretary of Defense approval (10 USC 980).

4. Is the research regulated by the FDA?

☐ No

☐ Yes, consent may not be waived (32 CFR 219.116)

5. Are you requesting to waive one or more elements of informed consent (use of deception)?

☐ No

☐ Yes, list the elements of informed consent you wish to waive. For a listing of elements see 32 CFR 219.116.

6. Waiver applies to the following subject populations:

Note: Please state if the waiver request is for all subjects or certain subject populations. For example, if your research involves use of pre-collected data (personnel records, training records, lessons learned, etc.) this request would be for all persons represented in the data set.

Waiver Criteria

To be approved waiver criteria found in 7a or 7b must be affirmative.

7a. Does the research meet the following conditions?

- The research involves no more than minimal risk to subjects. Minimal risk means that the probability and magnitude of harm or discomfort anticipated is not greater in and of themselves than those ordinarily encountered in daily life.
- The waiver or alteration will not adversely affect the rights and welfare of subjects by not obtaining consent.
- The research cannot practicably be carried out without the waiver or alteration.
- When appropriate, the subjects will be provided with additional pertinent information after participation.

☐ Yes, skip to Principal Investigator Statement of Assurance.

☐ No, continue to 7b.

7b. Does the research meet the following conditions?

- The research is conducted by or subject to the approval of state or local government officials.
- The research is designed to study, evaluate, or otherwise examine:
  - Public benefit or service programs
  - Procedures for obtaining benefits or services under those programs
  - Possible changes in methods or levels of payment for benefits or services under those programs
- The research cannot practicably be carried out without the waiver or alteration.

☐ Yes, skip to Principal Investigator Statement of Assurance.

☐ No, your research does not qualify for a waiver of documented consent.

Principal Investigator Statement of Assurance

I certify that the information provided in this application is complete and accurate.

I understand that as the Principal Investigator (PI), I have ultimate responsibility for the conduct of the study, the activities of all other investigators listed on the protocol, the ethical performance of the project, the protection of the rights and welfare of human subjects, and strict adherence to the study protocol.

I understand that human subject research activities, including recruitment, may not commence until the Institutional Review Board (IRB) completes its review and, if determined not to be exempt, the Institutional Official (IO) approves.

I will not implement changes to approved research without IRB and IO approval except when necessary to eliminate apparent immediate hazards to the subject and will submit an amendment to the IRB within 5 business days.

I will inform the IRB Chair or Vice Chair, and the Medical Monitor (if one is assigned) of any unanticipated problems involving risks to subjects or others (UPIRTSOs) within 24 hours. I will submit a UPIRTSO report form to the IRB within 5 business days.

I have no conflict of interest preventing me from performing this research.

I will maintain all research records on file. Records include but are not limited to: approved initial review protocol/amendments/ continuing reviews, CITI ethics training records for each member of the research team, correspondence with the IRB, research data and notes, consent forms, UPIRTSO reports, and research agreements.

I recognize that the IRB has the authority to observe (or have a third party observe) the consent process and the conduct of research, and to inspect all research records at any time.

I understand that a continuing review of the research must be reviewed by the IRB and approved by the IO before the expiration date or all research activities including interaction with subjects and personally identifiable data must stop.

I understand that I must submit a final report to the IRB upon expiration of the protocol for all non-exempt research.

I have read, understand, and agree to follow the NPS Instruction on the Protection of Human Subjects.

Principal Investigator Signature:

Date:

NPS Institutional Review Board
Scientific Review Form
(required for all new non-exempt research protocols)

Purpose:
Navy regulation require an independent review of research for scientific merit or scholarship prior to IRB review. A completed scientific review form is required in all applications for IRB review and approval of new research.

Form Instructions:
Please submit this Scientific Review Form to your Department Chair, Director, or Dean (if in GSBPP), along with your IRB application package. Scientific Reviewers will review the IRB application and research proposal when determining the scientific merit of the research. Reviewers can require investigators to revise their submissions if they find that the submission inadequately addresses the points below. Scientific Reviewers may not conduct a scientific review for their own studies. Scientific Reviewers must meet the CITI Ethics training requirements for Scientific Reviewers.

MAC users please use Adobe Reader for Macintosh. Do not use Apple Preview. Fee Adobe reader can be found here.

For questions regarding this form or process send an e-mail to IRB@nps.edu. Form Updated 3-8-17

A. Protocol Basics

Protocol #	
Title of the research:	**Study of Assembly Task**
Principal Investigator:	**Amela Sadagic**
Co-Investigators:	
Student Researchers:	**Douglas Yamashita de Moura**

B. Scientific Review Criteria

Research Team	**Yes**	**No**	**NA**
To the best of your knowledge, does the membership of the research team provide adequate expertise to perform all aspects of the proposed study? *See IRB application Q2.*	☒	☐	
Scientific Merit Review	**Yes**	**No**	**NA**
Does the proposal have a valid research hypothesis and/or appropriate objectives? *See IRB application Q6.*	☒	☐	
Does the protocol provide sufficient information to justify the conduct of the study?	☒	☐	
Is the study design adequate to achieve study objectives? *See IRB application Q7.*	☒	☐	
Is there a method to investigate the research question(s) that would not require the use of human subjects?	☐	☒	
Is the target subject group appropriate for this study? *See IRB application Q13-15.*	☒	☐	
Has the PI demonstrated careful consideration of subject inclusion and exclusion criteria? *See IRB application Q14.*	☒	☐	
Has the PI provided an adequate rationale for the stated sample size? *See IRB application Q15.*	☒	☐	
Is it likely that the PI will be able to meet his/her enrollment goals?	☒	☐	
If the study warrants a data safety monitoring plan is it appropriate? *See IRB application Q26.*	☒	☐	☐
In your opinion should the IRB review the research sooner than annually or monitor the process?	☐	☒	
Research Risks and Benefits	**Yes**	**No**	
Is the risk/benefit ratio favorable? *See IRB application Q19-Q22.*	☒	☐	
Have risks to subjects been minimized by employing sound scientific design? *See IRB application Q7 & Q21.*	☒	☐	
Could risk to subjects be further reduced in any way? If yes, please explain.	☐	☒	

Have all potential risks been accurately and fully described in the application and consent form? *See IRB application Q20 and IRB consent form.*	☒	☐	
Should this study be submitted to the safety office? If yes, please explain.	☐	☒	
Conflict of Interest	**Yes**	**No**	
Do you have a conflict of interest with the proposed research?	☐	☒	
Required Revisions	**Yes**	**No**	
If you have required revisions list them below.	☐	☒	

C. Reviewer Recommendation

☐	This research can be submitted to the IRB as currently written.
☐	This project requires the revisions described above and must be re-reviewed by the scientific reviewer prior to submission to the IRB.
☐	This project was revised, re-reviewed, and found acceptable to submit to the IRB.
☐	This project does not possess scientific merit.

Scientific Reviewer Name:

Scientific Reviewer Signature:

Date:

Naval Postgraduate School
Institutional Review Board
Conflict of Interest Disclosure Form

Update 8-13-12

Principal Investigator:	Dr. Amela Sadagic	Requeseted Start Date:	May 15, 2018
Title of Research	Study of Assembly Task		

About: The purpose of this form is to identify and evaluate potential conflicts of interest in research that may affect the rights and welfare of human subjects.

Instructions: The principal investigator must also identify each person involved with the project who, by the definition given below, qualifies as an "investigator." It is the principal investigator's responsibility to insure that each investigator has been advised of the conflict of interest disclosure (COID) policy and has reviewed and signed this COID form. Proposals will not be reviewed until all investigators have signed the form. The principal investigator must complete and submit to the IRB this COID statement along with the appropriate IRB application form in each of the following circumstances.

- Each time a new proposal is submitted to the IRB for review whether it is funding by a federal agency or a non- governmental (for-profit or not-for-profit) sponsor.

- Any time a new investigator is added to the project.

- Any time during the term of an ongoing award that an Investigator has a change in reportable Significant Financial Interest or acquires a new Significant Financial Interest that was not reported on the original disclosure form.

List all Research Sponsors:	1.		2.	
List all business entities owning or licensing technologies being tested:	1.		2.	
List any additional business entities involved in or potentially affected by the research project:	1.		2.	
	3.		4.	

List Investigators to be paid through NPS as employees:	List Investigators to be paid through NPS as consultants:	List Investigators to be paid through Subcontractors:	List Investigators not paid by NPS or unpaid investigators.
1.	1.	1.	1. Douglas Yamashita de Moura
2.	2.	2.	2.
3.	3.	3.	3.
4.	4.	4.	4.

Who designed the study?	Capt Douglas Yamashita de Moura
Where was the study designed?	Naval Postgraduate School
Where will the results be analyzed?	Naval Postgraduate School
Who will analyze the results?	Capt Douglas Yamashita de Moura and Amela Sadagic

Conflict of Interest Questions		Yes	No
a.	Do you, or does any member of your family have or receive a significant financial interest in or from the research sponsor or from a related business entity affected by the research?	☐	☒
b.	Do you, or does any member of your family receive income including any payment such as salary or consulting fees, royalty payments, reimbursement of expenses of $5,000 or more from the research sponsor or a related business entity?	☐	☒
c.	Do you or does any member of your family own or have any other financial interest in an entity that is proposed as a subcontractor, consortium member or lesser, that is involved in the project?	☐	☒
d.	Do you or does any member of your family have any agreement to receive financial benefit from the research beyond what is described in the proposal budget?	☐	☒
e.	Do you or does any member of your family have outside employment that could appear to cause a potential conflict with this research or raise questions about your professional commitments in undertaking the research, or your primary allegiance to NPS?	☐	☒
f.	Do you or does any member of your family have a position as a director, officer, partner, trustee, manager or employee of an outside entity that conducts business in an area related to the research?	☐	☒
g.	Do you or any member of your family have equity interests such as stocks, stock options or other ownership interests in the research sponsor or related business entities that represent more than 5% equity interest?	☐	☒
h.	Do you or any member of your family have an inventive or ownership interest in any intellectual property that will be utilized in this project?	☐	☒
i.	Have you assigned any student, postdoctoral fellow or other trainee, officer, support staff or other individual to a project sponsored by the research sponsor or related businesses entities?	☐	☒
j.	Do you serve on the Board of Directors or Scientific Advisory Board of the research sponsor or related business entities?	☐	☒
k.	Are you a member of a 'Speakers Bureau' or other list of approved speakers concerning the products or services of the research sponsor or related business entities?	☐	☒
l.	Have you taken any administrative action within the University which is likely to benefit the research sponsor or related business entities?	☐	☒
m.	Do you participate in research on a technology owned or contractually obligated (including by license or exercise of an option to license) to the research sponsor or related business entities?	☐	☒
n.	Have you participated in or otherwise influenced any university transaction to buy, sell, lease, or license real or intellectual property to or from the research sponsor or related business entities?	☐	☒

List below the names of each member of the research team and check whether they have answered "Yes" to the conflict of interest questions listed below. Each member with a conflict of interest will be contacted by the IRB for additional information.

Name of Research Member	Signature	Conflict of Interest Exists	
Amela Sadagic		Yes ☐	No ☒
Douglas Yamashita de Moura		Yes ☐	No ☒
		Yes ☐	No ☐
		Yes ☐	No ☐
		Yes ☐	No ☐
		Yes ☐	No ☐

Principal Investigator Certification:
I certify that all members of the study team have answered the conflict interest questions and those with a conflict of interest are listed above with the box "Conflict of Interest Exists" checked and a description of the conflict of interest is attached.

PI Signature Field.
Digital signatures now accepted.

Date _____

NPS Institutional Review Board

Amendment Form

(modify approved research)

Purpose:

All modifications to approved protocols must be reviewed and approved by the IRB and NPS President before implementation except when necessary to eliminate apparent immediate hazards to subjects. **Examples of protocol modifications include but are not limited to changes in: subject populations, sample size, recruitment procedure, consent procedure, research design, data collection tools and research team personnel.**

Form Instructions:

To request modifications to an approved protocol submit the following to IRB@nps.edu. A IRB administrator will contact you if additional information is required. MAC users please use Adobe Reader for Macintosh. Free Adobe reader can be found here.

1. A completed Amendment Form signed by Principal Investigator (PI).
2. Copies of new or revised protocol documents (if any) such as recruitment scripts, consent forms, data collection tools, etc.
3. Copy of CITI Ethics training completion reports (if new researchers are added).

For questions regarding this form or process send an e-mail to IRB@nps.edu. Form Updated 2-27-17

A. Protocol Basics

Protocol #	**NPS-2018.0065-EP7-A**
Protocol Title:	**User Study on The Role of Stereoscopic Depth Cue and Immersion in Maintenance Task**
Department:	**MOVES Institute**
Principal Investigator:	**Dr. Amela Sadagic**
Co-Investigators:	
Student Investigators:	**Capt Douglas Yamashita de Moura**

B. Protocol Modifications

1. Is there a change to research team personnel? Please remember to remove all student researchers who have graduated.

☒ No, skip to question 2.

☐ Yes, complete the table below.

Investigators Added Attach CITI ethics training completion reports and revised Conflict of Interest Form for all investigators added to the protocol.			
Name	Title	Dept.	Roles and Responsibilities in the Research
Investigators Removed The investigator must return all research data and consent forms to the PI.			

2. Is there a change to the sample size? Sample size is the number of individuals you will enroll in your study. You are required to report both increases and decreases.

☐ No, skip to question B3.

☒ Yes

2a. Approved sample size: 60

2b. Revised sample size: 80

2c. Provide justification for the change in sample size.

We would like to increase the significance level of our findings and ensure that we have opportunity to gather more subjective responser (comments) from subjects in study.

3. Is there a change to the research design?

☒ No

☐ Yes, describe the changes in the research design.

4. Are there changes to tasks subjects are asked to complete?

☒ No

☐ Yes, describe the changes and state if these changes will affect subjects willingness to participate. Attached a revised consent form.

5. Is there change to participant risks or benefits?

☒ No

☐ Yes, describe the change in risks and benefits to subjects and state if these changes will affect subjects willingness to participate.

6. Will the proposed modification require revision to or the addition of any of the following (check all that apply and attach revised documents)?

☐ Consent procedure, form, or script

☐ Recruitment procedure or script

☐ Data collection tools (e.g. questionnaires, survey, interview questions, etc)

☐ Conflict of Interest form and CITI Ethics Training Certificate (for added investigators)

☐ Other, describe below.

7. Additional comments.

No changes of any document or procedure would be needed.

C. Principal Investigator Statement of Assurance

I certify that the information provided in this application is complete and accurate.

I understand that as the Principal Investigator (PI), I have ultimate responsibility for the conduct of the study, the activities of all other investigators listed on the protocol, the ethical performance of the project, the protection of the rights and welfare of human subjects, and strict adherence to the study protocol.

