

Article

# Design of 3D Microgestures for Commands in Virtual Reality or Augmented Reality

**Guangchuan Li** <sup>1</sup> , **David Rempel** <sup>2,3</sup> , **Yue Liu** <sup>1,4,\*</sup>, **Weitao Song** <sup>1</sup> and **Carisa Harris Adamson** <sup>2,3,\*</sup><sup>1</sup> Beijing Engineering Research Center of Mixed Reality and Advanced Display, School of Optics and Photonics, Beijing Institute of Technology, Beijing 100081, China; 3120160305@bit.edu.cn (G.L.); swt@bit.edu.cn (W.S.)<sup>2</sup> Department of Medicine, University of California, San Francisco, CA 94143, USA; David.Rempel@ucsf.edu<sup>3</sup> School of Public Health, University of California, Berkeley, CA 94720, USA<sup>4</sup> Beijing Film Academy, Advanced Innovation Center for Future Visual Entertainment, Beijing 100088, China

\* Correspondence: liuyue@bit.edu.cn (Y.L.); carisaharris@berkeley.edu (C.H.A.)

**Abstract:** Virtual and augmented reality (VR, AR) systems present 3D images that users can interact with using controllers or gestures. The design of the user input process is crucial and determines the interactive efficiency, comfort, and adoption. Gesture-based input provides a device-free interaction that may improve safety and creativity compared to using a hand controller while allowing the hands to perform other tasks. Microgestures with small finger and hand motions may have an advantage over the larger forearm and upper arm gestures by reducing distraction, reducing fatigue, and increasing privacy during the interaction. The design of microgestures should consider user experience, ergonomic principles, and interface design to optimize productivity and comfort while minimizing errors. Forty VR/AR or smart device users evaluated a set of 33 microgestures, designed by ergonomists, and linked them to 20 common AR/VR commands based on usability, comfort, and preference. Based primarily on preference, a set of microgestures linked to specific commands is proposed for VR or AR systems. The proposed microgesture set will likely minimize fatigue and optimize usability. Furthermore, the methodology presented for selecting microgestures and assigning them to commands can be applied to the design of other gesture sets.



**Citation:** Li, G.; Rempel, D.; Liu, Y.; Song, W.; Adamson, C.H. Design of 3D Microgestures for Commands in Virtual Reality or Augmented Reality. *Appl. Sci.* **2021**, *11*, 6375. <https://doi.org/10.3390/app11146375>

Academic Editor: Heecheon You

Received: 10 June 2021

Accepted: 7 July 2021

Published: 9 July 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

VR and AR techniques can render a virtual environment or superimpose a virtual object onto a physical target; thus, the interaction experience with VR/AR is different from conventional displays that include monitors, tablets, and projectors. VR/AR devices are used for sports training, entertainment, tourism, manufacturing, warehouse work, and medical assistance [1–3]. They have the potential to change how we learn, recreate, and work [4,5]. The modes of communication between humans and VR/AR systems have a large impact on their implementation, use, and acceptance. The human interface requirements for VR/AR devices are different from other human–computer systems; the visual demands are greater, the command set is different, and the user may be sitting, standing, or walking. Therefore, researchers are designing tracking, interaction, and display techniques to improve comfort, and efficiency particularly [6], then applying those techniques to the implementation of AR systems.

Integrating a human factor approach to the design of gestures with gesture recognition that optimizes latency and accuracy is vital to facilitating effective interactions between humans and VR/AR systems [7–11]. Users are familiar with touch gestures and voice control due to their pervasive use of smart devices and these input methods have been evaluated by users interacting with VR/AR systems. FaceTouch (FT), a touch–input interface mounted on the backside of a head-mounted display (HMD) [12] was accepted by participants because of its low error rate, short selection time, and high usability. However,

it could be only used for short utilitarian purposes because users' arms tend to fatigue easily. The use of pre-programmed voice commands to manipulate objects in VR demonstrated the usefulness of voice control as a practical interface between users and systems [13,14]. However, the use of voice commands can reveal user actions and disturb others in public. Office workers are accustomed to using a keyboard and mouse. A virtual keyboard to execute commands has been evaluated in an immersive virtual environment by office workers [15,16]; however, the accurate detection of fingertips and the provision of feedback to users while typing on the virtual keyboard were difficult and limited usability and interaction efficiency.

Another common user interface approach when interacting with VR head-mounted displays (HMDs) is the hand-held controller, similar to a game controller, which provides accurate input and feedback while minimizing latency (Controller, HTC VIVE, HTC Inc., New Taipei City, Taiwan, 1997). However, the visual demands to identify buttons on a controller can distract the users and negatively impact their performance. Additionally, the number of buttons that can reasonably fit onto a controller is limited which constrains input efficiency. Furthermore, the use of hand controllers has been shown to increase motion sickness during VR HMD use [17] and the sustained grasp of a hand-held controller may increase the muscle load on the upper extremities. Interaction with VR or AR systems that rely on extra hardware like a controller or touch screen prevents users from performing other tasks with their hands such as manipulating physical objects.

Due to these constraints, mid-air, non-contacting, and 3D hand gestures have emerged as an alternative to a controller, touch screen, or voice control while interacting with high-resolution displays, computers, and robots [18–21]. Importantly, the human proprioception system allows users to perceive the spatial position of the hands relative to the trunk with high precision. The use of VR/AR HMDs is becoming widespread and user experience with HMDs is different from conventional smart devices such as phones, tablets, and TV. In response, the design and use of hand gestures for interacting with VR/AR applications have been investigated [22–24]. Navigation, object selection, and manipulation commands are commonly used in VR/AR systems and have been evaluated by researchers. The conversion of fingertip or hand position movement to the control of the velocity of self-movement in the virtual environment (VE) has been demonstrated [25,26] as has the use of gestures for the point and selection of virtual objects [27]. For example, to select virtual objects, users preferred the index finger thrust and index finger click gestures [27]. However, the lack of feedback can have a negative impact on user input confidence and efficiency while performing gestures [28,29]. Therefore, researchers have designed hardware, like a flexible nozzle [30] and ultrasound haptics [31], to provide feedback while using gestures in different application scenarios. Physical surfaces like tables and walls are additional ways to provide feedback for VR/AR systems, and, being ubiquitous, can provide feedback to users by mimicking interactions with a touch screen. Researchers developed the MRTouch system which affixed virtual interfaces to physical planes [32] and demonstrated that feedback from physical surfaces can improve input accuracy.

Gestures can provide an unconstrained, natural form of interaction with VR/AR systems. However, gestures can differ widely in how they are performed. For example, performing large gestures that require whole arm movement can lead to shoulder fatigue and are difficult to use for extended durations [33]. Sign language interpreters who perform mid-air gestures for a prolonged period experience pain and disorders of the upper extremities [34]. To avoid fatigue, studies suggest that repeated gestures should not involve large arm movements and should avoid overextension or over-flexing of the hand, wrist, or finger joints [34]. Microgestures, defined by utilizing continuous small hand or finger motion, have been developed to reduce the muscle load on the upper extremities associated with large and/or mid-air gestures. Microgestures can be inconspicuous and less obvious to nearby people and are more acceptable in public settings compared to large hand or arm motions [35,36]. Single-hand microgesture sets have been proposed to interact with miniaturized technologies [37] and to grasp hand-held objects with geometric

and size differences [38]. However, some of the microgestures studied could not be easily performed. For example, the gesture to use the thumb tip to draw a circle on the palm of the same hand was found to be very difficult and uncomfortable to perform. Additionally, the proposed microgesture set was designed with the forearm resting on the tabletop in a fully supinated (palm up) position, a posture that is associated with discomfort and should be avoided if performed repeatedly [34].

Although the use of microgestures for VR/AR is appealing, as yet, there is no universally accepted microgesture set developed for VR/AR systems designed with a human-centered approach [37–45]. Therefore, the design of a microgesture set for common VR/AR commands that are intuitive, easily recalled, and minimize fatigue and pain is warranted. The primary purpose of this study was to design a microgesture set for VR/AR following human factors and ergonomic principles that consider user interaction habits and gesture designs that minimize hand fatigue and discomfort. Additionally, to improve the wide acceptance and usability, participants from different cultural backgrounds were recruited to build mappings between microgestures and commands for the VR/AR system.

The paper is organized as follows: We first describe the methodology of the study including the selection of VR/AR commands, the design of the microgestures, the pre-assignment of microgestures to the commands by experts, and the design of software for participants to assign microgestures to the commands and rate the microgesture-command sets. Next, we describe the data analyses used. Then, we present study findings by first presenting the proposed microgesture set for VR/AR commands based on popularity and the user preference for microgestures' characteristics. Finally, in the discussion, we compare the proposed microgestures to prior studies and discuss the broader implications and limitations of this work.

## 2. Methodology

### 2.1. Participants

A sample of convenience of participants between the ages of 18–65 years who had experience using touch (2D) or mid-air (3D) gestures to interact with smart devices including phones, tablets, AR, or VR devices was recruited through individual emails of participants primarily from prior studies at our laboratories. Forty participants completed experiments, and they were compensated for the time spent on the experiment. The study was conducted during the COVID-19 pandemic; therefore, the study was conducted online using the participant's personal computer in a quiet setting where they were not disturbed or in the view of others. The study was approved by the Institutional Review Board of the University of California, San Francisco (IRB#10-04700).

### 2.2. Selection of Common Commands for VR/AR

Having a clear understanding of the usage context of the VR/AR system is important in designing a microgesture set for specific commands. We used the following premises to balance efficiency and mental load associated with using gestures: (1) the tasks completed should mimic shortcuts used for computers; (2) the gestures should be intuitive and follow current gesture lexicons common to touch screens; (3) commands with opposite purposes were linked together with similar gestures; (4) the gestures for the command that turns on gesture recognition, gesture on/off, should be clearly recognizable and differentiated from other common hand gestures, and (5) the gestures involved with several sequential commands for a task should follow the canonical interaction pattern that users are familiar with. For example, the sequence of the commands selection, translation, and rotation are frequently performed together.

