

THESIS

COMPARING MEMORABILITY OF GESTURE SETS IN AN EXTENDED REALITY
APPLICATION

Submitted by

Ethan J. Holen

Department of Computer Science

In partial fulfillment of the requirements

For the Degree of Master of Science

Colorado State University

Fort Collins, Colorado

Fall 2024

Master's Committee:

Advisor: Francisco R. Ortega

Sarath Sreedharan

Matthew Rhodes

Copyright by Ethan J. Holen 2024

All Rights Reserved

ABSTRACT

COMPARING MEMORABILITY OF GESTURE SETS IN AN EXTENDED REALITY APPLICATION

In free-form gesture sets, memorability is an important yet often under-explored metric, despite evidence that the usability of interfaces improves when designed with more memorable input gestures [1]. This study examines the memorability of three free-form gesture sets in the HoloLens 2: user-defined, elicitation-defined, and expert-defined. In addition, we examine gestures selected by the participants using common techniques from previous elicitation studies. We found that the user-defined gesture set was the most memorable, with an 88.57% recall rate. And was significantly more unforgettable than the expert-defined (72.73% recall) and the elicitation-defined (59.87% recall). This study also analyzed the user-defined gestures from this experiment. Although this was not an elicitation study, many of the methods commonly used in elicitation studies were used here. This analysis found a higher agreement rate when users were primed with a single gesture set before creating their own and a decrease in agreement when showing them two gesture sets beforehand. Given these results, we propose that designing systems with user-defined gestures will result in the most memorable sets; however, expert-defined gesture sets are also highly memorable and may better suit application design constraints.

ACKNOWLEDGEMENTS

I would like to extend my deepest appreciation and thanks to my advisor, Dr. Francisco Ortega whose advice and encouragement have been foundational to my education. I would also like to thank the rest of my committee, Dr. Matthew Rhodes and Dr. Sarath Sreedharan, for agreeing to be a part of this defense and for their advice along the way. Finally, I would like to thank my Fiancé Cassie, my parents Jeff and Molly, my brother Luke, and the rest of my family, without whose love and support this thesis would not have been possible.

DEDICATION

I would like to dedicate this thesis to my Fiancé Cassie Roth without whose constant encouragement, support, and love this work would not have been possible

I Love You

TABLE OF CONTENTS

	ABSTRACT	ii
	ACKNOWLEDGEMENTS	iii
	LIST OF FIGURES	vi
Chapter 1	Introduction	1
Chapter 2	Related Work	3
2.0.1	Memorability	4
Chapter 3	Methodology	7
3.1	Participants	7
3.2	Apparatus	7
3.3	Design	8
3.3.1	Gesture Sets	9
3.3.2	Within Subjects Study	10
3.3.3	Data Analysis	11
3.4	Procedure	16
Chapter 4	Results	19
4.1	User-Defined Gesture Comparisons	21
Chapter 5	Discussion & Future Work	25
5.1	Discussion	25
5.1.1	User-Defined Gesture Discussion	26
5.2	Limitations of the Study and Future Work	26
Chapter 6	Conclusion	28
	Bibliography	29
Appendix A	A look at the Winning gestures of used-defined gestures in this experiment and the ones by Zhou et al.	38
Appendix B	Demographic survey sent to participants after completion of the study.	43
Appendix C	Discussion of strict and relaxed participant video data analysis	49
C.1	Strict Analysis	49
C.2	Relaxed Analysis	50

LIST OF FIGURES

3.1	Picture of the experiment setup from the GoPro with one of the researchers in the subject position (closest to the camera) and experimenter (farthest from the camera) in the frame.	8
3.2	Participant view of the application during the “destroy cube” test scene.	17
4.1	Average percent of gestures correctly recalled per gesture type with standard error. . .	19
4.2	Average percent correct per gesture across gesture types	20
4.3	Average percent correct, partially correct, and incorrect per gesture type	21
4.4	Average memorability of each gesture referent inside each gesture group	22
4.5	Agreement rates for all referents graphed with confidence intervals.	24
4.6	Overall agreement rates for all groups along with separate agreement rates for each group of participants	24
5.1	Percent correct, partially correct, and incorrect organized by group and gesture type . .	26

Chapter 1

Introduction

Extended Reality (XR) is a promising medium that combines real and virtual environments to create immersive experiences, encompassing technologies such as Augmented Reality (AR), Virtual Reality (VR), and Mixed Reality (MR) [2]. XR systems, such as the Meta Quest 3 and Apple Vision Pro Headset, have continued to provide improved hand interaction options. One common interaction technique in research is gesture elicitation [3,4] to derive a gesture set where a gesture is mapped to a referent (e.g., a command or an action). For instance, a two-finger swipe from right to left may be mapped to translate a virtual object from its current position to the left. While other methods exist, such as semi-structured interviews [5,6], elicitation methodology continues to be popular.

A key question arises from these elicitation-defined sets: Are they more memorable than previously developed sets? This question was previously asked by Nacenta et al. [7] regarding author-defined versus user-defined gestures. The user-defined gestures in Nacenta et al. [7] were defined as gestures created by the user and recalled by the same user. While the author-defined gestures set was created by Nacenta and colleagues [7], these gestures complied with a set of conditions set by the authors that were outlined in the paper [7]. However, until now, no one has investigated whether participatory design gestures, such as those collected through gesture elicitation, are more or less memorable than other sets.

In this particular thesis, we compare three different gesture creation types. Firstly a set that users define for themselves during the experiment (user-defined), another using gestures from an elicitation study conducted by Zhou et al. [8] (elicitation-defined), and finally a set designed by interviewing experts in the field (expert-defined). This thesis expands on the work done by Nacenta et al. [7] by examining the memorability of all three gesture sets. The work is novel given that no one has looked at this problem to answer the question about gesture elicitation and its relationship with memorability, nor how we created the expert data set. It is also novel as this is the first

approach in this domain problem using augmented reality (AR) head-mounted displays. To also look at replication, we used the user-defined set as Nacenta et al. [7] used, and for the experiment, we used the same interface and approach described by Zhou et al. [8].

Designing gesture interfaces continues to be important, and this thesis adds to the body of science to further understand which type of gesture design (e.g., elicitation vs. expert design) may provide greater memorability. This thesis found that a well-defined expert gesture set provided greater recall.

Chapter 2

Related Work

A variety of methods for constructing gesture sets for 2D and 3D applications have been proposed [8–12] and compared over the years including gesture elicitation from users, expert design of gesture sets, allowing users to invent their own gestures, semi-structured interviews [6], and various combinations of the above methods. This study aims to explore and expand on one specific comparison point: the memorability of three specific gesture-creation methods. Usability of interfaces improves when designed with more memorable input gestures [1]. Therefore, identifying gestures that users can quickly and easily learn and retain is crucial for developing any gesture-based input system.

This work focused on comparing memorability across three separate gesture sets and is inspired by various related studies. Technical limitations in hand tracking and computer vision have restricted free-form gestural input from the mass market except in very limited scopes in some devices like the HoloLens, Meta Quest, and Apple Vision Pro. However, the attractive concept of employing a user’s hands as input for computers has proved an enticing area of research for many [9–11, 13]. Recognition also continues to advance, as evident by improvements in hand tracking in devices such as Meta Quest 3 hand tracking and the new Ultra Leap (version 2), which offer sensitive finger movement detection.

Gestural input, especially for 2D gestures, has been extensively studied in the field of HCI. In his work, Zhai et al. [14] lays out a comprehensive review of possible 2D gesture interfaces (see also [15–17]) However, to effectively use these interfaces, gestures and gesture sets must be designed and tested. In order to design gesture sets a variety of methods have been employed including Proton [18], gdt [19], and \$N-protractor [20]. In addition to these gesture design tools, other methods of gesture creation have been proposed such as elicitation mentioned later in this chapter.

Creating gesture sets alone does not necessarily guarantee user adoption or usability of the gestures. This is emphasized by gesture sets like the Graffiti gesture set, which Morris et al. found that users could learn in 5 minutes [21]. Wobbrock et al. aided in this by popularizing the concept of “guessability” for gestures [22], allowing for a method to measure each gesture individually. As mentioned later in this chapter, there have been substantial contributions to gestural input research.

While a substantial amount of research has been done in developing methods to generate and test gesture sets, far less has been done to compare these different methods, specifically to compare their memorability.

2.0.1 Memorability

A notable study addressing the memorability of gestures is the research conducted by Nacenta et al. [7] serves as a foundation for this thesis. They compared two gesture sets, namely pre-designed (gestures defined by the authors of the study) and user-defined, along with a random gesture set used as their control group.

In their study, two authors designed each gesture in their pre-designed set based on various criteria including distinguishability from other gestures in the set and specificity of the gestures as they related to the actions performed. User-defined gestures in this study were created by participants during the training phase and tested the following day.

The findings showed that user-defined gestures were significantly more memorable than both the pre-designed and randomly-defined gestures, with recall rate of 79% for user-defined gestures, 55% for pre-designed gestures, and 25% for random gestures. The study also ran a subjective questionnaire after the experiment was completed. This found that users generally enjoyed creating and recalling the user-defined gesture set more than the others and that they subjectively thought it was easier to recall later. Overall, Nacenta et al. [7] found that user-defined gestures were more memorable and enjoyable than the other gesture sets that they were compared to in their study.

Other notable research on gesture memorability includes the work by Bau et al. [23]. This paper focused on the development of a system meant to boost the memorability and ease of learning

gesture sets. A similar study by Nielsen et al. [1], aimed to create more intuitive gesture interfaces. Additionally, Jansen [24] explored different methods for teaching user-elicited gesture sets to assess the memorability of these different sets. They found that 61-71% of the gestures could be recalled, depending on the teaching style used.