I understand that human subject research activities, including recruitment, may not commence until the Institutional Review Board (IRB) completes its review and, if determined not to be exempt, the Institutional Official (IO) approves.

I will not implement changes to approved research without IRB and IO approval except when necessary to eliminate apparent immediate hazards to the subject and will submit an amendment to the IRB within 5 business days.

I will inform the IRB Chair or Vice Chair, and the Medical Monitor (if one is assigned) of any unanticipated problems involving risks to subjects or others (UPIRTSOs) within 24 hours. I will submit a UPIRTSO report form to the IRB within 5 business days.

I have no conflict of interest preventing me from performing this research.

I will maintain all research records on file. Records include but are not limited to: approved initial review protocol/amendments/ continuing reviews, CITI ethics training records for each member of the research team, correspondence with the IRB, research data and notes, consent forms, UPIRTSO reports, and research agreements.

I recognize that the IRB has the authority to observe (or have a third party observe) the consent process and the conduct of research, and to inspect all research records at any time.

I understand that a continuing review of the research must be reviewed by the IRB and approved by the IO before the expiration date or all research activities including interaction with subjects and personally identifiable data must stop.

I understand that I must submit a final report to the IRB upon expiration of the protocol for all non-exempt research.

I have read, understand, and agree to follow the NPS Instruction on the Protection of Human Subjects.

Principal Investigator Signature:

Date:

NPS IRB
Amendment Reviewer Checklist
last updated 2-27-17

The amendment checklist is completed when amendments to currently approved research are requested. Reviews of minor changes are completed by the IRB Chair or Vice Chair. Major changes that involve an increase in risks/discomforts or a decrease in benefit to subjects must be reviewed by the convened IRB.

Protocol #:	NPS-2018.0065-IR-EP7-A	Updated #:	NPS-2018.0065-AM01-EP7-A
Protocol Title:	User Study on The Role of Stereoscopic Depth Cue and Immersion in Maintenance Task		
Principal Investigator:	Dr. Amela Sadagic	Co-Investigators:	
Student Investigators:	Capt Douglas Yamashita de Moura		

Approval Criteria 32 CFR 219.111				
		Yes	No	NA
1.	Risks to subjects are minimized whenever appropriate, by using procedures already being performed on the subjects for diagnostic or treatment purposes.	☒	☐	☐
2.	Risks to subjects are reasonable in relation to anticipated benefits, if any, to subjects, and the importance of the knowledge that may reasonably be expected to result. In evaluating risks and benefits, the IRB should consider only those risks and benefits that may result from the research. The IRB should not consider possible long-range effects of applying knowledge gained in the research as among those research risks that fall within the purview of its responsibility.	☒	☐	
3.	Selection of subjects is equitable.	☒	☐	
4.	Informed consent will be sought from each prospective subject or the subject's legally authorized representative.	☒	☐	☐
5.	Informed consent will be appropriately documented (unless waived under 32 CFR 219.116-117).	☒	☐	☐
6.	When appropriate, the research plan makes adequate provision for monitoring the data collected to ensure the safety of subjects.	☒	☐	☐
7.	When appropriate, there are adequate provisions to protect the privacy of subjects and to maintain the confidentiality of data.	☒	☐	☐
8.	When some or all of the subjects are likely to be vulnerable to coercion or undue influence, additional safeguards have been included in the study to protect the rights and welfare of these subjects.	☐	☐	☒

Additional Criteria for Review				
		Yes	No	Reviewer Comments
1.	Do the requested changes increase the risk to research subjects? If yes, list the new risk level.	☐	☒	
2.	Will the amendment affect the current consent process? If yes, does the revised consent procedure and document reflect the new process?	☐	☒	

		Yes	No	Reviewer Comments
3.	Does the amendment require the protocol to be reviewed more frequently?	☐	☒	
4.	Have significant new findings occurred that might relate to subjects' willingness to continue participating? If yes, how will the new findings be communicated to subjects?	☐	☒	
5.	Have changes to the approved protocol been made to eliminate immediate hazards to subjects? If yes, were the changes reported to the IRB and was the change consistent with ensuring the subjects' continued welfare?	☐	☒	

Changes to the following are requested:					
☐	Informed Consent	☐	Research Personnel	☐	Research Site
☐	Recruitment Materials	☐	Data Collection Tools	☒	Sample Size

Additional Reviewer Comments:

Exempt Review	Expedited Review	Review by Convened IRB
☐ Approve	☒ Approve	☐ Approve
☐ Modifications Required	☐ Modifications Required	☐ Modifications Required
☐ Determined Expedited	☐ Defer to Convened IRB	☐ Back to Chair or Vice Chair (eligible for expedited review)
		☐ Defer
		☐ Dissaprove

IRB Chair / Vice Chair Signature:

Date:

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX B. CHECKLIST

USER STUDY CHECKLIST

LONG BEFORE PARTICIPANT COME:

- ☐ Make sure you have copies of all forms that you need (including backup copies in case electronic fail to work)
- ☐ Make sure controllers are fully charged!
- ☐ Make sure 3D printed pieces and printed mini-map are ready/available
- ☐ Make sure you have extra bottles of water in case participant needs water
- ☐ Assign participant to proper study condition

BEFORE PARTICIPANT COME:

- ☐ Put up the sign "DO NOT DISTURB - STUDY IN PROGRESS"
- ☐ Check videorecorder and tripod: it is staged in predetermined position, and ready
- ☐ Check if printed consent form and pen are ready for participant – position them on the table
- ☐ Check if electronic version of surveys for participant is ready (loaded on laptop)
- ☐ Make sure unity scene for training session is loaded for that particular study condition
- ☐ Make sure fresh/new protective liner for HMD is attached in the headset (if they are to use HMD)

EXPERIMENTAL SESSION:

- ☐ Welcome participant, offer them water if they need it
- ☐ Ask participant to complete informed **consent document**
- ☐ Ask participant if he/she has any question
- ☐ Make note if the participant wears corrective lenses or glasses
- ☐ Make note if the participant who normally wears glasses used them during experiment
- ☐ Ask participant to **fill initial SSQ** (baseline - 1st SSQ)
- ☐ Check if all questions in SSQ were answered
- ☐ Brief participant about training environment (introduce controller use, task and time limit using written briefing)
- ☐ Help participant put on HMD/hat and take seat: Adjust participant position and set up with controllers and HMD/hat
- ☐ Start/play **training environment** and allow participant familiarization period (10 min max). Start stopwatch.
- ☐ Walk participant through set of checks in training environment ("grab… release… move head…")
- ☐ Check stopwatch: Let participant know when training session is over (after 10 min)
- ☐ End training session and ask participant to **fill 2nd SSQ**
- ☐ Check if all questions were answered in SSQ, and save file
- ☐ Ask participant to read instructions that introduce experimental environment and objective
- ☐ Turn **ON** recording equipment (camcorder)
- ☐ Load and **start task #1** environment
- ☐ Signal to the participant that they can start the session and start stopwatch
- ☐ Keep checking stopwatch and make necessary notes. Signal when 10 min is reached, and stop the session
- ☐ Ask subject to **fill 3rd SSQ** (the end of 1st session – 3rd SSQ)
- ☐ Ask participant to **fill post-task #1 questionnaire**
- ☐ Check if all questions were answered in SSQ and in post-task #1 questionnaire, and save file
- ☐ Load and **start task #2** environment
- ☐ Signal to the participant that they can start the session, and start stopwatch
- ☐ Keep checking stopwatch and make necessary notes. Signal when 10 min is reached, and stop the session
- ☐ Ask subject to **fill 4th SSQ** (the end of 2st session – 4th SSQ)
- ☐ Ask participant to **fill post-task #2** questionnaire and **SUS questionnaire**
- ☐ Check if all questions were answered in SSQ and in post-task #2 questionnaire, and save file
- ☐ Introduce task #3 (including 10 min time limit)
- ☐ Signal to the participant that he/she can start the session and start stopwatch
- ☐ Keep checking stopwatch and make necessary notes. Signal when 10 min is reached, and stop the session
- ☐ Ask participant to **fill post-task #3** questionnaire
- ☐ Turn **OFF** the camcorder
- ☐ Check if all questions were answered in post-task #3 questionnaire, and save file
- ☐ Ask participant to complete **demographic survey**
- ☐ Check if all questions were answered in demographic survey, and save file

DEBRIEFING:

- ☐ Conduct final debriefing and answer any question that participant may have

AFTER DEBRIEFING, IF POSSIBLE, OR END OF DAY:

- ☐ Copy questionnaires and surveys to NPS secure server
- ☐ Fill electronic version of notes and save to NPS secure server
- ☐ Copy video recordings to NPS secure server and delete them from camcorder

Metrics to be reported by observer (tally during experimental environment)

1. List main participant's problems during Task #1

2. Check number of correct connections during Task #1

I I I I I I I I I I

3. Additional observations during Task #1

4. List main participant's problems during Task #2

5. Check number of correct connections during Task #2

I I I I I I I I I I I I I I

6. Additional observations during Task #2

7. List main participant's problems during Task #3

8. Check number of correct connections during Task #3

I I I I I I I I I I I I I I

9. Additional observations during Task #3

10. General observations during experiment

APPENDIX C.SIMULATOR SICKNESS QUESTIONNAIRE (SSQ)

No_____ Date_____

SIMULATOR SICKNESS QUESTIONNAIRE

Kennedy, Lane, Berbaum, & Lilienthal (1993)***

Instructions : Circle how much each symptom below is affecting you right now.

1. General discomfort	None	Slight	Moderate	Severe
2. Fatigue	None	Slight	Moderate	Severe
3. Headache	None	Slight	Moderate	Severe
4. Eye strain	None	Slight	Moderate	Severe
5. Difficulty focusing	None	Slight	Moderate	Severe
6. Salivation increasing	None	Slight	Moderate	Severe
7. Sweating	None	Slight	Moderate	Severe
8. Nausea	None	Slight	Moderate	Severe
9. Difficulty concentrating	None	Slight	Moderate	Severe
10. « Fullness of the Head »	None	Slight	Moderate	Severe
11. Blurred vision	None	Slight	Moderate	Severe
12. Dizziness with eyes open	None	Slight	Moderate	Severe
13. Dizziness with eyes closed	None	Slight	Moderate	Severe
14. *Vertigo	None	Slight	Moderate	Severe
15. **Stomach awareness	None	Slight	Moderate	Severe
16. Burping	None	Slight	Moderate	Severe

* Vertigo is experienced as loss of orientation with respect to vertical upright.

** Stomach awareness is usually used to indicate a feeling of discomfort which is just short of nausea.

Last version : March 2013

***Original version : Kennedy, R.S., Lane, N.E., Berbaum, K.S., & Lilienthal, M.G. (1993). Simulator Sickness Questionnaire: An enhanced method for quantifying simulator sickness. *International Journal of Aviation Psychology, 3*(3), 203-220.

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX D. POST-TASK #1 QUESTIONNAIRE

1. Did you successfully complete the task within allotted time?

   YES NO

2. If 'NO':

Why were you not able to complete it? List anything that hindered you from completing the task:

_____

_____

3. How was your experience with following actions?

	It was very difficult	It was moderately difficult	It was slightly difficult	Neither difficult nor easy	It was slightly easy	It was moderately easy	It was very easy
a. Selecting objects?	○	○	○	○	○	○	○
b. Manipulating objects?	○	○	○	○	○	○	○
c. Assembling (putting together) parts?	○	○	○	○	○	○	○

4. How was your overall experience with this assembly task?

It was very difficult	It was moderately difficult	It was slightly difficult	Neither difficult nor easy	It was slightly easy	It was moderately easy	It was very easy
○	○	○	○	○	○	○

5. How would you rate your own performance in this task?

Strongly dissatisfied	Moderately dissatisfied	Slightly dissatisfied	Neither dissatisfied nor satisfied	Slightly satisfied	Moderately satisfied	Strongly satisfied
○	○	○	○	○	○	○

6. To what extent did you have the sense of being in that room which has table with all parts of that object laid out on it. (For example, if you were asked this question about the room you are in now, you would give a score of 7. However, if you were asked this question about whether you were sitting in a room at home now, you would give this a score of 1).

In the last session, the extent to which I had a sense of being in the room which has the table with all parts of that object on it, was:

Not at all being there	Moderately not being there	Slightly not being there	Neutral	Slightly being there	Moderately being there	Very much so being there
○	○	○	○	○	○	○

7. Think about the session that you just completed, and imagine that you are doing it now. To what extent in your imagination can you interact with the parts of that object?

I can think about myself back into that room interacting with parts of that object:

Strongly not interacting	Moderately not interacting	Slightly not interacting	Neutral	Slightly interacting	Moderately interacting	Strongly interacting
○	○	○	○	○	○	○

8. How well did you fell the virtual environment accurately portrayed the experience of manipulating parts of the objects?

	Strongly Inaccurate	Moderately Inaccurate	Slightly Inaccurate	Neither inaccurate or accurate	Slightly Accurate	Moderately Accurate	Strongly Accurate
a. Visual representation of individual parts	○	○	○	○	○	○	○
b. Interaction with individual parts (grabbing and manipulating)	○	○	○	○	○	○	○
c. Interaction between the parts	○	○	○	○	○	○	○
d. Real time visualization of my hands and arms movements	○	○	○	○	○	○	○
e. Real time visualization of my head movements	○	○	○	○	○	○	○

9. How would you rate this assembly task?

It was very difficult	It was moderately difficult	It was slightly difficult	Neither difficult nor easy	It was slightly easy	It was moderately easy	It was very easy
○	○	○	○	○	○	○

10. How well were you able to distinguish position of the parts in space (parts that were closer to you vs parts that were further away from you)?

It was very difficult	It was moderately difficult	It was slightly difficult	Neither difficult nor easy	It was slightly easy	It was moderately easy	It was very easy
○	○	○	○	○	○	○

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX E. POST-TASK #2 QUESTIONNAIRE

1. Did you successfully complete the task within allotted time?

   YES NO

2. If 'NO':

Why were you not able to complete it? List anything that hindered you from completing the task:

_____

_____

3. How was your experience with following actions?

	It was very difficult	It was moderately difficult	It was slightly difficult	Neither difficult nor easy	It was slightly easy	It was moderately easy	It was very easy
a. Selecting objects?	○	○	○	○	○	○	○
b. Manipulating objects?	○	○	○	○	○	○	○
c. Assembling (putting together) parts?	○	○	○	○	○	○	○

4. How was your overall experience with this assembly task?

It was very difficult	It was moderately difficult	It was slightly difficult	Neither difficult nor easy	It was slightly easy	It was moderately easy	It was very easy
○	○	○	○	○	○	○

5. How would you rate your own performance in this task?

Strongly dissatisfied	Moderately dissatisfied	Slightly dissatisfied	Neither dissatisfied nor satisfied	Slightly satisfied	Moderately satisfied	Strongly satisfied
○	○	○	○	○	○	○

6. To what extent did you have the sense of being in that room which has table with all parts of that object laid out on it. (For example, if you were asked this question about the room you are in now, you would give a score of 7. However, if you were asked this question about whether you were sitting in a room at home now, you would give this a score of 1).

