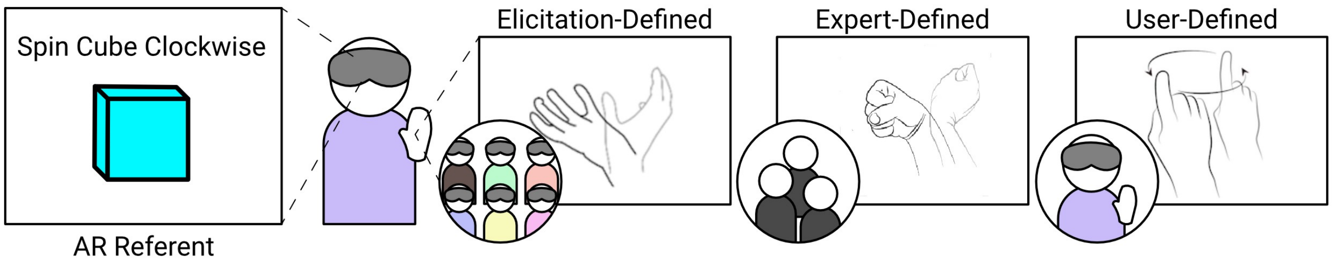
Recall This Gesture? Investigating Gesture Memorability in Extended Reality

Zahra Borhani[1], Ethan Holen[1], Adam Sinclair Williams[1], Mayra Donaji Barrera Machuca[2], Anil Ufuk Batmaz[3], Brendan Kelley[1], Xiaoyan Zhou[1], Matthew G. Rhodes[4], and Francisco Raul Ortega[1]



***Abstract*—Gesture interaction is standard in modern extended reality head-mounted displays, e.g., Microsoft Hololens 2 and Apple Vision Pro. These gestures come from multiple sources, such as elicitation studies, the application designer, or an expert. However, it is unclear if the source of the gesture affects the memorability of gestures (i.e., level of recall). In a user study, 35 participants used the HoloLens 2 to test three gestures set from different sources: 1) expert-defined, whereby we asked four experts to define gestures; 2) user-defined, whereby each participant defined their gestures; and 3) elicitation-defined, whereby we used gestures proposed by previous work. Our results showed that expert-defined gestures (72.73% recall) were more likely to be remembered than elicitation-defined (59.87% recall) gestures. Further, user-defined gestures (88.75% recall) were the most memorable. These data suggest that the source of gestures has a substantial impact on their memorability.**

***Index Terms*—Memorability, Gestures, User Study, Extended Reality, Augmented Reality, Input Methods**

I. Introduction

In recent years, extended reality (XR) has shown significant technological advances in hand-tracking with off-the-shelf gesture interaction available in head-mounted displays (HDMs), such as Meta Quest 3, Hololens 2, and Apple Vision Pro. Using hands to interact with virtual content feels more natural [27] and allows for familiar gestures like pinch, swipe, point, and rotate—many of which come from smartphones. However, the selection of gestures for a given device can vary. For instance, a developer or designer might assign different inputs to specific gestures depending on their system's requirements or choose to mimic familiar gestures from smartphones. They can also collaborate with external experts to refine gesture sets. Another option, which is currently unavailable off-the-shelf from HMDs, would allow users to create their own gestures.

Despite the different ways that exist to create a gesture set, research on gesture set memorability—how well users can recall gestures from long-term memory—remains limited. Yet, **memorability** is a key factor in evaluating the usability of a gesture set within any application [9]. Past work by Nacenta et al. [22] studied the memorability of three gesture sets (user-defined, author-designed, and random gesture sets) on tabletop systems. However, these sets do not account for expert knowledge in gesture design or natural user tendencies captured through elicitation studies, both of which can influence long-term recall in XR interactions [37]. Moreover, unlike tabletop systems, which allow users to rest their hands and receive tactile feedback, XR interactions occur entirely in mid-air, requiring users to rely on proprioception and visual feedback, making memorability even more critical. Thus, we extend this work by focusing on other potential gesture sets, including expert-defined (XD) [1] and elicitation-defined gesture (ED) sets [22]. We also focused on gestures performed in immersive XR environments and not on a tabletop device, which provides haptic feedback, as well as reducing the detrimental effects of mid-air gestures such as "gorilla arm effect" [13]. As XR glasses evolve [6], the importance of hand gestures in everyday XR applications becomes critical to the usability of these devices, making it essential to study gesture memorability in XR contexts.

This study examines three research methods for gesture design: a) User-Defined (UD) gestures that are customized by individual user, b) Expert-Defined (XD) gestures that are redefined by domain experts based on their knowledge, and c) Elicitation-Defined (ED) gestures that are derived from user studies whereby participants suggest intuitive gestures

E-mail: zahra.borhani@colostate.edu, eholen@rams.colostate.edu, adam6457@gmail.com, mbarrera@ucalgary.ca, ufuk.batmaz@concordia.ca, brendan.kelley@colostate.edu, xiaoyan.zhou@colostate.edu, matthew.rhodes@colostate.edu, f.ortega@colostate.edu.
[1]Colorado State University, United States. [2]University of Calgary, Canada. [3]Concordia University, Canada. [4]Colorado State University, United States.
Manuscript received April 19, 2021; revised August 16, 2021.

[1]A similar concept was called author-defined in [22], yet in our research, we used experts that were not authors in the publication

for specific tasks. The key question of our study is whether the ED gesture set sourced from Zhou et al. [43] is more memorable than the XD gesture set. Although prior research in cognitive psychology suggests that UD gesture sets should be most memorable [4, 15, 21, 32], we also investigate whether ED and XD gesture sets promote similar levels of retention.

Our findings provide valuable insight into the memorability of gestures in XR HMDs, such as whether XD gestures were more memorable than ED gestures and whether the most memorable gestures are those created by each user. In other words, when it comes to memory, well-designed gestures will yield better recall than elicitation gestures. In summary, the contributions of our work are as follows:

- We present a quantitative study comparing the memorability of UD, ED, and XD gesture sets to determine the source of the gesture affects memorability in XR.
- We verified the results of previous research on cognition on 2D displays that user-defined gestures are the most memorable, and extended them to XR.
- We provide a gesture set for XD consisting of 22 memorable pre-designed gestures.
- Based on our results, we provide a set of recommendations for conducting gesture elicitation studies. In addition, we show that elicitation study findings replicate in terms of gestures set by comparing the ED and commonly UD gesture sets.

II. Related Work

Creating gesture sets alone does not guarantee user adoption or usability of the gestures. Nielsen et al. [23, 24] listed five key principles that are important for evaluating a gesture set based on previous usability research: learnability (i.e., how quickly users can achieve a certain level of performance), efficiency (i.e., the performance of users over time), memorability (i.e., how easily users remember the system), errors (i.e., the frequency of minor and major errors), and satisfaction (i.e., the extent to which the system is pleasant to use and users are subjectively happy with it). Another study also considered factors such as customizability, consistency, collaborative awareness, recognizer discriminability, and empirical evaluation in the development of gesture sets [22]. Both these studies and several others [9, 10] highlighted memorability as an important factor in evaluating gesture sets. Each gesture set offers unique advantages and disadvantages in terms of these factors. For instance, technical gesture sets feature gestures that are easy for the system to recognize, but some gestures might be challenging or even impossible for certain users to perform [24]. UD gestures may be easier for individuals to learn and remember, but they are likely to be more challenging for collaborators to interpret [22]. Although selecting the ideal gesture set depends on balancing these factors for the specific system, the main focus of this paper is on identifying which gesture set yields the most memorable gestures.