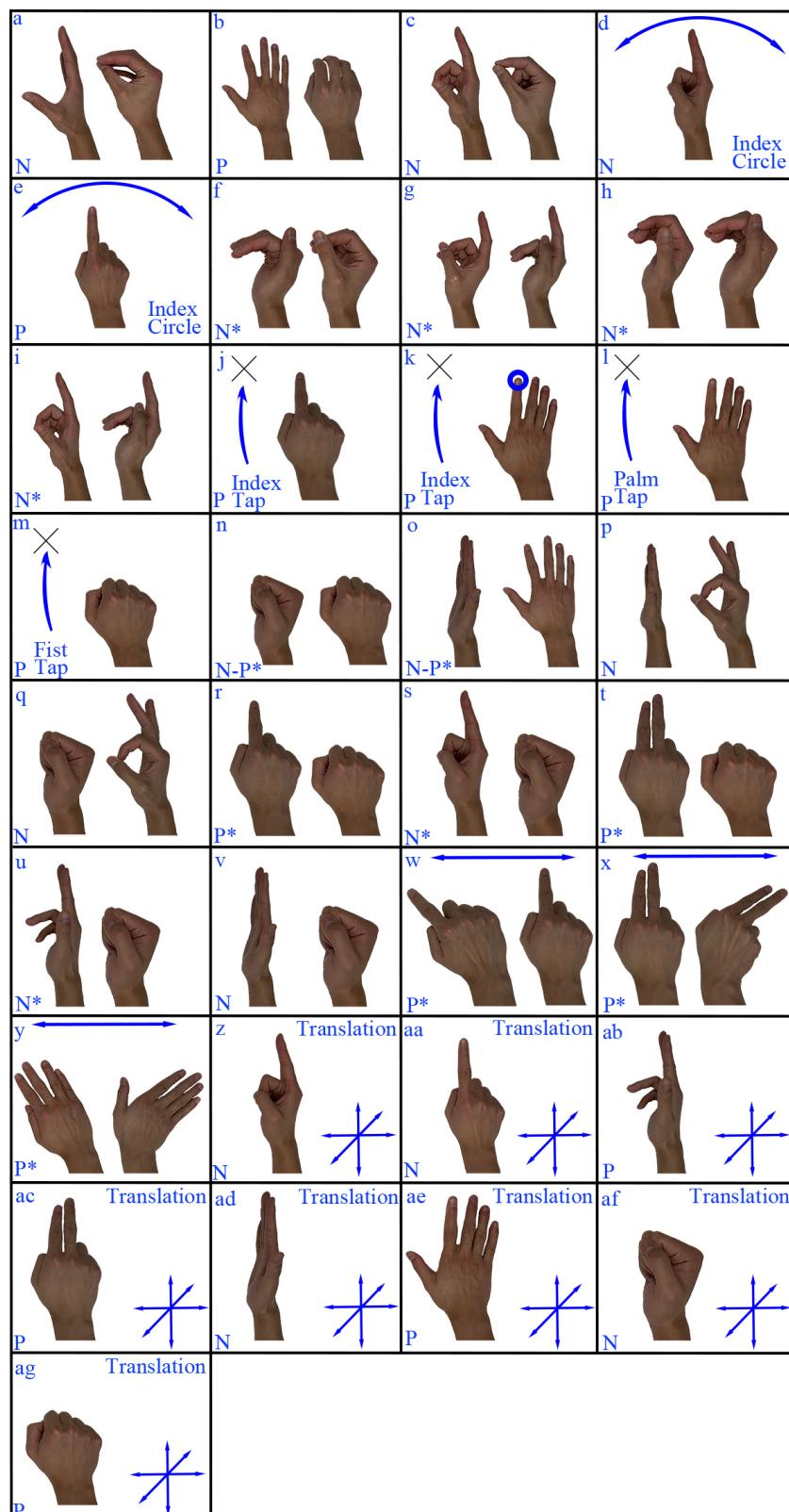
As a first step, four developers of VR/AR were invited to rate 33 commands identified for VR/AR systems from prior studies and commercial devices on their importance for interacting with VR/AR systems using a 5-point Likert scale: 1 (least important), 5 (most important) (Table 1). The 20 top-rated commands were selected for further study.

**Table 1.** Thirty-three commands with mean (*SD*) ratings by experts on their importance for using AR or VR systems (Likert Scale: 1 = least important, 5 = most important).

Task List	Score
1A. Gesture on	1B. Gesture off
2A. Open menu bar	2B. Close menu bar
3A. Scroll left	3B. Scroll right
4A. Scroll up	4B. Scroll down
5A. Previous	5B. Next,
6A. Previous on a circle	6B. Next, on a circle
7A. Play	7B. Pause
8. Accept call	
9. Decline call	
10A. Mute	10B. Undo mute
11A. Volume up	11B. Volume down
12A. Enable marker	12B. Disable marker
13A. Confirm selection	13B. Cancel selection
14A. Shrink	14B. Enlarge
15A. Initialize	15B. Restore
16. Cursor	
17. Translation	
18. Rotation	
19. Duplicate	
20. Delete	
21A. Add label	21B. Remove label
22. Back to center	
23. Screen shot	
24A. Fast forward	24B. Fast backward
25. Take picture	
26. Take a video	
27. Search	
28A. Brightness up	28B. Brightness down
29. Battery save	
30A. Enable keyboard	30A. Disable keyboard
31. Sleep mode	
32. Shut down	
33. Reset settings	

### 2.3. 3D Microgestures' Design

The designed 3D microgestures were guided by prior research studies and by currently used gestures [37,41–43] (HoloLens, Microsoft, Albuquerque, NM, USA) (Magic Leap One, Magic Leap Inc., Florida, MIA, USA). To prevent computer interaction gestures from being mistaken for static non-computer gestures, only dynamic microgestures were considered. Two certified professional ergonomists with expertise in hand biomechanics and the design of tools and gestures were consulted to design microgestures that could optimize comfort and reduce the risk of hand fatigue and pain. For example, based on prior research that has identified sustained supination (palm up) of the forearm as a risk factor for pain and discomfort [34], microgestures that promoted pronated (palm down) to neutral (thumb-up) forearm postures were selected. Based on the design criteria described, a library of 33 microgestures was created and described pictorially (Figure 1) and literally (Table 2).



**Figure 1.** The graphical description of 33 designed microgestures. Microgestures with N, P, or N-P indicated that the forearm in the neutral, pronated position, or rotating from neutral to pronated position while microgestures with \* were identified as unfamiliar gestures. The instruction on how to perform each microgesture could be checked in Table 2, for example, microgesture *a* was performed from extended fingers to close fingertips or the reversed one.

**Table 2.** The literal description of 33 designed microgestures.

Gesture Description
a. Grab/expand with the neutral forearm
b. Grab/expand with the pronated forearm
c. Pinch/expand with the neutral forearm
d. Index finger circles CW/CCW with the pronated forearm
e. Index finger circles CW/CCW with the neutral forearm
f. Thumb slides on index finger with the neutral forearm
g. Thumb slides on the middle finger with the neutral forearm
h. Thumb taps on the index finger with the neutral forearm
i. Thumb taps on the middle finger with the neutral forearm
j. Index fist taps on the table with the pronated forearm
k. Index finger taps on the table with palm down on the table
l. Palm taps on the table with palm down
m. Fist taps on the table with the pronated forearm
n. Fist posture rotates from the pronated forearm to the neutral forearm
o. Palm posture rotates from neutral forearm to pronated forearm
p. Forming okay posture from palm posture, with the neutral forearm
q. Forming okay posture from fist posture, with the neutral forearm
r. Index finger scratches toward/away with the pronated forearm
s. Index finger scratches toward/away with the neutral forearm
t. Index and middle fingers scratch toward/away, pronated forearm
u. Index and middle fingers scratch toward/away with the neutral forearm
v. Palm scratches toward/away with the neutral forearm
w. Index finger scrolls left/right with the pronated forearm
x. Index and middle fingers scroll left/right with the pronated forearm
y. Palm scrolls left/right with the pronated forearm
z. Index finger points forward with the neutral forearm
aa. Index finger points forward with the pronated forearm
ab. Index and middle fingers point forward with the neutral forearm
ac. Index and middle fingers point forward with the pronated forearm
ad. Curled palm with the neutral forearm
ae. Curled palm with the pronated forearm
af. Fist posture with the neutral forearm
ag. Fist posture with the pronated forearm

#### 2.4. Initial Expert Pre-Selection of Gestures to Match Commands

For each command (Table 3), three researchers pre-selected eight of the designed microgestures (Figure 1, Table 2) that best matched the command metaphorically. The purpose of pre-selecting eight gestures was to reduce the cognitive demands on participants, so they would not have to review all 33 microgestures when assigning microgestures to a command. Differences between experts were resolved with discussion.

Microgestures were mapped to 20 commands (Table 3) based on existing lexicons; mapping was not based on physical (i.e., shaking the HMD) [43], symbolic (i.e., drawing a letter O with fingertips) [46], or abstract (i.e., arbitrary gesture) mappings that could hinder recall or reliable performance. Commands considered to be opposites of each other were mapped to a similar gesture performed in opposite directions [43]. For example, a flick to the right was the most common gesture for the “next” command while a flick to the left was the gesture mapped to the “previous” command. Furthermore, depending on the command, microgestures could be performed in discrete or continuous movements. For example, the “duplicate” command was designed as a discrete movement while the “adjusting volume” command was designed as a continuous movement. The eight microgestures pre-selected for each command are listed in Table 3.