In the last session, the extent to which I had a sense of being in the room which has the table with all parts of that object on it, was:

Not at all being there	Moderately not being there	Slightly not being there	Neutral	Slightly being there	Moderately being there	Very much so being there
○	○	○	○	○	○	○

7. Think about the session that you just completed, and imagine that you are doing it now. To what extent in your imagination can you interact with the parts of that object?

I can think about myself back into that room interacting with parts of that object:

Strongly not interacting	Moderately not interacting	Slightly not interacting	Neutral	Slightly interacting	Moderately interacting	Strongly interacting
○	○	○	○	○	○	○

8. How well did you fell the virtual environment accurately portrayed the experience of manipulating parts of the objects?

	Strongly Inaccurate	Moderately Inaccurate	Slightly Inaccurate	Neither inaccurate or accurate	Slightly Accurate	Moderately Accurate	Strongly Accurate
a. Visual representation of individual parts	○	○	○	○	○	○	○
b. Interaction with individual parts (grabbing and manipulating)	○	○	○	○	○	○	○
c. Interaction between the parts	○	○	○	○	○	○	○
d. Real time visualization of my hands and arms movements	○	○	○	○	○	○	○
e. Real time visualization of my head movements	○	○	○	○	○	○	○

9. How would you rate this assembly task?

It was very difficult	It was moderately difficult	It was slightly difficult	Neither difficult nor easy	It was slightly easy	It was moderately easy	It was very easy
○	○	○	○	○	○	○

10. How well were you able to distinguish position of the parts in space (parts that were closer to you vs parts that were further away from you)?

It was very difficult	It was moderately difficult	It was slightly difficult	Neither difficult nor easy	It was slightly easy	It was moderately easy	It was very easy
○	○	○	○	○	○	○

11. How would you compare this task #2 to task #1 (your previous task)?

Task #2 was _____ than task #1. (select the answer bellow)

Much more difficult	Moderately more difficult	Slightly more difficult	Neither more difficult nor easier	Slightly easier	Moderately easier	Much easier
○	○	○	○	○	○	○

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX F. POST-TASK #3 QUESTIONNAIRE

1. Did you successfully complete the task within allotted time?

   YES NO

2. If 'NO':

Why were you not able to complete it? List anything that hindered you from completing the task:

_____

_____

3. How was your experience with following actions?

	It was very difficult	It was moderately difficult	It was slightly difficult	Neither difficult nor easy	It was slightly easy	It was moderately easy	It was very easy
a. Selecting objects?	○	○	○	○	○	○	○
b. Manipulating objects?	○	○	○	○	○	○	○
c. Assembling (putting together) parts?	○	○	○	○	○	○	○

4. How was your overall experience with this assembly task?

It was very difficult	It was moderately difficult	It was slightly difficult	Neither difficult nor easy	It was slightly easy	It was moderately easy	It was very easy
○	○	○	○	○	○	○

5. How would you rate your own performance in this task?

Strongly dissatisfied	Moderately dissatisfied	Slightly dissatisfied	Neither dissatisfied nor satisfied	Slightly satisfied	Moderately satisfied	Strongly satisfied
○	○	○	○	○	○	○

6. How would you rate this assembly task?

It was very difficult	It was moderately difficult	It was slightly difficult	Neither difficult nor easy	It was slightly easy	It was moderately easy	It was very easy
○	○	○	○	○	○	○

7. How would you compare this task #3 to task #2 (your previous task)?

Task #3 was _____ than task #2. (select the answer below)

Much more difficult	Moderately more difficult	Slightly more difficult	Neither more difficult nor easier	Slightly easier	Moderately easier	Much easier
○	○	○	○	○	○	○

APPENDIX G. DEMOGRAPHIC SURVEY

1. Year of birth: _____
2. Sex:

   FEMALE MALE
3. Do you wear regular glasses?

   YES NO
4. If 'YES':
   a. What is your vision? _____
   b. Did you wear your glasses with the VR headset or shutter glasses during the experiment?

      YES NO
5. What hand do you use to manipulate computer mouse? (circle one that applies)

   RIGHT LEFT EITHER
6. What is your occupation? _____
7. Does your daily activity or your hobbies involve assembly of objects?

   YES NO
8. If 'YES':
   a. What type of tasks do you perform? _____

      _____

      _____
   b. How do you feel about the importance of assembly job in those tasks?

Strongly not important	Moderately not important	Slightly not important	Neither	Slightly important	Moderately important	Strongly important
○	○	○	○	○	○	○

9. How do you feel about your ability of assembling the objects?

Strongly Dissatisfied	Moderately Dissatisfied	Slightly Dissatisfied	Neither	Slightly Satisfied	Moderately Satisfied	Strongly Satisfied
○	○	○	○	○	○	○

10. Do you play video games?

    YES NO

11. If ‘YES’:

a. How often? (circle one that applies)

Less than 2 hrs/wk 2-4 hrs/wk 4-8 hrs/wk More than 8 hrs/wk

b. What percentage of game types do you play? (Ensure that all values add up to 100%)

Single-player: _____ % Multi-player: _____ %

c. What percentage of game types do you play? (Ensure that all values add up to 100%)

First-person view: _____ % Third-person view: _____ %

12. Have you used a virtual reality head mounted display before?

YES NO

13. If ‘YES’:

a. What kind? (circle all that apply)

HTC VIVE Oculus Rift Gear VR Google Cardboard-style Hololens
Other:

_____

b. How many times in last 3 years? (circle one that applies)

Only once Less than 5 times Between 5 and 10 times More than 10 times

c. When was the last time you used it? (circle one that applies)

Within last 30 days Within last 6 months Within the last year More than a year ago

14. Please leave any comments and suggestions about the experiment:

_____

_____

_____

_____

_____

APPENDIX H. SYSTEM USABILITY SCALE (SUS)

		Strongly Disagree	Moderately Disagree	Slightly Disagree	Neither	Slightly Agree	Moderately Agree	Strongly Agree
1.	I think that I would like to use this system frequently	○	○	○	○	○	○	○
2.	I found the system unnecessarily complex	○	○	○	○	○	○	○
3.	I thought the system was easy to use	○	○	○	○	○	○	○
4.	I think that I would need the support of a technical person to be able to use this system	○	○	○	○	○	○	○
5.	I found the various functions in this system were well integrated	○	○	○	○	○	○	○
6.	I thought there was too much inconsistency in this system	○	○	○	○	○	○	○
7.	I would imagine that most people would learn to use this system very quickly	○	○	○	○	○	○	○
8.	I found the system very cumbersome to use	○	○	○	○	○	○	○
9.	I felt very confident using the system	○	○	○	○	○	○	○
10.	I needed to learn a lot of things before I could get going with this system	○	○	○	○	○	○	○

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX I. SHAPIRO-WILK TEST FOR ASSEMBLY TIMES

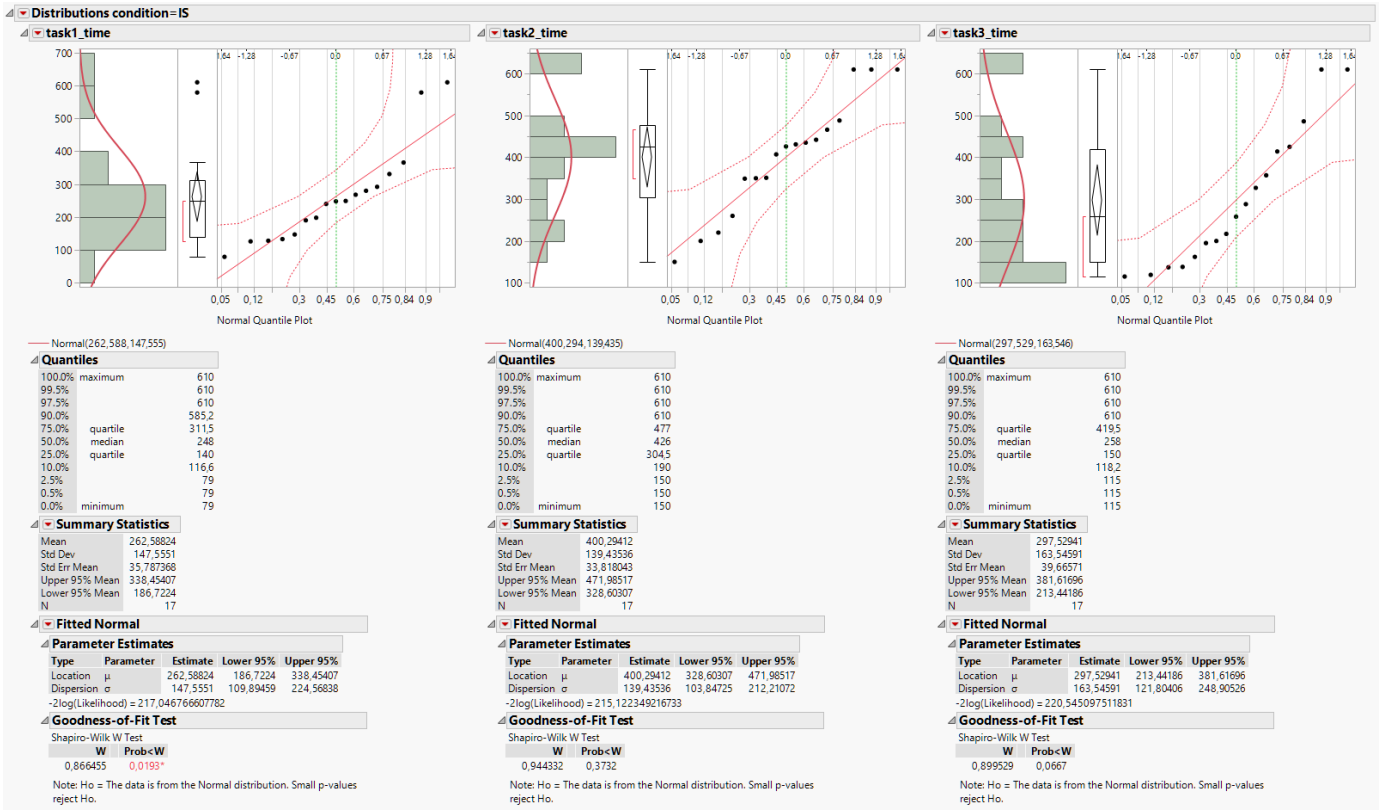


Figure 72. Shapiro-Wilk test of assembly times for Tasks 1, 2 and 3 in IS condition

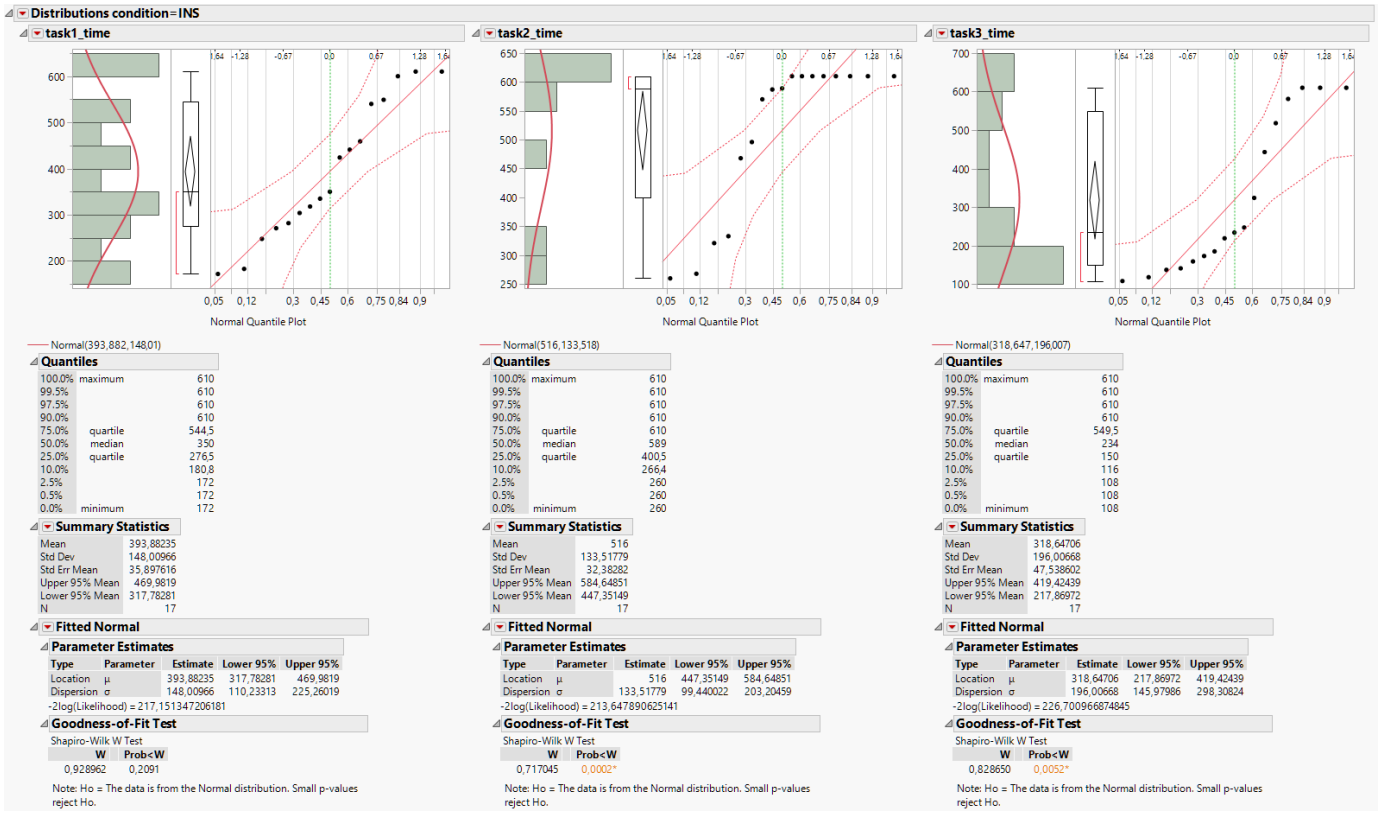


Figure 73. Shapiro-Wilk test of assembly times for Tasks 1, 2 and 3 in INS condition