A. *Enhancing Gesture Memorability*

Gesture memorability refers to how well a given gesture can be recalled, which is a vital component for creating intuitive and user-friendly interactive systems [23, 24]. Li et al. [18] observed that increasing the number of referents (memory load) reduced recall accuracy in both younger and older users. This comports with the findings of another study that showed only 2-5 gesture command associations could be held in working memory at one time depending on conditions [8, 29]. Consequently, several elicitation studies incorporate memory tests to evaluate and refine their proposed gesture sets [9, 10]. High memorability ensures that gestures are consistently recognized and executed by users, enhancing overall usability and user experience [23, 24].

Several previous studies have explored different strategies to enhance gesture memorability. For example, Li et al. [18] reported that younger users could significantly improve their recall of gesture-letter pairs with practice, especially when learning larger sets of gestures (9, 18, and 22 pairs). Another method introduced by Bergstrom-Lehtovirta et al. [2] explored the effects of placing and recalling virtual items on the skin. Participants used anatomical landmarks, personal landmarks (like scars and tattoos), and personal associations to place items. They found that personal landmarks and associations improved recall compared to anatomical landmarks. Additionally, personal associations and grouping important items, such as family members in specific areas, enhanced recall [2].

Combining modalities can make information easier to remember by creating multiple representations of to be remembered information. For instance, Church et al. [5] found that speech that was paired with a gesture was easier for participants to recall than speech with no gestures. Fruchard et al. [12] suggested another memorization strategy using body touched interactions. They compared BodyLoci (on-body hand interactions requiring users to memorize specific locations on their body) with MarkingMenus (directional gestures to select commands from a virtual menu) in VR. Their findings showed that MarkingMenus gestures were perceived as more effective for recall. In another gesture memory study, Fruchard et al. [11] compared the memorability of directional gestures (mapped to specific trajectories) and pointing gestures (associated with unique spatial positions) for selecting commands on a tablet. Pointing gestures showed a trend toward better recall, although the difference was not statistically significant.

B. *Gesture Sets Memorability*

The source of gesture generation can impact the efficacy of memorization. Nielsen et al. [24] examined three gesture sets defined by three different groups of users (engineers, engineering students, and architects) to explore the effect of cultural differences in terms of semantics, stress, and memory. Their findings on memory revealed that architects generally more easily recalled gestures, possibly due to their experience working with visual objects. In another study, Nacenta et al. [22] compared three gesture sets: PD gestures (defined by two authors of the study), UD gestures (created by participants), and a random gesture set (RG) as the control group. A later test showed that recall was best for UD gestures (79%), compared to 55% for PD gestures and 25% for RG. Additionally, participants reported that creating and recalling

TABLE I: Summary of prior research on **gestural memorability** across different gesture sources, technologies, and gesture types. Gesture sets with higher memorability performance are bolded. **Directional** gestures were mapped to a specific direction, while **positional** gestures were associated with a position in space. **BodyLoci** maps commands to different parts of the body, while **MarkingMenus** use midair directional gestures to select commands from a virtual menu. **UD**: User-defined, **XD**: Expert-defined, **ED**: Elicitation-defined, **RG**: Random gestures, **PD**: Pre-designed by authors; **UD Sci-fi** gestures created by participants who were primed by watching a video of sci-fi gestures before designing their own. **UD Creative Mindset** gestures were followed by the practice of creative mindset [31]. **Eng**: Engineers, **Eng Ss**: Engineering students, **Arch**: Architects.

Ref	Year	Gestural Sets	Technology	Display	Gesture type	Referents	Participants
Ours	2024	**UD**, ED, XD	AR-HMD	HoloLens 2	mid-air	23	35 (all)
[18]	2022	**young (19-38)**, old users (45-63 years)	2D touched	smartphone	touch gesture	6, 9, 18, 22	12, 12
[1]	2021	UD, **UD Sci-fi**, UD Creative Mindset	for AR/VR	2D display	mid-air	12	18 (both)
[11]	2018	directional, **positional**	2D touched	Galaxy Tablet	touch gesture	8	16 (both)
[12]	2018	BodyLoci , **MarkingMenus**	VR-HMD	HTC Vive	body-touched, directional	Max (144)	24 (both)
[22]	2013	RG, PD, **UD**	2D touched	Surface 1.0	touch gestures	16	18 (all)
[24]	2004	Eng, **Arch**, Eng Ss.	paper	No display	touched and mid-air	13	8, 5, 11

UD gestures was more enjoyable and easier to remember. This level of recall for the UD gesture set is in line with memory research that demonstrates a higher level of recall for items that one creates compared to what someone else created [4, 15, 21, 32]. In another study, the authors investigated the effect of priming for UD gestures across three different priming groups: no-priming as the control group, a sci-fi priming group who watched a video of gestures from sci-fi movies before designing, and creative mindset priming, which reflected the practice of creative mindset. Results showed that users memorized UD Sci-fi gestures better than the other two groups [1]. Table I presents previous studies on gestural memorability, comparing different target users, techniques, and gesture set sources in terms of memorability. The gesture set with higher memorability is bolded in the Gestural set column.

III. Experiment Gesture Sets

This paper aims to investigate gesture set memorability in AR-HMDs. Our research question (*RQ*) is: do elicitation-defined (XD) or expert-defined (ED) gestures yield higher recall (i.e., memorability) in HMDs? As a control group, we utilized UD gestures as past work on memorability already identifies UD gestures as being the most memorable [4, 15, 21, 32]. Our research extends past work by Nacenta et al. [22] that compared UD gestures versus gestures created by authors. In addition, they focused on multi-touch Surface 1.0 (allowing users to make gestures above the tabletop display) and not on XR devices.

In our study, different sources designed three gesture sets, described as follows.