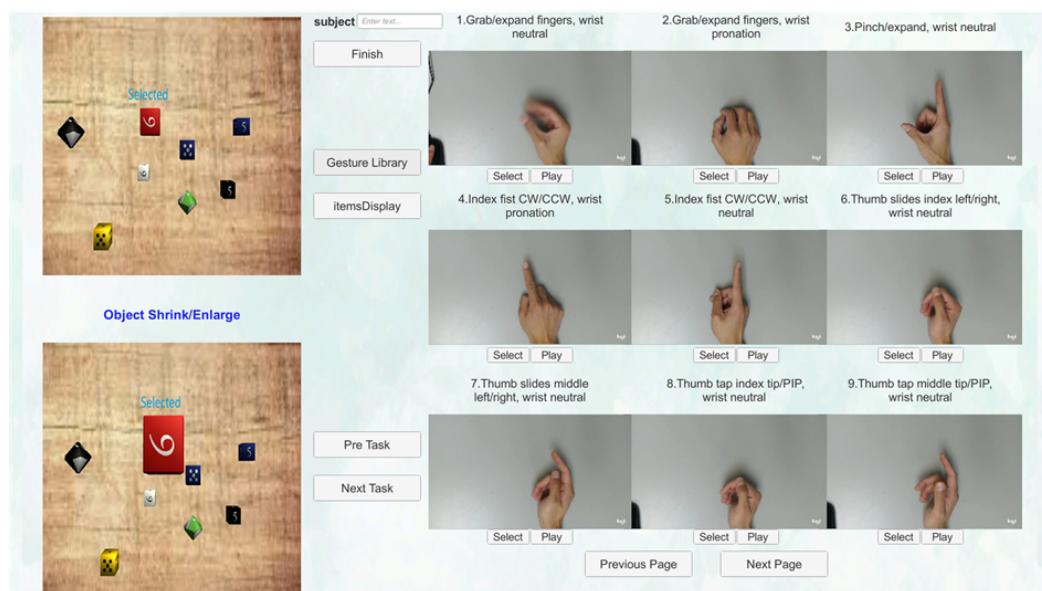
**Table 3.** Twenty commands with eight pre-selected microgestures (gesture numbers from Figure 1).

Commands	Eight Microgestures Candidates							
Gesture on/off	a	b	c	d	e	h	p	q
Open/close menu bar	a	c	e	i	l	q	t	w
Scroll left/right	f	g	h	u	v	w	x	y
Scroll up/down	d	g	h	r	s	t	u	v
Previous/next	f	g	h	s	v	w	x	y
Previous/next on a circle	d	e	f	g	i	w	x	y
Play/pause	a	b	c	e	h	j	l	m
Accept call	a	c	e	f	h	p	v	x
Decline call	b	c	d	g	i	q	v	y
Mute/unmute	a	b	c	e	h	m	o	v
Volume up/down	d	e	f	g	n	r	t	v
Marker open/close	a	e	f	h	p	r	t	y
Confirm/cancel selection	a	c	e	h	i	k	l	m
Shrink/enlarge	a	b	c	d	e	f	g	u
Initiate/restore	d	e	f	h	i	n	q	x
Cursor	z	aa	ab	ac	ad	ae	af	ag
Translation	z	aa	ab	ac	ad	ae	af	ag
Rotation	z	aa	ab	ac	ad	ae	af	ag
Duplicate	a	c	e	f	h	k	l	v
Delete	a	c	e	f	i	l	u	w

## 2.5. User Assignment of Gestures to Commands

### 2.5.1. Training and Selection Interface

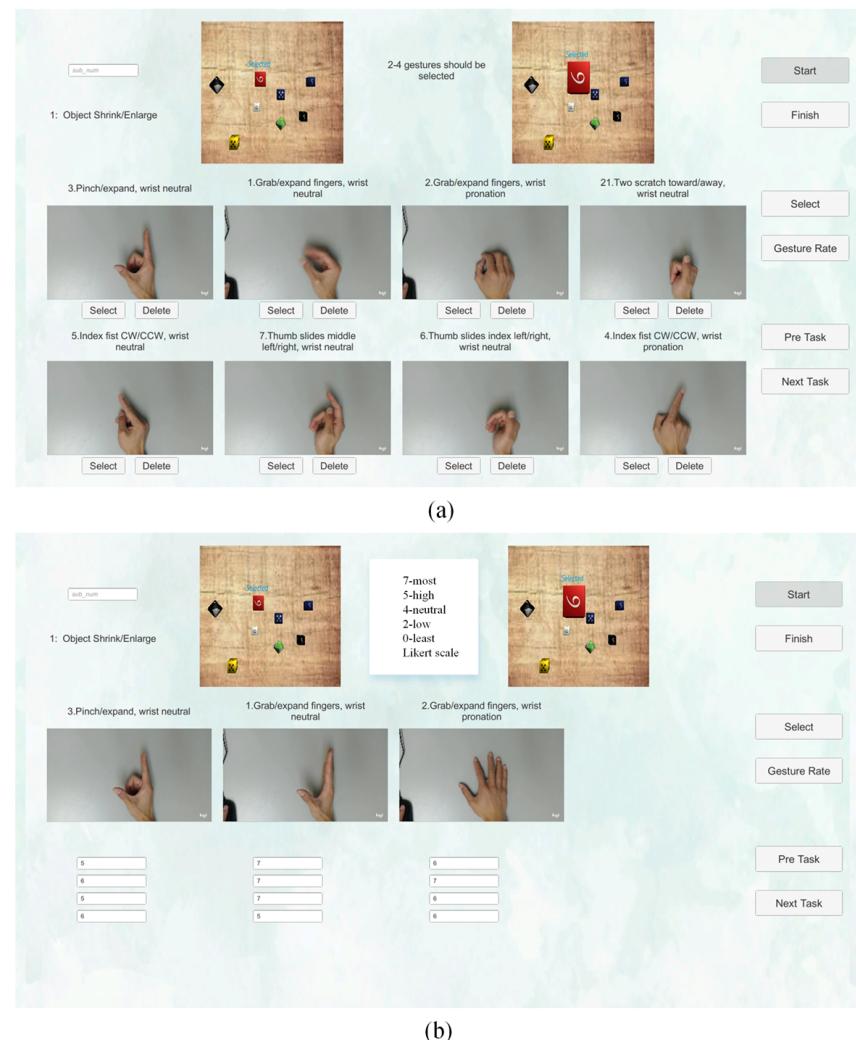
Two interfaces were developed using Unity3D (Unity Technologies, San Francisco, CA, USA). One was for training the participants on the commands and the microgestures (Training Interface, Figure 2). The training interface consisted of two parts: (1) pictures were displayed on the left showing a before and after screen image to demonstrate the purpose of the command, and (2) videos of nine microgestures with the other 24 microgestures displayed by clicking the button previous or next page.



**Figure 2.** Interface for training participants on the command (task) and all the designed microgestures displayed on the computer monitor.

Once the participants were adequately familiar with the 20 commands and 33 microgestures, they interacted with a second interface that allowed them to review each command, browse the eight pre-selected microgestures, and then select the 2 to 4 microges-

tures that best matched the command (Figure 3a). The selection was based on the personal experience of the participant and their interaction habits with prior smart devices (phones, tablets, etc.). For example, for the command “scroll left/right”, the highest-ranked gestures matched the direction of the command, such as gesture w (index finger scrolls left/right with the pronated forearm and the hand in index fist posture, Figure 1). After selecting the 2 to 4 microgestures for a command, participants rated each selected microgesture on four characteristics, preference, match, comfort, and privacy, using an 8-point Likert scale (0 = low, 7 = high, Figure 3b). Preference was used to rate the microgesture from most to least preferred. Match indicated how suitable the microgesture was to complete the command. Comfort indicated how easy or comfortable the microgesture was to perform repeatedly. Privacy indicated how easily bystanders could notice the microgesture if performed in a public setting, with higher scores associated with a higher level of privacy. The four characteristics were defined for each participant before the microgestures were rated. The ratings for comfort and privacy were for the microgesture and not the command; therefore, theoretically, these ratings should be the same when selected for different commands. Participants were also encouraged to demonstrate their own unique microgesture for a command or they could select a microgesture from the 33 microgestures if it was not represented among the eight pre-selected microgestures. If they demonstrated a new microgesture, it was video recorded via Zoom<sup>TM</sup> (San Jose, CA, USA).



**Figure 3.** The selection interface for each command consisted of two screens, (a) the first screen showed the pre-selected eight microgestures for the command and participants selected the 2 to 4 microgestures that best matched the command; (b) the second screen was for rating the selected microgestures on preference, match, comfort, and privacy.

## 2.6. Initial and Final Questionnaires (Appendix A.1)

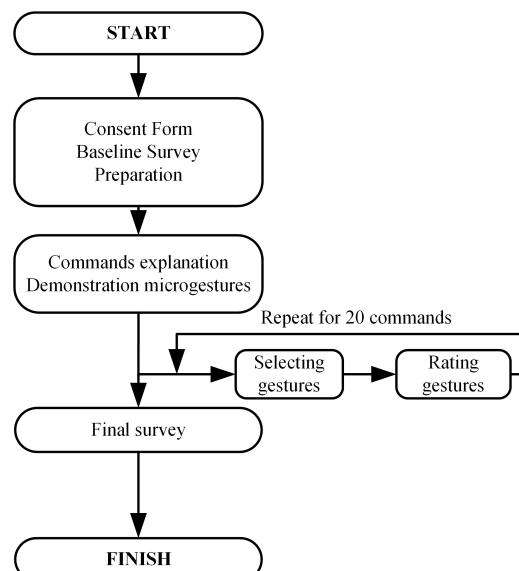
At the start of the study, participants completed an initial questionnaire (Qualtrics, Seattle, WA, USA) to collect demographic information and prior hand–computer interaction experience. Participants were asked to rate the ease of use of different input methods including touch, voice, controller, and hand gestures using the question “How easy is it for you to interact with smart devices through this mode” [1 (most difficult) to 10 (least difficult)].