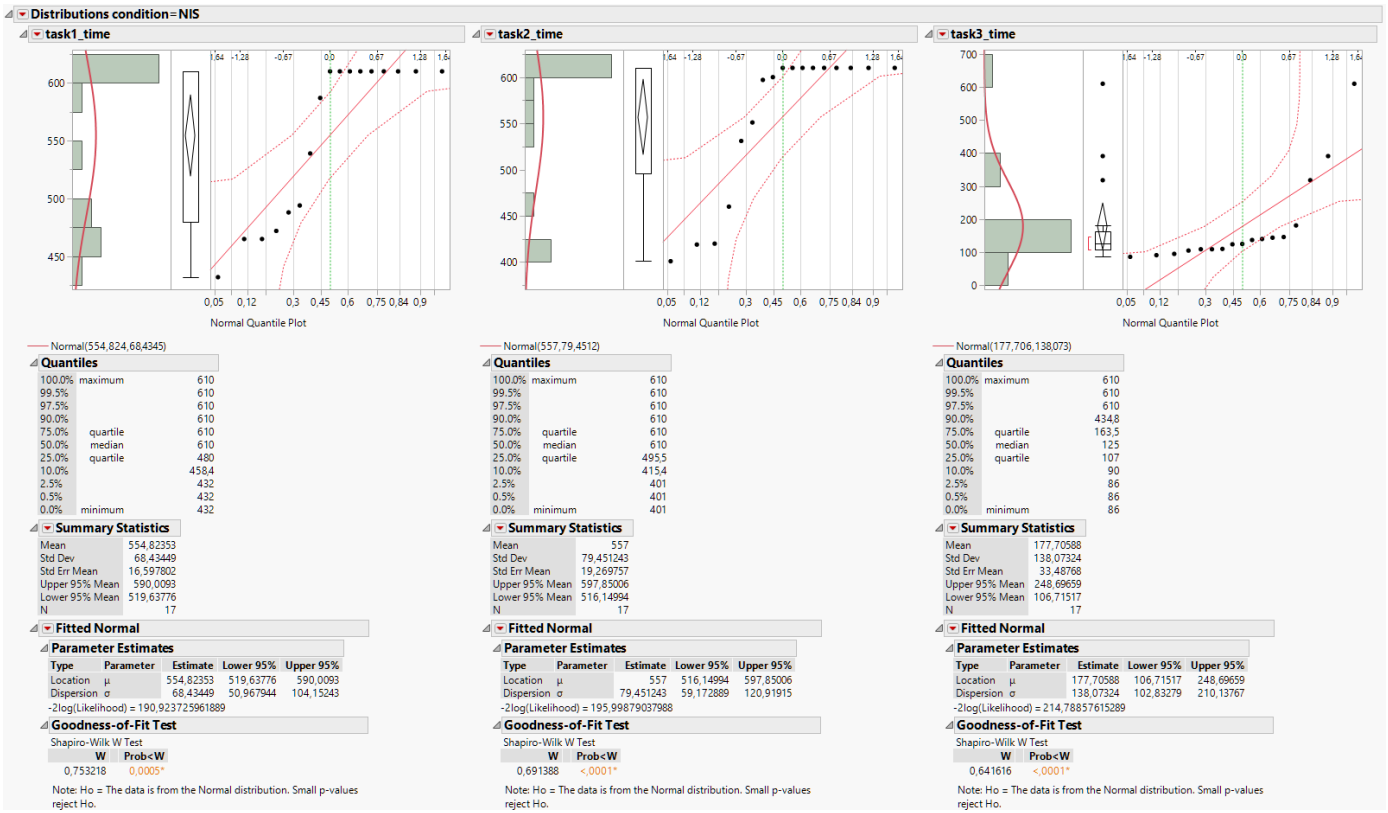


Figure 74. Shapiro-Wilk test of assembly times for Tasks 1, 2 and 3 in NIS condition

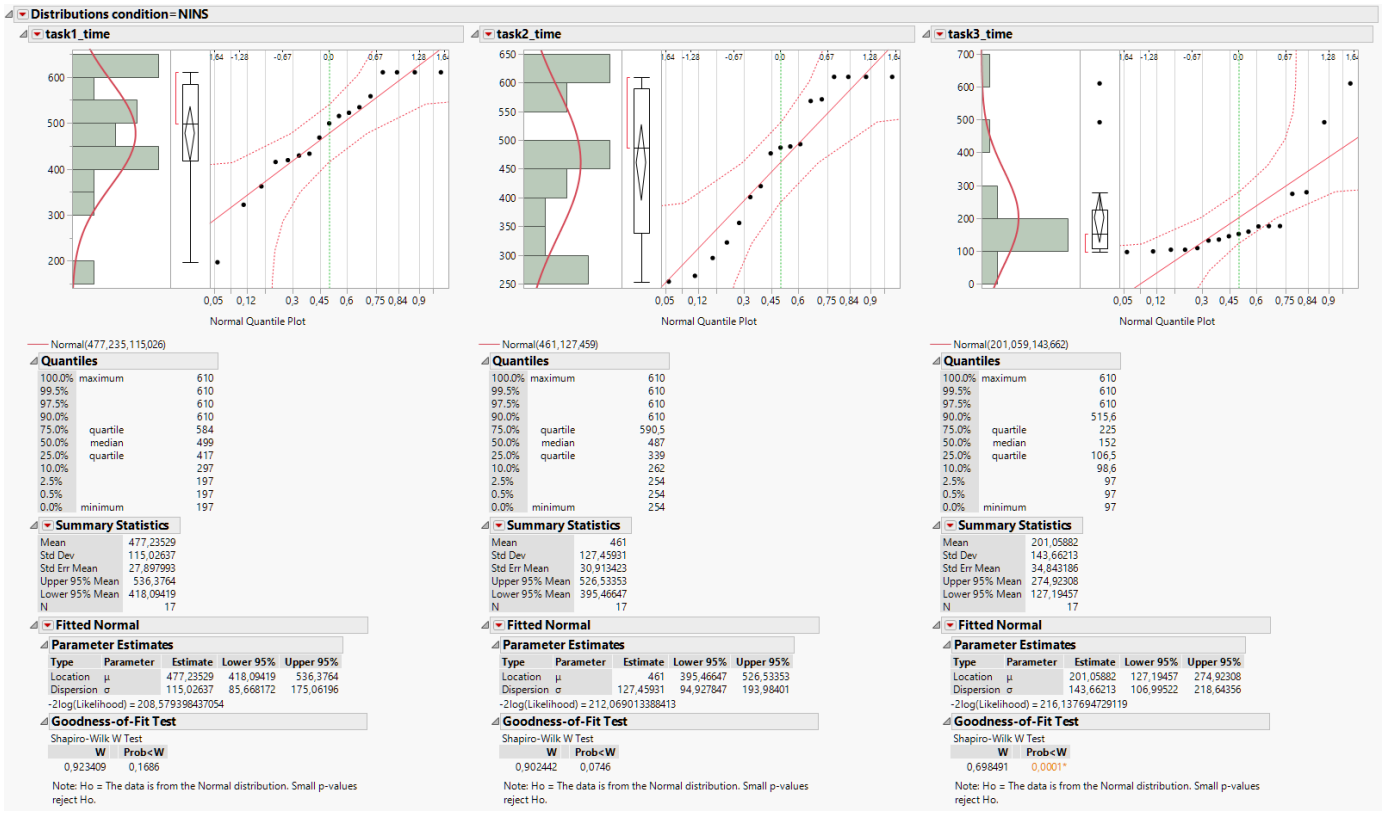


Figure 75. Shapiro-Wilk test of assembly times for Tasks 1, 2 and 3 in NINS condition

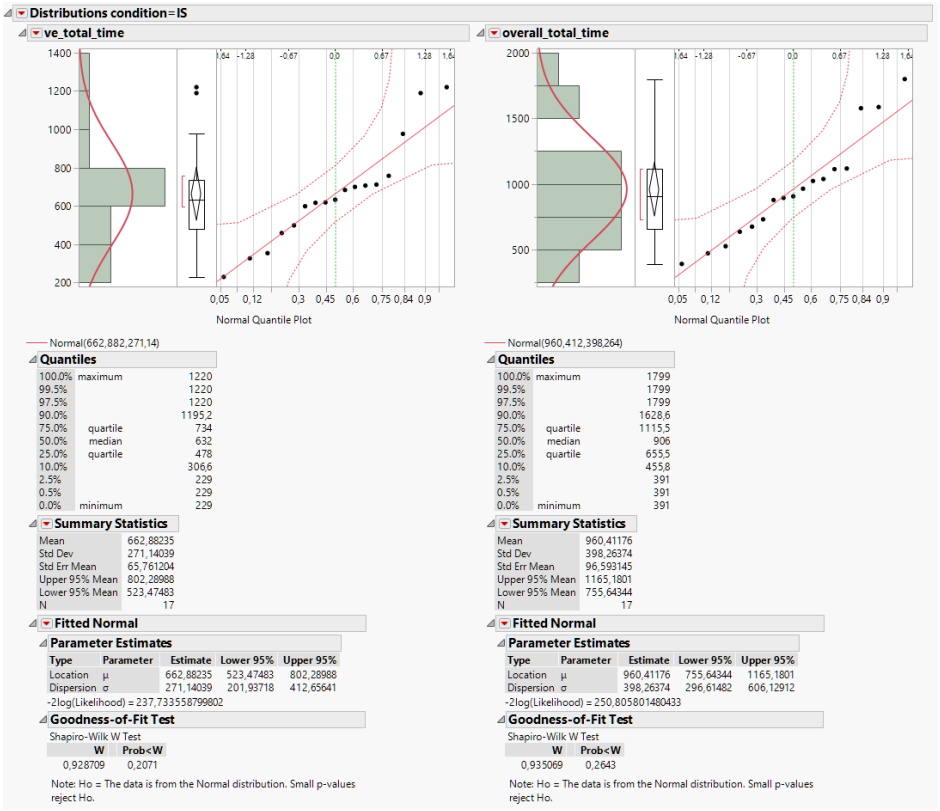


Figure 76. Shapiro-Wilk test of assembly times for VE Total and Overall Total in IS condition

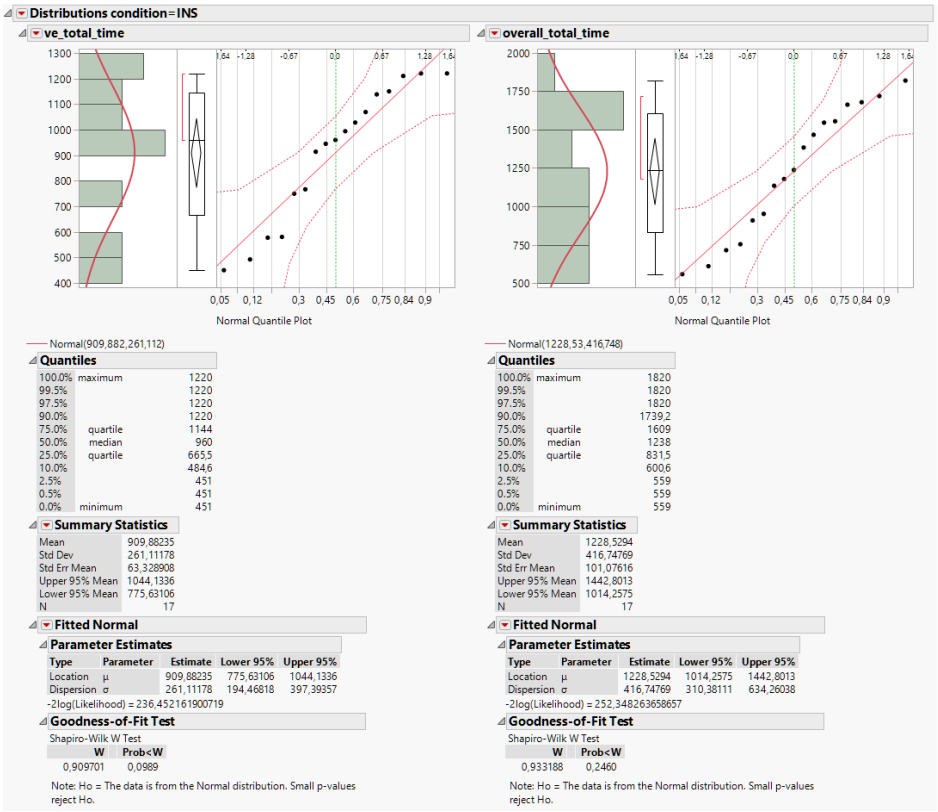


Figure 77. Shapiro-Wilk test of assembly times for VE Total and Overall Total in INS condition

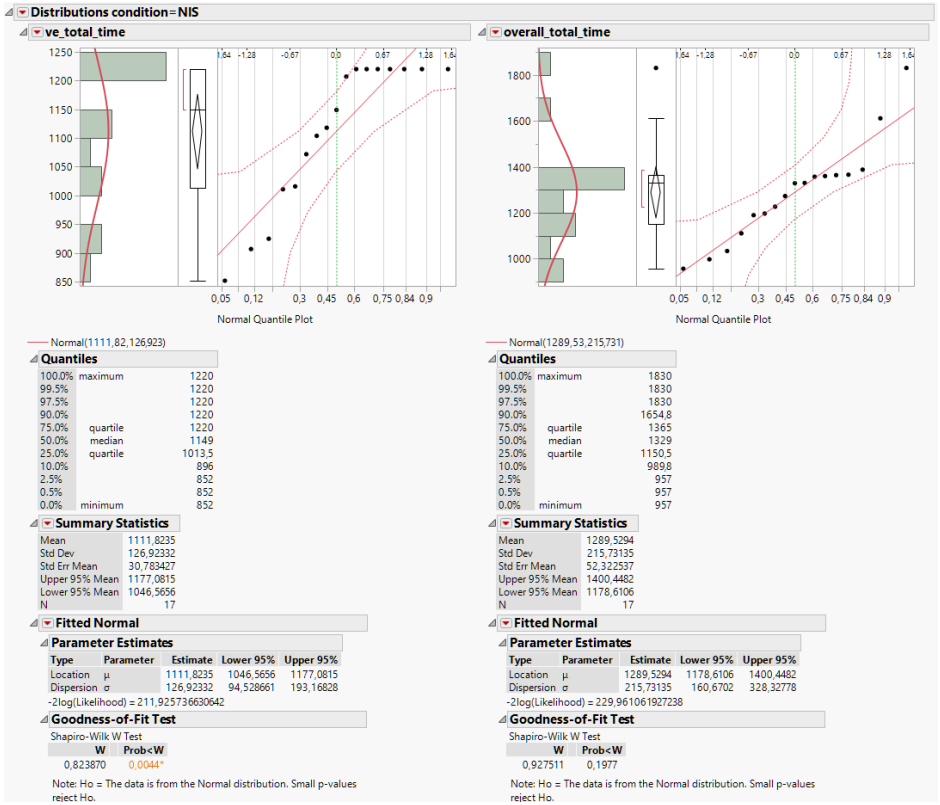


Figure 78. Shapiro-Wilk test of assembly times for VE Total and Overall Total in NIS condition