- **UD gestures** are unique to each user, as users were asked to create one per referent. In the study, participants were instructed to create their own gestures for each referent. When each referent was prompted, the user would create their own gesture and then be asked to repeat it.
- **ED gesture** came from the work by Zhou et al. [43]. We used the 22 referents and replicated the environment from Zhou et al. [43]. We chose to build on Zhou et al. because their study investigated gesture elicitation in a multi-object environment, offering a different perspective compared to most earlier elicitation studies focused on single-object scenarios [28, 38]. Zhou et al. found that user behavior and preferences can differ significantly when multiple objects are involved. Their setup better reflects real-world, multi-object applications. To ensure consistency, we replicated Zhou et al. environment and extended the work to memory and used updated elicitation methodology from [34, 40]. In this case, each referent was assigned a gesture, known as the "winning gesture". This is the gesture that had the highest agreement among pairs (of participants) in the elicitation study of Zhou et al. [43]. For each referent, the ED gesture was recorded in a video prior to the experiment. During the experiment, the participant watched the video and then was asked to perform the gesture.
- **XD gestures** was a set derived from four experts. Experts worked independently, without consulting each other or having any prior exposure to the UD Gesture set. None of the experts had conducted elicitation gesture studies in the past five years, minimizing potential biases from their own user elicitation research. The experts were selected to have varied levels of expertise in HCI and XR. None of the experts were part of the study or authors in this paper. Expert 1 (X1) was a professor with extensive knowledge in VR applications and HCI with over 3300 citations. Expert 2 (X2) is a VR/AR researcher with over 20 years of experience, more than 48,000 citations, and an h-index above 90. Expert 3 (X3) was a doctoral student researching context-aware XR, with around 20 citations, under the guidance of a well-known professor specializing in 3D user interaction, with industry experience as an AR/VR researcher at Apple and Adobe. Expert 4 (X4) was a doctoral student specializing in Human-Computer Interaction and educational VR, with around 200 citations. We searched for the experts in 3dui-g@vt.edu and CHI-ANNOUNCEMENTS@list-serv.acm.org mailing list. The people who reached out were only self-identified as male, with no one either non-binary or female applying to be part of the process. Each expert received a $50 (US dollars) gift card for a one hour interview conducted using the Zoom video conference tool. The expert interviews were approved by IRB protocol number 4118. During each interview, a brief description of the study was provided and the

experts were informed that the gestures would be used in AR-HMDs. Additional questions were asked. For each referent, an animation of a virtual object in a PowerPoint slide was shown and the experts were prompted to perform a gesture that they felt best represented that specific referent. The prompt was in the slide and was repeated by the experimenter. In total, 22 referents were presented to each expert. The final gestures for the XD gesture set were chosen from the pool of expert-recommended gestures, with the selected gestures being those on which the experts reached a consensus. In cases where consensus could not be achieved, a gesture designed by X2 was selected as this expert was considered to have the most knowledge about AR and gestures. In addition, X2 has been using HMDs for decades.

Gestures corresponding to each gesture set are provided in Appendix D, Table V, VII, IV, and VI.

IV. Methodology

We conducted a control experiment using the Wizard of Oz (WoZ) methodology [3], whereby the experimenter manually executed certain actions while allowing participants to believe that the system was responding automatically to their gestures. This study was approved by IRB protocol 4118.

Participants were presented with text-based referents (i.e., an action or command), such as "move left," [34, 41, 42, 43]. The participant then provided a gesture input corresponding to that referent. Behind the scenes, the experimenter triggered the recognition of that input.

The experiment consisted of three blocks. The blocks were counter-balanced using Latin squares (Table II). Each block involved a training phase, a reinforcement phase, and a testing phase. For ED and XD gesture sets, training refers to the first time the user either watched a video of the gesture or performed the gesture. For the UD gesture sets, during training, participants were instructed to create a gesture for a given referent. In the reinforcement phase, we asked the user to repeat the gesture performed in training. In the testing phase, the user attempted to recall each gesture. Thus, within a block, users engaged in the training and reinforcement phase, followed by five minutes of drawing. The block culminated with a memory test whereby the user was asked to recall each of the gestures in the set completed before the break. Detailed information about the procedure can be found in § IV-3.

The five-minute break was designed to prevent participants from practicing (i.e., mentally rehearsing) the gestures, which could turn the final test into a short-term memory task instead of a long-term memory test. To ensure they did not rehearse the gestures, we filled the break with a drawing activity. Any distracting activity could work, but it was important that the interim task was not too similar to the gesture-learning task to avoid interference [26], following guidelines from prior work regarding the capacity limits of memory [7].

1) Experimental Design: This within-subjects study employed a 3 (gesture sets: ED, XD, UD) x 22 (referents) design. Gesture set order was counterbalanced across participants using a Latin square. Each participant performed each gesture three times per referent: training (ED and XD) or creating (UD), reinforcement, and recall, resulting in a total of 198 trials per participant. The hypotheses for the experiment are defined as follows:

TABLE II: 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

- H_1: **UD gestures in AR-HMDs have a higher recall rate than XD and ED gestures.** This hypothesis is grounded in previous findings by Nacenta et al. [22] and cognitive memory research [4, 15, 21, 32]. H_1 is accepted by our analyses.
- H_2: **XD gesture recall rates are higher than ED gestures in AR-HMDs.** Although this was not grounded by previous studies, there have been preference differences between XD and ED gestures [36]. Our analysis supports H_2.

2) Participants: A total of 35 participants were recruited to achieve power for a medium effect size (d = 0.50). They were recruited through mailing lists and word of mouth. Participants included 17 females and 18 males with ages ranging from 20 to 60 years (Mean = 25.89, SD = 6.73). The majority (63%) were computer science majors, while the remainder came from diverse fields such as economics, engineering (chemical, computer, electrical), natural resources, tourism, neuroscience, physics, and epidemiology. Sixty-six percent of the participants had experience with VR-HMDs, and 51.4% had used AR-HMDs. Regarding usage frequency, 37.1% had never used VR, and 54.3% had never used AR. For VR (AR), 48.6% (34.3%) used it rarely (2-3 times a year), 2.9% (2.9%) used it sometimes (once every 2-3 months), 2.9% (5.7%) used it often (1-2 times a month), 5.7% (0%) used it weekly (1-3 times a week), and 2.9% (2.9%) used it daily (4-7 times a week). Among the 35 participants, three reported that they had been involved in XR research for less than one semester (8.6%), 11.4% for less than two years, and 8.6% for 3 to 7 years. Additionally, 11.4% of participants were familiar with XR research via participation in XR experiments.

3) Procedure: Participants entered the experiment room, read the consent form, and were seated facing a blank wall for optimal HoloLens 2 visibility. The experimenter was seated behind and to the left of the participant as can be seen in Figure 1. Next, the experimenter started the recording and instructed the participant to wear and adjust the headset. After completing the eye calibration tutorial, the experimenter entered the participant ID into the application.

Once all this was completed, participants either started with UD, ED, or XD conditions based on their assignment to each group (see Table II. For the ED and XD conditions, participants were shown a video of the gesture and asked to repeat it (training phase). For the UD condition, participants were given the referent in the HMD prompting them to create a

Fig. 1: GoPro image of the experimental setup, the experimenter in the background (blurred for anonymity).