Memorability of gestures has also been studied in the field of communication. Church et al. [25] found that when watching videos of people having conversations with both gestures and speech, the gestures were easier to recall. In addition, speech that was paired with a gesture was also easier for participants to recall than speech with no gestures. Gao et al. [26] explored working memory capacity or the number of gestures that participants can hold in their working memory at one time. They found that only 2-5 gesture command associations could be held in working memory at one time depending on conditions.

Further research by Kühnel et al. [27] focused in discovering the ability of users to recall smart home gestures. The gestures that the users were asked to recall in this study were elicitation gestures collected using the work by Wobbrock et al. [9]. This study also indicated in their findings that gestures that were more frequently requested by the system and less complex tended to be easier for participants to recall. However, they did not report on the percentage of gestures that the users remembered.

While this is not a gesture elicitation study like the one popularized by Wobbrock et al. [3], many of the methods used in those studies are employed in this paper. Therefore, a brief discussion of work related to elicitation studies is appropriate here. One of the most popular elicitation methodologies, as mentioned earlier, was popularized by Wobbrock et al. [3]. This work outlined a method for collecting proposed gestures from participants to create improved gesture sets.

This study has been built off of and developed in different studies such as in Dong et al. [10] in creating midair gesture sets for smart TVs as well as by Zhou et al. [8] in eliciting gesture and speech interactions in AR. Various work has also been done on priming elicitation before users invent their gestures as in the work done by Ali et al. [11]. While our study does not follow

the same method for collecting gestures or focus specifically on the collected gestures a post-hoc analysis of the gestures chosen by the participants was run.

This thesis focuses on comparing the memorability of three different gesture sets, elicitation-defined, user-defined, and expert-defined. To goal is to further our understanding of user interface adoption and usage.

Chapter 3

Methodology

In this study, we employed a Wizard of Oz (WoZ) methodology, where the experimenter manually executed certain actions while allowing participants to believe that the system was responding automatically to their inputs. In the context of our experiment, participants were shown a command to execute, such as “move left,” which is referred to as a referent. The participant then provided a gesture input corresponding to that referent. Behind the scenes, the experimenter triggered the recognition of that input. For example, when a participant performed a gesture to move a virtual object to the left, the experimenter would observe the gesture and then trigger the object’s movement by pressing a key on the keyboard.

3.1 Participants

A total of 35 participants were recruited to achieve an effect size of 0.5. Participants were gathered from a variety of sources including CSU mailing lists, NuiLab mailing lists, as well as word of mouth. The participants included 17 female and 18 male with ages ranging from 20 to 60 years (Mean=25.89, SD=6.73).

3.2 Apparatus

A Microsoft HoloLens 2 was used to replicate the environment from Zhou et al.’s study [8], which serves as the foundation for this study. Their elicitation gesture set is also used in our experiment. The Microsoft HoloLens 2 is running version 22621. The AR environment was developed using Unity Engine ¹ version 2019.f.4.F1 and the Microsoft Mixed Reality Toolkit (MRTK) ² version 2.4.0. The computer used for development was running Windows 11 Pro 64-bit, equipped

¹<https://unity.com/>

²<https://github.com/microsoft/MixedRealityToolkit-Unity>



Figure 3.1: Picture of the experiment setup from the GoPro with one of the researchers in the subject position (closest to the camera) and experimenter (farthest from the camera) in the frame.

with an Intel Core i9 9880H processor at 2.30GHz, 64GB of memory, and a 4095MB NVIDIA Quadro RTX 5000 graphics card.

During the experiment, the Microsoft HoloLens 2 was connected to a Bluetooth number pad, allowing the experimenter to trigger the different states (and their animations) of the experiment. All video and audio of the participants' gestures were recorded using a GoPro Hero7 Black camera version 1.90.

3.3 Design

The study employed a within-subjects design, meaning each participant experienced all the different conditions of the experiment. The independent variable being tested was the type of the gesture set (Elicitation-Defined (ED), User-Defined (UD), and Expert-Defined(XD)). Between each condition, participants took a five-minute break to draw. The purpose of these short breaks was to prevent participants from rehearsing the gestures they had learned. Any distracting task

Table 3.1: Latin square for determining the order of gestures each participant was presented with. Gesture sets labeled as elicitation-defined (ED) user-defined (UD) and expert-defined (XD).

Group	Block 1	Block 2	Block 3
Group 1	ED	UD	XD
Group 2	XD	ED	UD
Group 3	UD	XD	ED

could work, but it was important that the interim task wasn't too similar to the gesture-learning task to avoid interference [28]. The conditions were counter-balanced using Latin squares (Table 3.1). The dependent variable was the number of correctly recalled gestures for each gesture set (ED, UD, XD).

3.3.1 Gesture Sets

The study aimed to determine whether participants could better memorize and retain gesture sets that are defined by an expert (XD: expert-defined), an individual user (UD: user-defined), or gesture sets that are defined by a group of users in an elicitation study (ED: elicitation-defined). Each gesture set was compiled as follows:

Elicitation-Defined Gestures

The ED gestures chosen for this study were taken with permission from an elicitation study by Zhou et al. [8]. UD and XD gesture sets created for this study were also modeled around the referents defined in this paper. For each referent, the “winning gesture” as defined in the paper as the gesture with the highest agreement rate, was used as the gesture for the final elicitation gesture set. In the case of multiple winning gestures, the gesture that best fit the gesture set as a whole was selected, and this selection process was completed by the researchers.

Expert-Defined Gestures

For the XD gesture set, expert interviews were conducted to construct a gesture set based on the gestures selected by experts in the field. Interviews were conducted with four experts, during which time each referent and animation was presented to the expert who was prompted to give their

idea of the best-fitting gesture for that specific referent. After compiling a list of gestures from each participant, a single gesture was selected for each referent from the provided options. The gestures in this set were selected from the pool of possible recommended expert gestures based upon the cohesion that it had with the rest of the gesture set.

To evaluate potential experts, their backgrounds were assessed, and they were asked to provide their Google Scholar. Experts were also required to discuss their experience with gestures and interaction techniques, including details about relevant research and published papers.

User-Defined Gestures

The final gesture set for this study was the UD gesture set. This set was defined during the experiment by the participant. During each training scene described in Section 3.4, participants were asked to record a gesture they thought best fit the referent, instead of watching a training video as done for the XD and ED sets. After the gesture was recorded they were asked to repeat the gesture in the following scene which constituted their training on that gesture.

Each gesture set was unique to the particular participant. After the completion of the experiment, the gestures chosen by the participants were cataloged and analyzed using various methods discussed in section 4.

3.3.2 Within Subjects Study

As this experiment used a within-subjects design, each participant's session was divided into different blocks for training, reinforcement, and testing for each individual gesture set. In between each of these blocks, the participant was asked to draw their favorite animal for five minutes. Each block took the form of training and reinforcement of one of the gesture types during a single session, and then testing in the second block followed by the training and reinforcement of the next gesture type. This breakdown of the schedule of the blocks can be seen in Table 3.2.

The order of the gesture sets was also counterbalanced to avoid a bias towards any particular order in which the gesture sets were introduced. The counterbalancing method used was a Latin

Table 3.2: Outline of the block schedule followed when administering the experiment for group 1. Gesture sets labeled as elicitation-defined (ED) user-defined (UD) and expert-defined (XD).

Block Number	Contents
Block 1	EDtrain, EDreinforce
Drawing	5 minute drawing session
Block 2	EDtest, UDtrain, UDreinforce
Drawing	5 minute drawing session
Block 3	UDtest, XDtrain, XDreinforce
Drawing	5 minute drawing session
Block 4	XDtest

square where each participant was placed into one of three groups based on their participant ID number. These groups can be seen in Table 3.1.

3.3.3 Data Analysis

The collection of data consisted of recording the participants as they were trained and tested on the various gesture sets. This video was recorded on a GoPro beginning immediately after the participant consented to the experiment and ending after the final test was completed by the participant.

Analysis of the video data collected in this experiment consisted of the researcher watching each of the participant videos and manually verifying the correctness of each gesture. More discussion of the specific method of analysis can be found in the appendix section C. During the first pass, the testing portions of the videos were scrubbed through, and each gesture was marked with either a one or a zero, denoting whether the participant had correctly or incorrectly recalled the gesture that they were being prompted for. On the second pass through the video, the experimenter marked each gesture in the testing portion as either correct, partially correct, or incorrect allowing for a more fine-grained analysis of the performed gestures.