After completing the matching of microgestures to all the commands, participants were asked to estimate the fatigue in their neck/shoulder/arm, forearm, and wrist/hand regions on an 11-point Likert scale with the verbal anchors of “no fatigue” to “worst imaginable fatigue” after repeatedly performing the 3D finger microgestures during the study.

At the end of the study, based on prior experience using VR/AR HMDs and/or watching a video on “how to control AR HMDs by performing hand gestures” (HoloLens), participants ranked their preferred method (3D microgestures, controller, voice, or keyboard and mouse) of interacting with VR/AR HMDs from most (1) to least (4) preferred.

## 2.7. Experimental Procedures

Due to the COVID-19 pandemic, the experiment was conducted via Zoom with the use of screen sharing and remote control function, where users can share their screen with other users in the same virtual meeting and allow others to control their computer remotely, to execute the study while participants used their personal computers to complete all questionnaires and tasks. A flow chart of the steps of the experiment is provided in Figure 4. After providing informed consent and completing the initial questionnaire, participants were asked to watch a video demonstration on how to manipulate a virtual object with hand gestures while using an AR HMD (Hololens). Next, the researcher remotely shared their computer screen with the participants and explained the various commands. After the commands were reviewed, participants were asked to perform the 33 microgestures while resting the forearm and hand on the table. They followed along with the training interface (Figure 2) while additional verbal instructions were provided by the researcher. Participants were asked to place the hand within the capture field of the camera mounted to their computer so the researcher could ensure the microgestures were performed properly; corrective instruction was provided as needed. Once participants could perform all of the microgestures correctly and demonstrated the ability to control the researcher’s host computer successfully, they proceeded to selecting and rating the microgestures for each command (Figure 3). The selected gestures and their ratings were automatically saved by the interface and the screen was recorded throughout the experiment.



**Figure 4.** Flowchart of the experimental steps.

### 3. Data Analysis

The data were processed with MATLAB 9.4, and statistical analysis was conducted with the R language. The rating scores for the gesture–command combinations were normalized to a value with a mean of 10 and a standard deviation of 1 across participants to adjust for differences in rating scales.

The ultimate assignment of a gesture to a given command, to build the proposed gesture-command set, was primarily determined by its popularity among participants.

The agreement score reflects the consensus among users for selecting the same micro-gesture for a given command. For this study, a modified agreement Equation (1) [47] was used to calculate the agreement score:

$$AR(r) = \frac{|P|}{|P| - 1} \sum_{P_i \subseteq P} \left( \frac{|P_i|}{|P|} \right)_2 - \frac{1}{|P| - 1}, \quad (1)$$

where  $P$  is the number of different gestures selected for command  $r$ , and  $P_i$  is the number of participants who selected the identical gesture for command  $r$ . As an example of an agreement score calculation, the command *shrink/enlarge* had seven different gestures selected by participants and the number of participants selecting each of the seven gestures was 37, 34, 32, 3, 2, 1, and 1. Therefore, the agreement score for the command *shrink/enlarge* was:

$$AR(\text{shrink/enlarge}) = \frac{7}{6} \left( \left( \frac{37}{110} \right)^2 + \left( \frac{34}{110} \right)^2 + \left( \frac{32}{110} \right)^2 + \left( \frac{3}{110} \right)^2 + \left( \frac{2}{110} \right)^2 + \left( \frac{1}{110} \right)^2 + \left( \frac{1}{110} \right)^2 \right) - \left( \frac{1}{6} \right) = 0.29$$

Differences in comfort and preference between microgestures were analyzed using a repeated-measures ANOVA. For example, estimated comfort ratings between microgestures performed with a pronated forearm (palm down forearm) versus a neutral forearm (thumb-up forearm) were compared as were differences between familiar (gestures  $p$  and  $q$  from the *okay posture*) versus unfamiliar (gesture  $f$  *thumb tip slides on the index finger*) gestures, identified by a “\*” in Figure 1.

The preference, match, comfort, and privacy ratings for each gesture–command combination reflect different attributes or dimensions of the combination. The agreement scores indicated the consensus of the groups of gestures and commands among users. The importance score was the rating by VR/AR developers on the importance of that command to VR/AR systems (Table 1). The extent that the different dimensions were correlated was evaluated using Pearson’s correlation coefficient. The preference of interaction methods for VR/AR systems was evaluated using the Skillings–Mack test.

## 4. Results

### 4.1. Participants

Forty participants with a mean age of 26.4 ( $SD = 4.8$ ) years completed the study; 19 were female, and 22, 15, and 3 were from China, the US, and Europe, respectively. The average time participants reported spending on smartphones and tablets was 32.9 ( $SD = 15.8$ ) hours/week and 11.4 ( $SD = 13.2$ ) hours/week, respectively. Twenty-eight participants reported having experience using VR or AR devices, and nine had previous experience controlling an AR device (HoloLens) with hand gestures. Their ratings on ease of use for touch, voice, controller, and hand gestures when interacting with smart devices (phones, tablets, VR, AR, etc.) are presented in Table 4. The touchscreen was rated the easiest to use while hand gestures were the least easy to use.

**Table 4.** Mean ( $SD$ ) rating of ease <sup>a</sup> of use for the four methods of interacting with smart devices.

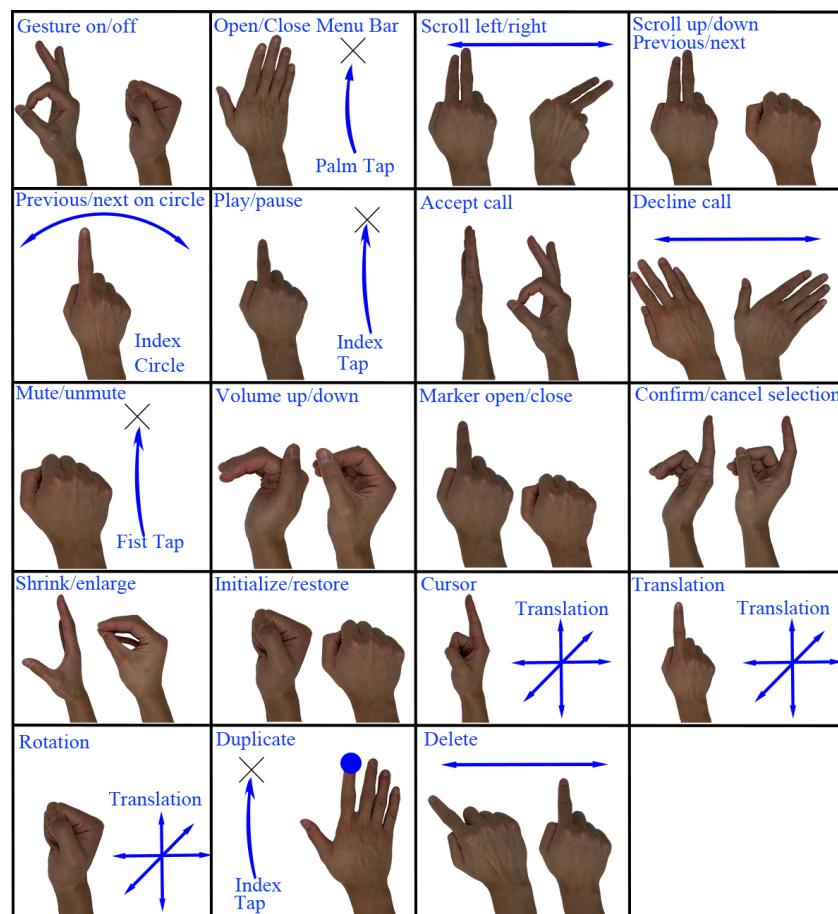
Touch	Voice	Controller	Hand Gesture
8.8 (1.7)	5.2 (2.4)	7.0 (1.9)	5.0 (2.2)

<sup>a</sup> rating score from 0 (most difficult) to 10 (least difficult).

#### 4.2. The Mapping between the Proposed Microgestures and Commands

A total of 2113 microgesture-command combinations (40 participants \* (2 to 4) microgestures \* 20 commands) were selected by participants for the 20 commands. Only two new microgestures were proposed by the participants, and, since they were so few, the two new microgestures were not considered for further analysis.

A final set of 19 microgestures linked to 20 commands is proposed (Figure 5) primarily based on popularity. Gesture popularity was determined by the number of participants who assigned the same microgesture to a command. For example, gesture  $x$  (*index and middle finger swipe left/right with the pronated forearm*) was assigned to two commands *previous/next* and *scroll left/right* because of its high popularity, with values of 27 and 34, respectively.



**Figure 5.** The assignment of microgestures to 20 commands.

The commands *accept/decline a call* were assigned different gestures. Gesture  $p$  (*form okay posture from the palm posture with the forearm in neutral position*) was preferred by 24 participants for command *accept a call* and gesture  $y$  (*palm swipes on the table back and forth with the pronated forearm*) was most preferred by 21 participants for command *decline a call*. Similarly, the commands *duplicate* and *delete the object* were assigned different gestures. Gestures  $k$  (*index finger taps on the table with the palm down*) was selected by 34 participants for command *duplicate the object* and gesture  $w$  (*index finger swipes back and forth with the pronated forearm and the hand in index fist posture*) was selected by 22 participants for command *delete the object*.