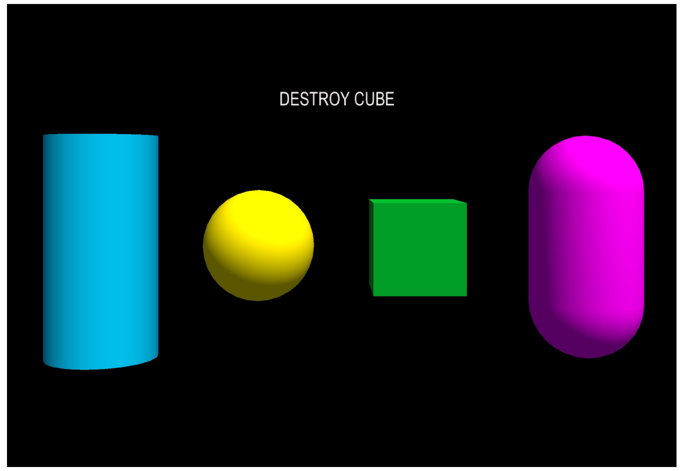


Fig. 2: Participant's AR view during the "destroy cube" test scene.

gesture (training phase) for a given referent. For all conditions, following training for a given referent, they were shown the prompt of the reference asking them to perform the same gesture (reinforcement phase). When all referents for a given condition were completed, participants were asked to remove their HMD and draw their favorite animal with crayons for five minutes. An alarm signaled the end of the 5-minute drawing period and participants were asked to put on the HMD. They confirmed whether they were still on the stop scene and the experimenter transitioned to the instructions for the testing phase. 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. For each trial of the testing phase, participants were given the referent (e.g., "Create Cube") and asked to perform the gesture, with this process continuing until all referents were completed. The "destroy cube" view of this scene can be seen in Figure 2. Following the testing phase, participants then moved to the next block (i.e., UD, ED, or XD, depending on assignment to condition) and completed the training, reinforcement, and testing phase again until all of the conditions were completed. Figure 3 provides a diagram of the procedure for a participant starting with the ED gesture set. The experiment took no more than 60 minutes to complete. Note that during the UD condition, two participants decided to change their initial gesture for one referent. For that particular case, the experimenter took the second creation as the one that should be tested for recall.

4) Apparatus: A Microsoft HoloLens 2 (version 22621) was used. This device was chosen for two reasons: to replicate the environment of Zhou et al. [43], whose referents and ED gesture set were used, and because its optical-see-through design allows the user to see their hand without any delay or distortion. The AR environment was developed using Unity Engine [2] version 2019.f.4.1f1 and the Microsoft Mixed Reality Toolkit (MRTK) 2.4.0 [3]. The computer used for development was running Windows 11 Pro 64-bit, equipped 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 of the experiment. All video and audio of the participants' gestures were recorded using a GoPro Hero7 Black camera version 1.90.

[2]https://unity.com/
[3]https://github.com/microsoft/MixedRealityToolkit-Unity

A. Data Analysis

Participant gestures and interactions during training and testing were recorded via GoPro. Recordings began immediately after participant consent and continued for the duration of the experiment. For data analysis, we used Microsoft Excel and Python and each gesture was coded as either correct (1) or incorrect (0).

For ED and XD gestures, scoring was based on the participant's interpretation of the provided video, 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 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, although it may have been incorrect compared to the initial tutorial video that was shown. UD gestures were verified by comparing the video of the participant's initial gesture creation with the video of their recall attempt.

A random subset of scored videos was reviewed by a second author, confirming the primary scorer's judgments. During the process of verifying the correctness of gesture responses from participants in the testing phase, there were many cases in which a gesture repeated by the participant was somewhat ambiguous as to its correctness. In all such cases, to ensure consistency in 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.

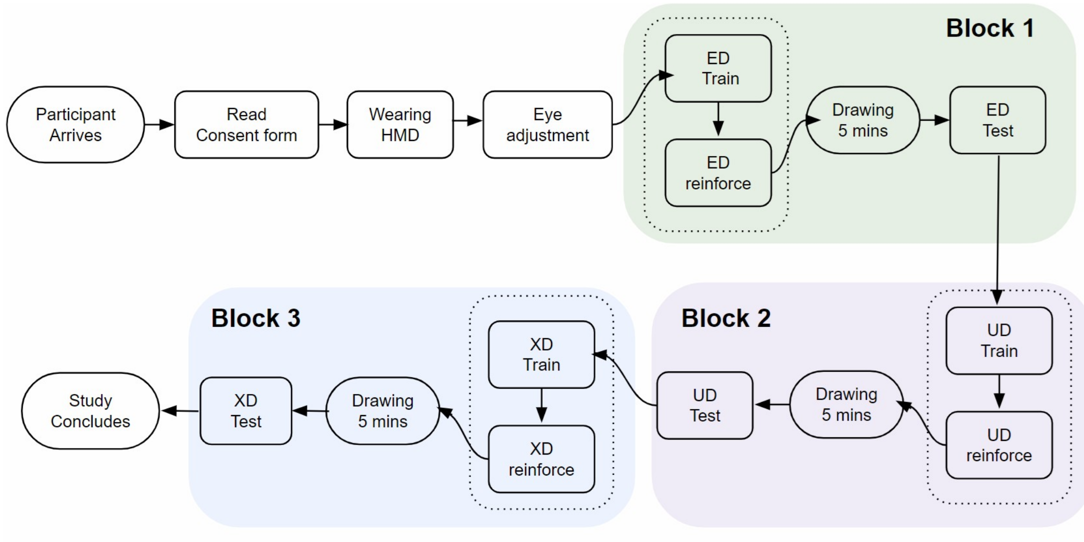


Fig. 3: Each participant followed the script shown above throughout the study. The order of gesture sets (ED, UD, and XD) was randomized using a Latin square design.

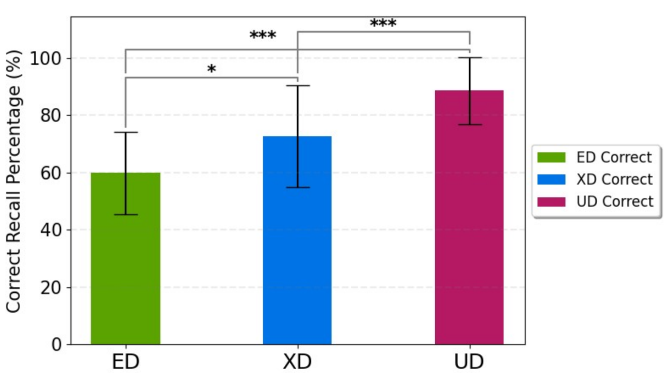


Fig. 4: Average percent of gestures correctly recalled per gesture type with standard error ($p < 0.05(*), p < 0.001(***)$)

V. Results

To evaluate the memorability of three gesture sets, we compared the average percentage of correct recall across the three conditions. Figure 4 shows the level of recall performance across UD, ED, and XD gesture sets. The UD gesture set is unique to each participant as opposed to the ED and XD gesture sets. During testing, participants correctly recalled 88% (Mean = 88.57, SD = 11.79) of UD, 59% (Mean = 59.87, SD = 14.31) of ED and 72% (Mean = 72.73, SD = 17.71) of XD gestures.