All gestures in both the ED and XD gesture sets were verified by reviewing the videos that were provided to the participant and assessing their interpretation of that gesture. This was determined by how the participant repeated the gesture in the reinforcement scene immediately after watching the video. This approach was necessary because some participants performed gestures that, when

Table 3.3: Referents and Their Matching Gestures for each Gesture creation type


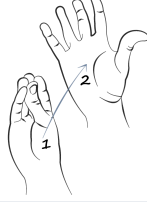


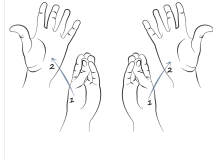
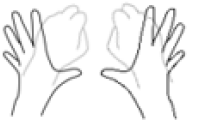

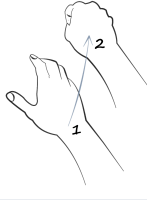
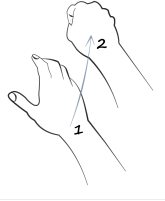
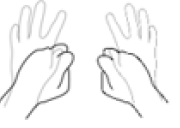
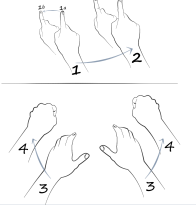
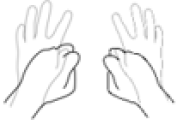




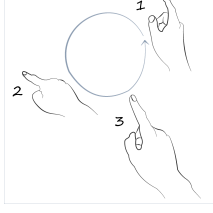


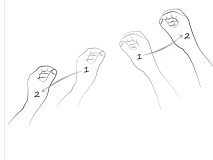

Referent	Elicitation-Defined	Expert-Defined	Most Common User-Defined
Create Cube			
Create Two Cubes			
Destroy Cube			
Destroy Both Cubes			
Select Cube			
Select Both Cubes			
Enlarge Cube			

Table 3.4: Referents and Their Matching Gestures for each Gesture creation type


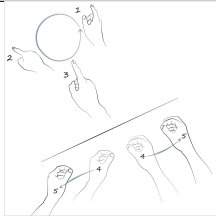


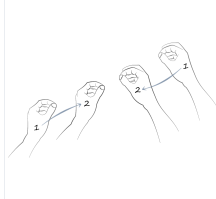

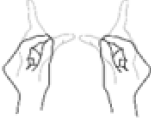
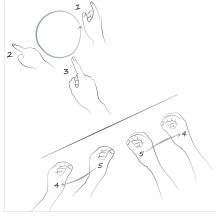
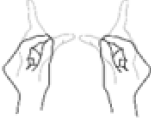

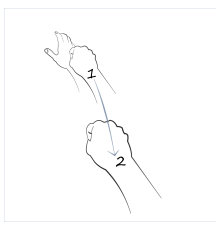


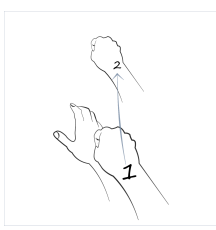




Referent	Elicitation-Defined	Expert-Defined	Most Common User-Defined
Enlarge Both Cubes			
Shrink Cube			
Shrink Both Cubes			
Move Cube Towards Yourself			
Move Cube Away			
Move Cube Up			

Table 3.5: Referents and Their Matching Gestures for each Gesture creation type


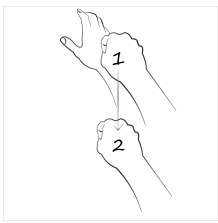


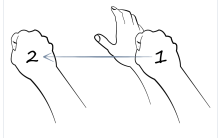
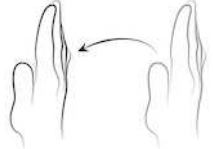

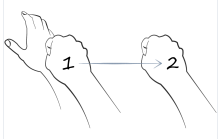


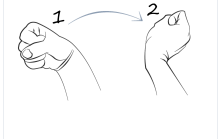

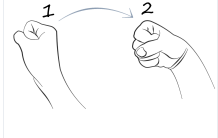
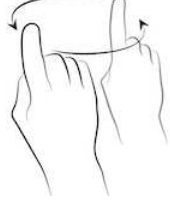

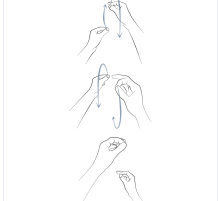


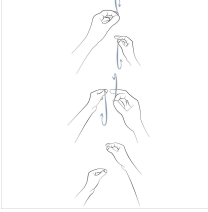


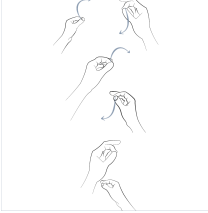


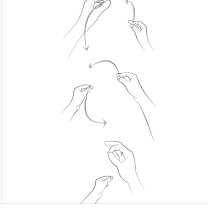

Referent	Elicitation-Defined	Expert-Defined	Most Common User-Defined
Move Cube Down			
Move Cube Left			
Move Cube Right			
Spin Cube Clockwise			
Spin Cube Counterclockwise			
Rotate Cube Up			

Table 3.6: Referents and Their Matching Gestures for each gesture creation type for each gesture creation type

Referent	Elicitation-Defined	Expert-Defined	Most Common User-Defined
Rotate Cube Down			
Roll Cube Clockwise			
Roll Cube Counterclockwise			

compared to the provided videos, could potentially be identified as incorrect. However, they were identical to how they had repeated the gesture in the earlier reinforcement scene. These gestures were marked as correct as the participant had correctly recalled the gesture as they interpreted it, even though it may have been incorrect compared to the initial tutorial video that was shown.

Finally, the user-defined gestures were verified by observing the video of the participant recording the gesture alongside the video in which they tested the gesture.

It is worth noting that during the process of verifying the correctness of gesture responses from participants in the testing, there were many cases in which a gesture repeated by the participant was somewhat ambiguous as to its correctness. In all such cases, to stay consistent in the data verification, these gestures were marked as correct as opposed to incorrect on the first pass. This ambiguity was then corrected in the second pass through the data.

3.4 Procedure

During each experiment, participants were welcomed into the room, asked to sit at a desk, and read the verbal consent form. Participants were seated at a desk looking at a blank wall to improve visibility in the HoloLens 2 headset. The experimenter was seated behind and to the left of the participant as can be seen in Figure 3.1.

After the participant provided verbal consent, the experimenter started the recording and instructed the participant to put on and adjust the HoloLens 2 headset. The Participant then completed the built-in eye adjustment tutorial for the HoloLens, after which the experimenter entered the participant ID number into the application.

Participants were then guided to the initial instruction scene, where they were asked to read aloud the written instructions. This helped the experimenter which set of gestures the participant would start with and clarified any potential confusion. Participants also had the opportunity to ask questions before the training period began. The application then guided them through a series of scenes designed to teach and reinforce one of three gesture sets.

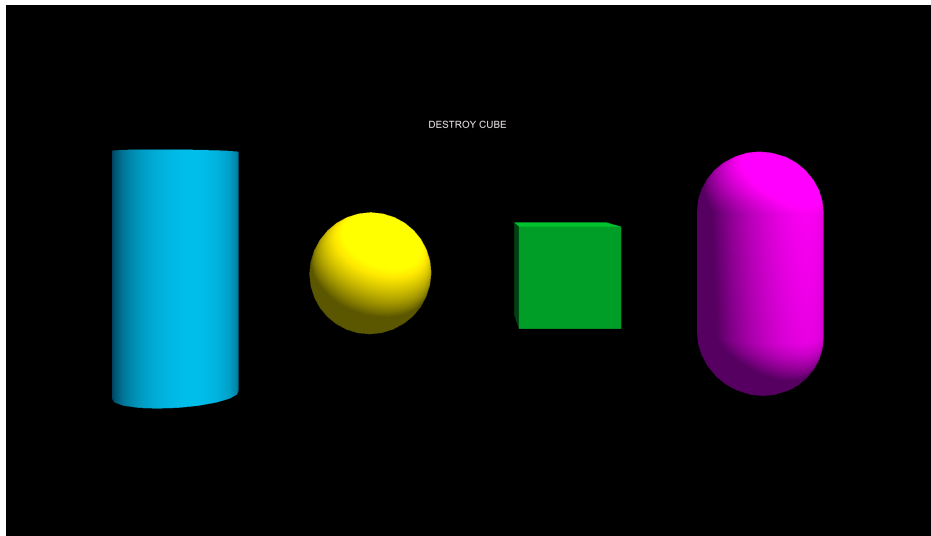


Figure 3.2: Participant view of the application during the “destroy cube” test scene.

The first scene involved a test in which participants were asked to perform gestures to manipulate a simple virtual object, such as a 3D block. These manipulations included common tasks like translation, rotation (along the x, y, z axes), creation, and destruction of objects.

The learned gesture set was then reinforced in the next scene, where the participant was presented with the same referent as the previous scene along with a cube on which they were asked to perform the gesture. Following this pattern, each gesture that the participant had learned was immediately reinforced in the following scene. This reinforcement consisted of having the participant repeat the gesture on the shape that they were interacting with.

After completing the training and reinforcement scenes, participants encountered a stop scene, which signaled them to inform the experimenter that they had reached the scene. Participants were then asked to remove their headset, place it to the side, and take drawing materials located to their left. They were instructed to draw for five minutes based on a given prompt.

An alarm signaled the end of the five minutes, and participants were asked to put the headset back on. They confirmed whether they were still on the stop scene, and the experimenter transitioned to the next instruction scene. In this scene, participants were informed that they would be tested on the gestures they had learned and were asked to read these instructions aloud. Any questions were addressed before the testing scenes began.

Testing consisted of replaying the reinforcement scenes, which included the referent (e.g., “Create Cube”) and the shape on which they were asked to perform the gesture. The “destroy cube” view of this scene can be seen in Figure 3.2.

After completing the testing, participants were automatically transitioned to the next training and reinforcement scenes for the next gesture set. This cycle repeated until all three gesture sets were trained, reinforced, and tested. This cycle is illustrated in Table 3.2 and in Table 3.1.

Chapter 4

Results

To evaluate the memorability of three gesture sets, we compared the average percentage of correct recall across the three conditions. Figure 4.1 shows the recall performance across UD, ED, and XD gesture sets. During testing, participants correctly recalled 88% (Mean=88.57, SD=11.79) of user-defined, 59% (Mean=59.87, SD=14.31) of elicitation-defined and 72% (Mean=72.73, SD=17.71) of expert-defined gestures. A detailed analysis of the memorability of each referent within each gesture group can be found in Figure 4.4. Notably, there is a 22.22% increase in memorability rate of user-defined gestures over expert-defined gestures and a 49.15% increase in memorability of user-defined gestures over elicitation-defined gestures. Additionally, expert-defined gestures had a 22.03% higher memorability rate over elicitation-defined gestures.