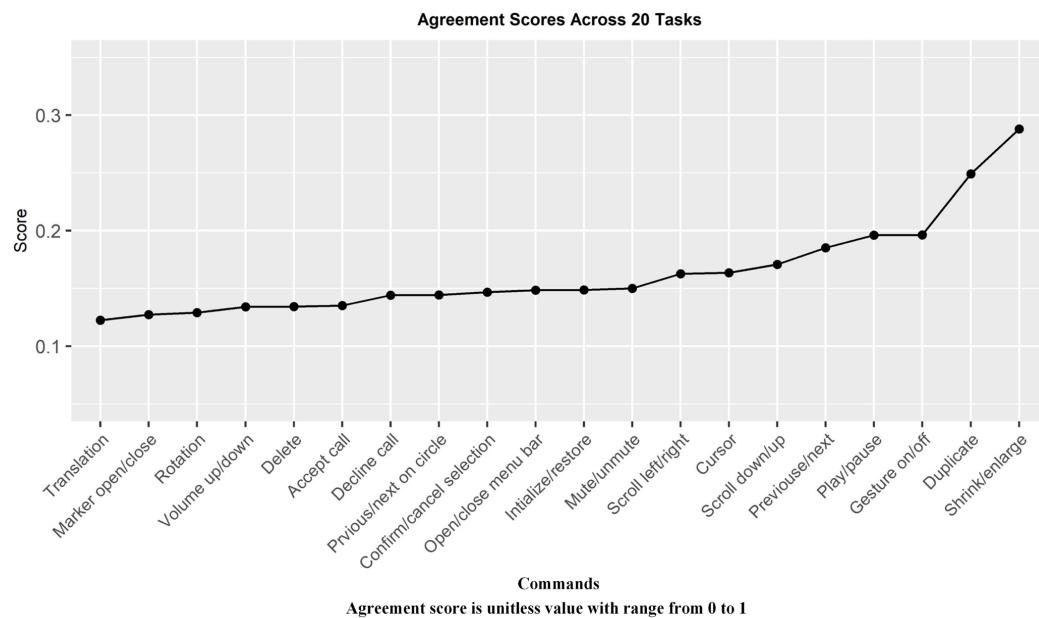
Twenty-nine participants preferred performing gesture  $q$  (*form okay posture from the fist posture or the reversed one*) to activate the interface to start hand gesture recognition to detect hand gestures as the input command or deactivate the input interface. To display or close the menu bar, gesture  $l$  (*palm taps on the table once or twice*) was selected and was the most popular among participants. The commands *scroll up/down* were used to select

different menu bars to change the system settings, and the most popular microgesture was gesture *t* (*index and middle finger scratch to the toward/away with the pronated arm*), selected by 28 participants. When the menu bar was designed in the form of a circle, 23 participants expressed that they would like to perform gesture *e* (*index finger circles CW/CCW with the pronated forearm*) to select the previous or next circled item. If participants were provided with VR to watch a movie, 29 of them selected gesture *j* (*index finger taps on the table once or twice with the pronated forearm*) to play or pause the video. To mute or unmute the system sound while interacting with the VR/AR, 26 participants preferred performing gesture *m* (*fist knocks on the table once or twice with the pronated forearm*). Moreover, 24 participants preferred using gesture *f* (*thumb slides on the index finger to the left/right with curled palm and a neutral forearm*) to turn up or turn down the volume. The popularity of gestures to activate or deactivate the marker was lower, with the most popular selection by 16 participants being gesture *r* (*index finger scratches forward/backward with the pronated arm*).

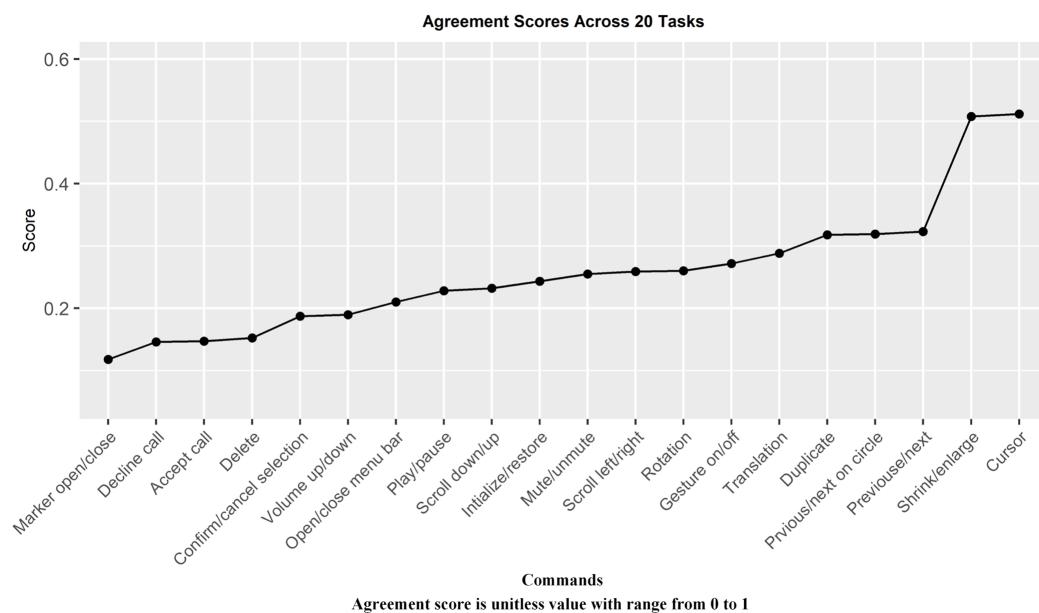
Moving the cursor is an important command. Twenty-five participants assigned gesture *z* (*the hand in index finger fist posture with the forearm in a neutral position*) to control the cursor with the conversion of the fingertip location to the movement of the cursor. Once users moved the cursor to the target, gesture *i* (*thumb tip taps on the side of the middle finger with a neutral forearm once or twice*) was selected by 18 participants to select the target or undo the selection. After confirming selecting the object, the users could manipulate the object with hand gestures. Thirty-seven participants preferred gesture *a* (*grab/expand with the forearm in the neutral position*) to shrink/enlarge the object. Moreover, popularity for the gesture *a* and command *shrink/enlarge* was the highest among all microgesture-command combinations. To translate an object in a VE, the most popular selection was gesture *aa* (*index finger fist posture with the pronated forearm*), with the object following the motion of the index fingertip, similar to controlling the cursor. To rotate the object, we designed the pattern that the conversion of hand translation to the rotation angles for objects. The most popular gesture for rotating the object was gesture *af* (*fist posture with the forearm in neutral position*) that conversion of fist distance moved to the rotation angles. Gesture *n* (*fist posture with the forearm rotates from the pronated position to a neutral position or vice versa*) was recommended by 22 participants to recover or initiate the settings for the target.

#### 4.3. Agreement Score

The agreement scores were calculated in two ways. First, all selected microgestures for each command (2 to 4 gestures were selected by each participant) were used to calculate the agreement score using Equation (1) (Figure 6). In this case, the command *shrink/enlarge* had the highest agreement score and *object translation* had the lowest score. However, prior studies [37,41,42,44] calculated the agreement score using the single most preferred gesture from each participant for a given command. Thus, a second method was used to calculate the agreement score by only using the most preferred microgesture for a command (Figure 7). In this case, the command *cursor* had the highest agreement score. The impact of the number of microgestures assigned to a command was of interest. Therefore, we performed an ANOVA analysis to compare differences in agreement scores between the two methods of calculating agreement and found a significant difference ( $p = 0.006$ ). However, the correlation between the two methods of calculating the agreement scores was strongly correlated ( $R = 0.62$ ).



**Figure 6.** The agreement scores for the 20 commands when calculated using the 2 to 4 microgestures preferred for a command.



**Figure 7.** The agreement scores for the 20 commands when calculated based on the most preferred microgesture assigned to a command.

#### 4.4. Preference and Comfort of Ratings for Unfamiliar and Familiar Microgestures

Participants had prior experience performing touch gestures on hand-held devices and using mid-air gestures to communicate with others in daily life. A microgesture was identified as a *familiar* one if it was similar to gestures that participants used before. Correspondingly, microgestures that participants never or rarely performed were identified as *unfamiliar* ones. The microgestures identified with “\*\*” in Figure 1 were unfamiliar to participants. ANOVA was performed to evaluate whether users’ past experience with a gesture had an impact on preference and comfort ratings while assigning microgestures to VR/AR commands. The ratings on preference and comfort were normalized to a value with a mean of 10 and a standard deviation of 1 across participants before conducting the statistical analysis. There was little difference in preference and comfort ratings between familiar and unfamiliar microgestures ( $p > 0.10$ , Table 5).