A repeated measures ANOVA was conducted to examine the relationship between gesture type and the average percentage of correctly recalled gestures. 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 difference was evident between all three conditions. Specifically, there was a large advantage for UD relative to ED gestures($d = 1.548, p < 0.001$), a medium advantage for UD compared to XD gestures ($d = 0.745, p < 0.001$), and a medium effect size reflecting poorer recall for ED than XD gestures (($d = -0.565, p < 0.012$). Thus, UD gestures were far more memorable than XD or ED gestures and XD gestures were more memorable than ED gestures.

The memorability of each gesture across all gesture sets can be observed in Figure 5, using strict analysis.

In the second pass through the testing video data, we employed a relaxed scoring scheme that classified gestures as correct, partially correct, and incorrect, with some gestures marked as correct and incorrect in the previous analysis categorized as partially correct with the new metric. This relaxed form is a common practice in memory research [22]. The results can be seen in Figure 8. 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. Because the relaxed analysis corroborated the strict-analysis findings, we report the full details in Appendix D.

A. UD Gestures Comparisons

Although deriving a gesture set from UD gestures was not the objective of this study, we used the elicitation methodology [34, 41] to extract a set from the 35 unique UD gestures. We believe sharing these results can benefit future research. To assess consensus among participants per referent [34], we calculated the agreement rate ($\mathcal{AR}$) as defined in Equation 1. For a sample size of 35, an $\mathcal{AR}$ above .3 is considered enough to make an inference of a gesture that generated consensus among the participants [33, 34, 41]. $\mathcal{AR}$ is the agreement of pairs among all participants, as shown in Equation 1. In order to calculate this metric, we used the AGATe 2.0 (AGreement Analysis Toolkit)[4] to aid in our statistical analysis. For a more in-depth explanation of elicitation studies and their analysis,

[4] http://depts.washington.edu/acelab/proj/dollar/agate.html

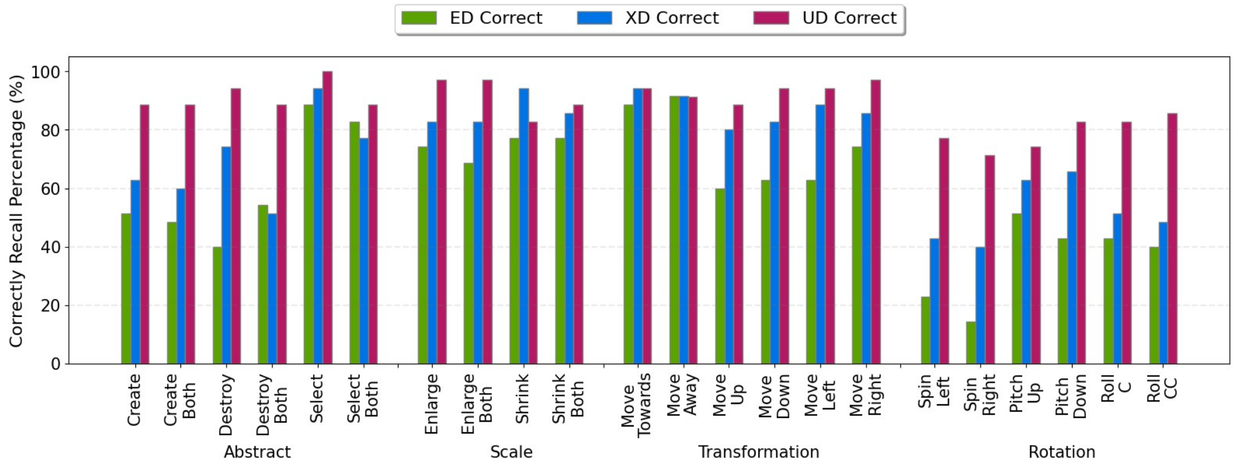


Fig. 5: Average percent correct per gesture across gesture types

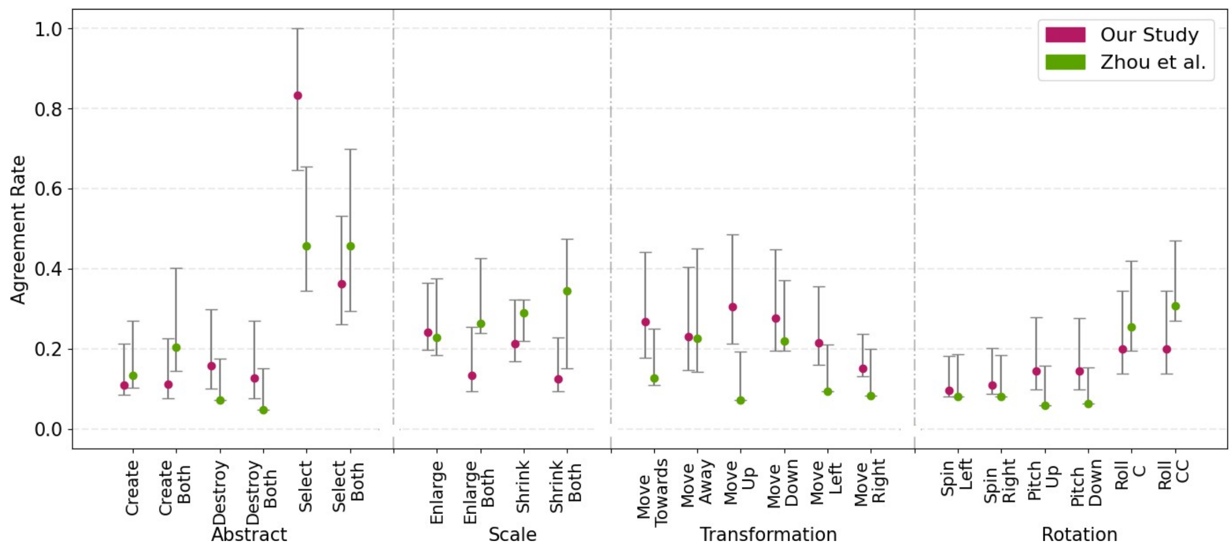


Fig. 6: The comparison of agreement rate ($\mathcal{AR}$) and confidence intervals ($\mathcal{CI}$) ranges for our study UD gesture set and Zhou et al. [43] ED gesture set. The $\mathcal{AR}$ value is denoted by a small circle and the $\mathcal{CI}$ range is shown with the bar. For example, our study $\mathcal{AR}_{create} = 0.109$ and $\mathcal{CI} = [0.084, 0.210]$.

refer to the book by Williams and Ortega [41] and Vatavu et al. [34, 35]. The gesture set derived from this study for the UD condition should be taken with the understanding of its limitation, which was not an elicitation study.

$$\mathcal{AR}(r) = \frac{\sum_{P_i \subseteq P} \binom{|P_i|}{2}}{\binom{|P|}{2}} \tag{1}$$

Participants were randomly assigned to different groups, as discussed in section § IV, which altered the order in which the gesture sets were presented. Overall, the average $\mathcal{AR}$ for all of the UD gestures of our study was $\mathcal{AR}_{all} = 0.216$, indicating a medium level of consensus. The confidence intervals ($\mathcal{CI}$) for these agreement rates for all the gesture referents are shown in Figure 6.