To examine the relationship between gesture type and the average percentage of correctly recalled gestures, a repeated measure ANOVA was conducted. The analysis revealed a significant effect of gesture type on recall performance ($F_{2,68} = 41.89, p < 0.001$). Post hoc analysis using Tukey's Test with Bonferroni Correction demonstrated a significant effect between all three conditions. Specifically, there was a large effect size between UD and ED ($d = 1.548, p < 0.001$),

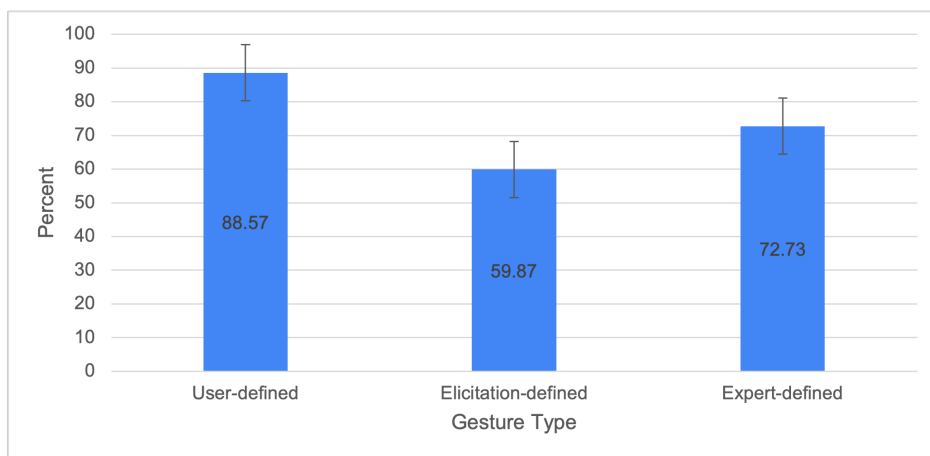


Figure 4.1: Average percent of gestures correctly recalled per gesture type with standard error.

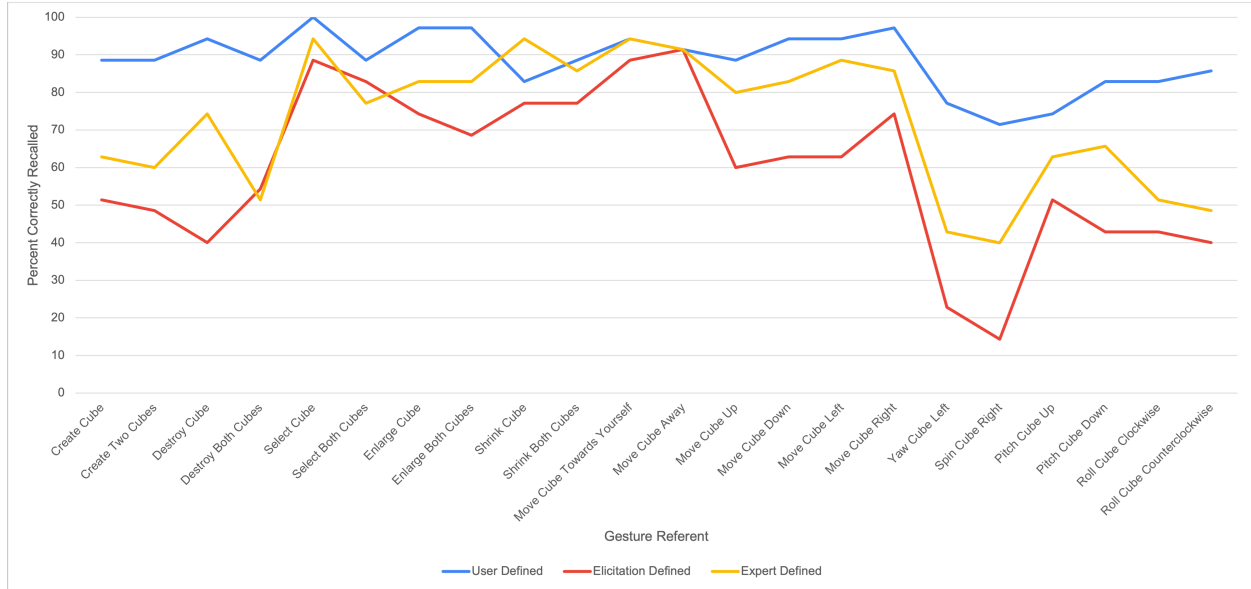


Figure 4.2: Average percent correct per gesture across gesture types

a medium effect size between UD and XD ($d = 0.745, p < 0.001$), and a medium effect size between ED and XD ($d = -0.565, p < 0.012$).

In the second pass through the testing video data, gestures were binned into correct, partially correct, and incorrect. The results of the data collected can be seen in Figure 4.3. This analysis showed that 86.49% of user-defined, 51.95% of elicitation-defined, and 67.4% of expert-defined gestures were correctly recalled during testing. This difference is because some gestures marked as correct and incorrect in the previous analysis were more accurately categorized as partially correct with the new metric.

This second pass was also analyzed using a repeated measures ANOVA test. To run the analysis, it was necessary to get the percentage of the correct gestures for each gesture type for each participant, to accomplish this, we considered all partially correct gestures as correct. Therefore, this second pass ANOVA test functions as a more relaxed analysis of the gestures with all partially correct gestures being categorized as correct. The results showed a significant effect of the conditions ($(F_{2,68} = 36.96, p < 0.001)$). Post hoc analysis using Tukey’s Test with Bonferroni Correction demonstrated a significant difference between all three conditions. A large effect was

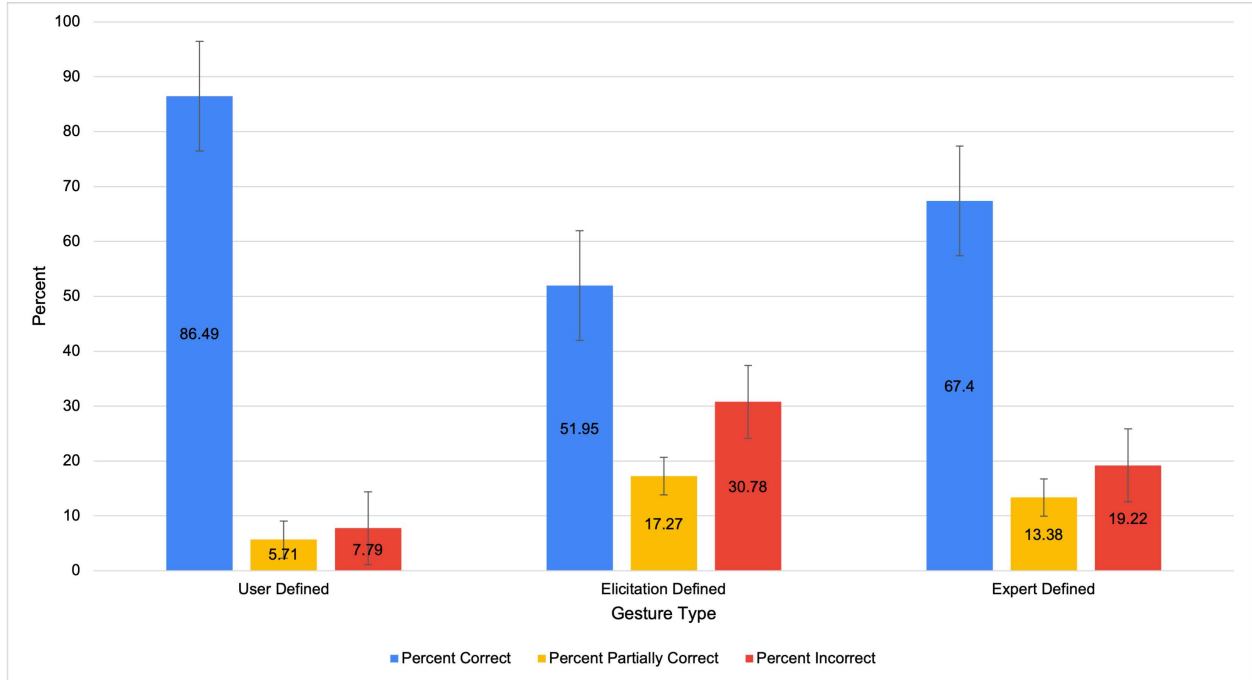


Figure 4.3: Average percent correct, partially correct, and incorrect per gesture type

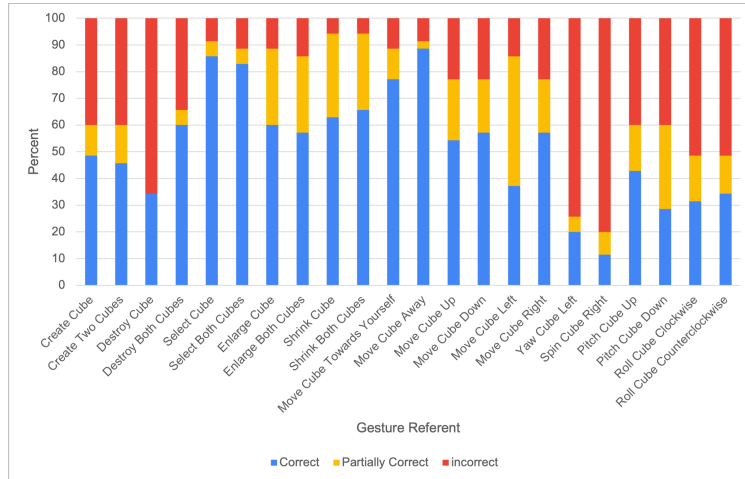
observed ($d = 1.551$) between UD and ED ($d = 1.551, p < 0.001$). A medium effect was observed between UD and XD ($d = 0.693, p < 0.002$), and between ED and XD ($d = -0.67, p < 0.002$).