**Table 5.** The mean (*SD*) ratings on comfort and preference for unfamiliar and familiar gestures.

Gestures	Unfamiliar ( <i>n</i> = 13)	Familiar ( <i>n</i> = 20)	<i>p</i> -Value
Preference	9.96 (0.22)	10.04 (0.22)	0.14
Comfort	10.02 (0.37)	9.93 (0.37)	0.91

#### 4.5. Comfort Ratings for Microgestures with Different Forearm Postures

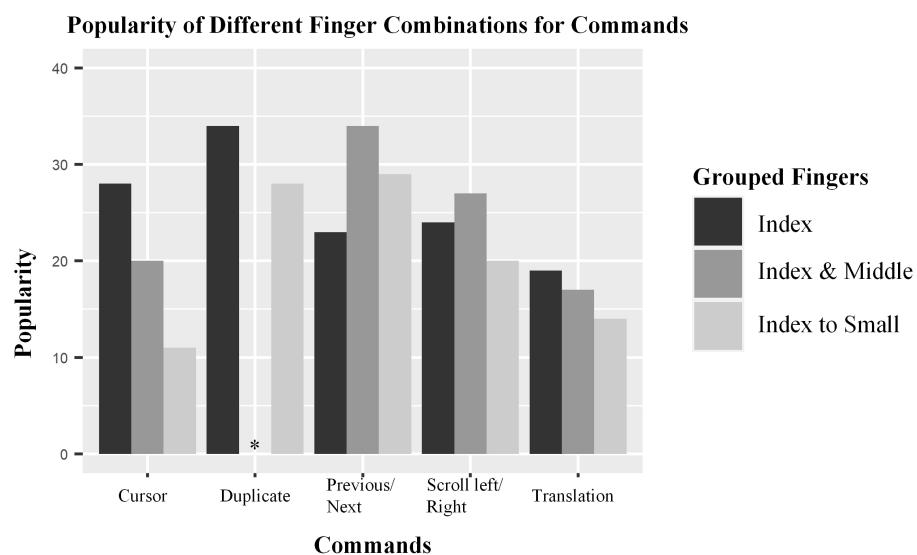
All microgestures (Figure 1) were designed with the forearm posture between a pronated position (palm down) to a neutral position (thumb-up). The impact of forearm posture on user comfort while performing microgestures was of interest. The 40 participants assigned 1106 microgestures with a pronated posture and 959 microgestures with a neutral forearm posture to the 20 commands. ANOVA analysis was used to compare the difference in comfort between microgestures with the different forearm postures. Comfort ratings were normalized to a mean value of 10 and a standard deviation of 1 and averaged across 40 participants. The difference in comfort ratings between microgestures performed with a pronated forearm and those performed with a neutral forearm was not significant (*p* = 0.78, Table 6).

**Table 6.** The comfort score and *p*-value of the neutral and pronated forearm postures.

Postures	Neutral ( <i>n</i> = 16)	Pronated ( <i>n</i> = 15)	<i>p</i> -Value
Comfort	9.99 (0.44)	9.94 (0.31)	0.78

#### 4.6. Preference of Different Fingers Combinations for Microgestures

Some microgestures can be formed by different combinations of fingers. For five commands, including *cursor*, *translation*, *scroll left/right*, *previous/next*, *duplicate*, participants selected a microgesture from a set of microgestures that were similar, except they involved either just the index finger; the index and middle finger; or they used all four fingers. The popularities for these three different finger combinations were compared across the five commands (Figure 8). There was no strong preference for any of the three different finger combinations.

**Figure 8.** The number of participants (popularity) who assigned gestures with different finger combinations to five commands, \* indicating that such finger combination was not available for this command.

#### 4.7. Correlation of Various Dimensions of the Proposed Microgesture–Command Set

The correlations between the different dimension ratings (e.g., match, comfort, privacy, popularity, agreement, importance) for the proposed gesture-command set is presented in Table 7. Popularity was strongly correlated with the agreement score, preference was strongly correlated with match and popularity, and, surprisingly, comfort was strongly correlated with privacy. However, privacy was negatively correlated with the match.

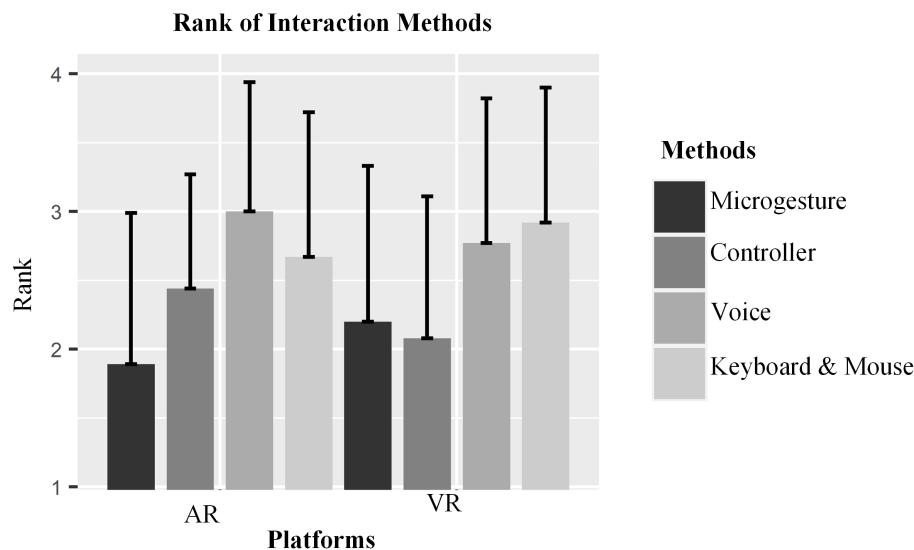
**Table 7.** The correlation ( $R$ ) between different ratings <sup>a</sup> of microgestures on the proposed microgesture–command set.

Correlation	Preference	Match	Comfort	Privacy	Popularity	Agreement Score
Match	0.51	-	-	-	-	-
Comfort	0.30	0.04	-	-	-	-
Privacy	-0.03	-0.26	0.75	-	-	-
Popularity	0.52	0.41	-0.01	-0.13	-	-
Agreement Score	0.26	0.15	0.05	-0.02	0.80	-
Importance Score	0.33	0.21	-0.03	-0.30	-0.05	-0.13

<sup>a</sup> the ratings for calculating the correlations among participants from China, and the US and Europe are in Tables A1 and A2 in the Appendix A.2.

#### 4.8. Ranking of Different Methods of Interacting with VR/AR

At the end of the study, participants completed a final questionnaire and rank-ordered four methods of interaction for VR and AR. Microgestures were ranked as the first or second preferred interaction method for AR and VR while voice and keyboard/mouse were the least preferred methods (Figure 9). However, the differences in ranking between methods were not significant (AR:  $p = 0.25$ , VR:  $p = 0.07$ ).



**Figure 9.** Ranking of four interaction methods for VR and AR systems averaged across 40 participants (1 = most preferred; 4 = least preferred). [In addition, see Table A3 in the Appendix A.3.]

#### 4.9. Differences and Correlations in Ratings on the Proposed Microgesture Set between Participants from Different Countries

To determine whether national background influenced microgesture selection, ANOVA was used to compare the differences in ratings between participants from China and those from US and Europe. Participants from China rated comfort and privacy, for the proposed microgestures, higher than participants from the US and Europe ( $p < 0.05$ , Table 8).

**Table 8.** Ratings <sup>a</sup> for the proposed microgestures for participants from the US and Europe and from China.

	USA & Europe ( <i>n</i> = 18)	China ( <i>n</i> = 22)	<i>p</i> -Value
Preference	9.46 (1.02)	9.52 (0.43)	0.81
Match	9.23 (1.01)	9.21 (0.46)	0.93
Comfort	9.43 (0.81)	10.27 (0.82)	0.01
Privacy	8.85 (1.06)	10.2 (0.73)	0.00
Popularity	12.1 (3.16)	12.5 (3.43)	0.53

<sup>a</sup> also see Tables A1 and A2 in the Appendix A.2.

The consistency of ratings on the proposed microgesture set between participants with different national backgrounds (China vs. US and Europe) was evaluated with Pearson's Correlation Coefficient (Table 9). The popularity of the proposed microgesture set was strongly correlated between participants from China and those from US and Europe ( $R = 0.65$ ).

**Table 9.** Correlations ( $R$ ) of ratings <sup>a</sup> for the proposed microgestures between participants from China and the US and Europe.

	Preference	Match	Comfort	Privacy	Popularity
Correlation	0.12	-0.01	-0.13	-0.14	0.65

<sup>a</sup> also see Tables A1 and A2 in the Appendix A.2.

## 5. Discussion

Based on the results of this study, a 3D microgesture set of 19 microgestures is proposed for 20 commands used in a VR or AR system, determined by popularity with adjustment. The microgestures proposed in this study were from a set of microgestures designed by ergonomists who have experience in designing gestures and tools for human-computer interaction that minimize discomfort and fatigue and optimize interaction efficiency. Users performed 3D microgestures with the forearm and hand on a table to reduce the loads to the neck and shoulder muscles [48]. Therefore, with these gestures, users should be able to perform the gestures repeatedly while interacting with VR/AR displays. Furthermore, the gesture-command combinations were selected by participants based on their prior experience and familiarity with gestures used for touch screens making the findings acceptable to others. Additionally, the assignment of microgestures to commands was completed by 40 participants with multinational backgrounds improving acceptability across cultures. There was a significant difference in comfort and privacy ratings on the proposed microgesture set between participants from China and those from the US and Europe ( $p < 0.05$ ) which demonstrates that nationality has some effect on user preference of microgestures. Therefore, the development and evaluation of universal gesture sets should include participants of different nationalities.

The commands assigned to microgestures will vary from application to application. It is likely that a particular application will just use a subset of the proposed microgesture-command set. For example, microgestures used to watch an interactive movie with a VR HMD are likely to be needed for just a few commands, such as *volume up/down* and *pause*. A larger set of command and microgestures are likely to be needed by CAD designers using VR/AR HMDs, who may perform gestures for a prolonged period while conceptualizing their ideas. More research may be needed to optimize the microgesture-command set for specific applications. However, the generalized microgesture-command set developed here can be a starting point for such research.

For the proposed gesture-command set, the correlation between the characteristics of preference and match was high, indicating that assigning gestures to commands based on popularity is a reasonable approach. Interestingly, comfort had a very strong correlation with privacy ( $R = 0.75$ ). This may be due to comfort including assessment of both physical and emotional comfort; gestures involving very small finger or hand motions are both less

physically demanding and are not easily noticed by others. The agreement score had a weak positive relationship with preference and comfort, which may have been influenced by the way the experiment was conducted. The agreement score calculated based on multiple gestures selected for a given command in this study was lower than observed in prior studies [37,42–44]. Recalculating the agreement score based on the most preferred microgesture for commands increased the score. In addition, relative to other studies, subjects had to select two or more gestures from a larger number of gestures for a given command and this reduced the agreement score.