VI. Discussion

This study found that UD gestures were significantly more memorable than both ED and XD gestures. Our experiment also revealed that gesture sets defined by experts were more memorable than those collected from participants using the elicitation methodology, which answers *RQ*. More importantly, XD gestures were only 22.22% less memorable than UD gestures. This is important because XD gestures exhibit input legacy bias [19] (i.e., influence from previous experience with another device); the reason is that experts have been using existing HMDs and smartphones and this alone produces input legacy bias. In addition, our finding provides a different perspective regarding XD gesture sets because Morris et al. [20] had shown that users gave ED gestures a higher rating for a given command. This provides an understanding that there is no "ideal" gesture set, as different properties warrant further study. For example, findings from Morris et al. [20]) showed that, in a 2D tabletop setup, the gesture set designed by three

HCI researcher was less preferred than UD gestures; however, our findings showed that the opposite is true when it comes to the level of recall (i.e., memorability). Thus, while the UD gesture set has the highest level of recall, the adoption and availability of such sets have not been observed in other domains of custom interfaces [17].

One possible explanation for why XD gestures outperformed ED gestures in memorability might be the expert-designed set functioned as a more coherent and discriminable "gesture language." Specifically, our expert created a consistent, rule-based vocabulary in which many commands shared a common structure (e.g., lasso-select for multiple objects' actions, then perform a command-specific action gesture). This systematic organization likely supported chunking, allowing participants to remember a small set of reusable rules rather than memorizing each command independently. In addition, our expert might intentionally design gestures to be more distinct, which can reduce similarity-based confusion and interference during recall. In contrast, the elicitation process sometimes produced overlapping intuitive gestures across different commands (e.g., select and destroy), which may feel natural initially but can be harder to recall reliably over time.

At the same time, we acknowledge that other factors could also contribute. For example, the "curse of knowledge" might predict the opposite pattern (experts may not match non-expert mental models [16]), while "wisdom of crowds" suggests that a larger contributor pool could yield more representative gestures [14]. Although our study cannot directly isolate these mechanisms, the observed memorability advantage for ED is consistent with the expert set's greater structure and distinctiveness, and it motivates future work to disentangle these explanations.

Overall, although UD gestures are significantly more memorable than XD 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 UD gestures are not tenable or desired, XD gestures can be implemented with only a moderate reduction in memorability. We also extend past work by Nacenta et al. [22] who also found that UD gestures were more easily remembered (up to 24%) than authors' gestures. However, memorability may not be the priority of the application; therefore, other properties may be more desired making ED gesture set also an option.

A. User-Defined Gesture Discussion

Although the UD gesture set was not collected in a traditional elicitation study [41], a more detailed examination of the agreement rate for each referent is provided in Figure 6 to highlight similarities and differences with other studies, including Zhou et al. [43]. This figure shows a much higher agreement rate for the "select cube" referent ($\mathcal{AR}_{select} = .834$) because most participants choose to point at the cube with their index finger, which is expected and aligned with previous studies (e.g., [40, 43]). In contrast, a gesture like "Spin Cube left" has a much lower agreement rate ($\mathcal{AR}_{spin_left} = .096$), indicating that participants created a more varied set of gestures for the referent. The difference in the agreement rates between Zhou et al. [43] and our UD condition should be treated with caution, as the UD condition was not an elicitation study. However, of note, 13 out the 22 referents resulted in very similar gestures among UD gesture set and Zhou et al. [43] gesture set. Other examples of such convergence can be found in (see Tables V, IV VI, VII). Thus, elicitation studies with different participants continue to reproduce similar gestures when the environment and populations are similar (which was the case for our study), highlighting the replicability of elicitation studies.

VII. Design Implications

Memorability is one factor in midair gesture design. The Apple iPhone released in 2007 set a path for gesture interaction to become pervasive. Typical gestures include selection (touching an object), pinch (bringing two fingers together or apart), and rotate with two fingers. This has created what is known as input legacy bias [19], which could be favorably used when transferring from one device to another [25], and may support memorability if the device is used frequently for a given user. There also exist evolutionary gestures, which are gestures that transfer to the same generation device (e.g., Bloom gesture from Microsoft Hololens 2 to Magic Leap AR-HMD) as described by Williams and Ortega [39]. This further highlights the impact of our study: existing gestures continue to offer a higher level of recall. Short of that, the gesture sets that we used for our experiment provide a path for developers and designers of systems and applications.

UD gesture sets are obviously more memorable and should be made available to users if the technology allows close to perfect gesture recognition. However, user preferences have yet to be determined and remain an open question for a future qualitative study (see §IX). Another option is to rely on the XD gestures set, as they are more memorable. However, due to their replicability across studies, ED gesture sets may not be discarded and may support high levels of memorability with frequent use. In addition, Morris et al. [20] showed that users rated ED gestures higher than XD gestures.

An important design implication of our study is that multiple metrics should be considered when evaluating a gesture set, such as workload, efficiency, error rate, learnability, errors, satisfaction, and **memory**. Nielsen [23] divides these dimensions in two, learnability and memorability, corresponding to different phases of user interaction with a system. Learnability helps new users quickly understand and use the system, while memorability ensures returning users can easily remember and reuse the system after a period of not using it. For this, we hope future research in gesture sets may investigate all of the metrics mentioned whenever possible, including measurements of those metrics in elicitation studies. In addition, while difficult due to the amount of experts in gestures and XR, we hope that more expert elicitation studies can be conducted.

In terms of gesture sets provided, the implication for researchers is that similar elicitation studies tend to produce the same gestures. Therefore, it is important to understand the specific contribution and to determine how they are different. In the case of Williams et al. [40] compared to Zhou et

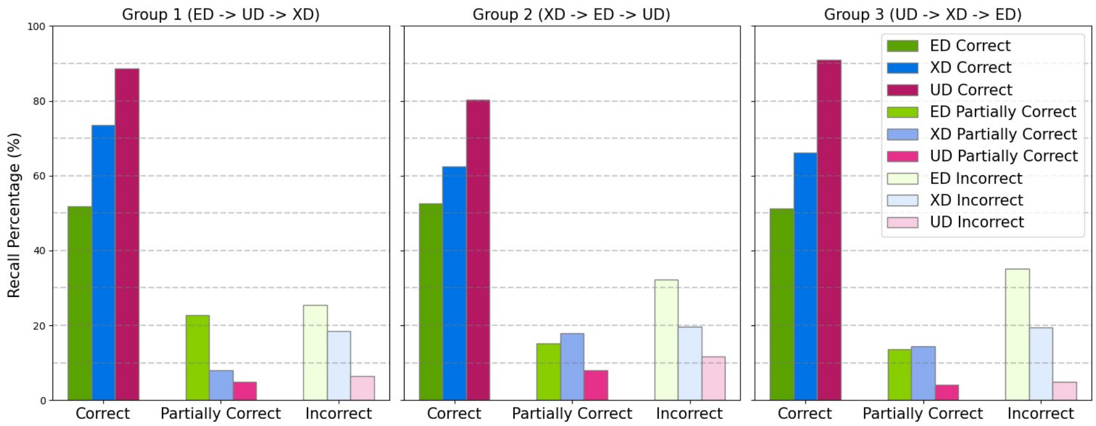


Fig. 7: Comparison of gesture memorability across three groups (Group 1, Group 2, and Group 3) for ED, XD, and UD gesture sets. Participants in each group performed the gesture sets in a different order determined by Latin square.

al [43], Zhou et al.'s environment was more complex due to its multimodal interactions which incorporated both gestures and speech, resulting in novel findings. However, our study shows how similar the ED gesture set is to the UD gesture set using the same environment (that we had to reproduce). In our case, this is not important as our study focuses on memorability and not elicitation. Yet, in the case of elicitation, this must continue to mark the progress forward.