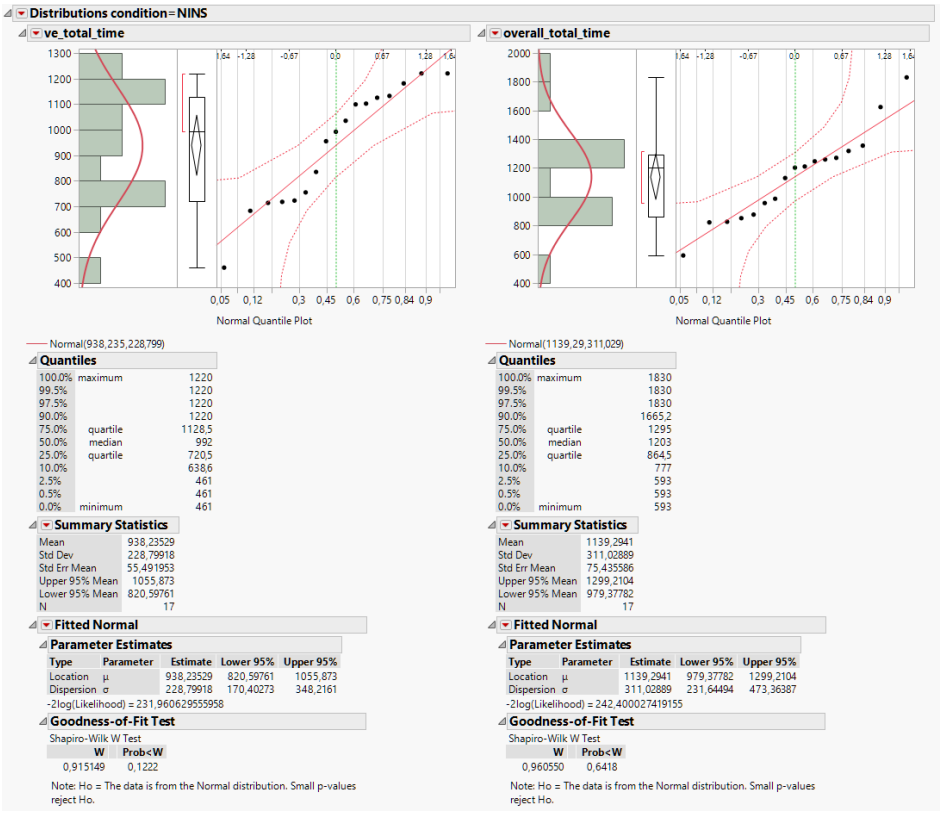


Figure 79. Shapiro-Wilk test of assembly times for VE Total and Overall Total in NINS condition

APPENDIX J. SHAPIRO-WILK TEST FOR CONNECTIONS/PARTS

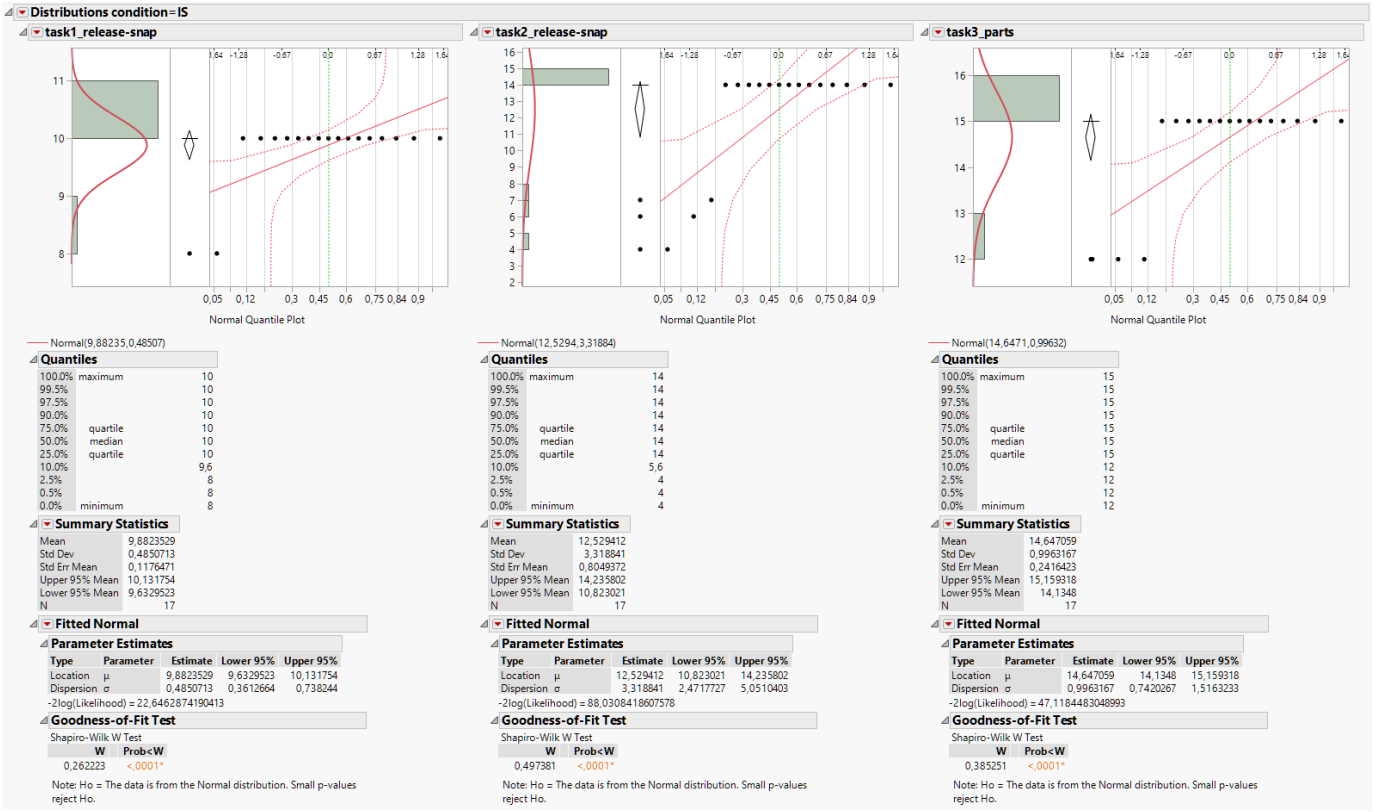


Figure 80. Shapiro-Wilk test of connections/parts for Tasks 1, 2 and 3 in IS condition

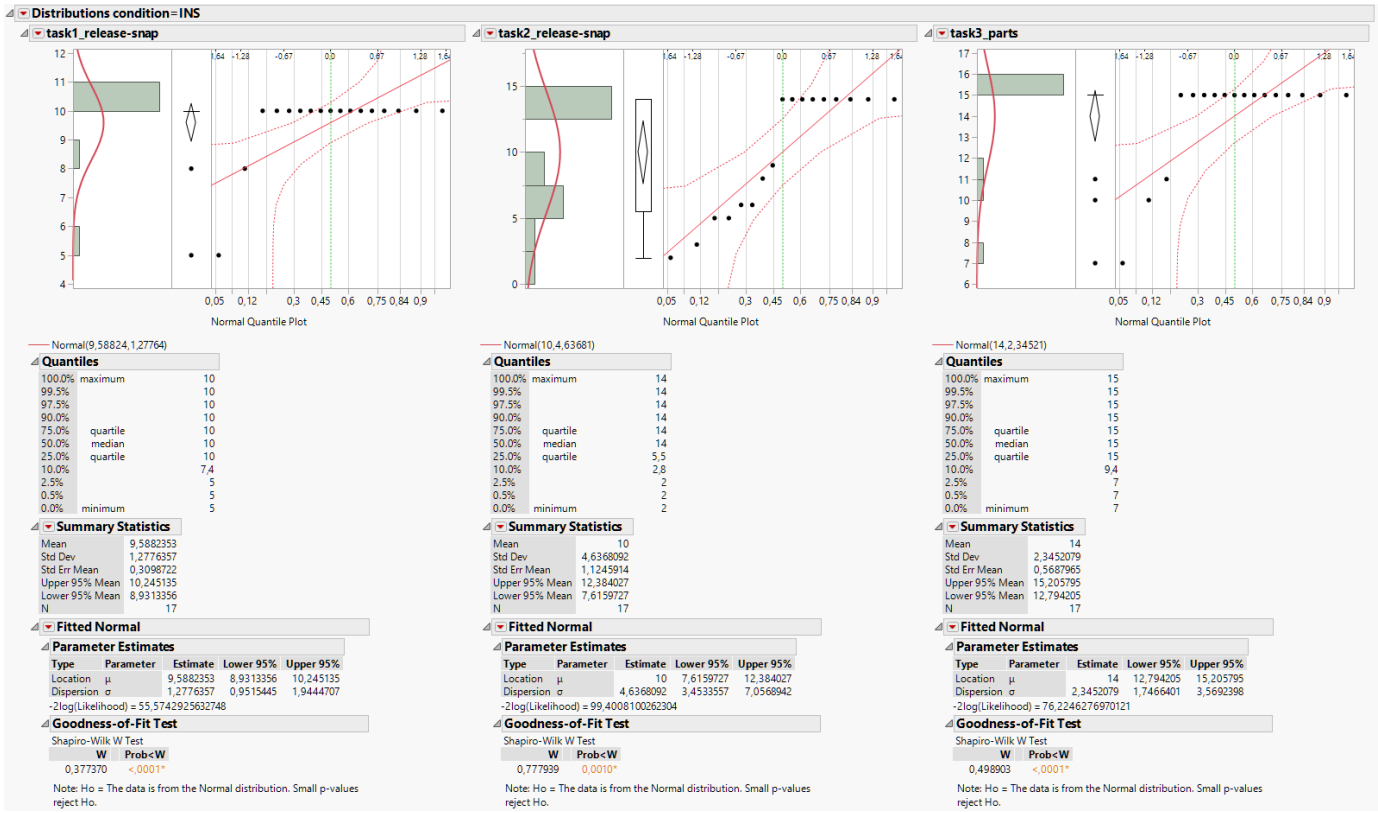


Figure 81. Shapiro-Wilk test of connections/parts for Tasks 1, 2 and 3 in INS condition

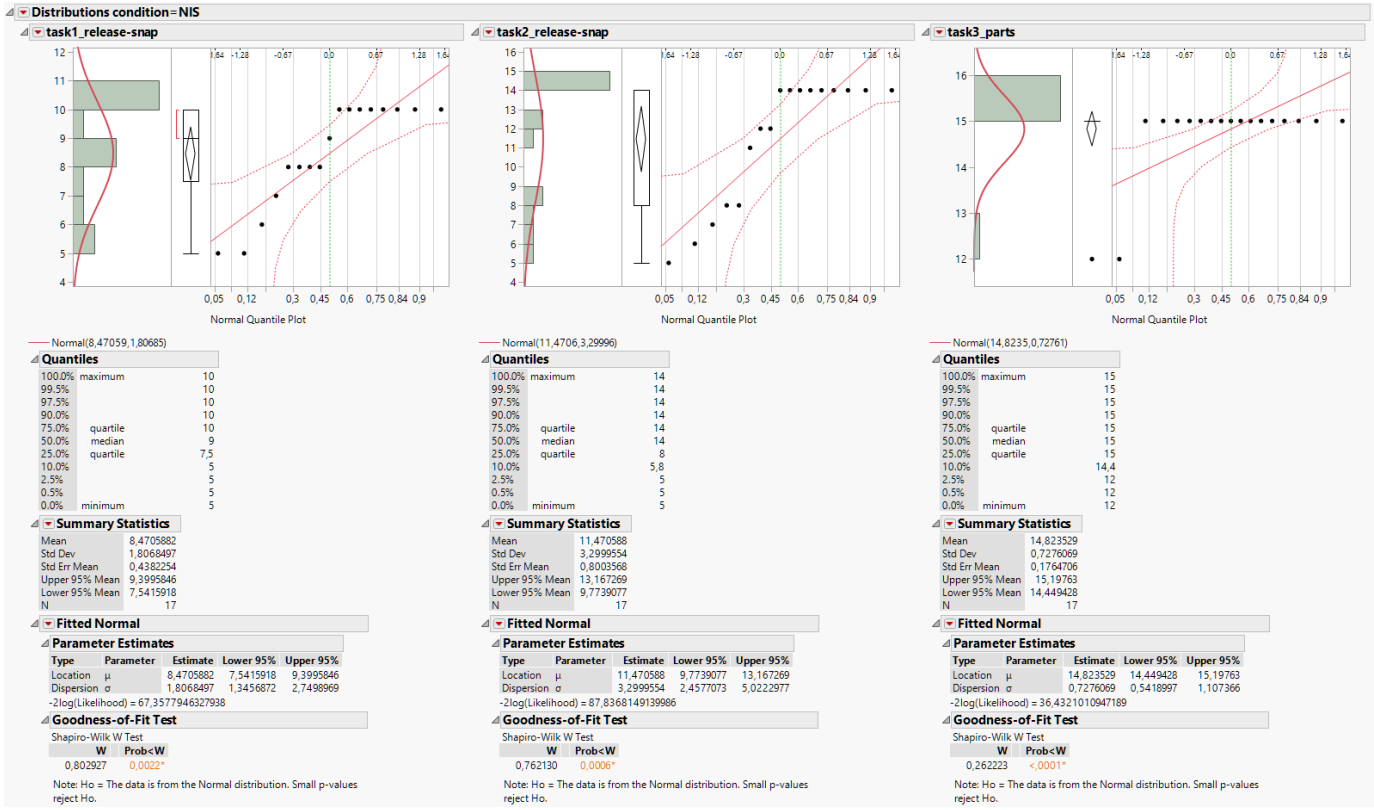


Figure 82. Shapiro-Wilk test of connections/parts for Tasks 1, 2 and 3 in NIS condition