The analysis revealed a significant effect of the conditions ($F_{2,68} = 36.96, p < 0.001$). Pairwise comparisons of the gesture types confirmed this, with significance observed for each comparison ($p < 0.002$). Specifically, the Comparison of the gesture type rendered the following, UD and ED had ($F_{2,68} = 22.99, p < 0.001$), UD and XD had ($F_{2,68} = 11.43, p < 0.002$), and XD and ED had ($F_{2,68} = 11.56, p < 0.002$).

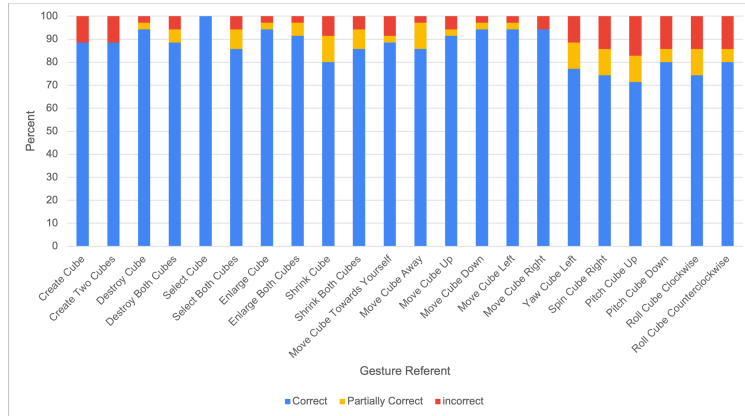
The memorability of each gesture across all gesture sets can be observed in Figure 4.2, using the correct and incorrect data. For a more detailed view of each of gesture type and the corresponding memorability of their gestures, see Figure 4.4.

4.1 User-Defined Gesture Comparisons

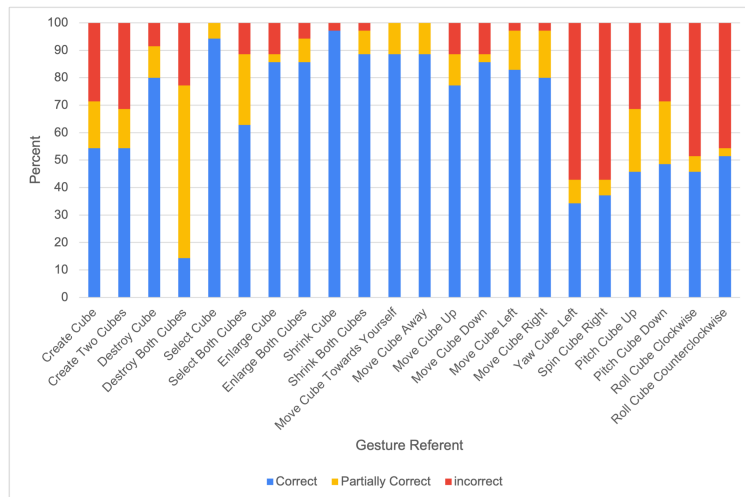
During the user-defined gesture phase of the experiment, each participant was asked to create their own gestures for twenty-two referents. These referents were displayed as text that partici-



(a) Elicitation-Defined



(b) User-Defined



(c) Expert-Defined

Figure 4.4: Average memorability of each gesture referent inside each gesture group

pants read before creating their gestures. As this approach is one of the methods for designing an elicitation study [8], all of the gestures proposed by users during the user-defined phase were collected and analyzed according to the principles of an elicitation study.

$$\mathcal{AR}(r) = \frac{\sum_{P_i \subseteq P} \binom{|P_i|}{2}}{\binom{|P|}{2}} \quad (4.1)$$

The agreement rate (\mathcal{AR}) was used to measure consensus among gestures proposed by the participants for a given referent [29], as defined in Equation 4.1. For a sample size of (35), an agreement rate above .3 is considered high, whereas an agreement rate below .1 would be regarded as low. If the agreement rate for a particular gesture is .3 or higher, then the majority of the participants selected the same gesture and vice versa. In order to calculate this metric, we used the AGATe 2.0 tool (Agreement Analysis Toolkit)³ to aid in our statistical analysis. For a more in-depth explanation of elicitation studies and their analysis, refer to Williams et al. [4].

The agreement rates for each gesture referent are shown in Figure 4.6. Overall, the agreement rate for all of the gestures that were selected by the participants was $\mathcal{AR} = .216$, indicating a medium level of consensus. The breakdown of agreement for each gesture across all participants is presented in Figure 4.5. Participants were randomly assigned to different groups, as discussed in the chapter 3, which altered the order in which the gesture sets were presented. Each group's agreement rate was analyzed, and the results are summarized in Table 4.6.

³Available at <http://depts.washington.edu/acelab/proj/dollar/agate.html>

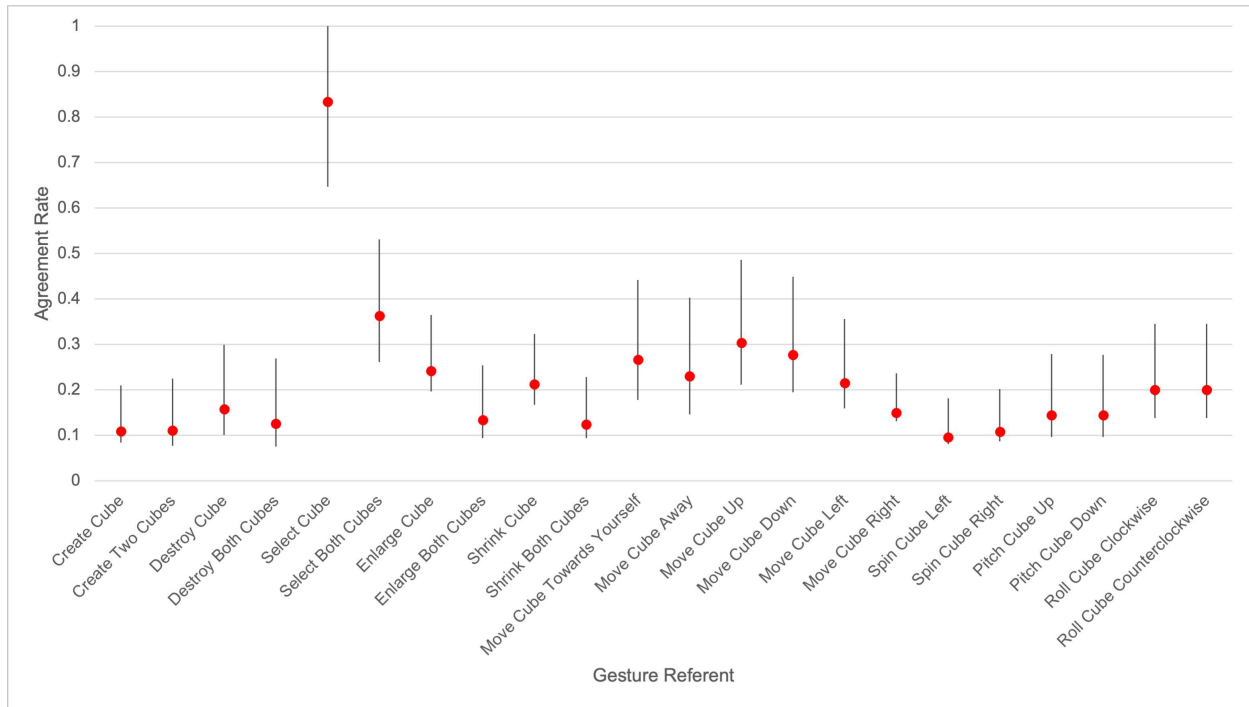


Figure 4.5: Agreement rates for all referents graphed with confidence intervals.

Referent	Create Cube	Create Two Cubes	Destroy Cube	Destroy Both Cubes	Select Cube	Select Both Cubes	Enlarge Cube	Enlarge Both Cubes	Shrink Cube	Shrink Both Cubes	Move Cube Towards Yourself	Move Cube Away	Move Cube Up	Move Cube Down	Move Cube Left	Move Cube Right	Spin Cube Left	Spin Cube Right	Pitch Cube Up	Pitch Cube Down	Roll Cube Clockwise	Roll Cube Counterclockwise	Average AR	Number of Participants
All Groups	0.109	0.111	0.158	0.126	0.834	0.363	0.242	0.134	0.213	0.124	0.267	0.23	0.304	0.277	0.215	0.15	0.096	0.108	0.145	0.145	0.2	0.2	0.216	35
Group 1 ED -> UD -> XD	0.227	0.318	0.318	0.318	0.833	0.545	0.318	0.242	0.379	0.303	0.424	0.439	0.47	0.364	0.212	0.167	0.167	0.167	0.318	0.318	0.379	0.379	0.346	12
Group 2 XD -> ED -> UD	0.197	0.197	0.273	0.136	1	0.47	0.333	0.212	0.182	0.106	0.273	0.167	0.212	0.212	0.439	0.121	0.182	0.182	0.091	0.091	0.121	0.121	0.242	12
Group 3 UD -> XD -> ED	0.2	0.2	0.091	0.055	0.655	0.4	0.145	0.055	0.145	0.055	0.109	0.091	0.236	0.236	0.2	0.091	0.055	0.091	0.109	0.109	0.4	0.4	0.188	11

Figure 4.6: Overall agreement rates for all groups along with separate agreement rates for each group of participants

Chapter 5

Discussion & Future Work

5.1 Discussion

This study found that user-defined gestures were significantly more memorable than both elicitation and expert-defined gestures. This result supports the finding in the work by Nacenta et al. [7] that user-defined gestures are more easily remembered. However, the study also revealed that gesture sets defined by experts were more memorable than anticipated compared to those collected from participants using the elicitation methodology. Expert-defined gestures were only 22.22% less memorable than user-defined gestures, whereas Nacenta et al. reported that user-defined gestures were 24% more memorable than their pre-designed gesture set.