VIII. Limitations

There are several important limitations to our study that should be considered when analyzing the results and its implications. First, we did not ask participants which gestures they preferred. This was by design because asking them afterward would have made the experiment longer than 60 minutes and might not have given real-time answers. Second, during the UD condition, participants were told that if they did not like the gesture they created, they could change it. We allowed this to reduce the stress for participants seeking perfection. Given the expectation that the level of recall for UD gestures would be higher, we expected that it would not affect the study. It is also important to note that only 2 participants out of 35 changed their initial gesture during the UD gesture set.

Third, our experiment aimed to evaluate the memorability of three gesture sets in an XR environment. The referents and UD gestures used in this study were adopted from the elicitation study by Zhou et al.[43] 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 XR device.

Fourth, Our study examined memory at a single retention interval and used a drawing task as the interim activity. Although any non-overlapping distractor could in principle serve this role, and we deliberately avoided tasks too similar to gesture learning to reduce interference [26], future work should test whether our findings replicate across different retention intervals and with alternative intervening tasks. Fifth, the three gesture sets were not strictly controlled for one-handed vs. two-handed input. Because we adopted gestures and criteria from prior work [43], the sets contain a mix of single and bimanual gestures, and expert-defined gestures in particular tend to be bimanual. Finally, while the authors tried to include people from the Global South in the studies (with some of the authors themselves being from the Global South), the experiment and the entire apparatus are from the Global North.

IX. Future work

We identify three directions for future work. First, extend memorability studies beyond hand-only gestures to multimodal interactions (e.g., speech+gesture, pen+gesture, gaze+gesture). Second, investigate microgestures and hybrid combinations—both individually and in pairings such as gaze+pinch or microgestures+controllers—to quantify their effects on recall. Third, examine task-specific gesture memorability, assessing how domain constraints shape what users remember. When expert input is needed, we recommend engaging two complementary pools—HCI/psychology experts and domain specialists (e.g., an artist for VR-sketching workflows), consistent with prior practice [30].

X. 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 found that user-defined gestures had the highest percentage of recalled gestures (88.5%). Interestingly, this study also found that expert-defined gestures had a higher-than-expected average recall at 72.73%. Using these data, we recommend that when designing gesture sets for applications, developers should 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 moderate 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 better memorability and user acceptance of gesture-based input in future technologies.

XI. ACKNOWLEDGMENTS

This work is based upon work supported by the National Science Foundation (NSF) under Grant Nos. 2439474, 2444132, and 2238313, and by the Office of Naval Research (ONR) under Grant No. N00014-24-1-2214.

REFERENCES

[1] 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.

[2] Joanna Bergstrom-Lehtovirta, Sebastian Boring, and Kasper Hornbæk. Placing and recalling virtual items on the skin. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, CHI ’17, page 1497–1507, New York, NY, USA, 2017. Association for Computing Machinery.

[3] Niels Ole Bernsen, Hans Dybkjær, and Laila Dybkjær. Wizard of oz prototyping: How and when. *Proc. CCI Working Papers Cognit. Sci./HCI, Roskilde, Denmark*, 1994.

[4] Sharon Bertsch, Bryan J Pesta, Richard Wiscott, and Michael A McDaniel. The generation effect: A meta-analytic review. *Memory & cognition*, 35:201–210, 2007.

[5] Ruth Breckinridge Church, Philip Garber, and Kathryn Rogalski. The role of gesture in memory and social communication. *Gesture*, 7(2):137–158, 2007.

[6] CNBC. Qualcomm says it’s working on mixed reality smart glasses with samsung and google. https://www.cnbc.com/2024/09/05/qualcomm-working-on-mixed-reality-smart-glasses-with-google-samsung.html, 2024. Accessed: 2024-09-11.

[7] Nelson Cowan. The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and brain sciences*, 24(1):87–114, 2001.

[8] Nelson Cowan. The magical mystery four: How is working memory capacity limited, and why? *Current directions in psychological science*, 19(1):51–57, 2010.

[9] 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.

[10] Elias Fares, Victor Cheung, and Audrey Girouard. Effects of bend gesture training on learnability and memorability in a mobile game. In *Proceedings of the 2017 ACM International Conference on Interactive Surfaces and Spaces*, ISS ’17, page 240–245, New York, NY, USA, 2017. Association for Computing Machinery.

[11] 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.

[12] Bruno Fruchard, Eric Lecolinet, and Olivier Chapuis. Impact of semantic aids on command memorization for on-body interaction and directional gestures. In *Proceedings of the 2018 International Conference on Advanced Visual Interfaces*, AVI ’18, New York, NY, USA, 2018. Association for Computing Machinery.

[13] Emilie Genty, Thomas Breuer, Catherine Hobaiter, and Richard W Byrne. Gestural communication of the gorilla (gorilla gorilla): repertoire, intentionality and possible origins. *Animal cognition*, 12:527–546, 2009.

[14] Ralph Hertwig. Tapping into the wisdom of the crowd—with confidence. *Science*, 336(6079):303–304, 2012.

[15] Larry L Jacoby. On interpreting the effects of repetition: Solving a problem versus remembering a solution. *Journal of verbal learning and verbal behavior*, 17(6):649–667, 1978.

[16] Colleen M Kelley and Larry L Jacoby. Adult egocentrism: Subjective experience versus analytic bases for judgment. *Journal of Memory and language*, 35(2):157–175, 1996.

[17] David M Lane, H Albert Napier, S Camille Peres, and Aniko Sandor. Hidden costs of graphical user interfaces: Failure to make the transition from menus and icon toolbars to keyboard shortcuts. *International Journal of Human-Computer Interaction*, 18(2):133–144, 2005.

[18] Bingxin Li, Tong Yang, Yanfang Liu, and Feng Du. Memory load differentially influences younger and older users’ learning curve of touchscreen gestures. *Scientific Reports*, 12(1):10814, 2022.

[19] Meredith Ringel Morris, Andreea Danielescu, Steven Drucker, Danyel Fisher, Bongshin Lee, m. c. schraefel, and Jacob O. Wobbrock. Reducing legacy bias in gesture elicitation studies. *Interactions*, 21(3):40–45, may 2014.

[20] 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.