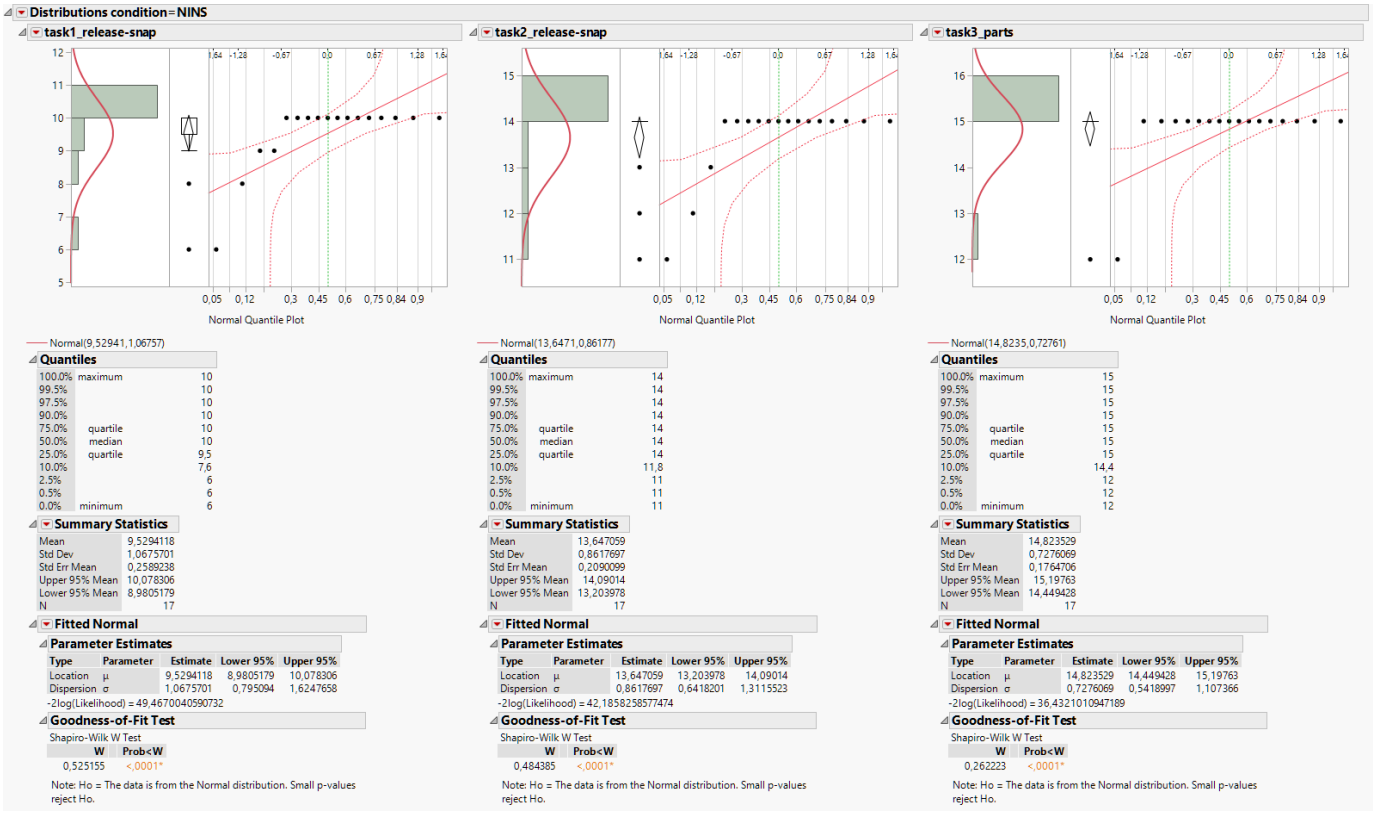


Figure 83. Shapiro-Wilk test of connections/parts for Tasks 1, 2 and 3 in NINS condition

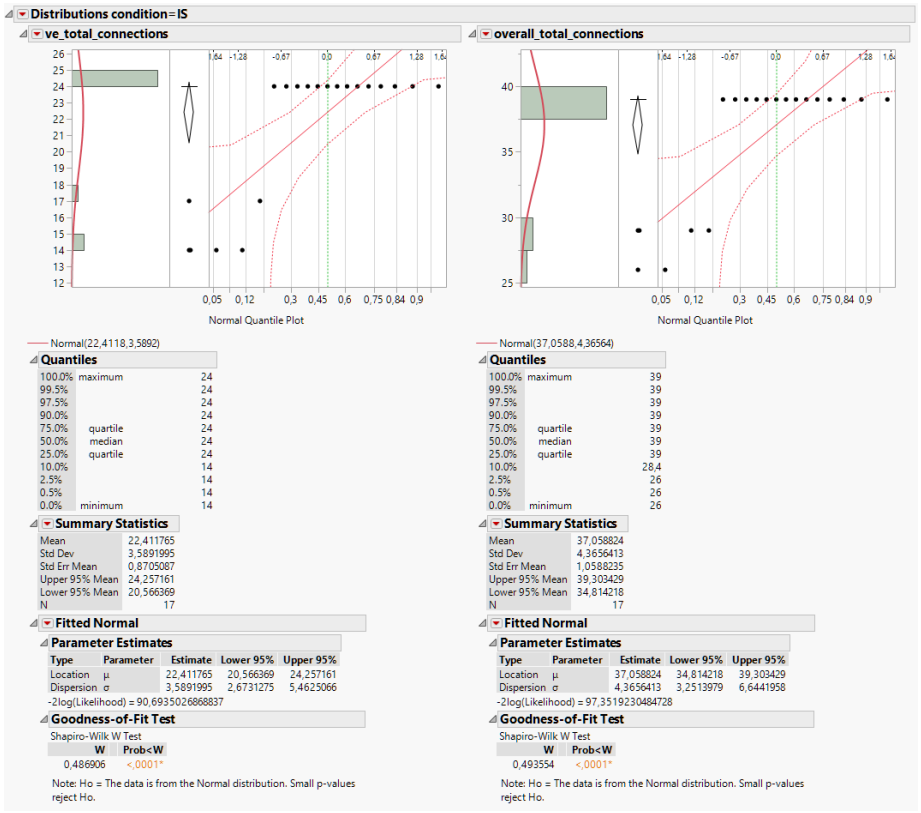


Figure 84. Shapiro-Wilk test of connections/parts for VE Total and Overall Total in IS condition

Figure 85. Shapiro-Wilk test of connections/parts for VE Total and Overall Total in INS condition

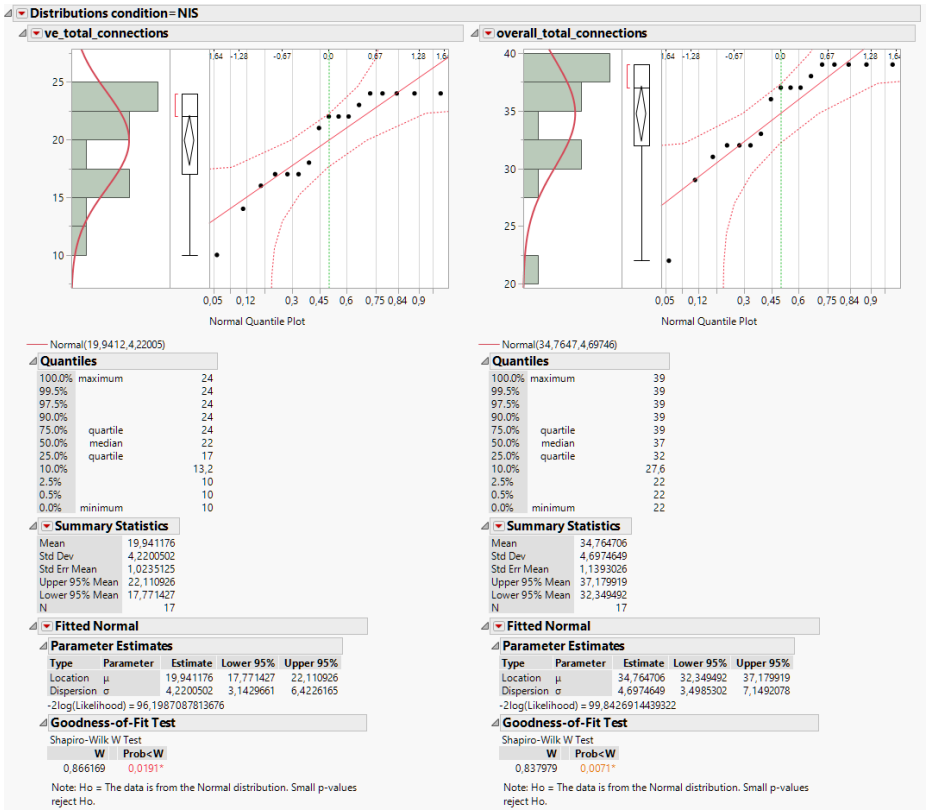


Figure 86. Shapiro-Wilk test of connections/parts for VE Total and Overall Total in NIS condition

Figure 87. Shapiro-Wilk test of connections/parts for VE Total and Overall Total in NINS condition

APPENDIX K. SHAPIRO-WILK TEST FOR COLLISIONS

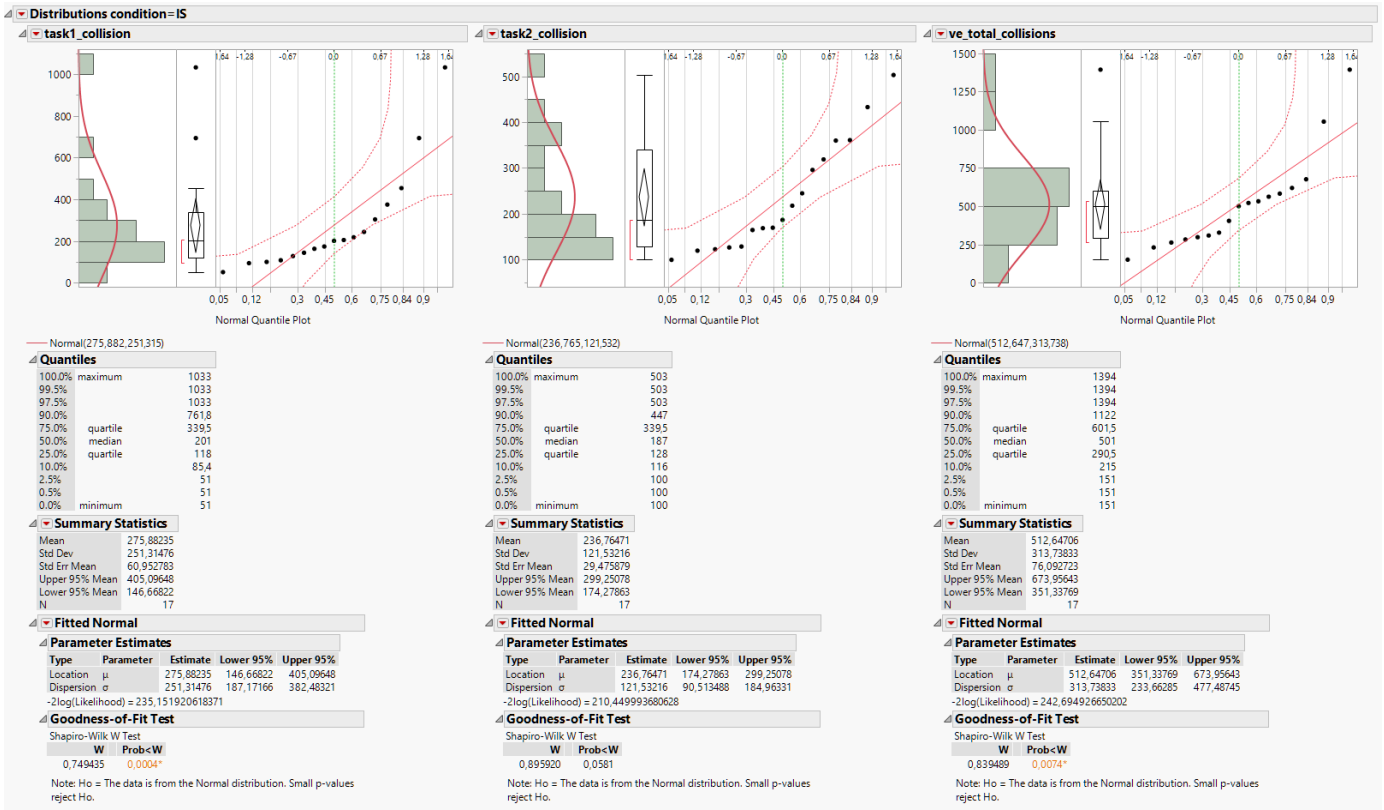


Figure 88. Shapiro-Wilk test of collisions for Tasks 1 and 2, and VE Total in IS condition

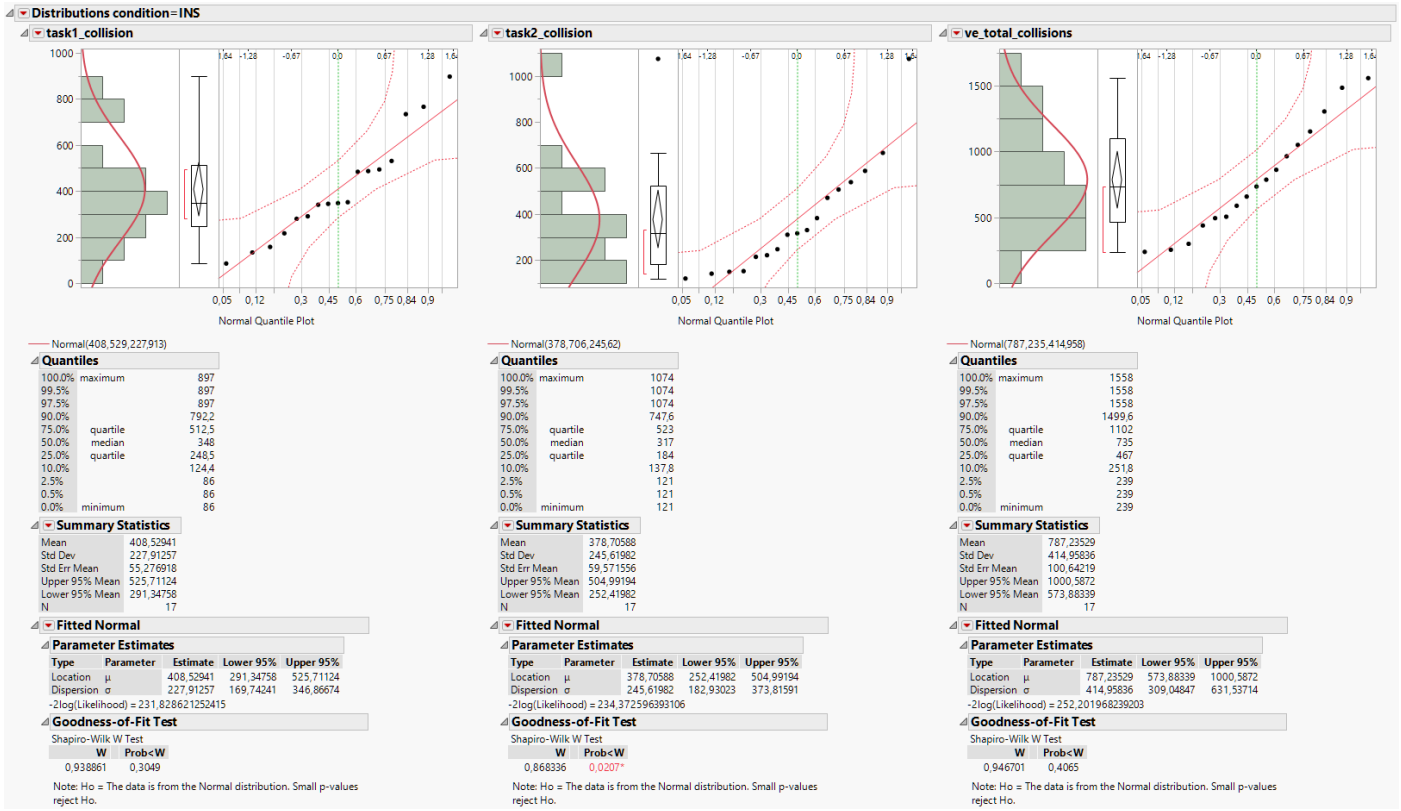


Figure 89. Shapiro-Wilk test of collisions for Tasks 1 and 2, and VE Total in INS condition