The difference in comfort while performing gestures with a pronated or neutral forearm indicated that users had no preference between the forearm postures ( $p > 0.10$ ). It may be difficult to note fatigue or discomfort when gestures are performed very briefly. During the matching, subjects were not required to repeatedly perform the gesture. However, future designs of 3D microgestures should avoid fully supinated postures (palm up) [34]. The learnability of familiar microgestures should be high, and, therefore, it would be expected that users could incorporate these gestures with VR/AR with minimal training. It was surprising to find that there was little difference in comfort and preference between familiar and unfamiliar gestures ( $p > 0.10$ , Table 5). During training, participants performed all of the microgestures properly, and the learnability for unfamiliar microgestures was similar to familiar microgestures. In addition, participants were open to performing unfamiliar microgestures to replace the existing gestures as long as the gesture matched the task and was easy and comfortable to perform. For example, gesture e (index finger circles CW/CCW) is familiar and intuitive to interpret as a command to adjust volume up or down and was selected by 10 participants. Surprisingly, 24 participants preferred gesture f (*thumb slides on the side of the index finger*) for the volume command, a gesture that is not widely used.

The comparison of the number of fingers to move while performing a microgesture, e.g., index, index and middle, vs. all fingers, did not reveal a strong preference. This finding may be useful for gesture designers—that any of the three types of finger movements can be used. A prior study found that, for individual digit movement, the thumb was most popular, while the little finger was least popular in performing microgestures [37].

Current interfaces are usually designed to require users to input parameters along with the commands through typing which may cost extra time and increase workload—such as users possibly wanting to change the mapping weights between the translation distance of hand in physical 3D spaces to the distance of virtual cursor in VEs based on their actual demand [25], similar to the value of dpi (dots per inch) for a conventional computer mouse, both of which impact the sensitivity of HCI input tools. Therefore, the number of fingers used in a microgesture can be an independent input parameter with some commands in order to skip the extra step of typing on the interfaces. Moreover, using a different number of fingers as a parameter for a command is intuitive to understand, thereby reducing memory load, improving interaction efficiency, and making interaction more natural. The conversion of numbers of fingers to parameters can be applied in various scenarios. Commands such as *acceleration*, *slow down*, and *fast forward* are difficult for users to execute while gaming or wandering in a VE. Users could perform a gesture with different numbers of fingers pointing forward to control the magnitude of navigation speed. For example, a hand pointing forward with the index finger, index and middle finger, or index to small finger extended could represent 1x, 2x, or 4x speed, respectively. The use of VR and AR devices can facilitate designers to conceptualize their ideas [22,49] but repeatedly creating components can be annoying. From the findings of this study, performing the tapping gestures with index, index and middle, or index to small fingers, respectively, could duplicate the component at different speeds which can accelerate the process of idea conceptualization.

In contrast to prior studies, the aim of recruiting participants with multi-cultural backgrounds was to build mappings between 3D microgestures and commands so they were not required to design gestures for a given command with a ‘think out loud’ strategy [41].

Thus, a comparison to the selection of microgestures for similar commands in other elicitation studies is vital to support the feasibility of our experimental design. Gesture *q* (*form the okay posture*) was assigned by 29 participants to activate the interface to accept hand gestures as input. The same gesture was preferred by users from China while gesture *thumb up* was preferred by users from the US to commit the command [42,44]. From our study, gesture *form the okay posture* can be a replacement for the gesture *thumb up* to execute commands like *confirmation*, *accept*, and *activation* for users from the US and Europe.

Studies [25–27] pointed out that the consensus on the importance of designing gestures for commonly used commands *navigation* and *selection*. The gesture *index finger points forward with the forearm in a neutral position* was adopted to control the 3D virtual cursor while the gesture *index finger points forward with the forearm in a pronated position* was utilized to translate the virtual target which was the same as prior studies [25,27,41]. Moving the cursor movement is usually required for selecting a virtual object, so the choices of gesture combinations for controlling the virtual cursor and for object selection are important so that users can execute the sequential commands continuously with fluid movements and transitions. The microgestures assigned to the commands' cursor and selection were similar to those of a prior study where users preferred the gestures *index finger points forward with the forearm in a neutral position* and *thumb taps on the side of the middle finger with the index finger pointing forward and forearm in a neutral position* [27]. Similarly, the gesture *index finger points forward with the palm down* and the gesture *index finger taps on the table with the palm down* were selected for the two commands, respectively, by more than 20 participants. Therefore, gesture developers may provide the above two combinations as choices for users to control the cursor and select an object.

The commands *confirm*, *reject*, and *undo* are commonly used in human-computer interaction. The microgesture *index fingertip taps thumb tip* was performed to complete commands *select*, *play*, *next*, and *accept*, while gesture *middle fingertip taps thumb tip* was assigned to complete *pause*, *previous*, and *reject* commands in the study [37]. However, such connections between gesture and commands may not be intuitive. In our study, participants preferred performing gestures *palm scrolls repeatedly with palm down* and *index finger scrolls repeatedly with palm down* to reject the call and delete the object, respectively. The same gesture was proposed to undo an operation while interacting through a touch screen [41]. For the command *rotation objects*, it has been shown that users preferred to match the rotation of the hand wrist with the rotation of the object in a metaphor [42]. However, the rotation of the object based on such mapping is limited by the dexterity of the hand/wrist, and over-extreme hand/wrist postures pose a health risk [34]. Importantly, the accurate detection of the hand/wrist angle for rotating an object is a challenge, and hand tremors are inevitable while performing mid-air gestures. To address the above problems, we converted the distance of hand translation to the amount of object rotation. Users could rotate an object with high accuracy while removing the negative consequences caused by hand tremors in mid-air.

In the future, users may work or engage in recreational activities with VR or AR HMD for long periods of time. Watching a movie with a virtual display created with a large size and high resolution is a potential replacement for cinemas. Thus, gestures to browse movies and play/pause a video are desired. Gesture *x* (*index and middle fingers scroll left/right with the pronated forearm*) was assigned to show the previous or next item. For the command *play/pause*, gesture *j* (*index finger taps on the table with the forearm in pronated position*) was the most popular gesture among users. The same gesture was proposed in another study to play the selected channel while watching TV [39].

From the proposed microgesture-command sets, we can acknowledge that tapping and swiping microgestures were popular among participants which was discovered while designing gestures to control mobile devices [43,50]. The preference for tapping and swiping gestures may be caused by the users' past experience with touch screens. The preference for microgestures revealed that understanding the users' interaction habits is vital to implementing an interface based on microgestures for VR or AR systems.

The proposed 3D microgestures are not limited to commands investigated in this study (Table 1). The microgestures designed to manipulate an object could be used to set parameters for the virtual scene. For example, microgestures with the purpose of enlarging or shrinking an object could be performed to zoom in or out of a virtual scene while no object was selected. Similarly, microgestures used to translate or rotate an object could control the coordinate systems in a VE. Although the proposed 3D microgesture set is designed for VR/AR systems specifically, the use of it can be extended to other platforms using similar context-based commands. Interfaces based on hand gestures have been developed to complete secondary tasks while driving a car so that drivers can pay attention to the road safely without distraction while manipulating a control panel through a touchscreen or keypad [44,51]. However, another study [52] found that an interface based on mid-air gestures takes longer to complete a task and takes a higher workload compared to a touch-based system. Perhaps, performing 3D microgestures with the hands resting on the steering wheel or one forearm resting on an armrest could allow drivers to reduce muscle load, cognitive workloads, and the duration to complete a task when performing secondary tasks. Similar to feedback provided to users when resting their arms on a table, drivers may receive feedback when resting their hands on the steering wheel when compared to mid-air gestures.

## 6. Limitations and Future Work

The proposed 3D microgestures are one-handed, but some users may prefer two-handed gestures [23,42]. Additionally, due to the Covid-19 pandemic, this study was conducted online preventing close observation of user habits while they assigned gestures to commands. If participants had actually used a VR or AR HMD during the experiment, they may have preferred different microgestures. Even though participants were encouraged to design different microgestures during the study, only two different microgestures were proposed. A face-to-face study rather than an online study might have led to more new gestures being proposed by participants. Although great care was taken to create a list of commands that would be representative of the commands users may perform while using VR or AR devices, it is possible that some important commands were excluded from this study.

People with cognitive or motor disabilities may have difficulty performing microgestures. Video capture systems and machine learning could be used to recognize the microgestures and intentions from individuals with disabilities and may improve their ability to interact with computers, especially if they have difficulty using hand-held controllers or other input devices. Further evaluation of this microgesture set with qualitative and quantitative measurements of comfort, efficiency, accessibility, and acceptability is essential. Microgestures should be compared to other modes of input (e.g., hand-held controller) on productivity, error, comfort, and other usability factors when subjects use VR/AR. It is likely that the acceptability of gesture input will be different for VR vs. AR.

## 7. Conclusions

A 3D microgesture set was designed by ergonomists as an alternative to the use of hand-held controllers and for interacting with VR or AR displays. The mappings between microgestures and commands were guided by user preference with users from multicultural backgrounds to increase international application and acceptance. The use of 3D microgestures for prolonged periods may reduce fatigue and discomfort associated with prolonged use when compared to holding a controller or using large arm mid-air gestures. An interesting finding was that participants were open to using new, unfamiliar gestures instead of more familiar ones for some commands. The findings of this study provide new insights into user preference of microgestures and the importance of users' past interaction experience on the selection of microgestures for VR and AR devices. The pattern of using a different number of fingers as input parameters while performing microgestures could improve the interaction efficiency and reduce workload. Further research is needed to

determine whether this or another microgesture set can improve interaction comfort and efficiency with VR, AR compared to using a controller or other input device.