[21] Neil W Mulligan and Jeffrey P Lozito. Self-generation and memory. *The psychology of learning and motivation: Advances in research and theory*, 45:175–214, 2004.

[22] 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.

[23] Jakob Nielsen. *Usability engineering*. Morgan Kaufmann, 1994.

[24] 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.

[25] Francisco R. Ortega, Alain Galvan, Katherine Tarre, Armando Barreto, Naphtali Rishe, Jonathan Bernal, Ruben Balcazar, and Jason-Lee Thomas. Gesture elicitation for 3d travel via multi-touch and mid-air systems for procedurally generated pseudo-universe. In *2017 IEEE Symposium on 3D User Interfaces (3DUI)*, pages 144–153, 2017.

[26] 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.

[27] Thammathip Piumsomboon, David Altimira, Hyungon Kim, Adrian Clark, Gun Lee, and Mark Billinghurst. Grasp-shell vs gesture-speech: A comparison of direct and indirect natural interaction techniques in augmented reality. In *2014 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 73–82, 2014.

[28] Thammathip Piumsomboon, Adrian Clark, Mark Billinghurst, and Andy Cockburn. User-defined gestures for augmented reality. In *CHI'13 Extended Abstracts on Human Factors in Computing Systems*, pages 955–960. 2013.

[29] 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.

[30] 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.

[31] Kai Sassenberg and Gordon B Moskowitz. Don't stereotype, think different! overcoming automatic stereotype activation by mindset priming. *Journal of Experimental Social Psychology*, 41(5):506–514, 2005.

[32] Norman J Slamecka and Peter Graf. The generation effect: Delineation of a phenomenon. *Journal of experimental Psychology: Human learning and Memory*, 4(6):592, 1978.

[33] Theophanis Tsandilas. Fallacies of agreement: A critical review of consensus assessment methods for gesture elicitation. *ACM Trans. Comput.-Hum. Interact.*, 25(3), jun 2018.

[34] 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.

[35] Radu-Daniel Vatavu and Jacob O. Wobbrock. Clarifying agreement calculations and analysis for end-user elicitation studies. *ACM Trans. Comput.-Hum. Interact.*, 29(1), jan 2022.

[36] Santiago Villarreal-Narvaez, Arthur Sluÿters, Jean Vanderdonckt, and Radu-Daniel Vatavu. Brave new ges world: A systematic literature review of gestures and referents in gesture elicitation studies. *ACM Comput. Surv.*, 56(5), jan 2024.

[37] Panagiotis Vogiatzidakis and Panayiotis Koutsabasis. Gesture elicitation studies for mid-air interaction: A review. *Multimodal Technologies and Interaction*, 2(4), 2018.

[38] Adam S. Williams, Jason Garcia, and Francisco Ortega. Understanding multimodal user gesture and speech behavior for object manipulation in augmented reality using elicitation. *IEEE Transactions on Visualization and Computer Graphics*, 26(12):3479–3489, 2020.

[39] Adam S Williams and Francisco R Ortega. Evolutionary gestures: When a gesture is not quite legacy biased. *Interactions*, 27(5):50–53, 2020.

[40] 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.

[41] Adam S. Williams and Francisco R. Ortega. A concise guide to elicitation methodology, 2021.

[42] 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.

[43] 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.

Zahra Borhani received her M.S. and Ph.D. degrees in computer science from Colorado State University. Her research focuses on interaction modalities in extended reality (XR), including gestures, annotations, and gaze-based techniques.

Ethan Holen received his MS degree in Computer Science from Colorado State University. He was a member of the Natural User Interaction Lab, where he defended his thesis on the memorability of gestural input in XR systems. He is currently working as a Cloud Kubernetes Platform Engineer at HPE, designing and building tools for system management in high-performance computing.

This article has been accepted for publication in IEEE Transactions on Visualization and Computer Graphics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVCG.2026.3678203

Adam Sinclair Williams is a machine learning engineer who also serves as an affiliate researcher in the computer science department at Colorado State University. His professional work focuses on addressing problems within the search domain. His research is centered on augmented/virtual reality, specifically exploring multimodal inputs, such as user-elicited speech and gestures, to create novice-friendly interactions within 3D learning environments.

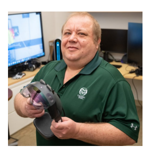
Mayra Donaji Barrera Machuca is an Assistant Professor at the Department of Computer Science at the University of Calgary. Her current research areas include using multimodal interactions for 3D sketching and understanding human spatial thinking in the design of 3D objects. She also proposes new user interfaces that improve the accuracy of people working inside virtual environments and studies skill transfers between virtual reality and real life. Dr. Barrera Machuca's work has generated considerable interest in the community, including winning the Dalhousie FCS Dean's Award - Research Excellence 2024, and the 2021 VGTC Best Dissertation Award.

Anil Aufuk Batmaz (Member, IEEE) received the PhD degree in biomedical engineering from Stras- bourg University, France. He is currently an assistant professor with the Department of Computer Science and Software Engineering, Concordia University, Gina Cody School of Engineering and Computer Science, Montreal, Canada. He leads the "EXtended Reality and Interaction Technologies" (EXIT) Lab. His research interests include virtual reality, augmented reality, training systems and simulators, human-computer interaction, and human–AI interaction.

Brendan Kelley is pursuing the Ph.D. degree with the Natural User Interface Laboratory (NuiLab). His research primarily focuses on cognitive cueing techniques, particularly visual methods, with the goal of improving the design and use of cues in 3D user interfaces such as AR and VR head-mounted displays (HMDs). He is particularly interested in their application in high-stress or potentially hazardous environments (e.g., first response, search and rescue, and military operations), where improvements in time and accuracy can help save lives.

Xiaoyan Zhou is a postdoctoral researcher at the Department of Media Technology and Interaction Design, KTH Royal Institute of Technology, Sweden. She received her Ph.D. in Computer Science from Colorado State University in 2024. Her research focuses on human–computer interaction and extended reality (XR), with an emphasis on immersive interaction and visualization. Her recent work explores micromobility simulation, climate change visualization, and misinformation in immersive environments. Her publications appear in leading venues such as ACM CHI, VRST, and ISS.

Matthew G. Rhodes is a Professor of Psychology at Colorado State University. His research focuses on memory, metacognition, and evidence-based approaches to learning and training. He is a Fellow of the American Psychological Association (Division 3) and the Association for Psychological Science and has received several awards for teaching and student mentoring. Dr. Rhodes is an author of the book A Guide to Effective Studying and Learning: Practical Strategies from the Science of Learning (2020; Oxford University Press).



Francisco R. Ortega is an Associate Professor at Colorado State University (CSU) and has been the Director of the Natural User Interaction Lab (NUILAB) since 2018. He earned his Ph.D. in Computer Science (CS), specializing in Human–Computer Interaction (HCI) and 3D User Interfaces (3DUI), from Florida International University (FIU). His research broadly focuses on multimodal and unimodal (gesture-centric) interaction, including gesture elicitation (e.g., participatory design), information access effort in augmented reality (AR), virtual reality (VR) forest bathing, and VR/AR training.