This result suggests that although user-defined gestures are significantly more memorable than expert-defined gestures and may require slightly more training at the outset, they still have a high rate of memorability. Therefore, this study suggests that in applications where user-defined gestures are not tenable, expert-defined gestures can be implemented with only a 22.22% reduction in memorability.

As discussed in Chapter 3, participants were assigned to different groups to randomize the order in which they were presented with the three gesture sets. This order can be seen in Table 3.1. Figure 5.1 displays the percentage of gestures correctly recalled and organized into these groups. As expected, memorability is higher for the first gesture set that presented to users. For example, in group 3 (UD -> XD -> ED), the UD gesture has the highest memorability score among all groups, with 90.91% of gestures correctly recalled. Conversely, this group had the lowest score for ED, with a recall rate of 51.24%.

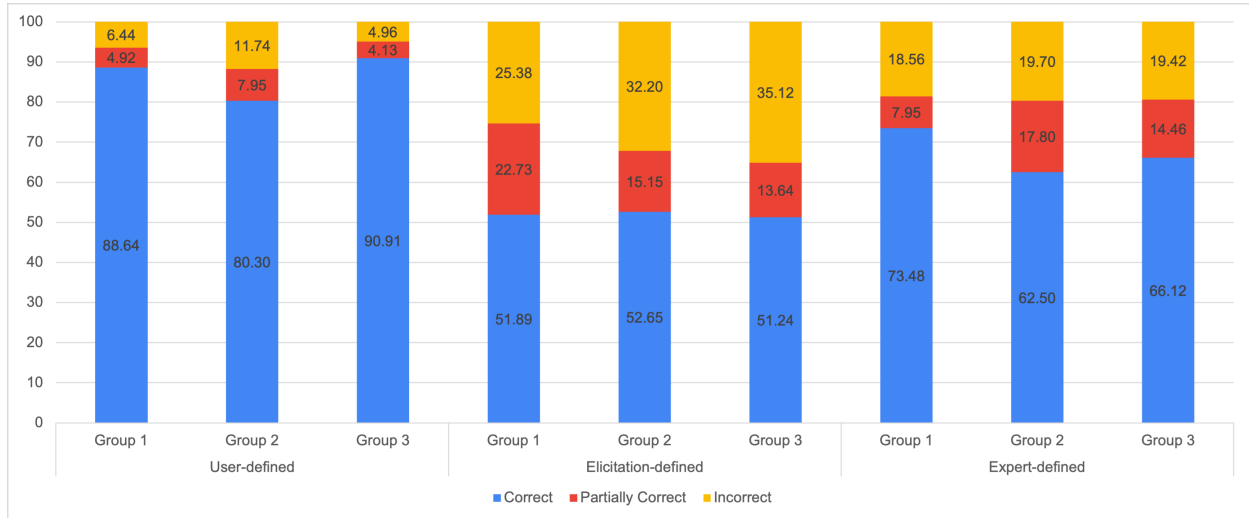


Figure 5.1: Percent correct, partially correct, and incorrect organized by group and gesture type

5.1.1 User-Defined Gesture Discussion

While the user-defined gestures were not considered to be an elicitation study, a more detailed look at the agreement rate for each referent is provided in Figure 4.5 to highlight similarities and differences with other studies. This figure shows a much higher agreement rate for the “select cube” referent $\mathcal{AR} = .834$ because most participants choose to point at the cube with their index finger, which is expected and aligned with previous studies (e.g., [8, 30]). In contrast, a gesture like “Spin Cube left” has a much lower agreement rate $\mathcal{AR} = .096$, indicating that the participants created a more varied set of gestures for the referent.

5.2 Limitations of the Study and Future Work

This experiment aimed to evaluate the memorability of three gesture sets in an augmented reality environment. The referents and UD gestures used in this study were adopted from the elicitation study by Zhou et al. [8] and intended to be a sampling of gestures that would be useful in general augmented reality tasks. While these referents and gestures succeed in their generalizability, they are not application-specific, and thus may not fully reflect daily user experience with an AR device.

One potential future study that could be conducted is comparing user-defined and expert-defined gestures in a specific AR application. This approach may benefit the experts by allowing

them to design a gesture set tailored to a particular application, rather than using the generalized gestures created for this stud.

While this experiment included elicitation-defined gestures as one of the three gesture sets, there are various methods and approaches that can potentially be used to gather an elicitation gesture set, and the one that was used in this study is only one of them. For example, Rodriguez et al. used a semi-structured interview with artists for VR sketching [5]. In the future, it could be valuable to compare the memorability of different elicitation-defined gesture sets. In addition, understanding how gestures transfer to specific application and investigating other metrics beyond memorability would be important areas for future research.

Chapter 6

Conclusion

Gestural input has the potential to become one of the primary ways that we interact with technology in the future. However, to achieve widespread adoption, it is crucial to lower the barrier to entry for these gesture sets, with an increase in memorability being a key component. This study has found that user-defined gestures have the highest percentage of recalled gestures, with an 88.5% recall rate.

Interestingly, this study also found that expert-defined gestures had a higher-than-expected average recall at 72.73%. Using this data we can recommend that when designing gesture sets for applications, developers will increase memorability and usability by allowing users to create their own gestures. However, if the application context does not allow for user-defined gestures, expert-defined gestures can be created with only a 22.22% decrease in memorability.

In conclusion, developers should prioritize user-defined gestural input wherever possible to maximize the memorability of gestural input systems. However, when application constraints prevent their use, well-developed expert-defined gestures can still achieve high levels of memorability, facilitating higher user adoption. Utilizing both gesture-creation methods can help ensure higher memorability and user acceptance of gesture-based input in future technologies.

Bibliography

- [1] Michael Nielsen, Moritz Störring, Thomas B. Moeslund, and Erik Granum. A procedure for developing intuitive and ergonomic gesture interfaces for hci. In Antonio Camurri and Gualtiero Volpe, editors, *Gesture-Based Communication in Human-Computer Interaction*, pages 409–420, Berlin, Heidelberg, 2004. Springer Berlin Heidelberg.
- [2] Satu Rantakokko. Data handling process in extended reality (xr) when delivering technical instructions. *Technical Communication*, 69(2):75–96, 2022.
- [3] Jacob O. Wobbrock, Meredith Ringel Morris, and Andrew D. Wilson. User-defined gestures for surface computing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1083–1092, New York, NY, USA, 2009. ACM.
- [4] Adam S. Williams and Francisco R. Ortega. A concise guide to elicitation methodology, May 2021.
- [5] Richard Rodriguez, Brian T. Sullivan, Mayra Donaji Barrera Machuca, Anil Ufuk Batmaz, Cyane Tornatzky, and Francisco R. Ortega. An artists’ perspectives on natural interactions for virtual reality 3d sketching. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*, CHI ’24, New York, NY, USA, 2024. Association for Computing Machinery.
- [6] Danielle Magaldi and Matthew Berler. *Semi-structured Interviews*, pages 4825–4830. Springer International Publishing, Cham, 2020.
- [7] Miguel Nacenta, Yemliha Kamber, Yizhou Qiang, and Per Ola Kristensson. Memorability of pre-designed and user-defined gesture sets. In *Proceedings of the SIGCHI Conference on human factors in computing systems*, pages 1099–1108. ACM, 2013.
- [8] Xiaoyan Zhou, Adam Sinclair Williams, and Francisco Raul Ortega. Eliciting multimodal gesture+speech interactions in a multi-object augmented reality environment. In *Proceedings*

- of the 28th ACM Symposium on Virtual Reality Software and Technology, VRST '22, New York, NY, USA, 2022. Association for Computing Machinery.*
- [9] Jacob O Wobbrock, Meredith Ringel Morris, and Andrew D Wilson. User-defined gestures for surface computing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '09*, pages 1083–1092, New York, NY, USA, 2009. ACM.
- [10] Haiwei Dong, Ali Danesh, Nadia Figueroa, and Abdulmotaleb El Saddik. An elicitation study on gesture preferences and memorability toward a practical hand-gesture vocabulary for smart televisions. *IEEE access*, 3:543–555, 2015.
- [11] Abdullah Ali, Meredith Ringel Morris, and Jacob O. Wobbrock. “i am iron man”: Priming improves the learnability and memorability of user-elicited gestures. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems, CHI '21*, New York, NY, USA, 2021. Association for Computing Machinery.
- [12] R. Kjeldsen and J. Kender. Toward the use of gesture in traditional user interfaces. In *Proceedings of the Second International Conference on Automatic Face and Gesture Recognition*, pages 151–156, Quarterly 1996.
- [13] Caroline Appert and Shumin Zhai. Using strokes as command shortcuts: cognitive benefits and toolkit support. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 2289–2298, New York, NY, USA, 2009. ACM.
- [14] Shumin Zhai, Per Ola Kristensson, Caroline Appert, Tue Haste Anderson, and Xiang Cao. Foundational issues in touch-surface stroke gesture design — an integrative review. *Foundations and Trends® in Human–Computer Interaction*, 5(2):97–205, 2012.
- [15] Santiago Villarreal-Narvaez, Jean Vanderdonckt, Radu-Daniel Vatavu, and Jacob O. Wobbrock. A systematic review of gesture elicitation studies: What can we learn from 216 studies? In *Proceedings of the 2020 ACM Designing Interactive Systems Conference, DIS '20*, page 855–872, New York, NY, USA, 2020. Association for Computing Machinery.