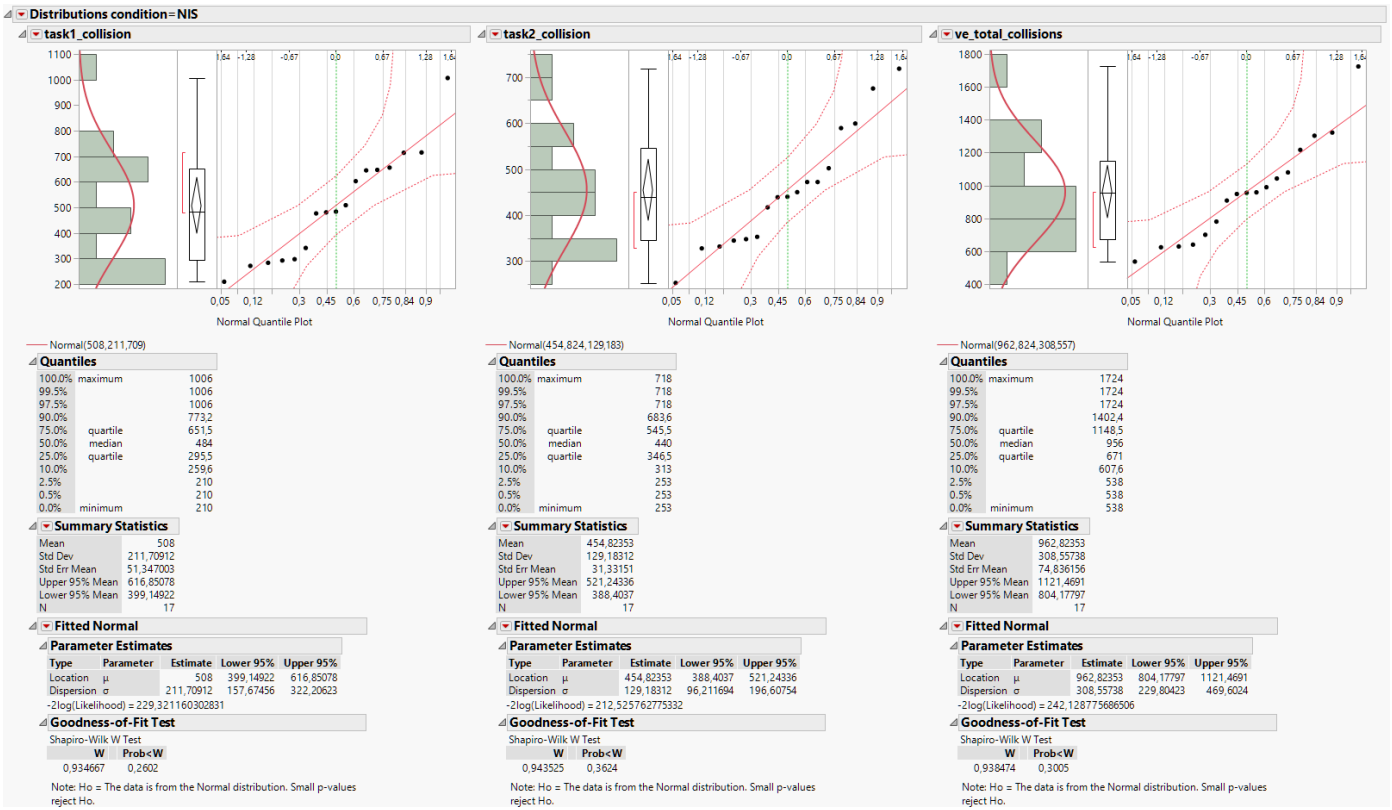


Figure 90. Shapiro-Wilk test of collisions for Tasks 1 and 2, and VE Total in NIS condition

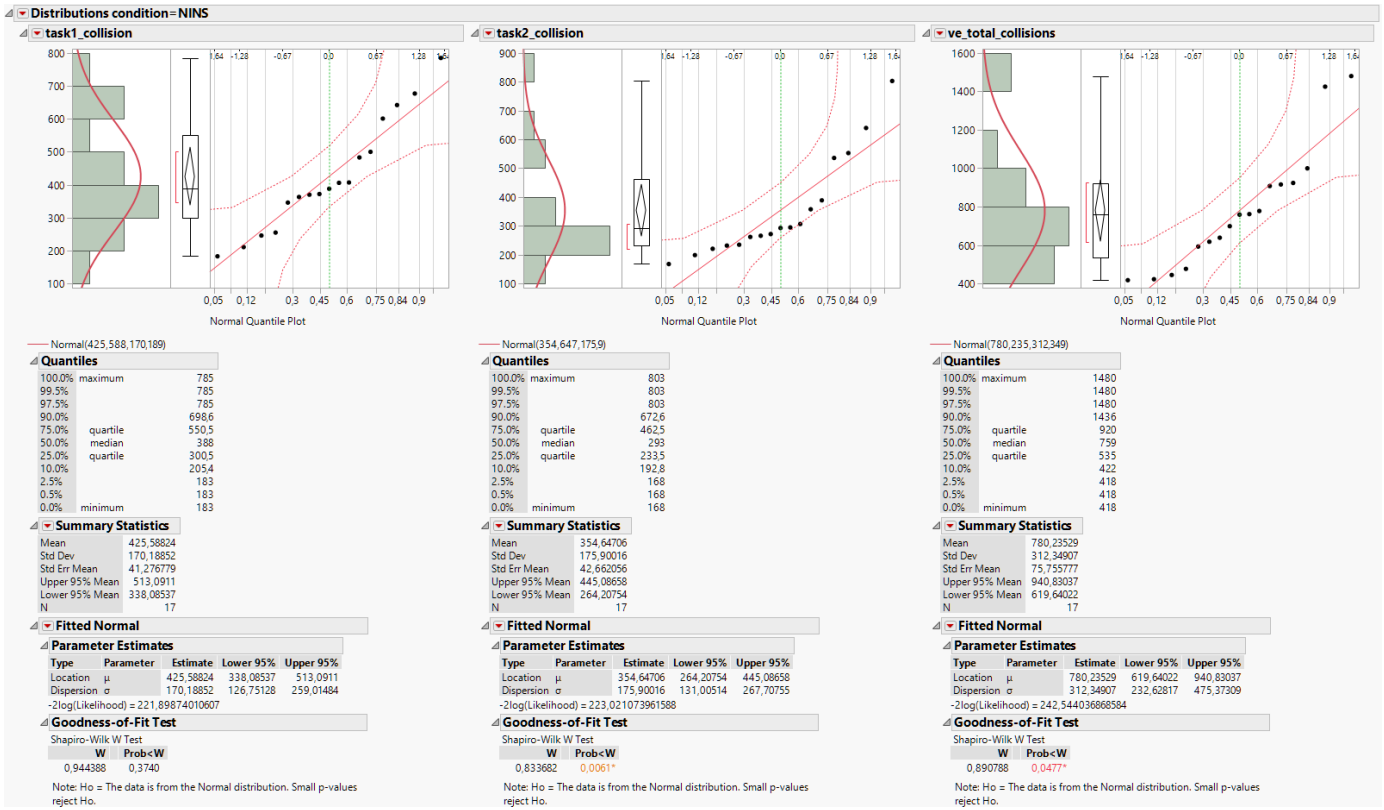


Figure 91. Shapiro-Wilk test of collisions for Tasks 1 and 2, and VE Total in NINS condition

APPENDIX L. SHAPIRO-WILK TEST FOR TASK RATINGS

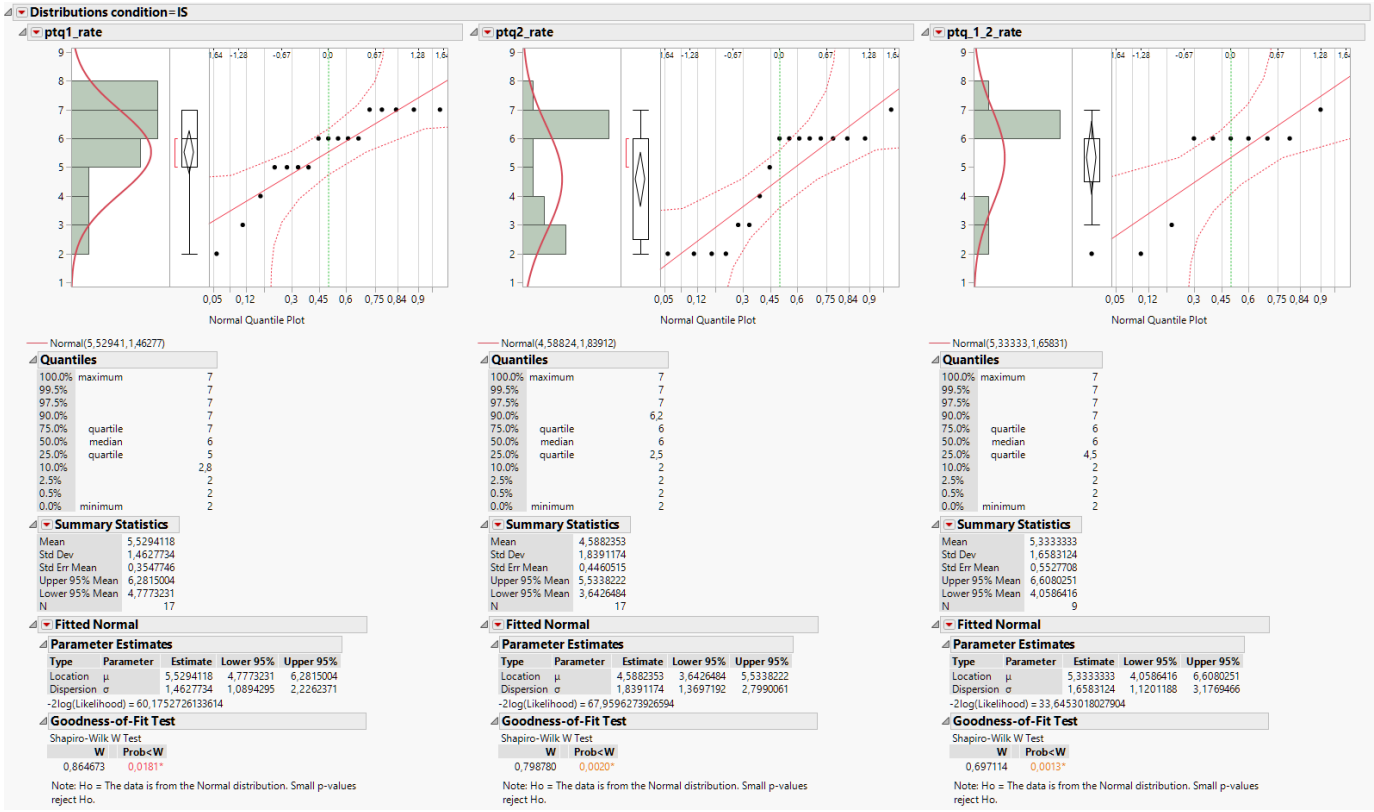


Figure 92. Shapiro-Wilk test of task ratings for Tasks 1 and 2, and VE Total in IS condition

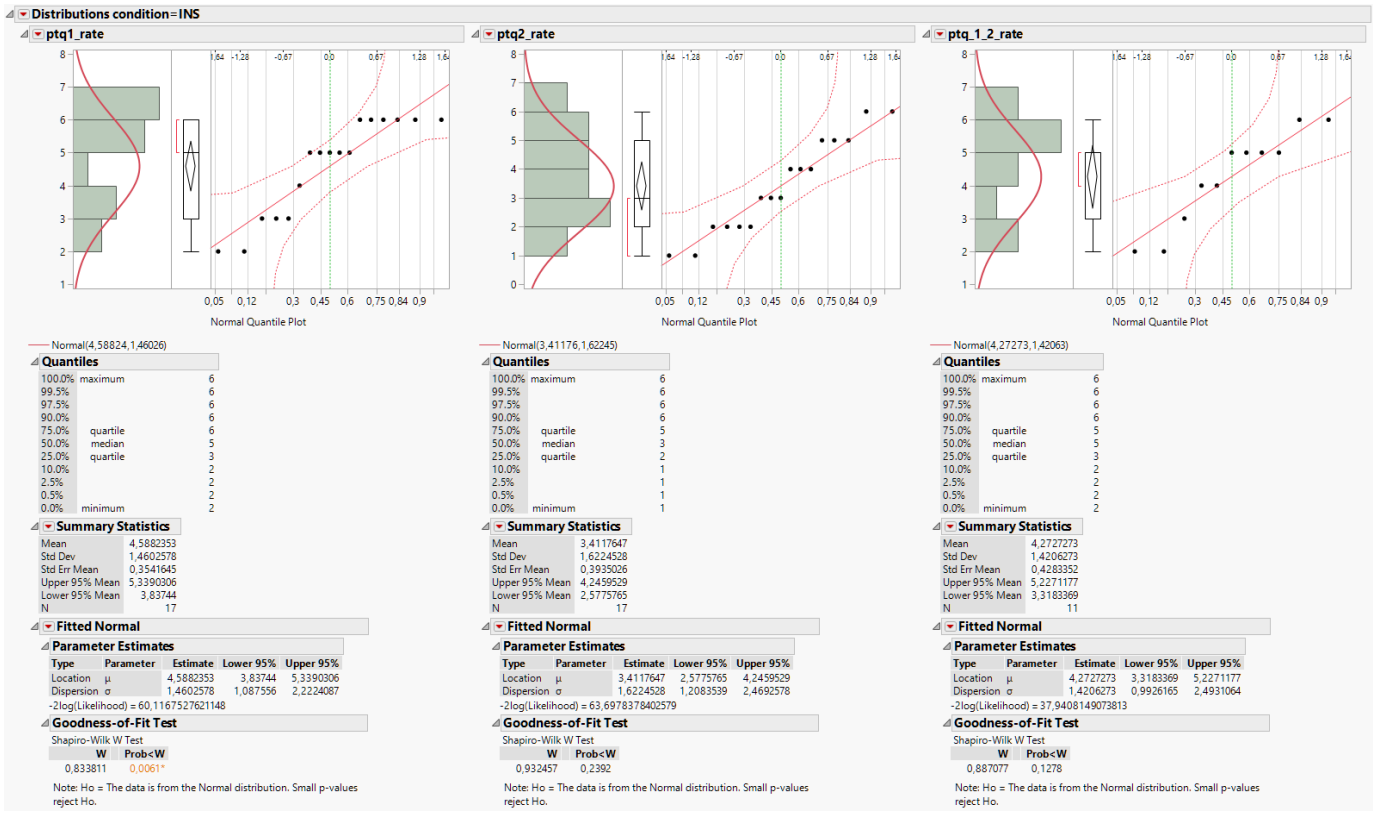


Figure 93. Shapiro-Wilk test of task ratings for Tasks 1 and 2, and VE Total in INS condition