- [16] Ruben Balcazar, Francisco R. Ortega, Katherine Tarre, Armando Barreto, Mark Weiss, and Naphtali D. Rishe. Circgr: Interactive multi-touch gesture recognition using circular measurements. In *Proceedings of the 2017 ACM International Conference on Interactive Surfaces and Spaces*, ISS '17, page 12–21, New York, NY, USA, 2017. Association for Computing Machinery.
- [17] Jacob O Wobbrock, Andrew D Wilson, and Yang Li. \$1 unistroke recognizer.
- [18] Kenrick Kin, Björn Hartmann, Tony DeRose, and Maneesh Agrawala. Proton: multitouch gestures as regular expressions. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '12, page 2885–2894, New York, NY, USA, 2012. Association for Computing Machinery.
- [19] Allan Christian Long, James A. Landay, and Lawrence A. Rowe. Implications for a gesture design tool. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '99, page 40–47, New York, NY, USA, 1999. Association for Computing Machinery.
- [20] Lisa Anthony and Jacob O Wobbrock. \$N-protractor: a fast and accurate multistroke recognizer. 2012.
- [21] Meredith Ringel Morris, Anqi Huang, Andreas Paepcke, and Terry Winograd. Cooperative gestures: multi-user gestural interactions for co-located groupware. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '06, page 1201–1210, New York, NY, USA, 2006. Association for Computing Machinery.
- [22] Jacob O. Wobbrock, Htet Htet Aung, Brandon Rothrock, and Brad A. Myers. Maximizing the guessability of symbolic input. In *CHI '05 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '05, page 1869–1872, New York, NY, USA, 2005. Association for Computing Machinery.

- [23] Olivier Bau and Wendy E. Mackay. Octopocus: a dynamic guide for learning gesture-based command sets. In *Proceedings of the 21st annual ACM symposium on User interface software and technology*, pages 37–46, New York, NY, USA, 2008. ACM.
- [24] Eline Jansen, Ir Raymond Cuijpers, Wijnand IJsselsteijn, Ir Ing Maurice van Beurden, Ir John de Vet, Philips Consumer Lifestyle, and BG TV Site Eindhoven. Teaching users how to interact with gesture-based interfaces.
- [25] Ruth Breckinridge Church, Philip Garber, and Kathryn Rogalski. The role of gesture in memory and social communication. *Gesture*, 7(2):137–158, 2007.
- [26] Quan Gu Qi Gao, Zheng Ma, Jiaofeng Li, and Zaifeng Gao. Working memory capacity for gesture-command associations in gestural interaction. *International Journal of Human-Computer Interaction*, 39(15):3045–3056, 2023.
- [27] Christine Kühnel, Tilo Westermann, Fabian Hemmert, Sven Kratz, Alexander Müller, and Sebastian Möller. I’m home: Defining and evaluating a gesture set for smart-home control. *International Journal of Human-Computer Studies*, 69(11):693–704, 2011.
- [28] David G Payne, Jeffrey S Anastasi, Jason M Blackwell, and Michael J Wenger. Selective disruption of hypermnesia for pictures and words. *Memory & cognition*, 22:542–551, 1994.
- [29] Radu-Daniel Vatavu and Jacob O. Wobbrock. Formalizing agreement analysis for elicitation studies: New measures, significance test, and toolkit. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, CHI ’15*, page 1325–1334, New York, NY, USA, 2015. Association for Computing Machinery.
- [30] Adam S. Williams and Francisco R. Ortega. Understanding gesture and speech multimodal interactions for manipulation tasks in augmented reality using unconstrained elicitation. *Proc. ACM Hum.-Comput. Interact.*, 4(ISS), nov 2020.

- [31] Haiwei Dong, Ali Danesh, Nadia Figueroa, and Abdulmotaleb El Saddik. An elicitation study on gesture preferences and memorability toward a practical hand-gesture vocabulary for smart televisions. *IEEE Access*, 3:543–555, 2015.
- [32] Bruno Fruchard, Eric Lecolinet, and Olivier Chapuis. How memorizing positions or directions affects gesture learning? In *Proceedings of the 2018 ACM International Conference on Interactive Surfaces and Spaces*, ISS '18, page 107–114, New York, NY, USA, 2018. Association for Computing Machinery.
- [33] A. Chris Long, James A. Landay, Lawrence A. Rowe, and Joseph Michiels. Visual similarity of pen gestures. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '00, page 360–367, New York, NY, USA, 2000. Association for Computing Machinery.
- [34] Uran Oh and Leah Findlater. The challenges and potential of end-user gesture customization. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, page 1129–1138, New York, NY, USA, 2013. Association for Computing Machinery.
- [35] Ovidiu-Andrei Schipor and Radu-Daniel Vatavu. Invisible, inaudible, and impalpable: Users' preferences and memory performance for digital content in thin air. *IEEE Pervasive Computing*, 17(4):76–85, 2018.
- [36] Ana Villanueva, Sujin Jang, Wolfgang Stuerzlinger, Satyajit Ambike, and Karthik Ramani. Advanced modeling method for quantifying cumulative subjective fatigue in mid-air interaction. *International Journal of Human-Computer Studies*, 169:102931, 2023.
- [37] Julie Wagner, Eric Lecolinet, and Ted Selker. Multi-finger chords for hand-held tablets: recognizable and memorable. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '14, page 2883–2892, New York, NY, USA, 2014. Association for Computing Machinery.

- [38] Adam S. Williams and Francisco R. Ortega. Understanding gesture and speech multimodal interactions for manipulation tasks in augmented reality using unconstrained elicitation. *Proc. ACM Hum.-Comput. Interact.*, 4(ISS), nov 2020.
- [39] Yulong Yang, Gradeigh D. Clark, Janne Lindqvist, and Antti Oulasvirta. Free-form gesture authentication in the wild. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI '16, page 3722–3735, New York, NY, USA, 2016. Association for Computing Machinery.
- [40] Per Ola Kristensson and Shumin Zhai. Command strokes with and without preview: using pen gestures on keyboard for command selection. In *Conference on Human Factors in Computing Systems: Proceedings of the SIGCHI conference on Human factors in computing systems; 28 Apr.-03 May 2007*, pages 1137–1146, New York, NY, USA, 2007. ACM.
- [41] Uran Oh and Leah Findlater. The challenges and potential of end-user gesture customization. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, page 1129–1138, New York, NY, USA, 2013. Association for Computing Machinery.
- [42] Thomas Baudel and Michel Beaudouin-Lafon. Charade: remote control of objects using free-hand gestures. *Commun. ACM*, 36(7):28–35, jul 1993.
- [43] Andrew Bragdon, Robert Zeleznik, Brian Williamson, Timothy Miller, and Joseph J. LaViola. Gesturebar: improving the approachability of gesture-based interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 2269–2278, New York, NY, USA, 2009. ACM.
- [44] Andy Cockburn, Per Ola Kristensson, Jason Alexander, and Shumin Zhai. Hard lessons: effort-inducing interfaces benefit spatial learning. In *Conference on Human Factors in Computing Systems: Proceedings of the SIGCHI conference on Human factors in computing systems; 28 Apr.-03 May 2007*, pages 1571–1580, New York, NY, USA, 2007. ACM.

- [45] Aurélie Cohé and Martin Hachet. Beyond the mouse: Understanding user gestures for manipulating 3d objects from touchscreen inputs. *Computers & Graphics*, 36(8):1119–1131, 2012. Graphics Interaction Virtual Environments and Applications 2012.
- [46] Dustin Freeman, Hrvoje Benko, Meredith Ringel Morris, and Daniel Wigdor. Shadowguides: visualizations for in-situ learning of multi-touch and whole-hand gestures. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces*, ITS '09, page 165–172, New York, NY, USA, 2009. Association for Computing Machinery.
- [47] Mathias Frisch, Jens Heydekorn, and Raimund Dachsel. Investigating multi-touch and pen gestures for diagram editing on interactive surfaces. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces*, ITS '09, page 149–156, New York, NY, USA, 2009. Association for Computing Machinery.
- [48] David Goldberg and Cate Richardson. Touch-typing with a stylus. In *Proceedings of the INTERACT'93 and CHI'93 conference on Human factors in computing systems*, pages 80–87, 1993.
- [49] Hao Lü and Yang Li. Gesture coder: a tool for programming multi-touch gestures by demonstration. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '12, page 2875–2884, New York, NY, USA, 2012. Association for Computing Machinery.
- [50] Christian Kray, Daniel Nesbitt, John Dawson, and Michael Rohs. User-defined gestures for connecting mobile phones, public displays, and tabletops. In *Proceedings of the 12th International Conference on Human Computer Interaction with Mobile Devices and Services*, MobileHCI '10, page 239–248, New York, NY, USA, 2010. Association for Computing Machinery.
- [51] Ekaterina Kurdyukova, Matthias Redlin, and Elisabeth André. Studying user-defined ipad gestures for interaction in multi-display environment. In *Proceedings of the 2012 ACM In-*

- ternational Conference on Intelligent User Interfaces*, IUI '12, page 93–96, New York, NY, USA, 2012. Association for Computing Machinery.
- [52] I Scott MacKenzie and Shawn X Zhang. The immediate usability of graffiti. In *Proceedings of Graphics Interface'97*, 1997.
- [53] Meredith Ringel Morris, Jacob O Wobbrock, and Andrew D Wilson. Understanding users' preferences for surface gestures. In *Proceedings of graphics interface 2010*, pages 261–268. 2010.
- [54] Tao Ni, Doug Bowman, and Chris North. Airstroke: bringing unistroke text entry to free-hand gesture interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, page 2473–2476, New York, NY, USA, 2011. Association for Computing Machinery.
- [55] Chris North, Tim Dwyer, Bongshin Lee, Danyel Fisher, Petra Isenberg, George Robertson, and Kori Inkpen. Understanding multi-touch manipulation for surface computing. In Tom Gross, Jan Gulliksen, Paula Kotzé, Lars Oestreicher, Philippe Palanque, Raquel Oliveira Prates, and Marco Winckler, editors, *Human-Computer Interaction – INTERACT 2009*, pages 236–249, Berlin, Heidelberg, 2009. Springer Berlin Heidelberg.
- [56] Dean Rubine. Specifying gestures by example. *ACM SIGGRAPH computer graphics*, 25(4):329–337, 1991.
- [57] Sebastian Schmidt, Miguel A. Nacenta, Raimund Dachselt, and Sheelagh Carpendale. A set of multi-touch graph interaction techniques. In *ACM International Conference on Interactive Tabletops and Surfaces*, ITS '10, page 113–116, New York, NY, USA, 2010. Association for Computing Machinery.
- [58] Rajinder Sodhi, Hrvoje Benko, and Andrew Wilson. Lightguide: projected visualizations for hand movement guidance. In *Proceedings of the SIGCHI Conference on Human Factors in*

Computing Systems, CHI '12, page 179–188, New York, NY, USA, 2012. Association for Computing Machinery.

- [59] Radu-Daniel Vatavu. User-defined gestures for free-hand tv control. In *Proceedings of the 10th European Conference on Interactive TV and Video*, EuroITV '12, page 45–48, New York, NY, USA, 2012. Association for Computing Machinery.
- [60] Jacob O. Wobbrock, Meredith Ringel Morris, and Andrew D. Wilson. User-defined gestures for surface computing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '09, page 1083–1092, New York, NY, USA, 2009. Association for Computing Machinery.
- [61] Jacob O. Wobbrock, Andrew D. Wilson, and Yang Li. Gestures without libraries, toolkits or training: a \$1 recognizer for user interface prototypes. In *Proceedings of the 20th Annual ACM Symposium on User Interface Software and Technology*, UIST '07, page 159–168, New York, NY, USA, 2007. Association for Computing Machinery.
- [62] M. Wu, Chia Shen, K. Ryall, C. Forlines, and R. Balakrishnan. Gesture registration, relaxation, and reuse for multi-point direct-touch surfaces. In *First IEEE International Workshop on Horizontal Interactive Human-Computer Systems (TABLETOP '06)*, pages 8 pp.–, 2006.
- [63] Shumin Zhai and Per-Ola Kristensson. Shorthand writing on stylus keyboard. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '03, page 97–104, New York, NY, USA, 2003. Association for Computing Machinery.

Appendix A

A look at the Winning gestures of used-defined gestures in this experiment and the ones by Zhou et al.

Table A.1 shows the winning gestures of both studies. The one on the left is the winning used-defined gestures for this thesis study, and the column on the right is the one by Zhou et al. [8]. While comparison among elicitation studies are not valid (see reasons in [4]), this provides a look at how the gestures do repeat in many cases when using a replication of the elicitation study. It is also important to note that the used-defined gestures cannot be considered an elicitation study as the objective and methodology differ from a traditional elicitation technique.

Table A.1: Referents and Their Winning Gestures between this study used-defined, and Zhou et al. [8] study



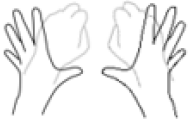
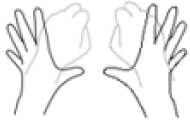
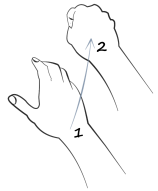

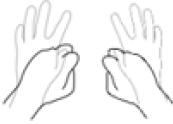
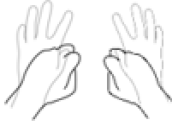






Referent	Most Popular User-Defined Current Study	Most Popular Elicited Zhou et al.
Create Cube		
Create Two Cubes		
Destroy Cube		
Destroy Both Cubes		
Select Cube		
Select Both Cubes		
Enlarge Cube		

Table A.2: Referents and Their Winning Gestures between this study used-defined, and Zhou et al. [8] study





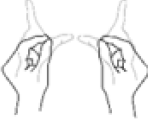
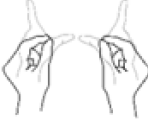








Referent	Winning Gesture User-Defined Current Study	Winning Gesture Elicited Zhou et al.
Enlarge Both Cubes		
Shrink Cube		
Shrink Both Cubes		
Move Cube Towards Yourself		
Move Cube Away		
Move Cube Up		
Move Cube Down		

Table A.3: Referents and Their Winning Gestures between this study used-defined, and Zhou et al. [8] study

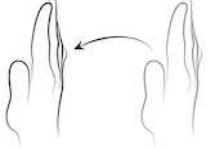



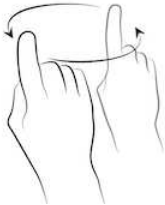

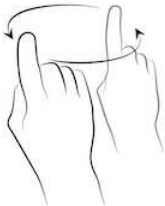









Referent	Most Popular User-Defined Current Study	Most Popular Elicited Zhou et al.
Move Cube Left		
Move Cube Right		
Spin Cube Clockwise		
Spin Cube Counterclockwise		
Rotate Cube Up		
Rotate Cube Down		
Roll Cube Clockwise		

Table A.4: Referents and Their Winning Gestures between this study used-defined, and Zhou et al. [8] study

Referent	Most Popular User-Defined Current Study	Most Popular Elicited Zhou et al.
Roll Cube Counterclockwise		

Appendix B

Demographic survey sent to participants after completion of the study.

Demographic Questionnaire

Thank you for participating in this study. If you have any question, please contact the PI at fortega@colostate.edu.

* Indicates required question

1. Token ID *

Token IDs are positive integers ranging from 1 to 35 that you received in your email.

Demographic

2. Please select your gender *

Mark only one oval.

Male

Female

Nonbinary

Prefer not to say

Other: _____

3. Please select your age *

4. Please select your ethnicity *

Mark only one oval.

- American Indian or Alaska Native
- Asian
- Black or African American
- Native Hawaiian or Other Pacific Islander
- White
- Hispanic or Latino
- Other: _____

5. Current academic standing (if applicable): *

6. What is your major (if applicable)? *

AR/VR Background

7. Have you ever used a virtual reality headset? *

Mark only one oval.

- Yes
- No
- Don't know
- Other: _____

8. How often do you use Virtual Reality? *

Mark only one oval.

- Never
- Rarely (2-3 times a year)
- Sometimes (once every 2-3 months)
- Often (1-2 times a month)
- Weekly (1-3 times week)
- Daily (4-7 times a week)

9. What devices or accessories have you used with Virtual Reality? Select all that apply

Check all that apply.

- Earphone with Microphone
- Headset (e.g. Oculus quest, Sony PlayStation, etc..)
- Hand Controllers/Joystick for VR
- Keyboard
- Monitor
- Haptic Gloves
- VR Pen
- None of the above
- Other: _____

10. Have you ever used an Augmented Reality headset? *

Mark only one oval.

- Yes
- No
- Don't know

11. How often do you use Augmented Reality? *

Mark only one oval.

- Never
- Rarely (2-3 times a year)
- Sometimes (once every 2-3 months)
- Often (1-2 times a month)
- 0 Weekly (1-3 times week)
- Daily (4-7 times a week)

12. What devices or accessories have you used with Augmented Reality? Select all that apply *

Check all that apply.

- Earphone with Microphone
- Headset (e.g. HoloLens 2 etc..)
- Smart Glasses (e.g. Google glasses etc..)
- Keyboard
- Monitor
- Haptic Gloves
- Hand Controllers/Joystick for AR
- None of the above
- Other: _____

13. For how long have you been involved in XR research or developing XR applications? (if applicable) *

14. Indicate the average number of hours per week you spend using computers – these include tablets, desktop computers, laptops, and phones (personal and work combined):

15. Anything else you would like to share with us

This content is neither created nor endorsed by Google.

Google Forms

Appendix C

Discussion of strict and relaxed participant video data analysis

During the analysis of the video data during this thesis, two separate passes through the data were completed. The first of which was a more strict analysis in which the researcher assigned a value of one to a correct gesture and zero to an incorrect gesture. The second pass through the data was a more relaxed approach categorizing its gesture as correct partially correct or incorrect. Examples of these categorizations can be found below.

C.1 Strict Analysis

During the strict analysis of the participant videos the researcher observed recordings of the reinforcement of the gesture alongside videos of the testing portion. This side-by-side analysis allowed the researcher to examine the correctness of each gesture as compared to the first time that the participant performed the gesture.

This method of comparison was used because some participants during the reinforcement portion immediately following the instruction video performed the gesture slightly incorrectly. If this slightly incorrect gesture was then repeated during the testing portion it was considered correct. In addition, if the gesture was initially repeated slightly incorrectly but was then corrected during the testing portion the gesture was also considered correct.

Barring the two exceptions listed above all other gestures were only considered correct if the participant performed the exact gesture from the training and reinforcement section during this analysis.

C.2 Relaxed Analysis

During the more relaxed analysis of the gestures, the researcher observed the reinforcement and testing portions of the videos side-by-side following the example set in the strict analysis section.

During the gesture verification section, the researcher categorized gestures into one of three categories based on the following factors. Correct gestures were verified following the same criteria as in the strict analysis. Incorrect gestures followed the same example as correct gestures in following the same criteria as in the strict analysis.

Partially correct gestures, however, were pulled from the preceding analysis's correct and incorrect gestures. These gestures were considered partially correct if they followed a similar pattern to the gesture that was supposed to be performed but was not precisely the same gesture. For example, if the participant was asked to perform a gesture where they were asked to pinch and rotate a cube using only their pointer finger and thumb, keeping the remaining fingers clenched to their palm. However, instead, they performed a pinch gesture with their thumb and remaining fingers in a claw shape, this would be considered partially correct. This was due to the participant correctly remembering some component of the gesture, namely the pinch while forgetting the exact performance of the gesture.

This analysis on both accounts is somewhat subjective and the researcher performing the analysis was not blind to the condition. Future work could be done to analyze this data using more graders and machine learning techniques. As well as having the graders be blind to the condition of the gesture they were grading.