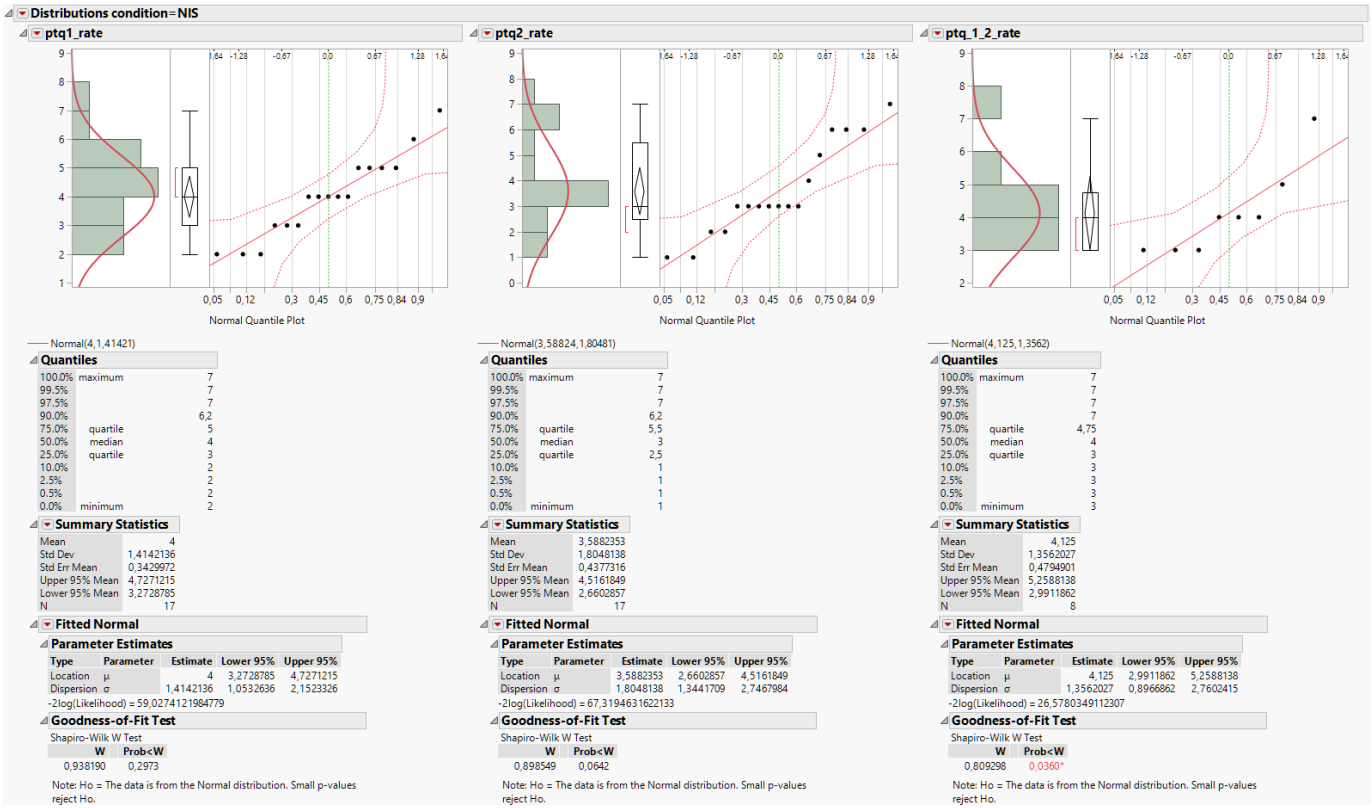


Figure 94. Shapiro-Wilk test of task ratings for Tasks 1 and 2, and VE Total in NIS condition

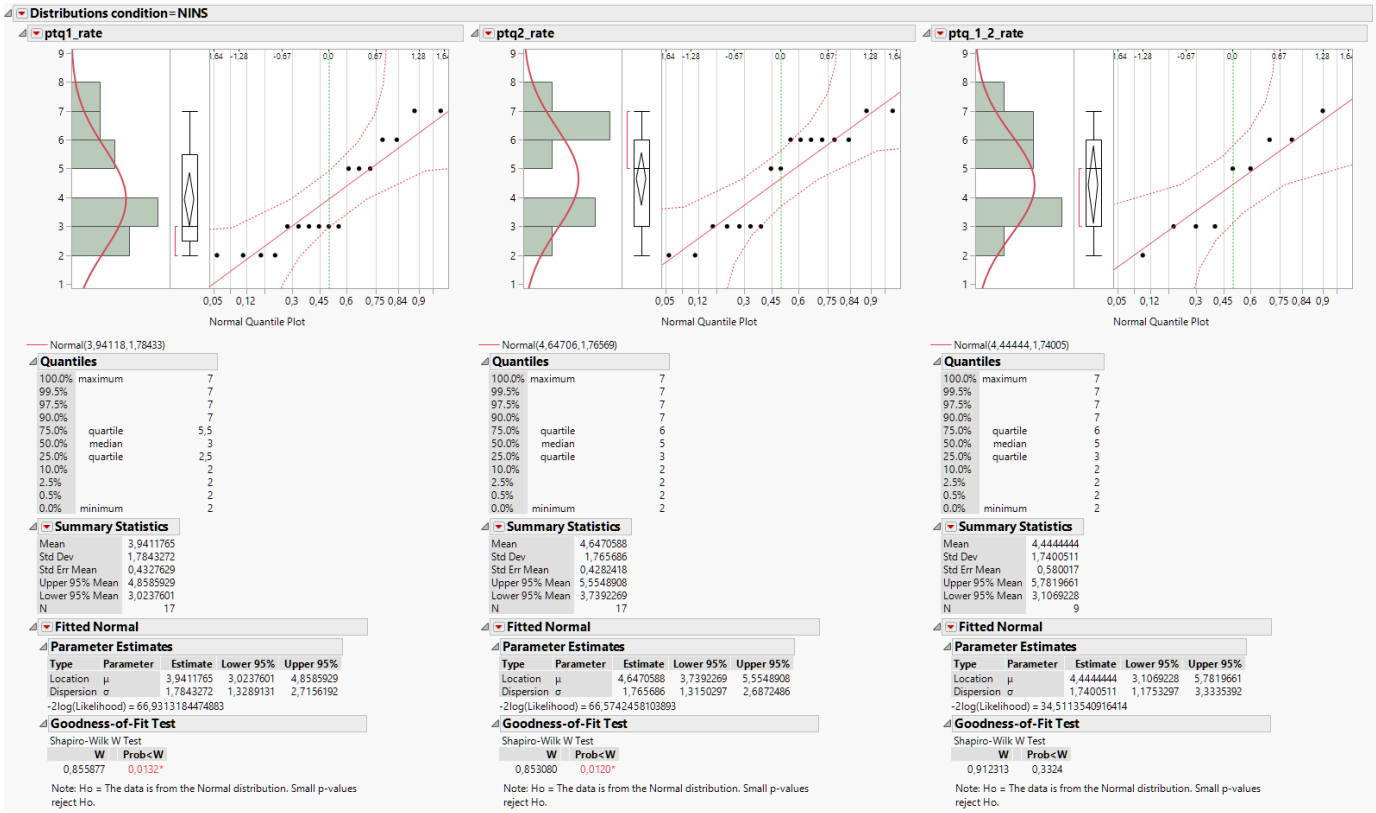


Figure 95. Shapiro-Wilk test of task ratings for Tasks 1 and 2, and VE Total in NINS condition

LIST OF REFERENCES

Bowman, D.A., and R.P. Mcmahan. 2007. "Virtual Reality: How Much Immersion Is Enough?" *Computer* 40, no. 7 (July): 36-43. https://doi.org/10.1109/MC.2007.257.

Brooke, John. 1995. *SUS: A Quick and Dirty Usability Scale*. London: Taylor and Francis.

Cheng, Xu De, Lei Lei, Jian Hu Zhang, Wei Peng, and Xin You Cui. 2011. "Research on Equipment Virtual Maintenance Technology." *Applied Mechanics and Materials; Zurich* 121–126 (October): 3546-3550. http://dx.doi.org.libproxy.nps.edu/10.4028/www.scientific.net/AMM.121-126.3546.

Costello, P J. 1997. "Health and Safety Issues Associated with Virtual Reality—A Review of Current Literature." *AGOCG Technical Report Series*.

Draper, John V., David B. Kaber, and John M. Usher. 1998. "Telepresence." *Human Factors* 40, no. 3 (September): 354–75. https://doi.org/10.1518/001872098779591386.

Duan, P., L. Pang, Y. Jin, Q. Guo, and Z. Jia. 2012. "Equipment Maintenance System Based on Virtual Reality Technology." In *2012 International Conference on Quality, Reliability, Risk, Maintenance, and Safety Engineering*: 1396–99. https://doi.org/10.1109/ICQR2MSE.2012.6246482.

Eva, Kevin. 2017. "Oculus Rift vs. HTC VIVE Specs." VRFocus. Last modified April 25, 2017. https://www.vrfocus.com/2017/04/oculus-rift-vs-htc-vive-specs/.

Facebook Technologies, LLC. 2018. "Oculus Utilities for Unity | Developer Center | Oculus." June 20, 2018. https://developer.oculus.com/downloads/package/oculus-utilities-for-unity-5/1.27.0/.

Gutiérrez, T., J. Rodríguez, Y. Vélaz, S. Casado, A. Suescun, and E. J. Sánchez. 2010. "IMA-VR: A Multimodal Virtual Training System for Skills Transfer in Industrial Maintenance and Assembly Tasks." In *19th International Symposium on Robot and Human Interactive Communication*: 428–33. https://doi.org/10.1109/ROMAN.2010.5598643.

Howard, Ian P., and Brian J. Rogers. 1996. *Binocular Vision and Stereopsis*. New York: Oxford University Press. http://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780195084764.001.0001/acprof-9780195084764.

Kennedy, Robert S., Norman E. Lane, Kevin S. Berbaum, and Michael G. Lilienthal. 1993. “Simulator Sickness Questionnaire: An Enhanced Method for Quantifying Simulator Sickness.” *International Journal of Aviation Psychology* 3, no. 3 (July): 203-220.

Kleiber, Michael, and Carsten Winkelholz. 2008. “Distortion of Depth Perception in Virtual Environments Using Stereoscopic Displays: Quantitative Assessment and Corrective Measures.” In *Proceedings of the Society of Photo-Optical Instrumentation Engineers* 6803: 68030C. https://doi.org/10.1117/12.766462.

Makerbot Industries, LLC. n.d. “Replicator+ 3D Printer - Desktop 3D Printer - Reliable 3D Printing.” Accessed August 9, 2018. https://www.makerbot.com/3d-printers/replicator/.

Murcia-Lopez, Maria, and Anthony Steed. 2018. “A Comparison of Virtual and Physical Training Transfer of Bimanual Assembly Tasks.” *IEEE Transactions on Visualization and Computer Graphics* 24, no. 4 (April): 1574-1583. https://doi.org/10.1109/TVCG.2018.2793638.

Natural Point Inc. n.d. “TrackIR.” Accessed August 9, 2018. http://www.naturalpoint.com/trackir/.

Pausch, Randy, Dennis Proffitt, and George Williams. 1997. “Quantifying Immersion in Virtual Reality.” In *Proceedings of the 24th Annual Conference on Computer Graphics and Interactive Techniques*: 13–18. https://doi.org/10.1145/258734.258744.

Ponce, Carlos R., and Richard T. Born. 2008. “Stereopsis.” *Current Biology* 18, no. 18 (September): R845–R850. https://doi.org/10.1016/j.cub.2008.07.006.

Reichelt, Stephan, Ralf Häussler, Gerald Fütterer, and Norbert Leister. 2010. “Depth Cues in Human Visual Perception and Their Realization in 3D Displays.” In *Proceedings of the Society of Photo-Optical Instrumentation Engineers* 7690: 76900B. https://doi.org/10.1117/12.850094.

Reinhart, William F., Robert J. Beaton, and Harry L. Snyder. 1990. “Comparison of Depth Cues for Relative Depth Judgments.” In *Proceedings of the Society of Photo-Optical Instrumentation Engineers* 1256: 12–22. https://doi.org/10.1117/12.19884.

Slater, Mel, Beau Lotto, Maria Marta Arnold, and Maria V Sanchez-Vives. 2009. “How We Experience Immersive Virtual Environments: The Concept of Presence and Its Measurement.” *Anuario de Psicología* 40: 193-210.

Slater, Mel, and Sylvia Wilbur. 1997. “A Framework for Immersive Virtual Environments (FIVE): Speculations on the Role of Presence in Virtual Environments.” *Presence: Teleoperators and Virtual Environments* 6, no. 6 (December): 603–16. https://doi.org/10.1162/pres.1997.6.6.603.

Unity Technologies. n.d. “Oculus Integration - Asset Store.” Accessed August 9, 2018. https://assetstore.unity.com/packages/tools/integration/oculus-integration-82022.

Wichansky, Anna M. 1991. “User Benefits of Visualization with 3-D Stereoscopic Displays.” In *Proceedings of the Society of Photo-Optical Instrumentation Engineers* 1457: 267–72. https://doi.org/10.1117/12.46322.

THIS PAGE INTENTIONALLY LEFT BLANK

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
   Ft. Belvoir, Virginia

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California