**Author Contributions:** Microgesture design, D.R., C.H.A.; Experiment design, G.L., D.R., and C.H.A.; Software Development, G.L.; Data analysis, G.L., D.R., and C.H.A.; Data interpretation, G.L., D.R., C.H.A., Y.L., and W.S.; Writing—Original Draft Preparation, G.L.; Writing—Review and Editing, D.R., C.H.A., Y.L., and W.S. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the Key-Area Research and Development Program of Guangdong Province Grant No. 2019B010149001, the National Natural Science Foundation of China Grant No. 61960206007, and the 111 Project Grant No. B18005.

**Institutional Review Board Statement:** The study was conducted according to the guidelines of the Declaration of Helsinki and approved by the Institutional Review Board of the University of California, San Francisco (IRB#10-04700).

**Informed Consent Statement:** Informed consent was obtained from all subjects involved in the study.

**Data Availability Statement:** Data sharing is not applicable to this article.

**Acknowledgments:** We wish to thank Alan Barr for his support in this study during the COVID-19 epidemic.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Appendix A

### Appendix A.1. Questionnaire

1. Please indicate which devices (smartphones, tablets, VR, and AR) you have experience using (choose as many as appropriate):
2. Approximately, how many hours do you spend on smart devices (smartphones, tablets, VR, and AR) per week?
3. How easy is it for you to interact with your smart devices (phones, tablets, VR, and AR HMDs) using the touch screen, voice, controller, and keyboard and mouse? (0: most difficult, to 10 least difficult)
4. Which methods (hand gestures, hand-held controllers, touch screen, keyboard and mouse) have you ever used to interact with Augmented Reality HMDs? (choose as many as appropriate)
5. Which methods (hand gestures, hand-held controllers, touch screen, keyboard and mouse) have you ever used to interact with Virtual Reality HMDs? (choose as many as appropriate)
6. When using an Augmented Reality HMD, please rank your preferred method (microgestures, controller, voice, keyboard and mouse) for interacting with the device. (1) most preferred to (4) least preferred
7. When using a Virtual Reality HMD, please rank your preferred method (microgestures, controller, voice, keyboard and mouse) for interacting with the device. (1) most preferred to (4) least preferred
8. Do you experience pain in your shoulder, elbow, wrist or hand that you think caused or worsened by performing the gestures tested in the study?
9. Please rate the WORST pain felt in your Neck/Shoulder caused by Performing gestures during the study. (0: no pain to 10 worst imaginable pain)
10. Please rate the WORST pain felt in your Elbow caused by Performing gestures during the study. (0: no pain to 10 worst imaginable pain)
11. Please rate the WORST pain felt in your Wrist/Hand caused by Performing gestures during the study. (0: no pain to 10 worst imaginable pain)

### Appendix A.2. Ratings and Popularity of the Proposed Microgesture Set

**Table A1.** The ratings and popularity of the proposed microgesture set among participants from the US and Europe ( $n = 18$ ).

Commands	Performance	Match	Comfort	Privacy	Popularity
1	8.44	8.67	8.98	8.58	14
2	8.17	8.50	8.95	8.25	11
3	9.33	9.09	9.23	9.02	12
4	7.86	7.73	9.33	7.61	13
5	9.97	9.99	8.44	8.34	17
6	8.09	8.18	7.94	8.08	13
7	10.09	9.28	9.85	9.27	13
8	9.40	9.57	9.60	7.72	13
9	10.49	10.30	9.03	8.63	11
10	8.27	8.53	8.73	8.03	13
11	8.75	7.86	9.94	7.61	14
12	10.35	11.13	9.02	9.21	5
13	10.15	10.33	10.08	10.13	8
14	10.77	10.88	9.00	9.08	17
15	9.90	10.34	10.15	10.23	8
16	9.49	8.70	9.44	8.15	15
17	9.46	9.97	9.91	10.39	7
18	12.04	7.84	12.00	11.76	10
19	8.95	9.12	9.48	8.86	16
20	9.23	8.60	9.58	7.99	12

**Table A2.** The ratings and popularity of the proposed microgesture set among participants from China ( $n = 22$ ).

Commands	Performance	Match	Comfort	Privacy	Popularity
1	10.15	9.67	10.87	9.85	15
2	9.14	8.86	10.11	10.68	11
3	9.72	8.76	11.32	11.28	15
4	9.40	9.33	11.56	11.56	15
5	9.23	8.95	10.35	10.48	17
6	9.37	8.97	10.34	9.87	10
7	9.56	8.93	9.38	9.76	16
8	9.94	10.32	11.30	11.18	11
9	9.55	9.83	11.30	11.10	10
10	9.58	9.31	9.61	9.57	13
11	9.65	9.14	10.70	9.36	10
12	9.25	8.94	9.61	10.27	9
13	9.70	9.41	9.59	10.00	10
14	10.33	9.05	10.68	10.86	20
15	8.80	8.27	9.45	10.11	14
16	8.40	8.79	8.68	10.05	10
17	9.44	9.63	10.57	10.44	8
18	9.77	9.54	10.16	9.55	8
19	9.44	8.84	8.91	8.97	18
20	9.97	9.60	10.92	8.99	10

### Appendix A.3. Ranking of Different Methods of Interacting with VR/AR

**Table A3.** The mean ( $SD$ ) rank of four interaction methods for VR and AR systems across 40 participants, respectively.

System	Microgestures	Controller	Voice	Keyboard & Mouse	p-Value
AR	1.89 (1.10)	2.44 (0.83)	3.00 (0.94)	2.67 (1.05)	0.25
VR	2.32 (1.17)	2.07 (1.03)	2.61 (1.05)	2.96 (0.94)	0.07

## References

- Walczak, K.; Wojciechowski, R.; Cellary, W. Dynamic interactive VR network services for education. In Proceedings of the ACM Symposium on Virtual Reality Software and Technology (VRST 2006), Limassol, Cyprus, 1–3 November 2006; pp. 277–286. [[CrossRef](#)]
- Chirico, A.; Lucidi, F.; Laurentiis, M.D.; Milanese, C.; Napoli, A.; Giordano, A. Virtual Reality in Health System: Beyond Entertainment. A Mini-Review on the Efficacy of VR During Cancer Treatment. *J. Cell. Physiol.* **2016**, *231*, 275–287. [[CrossRef](#)] [[PubMed](#)]
- De Pace, F.; Manuri, F.; Sanna, A. Augmented Reality in Industry 4.0. *AJCSIT* **2018**, *6*, 1–17. [[CrossRef](#)]
- Guo, J.; Weng, D.; Zhang, Z.; Jiang, H.; Liu, Y.; Wang, Y.; Duh, H.B. Mixed Reality Office System Based on Maslow’s Hierarchy of Needs: Towards the Long-Term Immersion in Virtual Environments. In Proceedings of the 2019 IEEE International Symposium on Mixed and Augmented Reality (ISMAR 2019), Beijing, China, 14–18 October 2019; pp. 224–235. [[CrossRef](#)]
- Arora, R.; Kazi, R.H.; Kaufman, D.M.; Li, W.; Singh, K. MagicalHands: Mid-Air Hand Gestures for Animating in VR. In Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology (UIST 2019), New Orleans, LA, USA, 20–23 October 2019; pp. 463–477. [[CrossRef](#)]
- Zhou, F.; Duh, H.B.; Billinghamurst, M. Trends in augmented reality tracking, interaction and display: A review of ten years of ISMAR. In Proceedings of the 7th IEEE/ACM International Symposium on Mixed and Augmented Reality (ISMAR 2008), Cambridge, UK, 15–18 September 2008; pp. 193–202. [[CrossRef](#)]
- Stern, H.I.; Wachs, J.P.; Edan, Y. Human Factors for Design of Hand Gesture Human-Machine Interaction. In Proceedings of the 2006 IEEE International Conference on Systems, Man and Cybernetics (SMC 2006), Taipei, Taiwan, 8–11 October 2006; pp. 4052–4056. [[CrossRef](#)]
- Wachs, J.P.; Kölsch, M.; Stern, H.; Edan, Y. Vision-Based Hand-Gesture Applications. *Commun. ACM* **2011**, *54*, 60–71. [[CrossRef](#)]
- Stern, H.I.; Wachs, J.P.; Edan, Y. Optimal Hand Gesture Vocabulary Design Using Psycho-Physiological and Technical Factors. In Proceedings of the 7th International Conference on Automatic Face and Gesture Recognition (FGR 2006), Southampton, UK, 10–12 April 2006; pp. 257–262. [[CrossRef](#)]
- Speicher, M.; Nebeling, M. GestureWiz: A Human-Powered Gesture Design Environment for User Interface Prototypes. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI 2018), Montreal, QC, Canada, 21–28 April 2018; pp. 1–11. [[CrossRef](#)]
- Mo, G.B.; Dudley, J.J.; Kristensson, P.O. Gesture Knitter: A Hand Gesture Design Tool for Head-Mounted Mixed Reality Applications. In Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI 2021), Yokohama, Japan, 8–13 May 2021; pp. 291:1–291:13. [[CrossRef](#)]
- Gugenheimer, J.; Dobbelstein, D.; Winkler, S.; Haas, G.; Rukzio, E. FaceTouch: Enabling Touch Interaction in Display Fixed UIs for Mobile Virtual Reality. In Proceedings of the 29nd Annual ACM Symposium on User Interface Software and Technology (UIST 2016), Tokyo, Japan, 16–19 October 2016; pp. 49–60. [[CrossRef](#)]
- Kranzlmuller, D.; Reitinger, B.; Hackl, I.; Volkert, J. Voice controlled virtual reality and its perspectives for everyday life. *ITG-Fachbericht* **2001**, 101–107. [[CrossRef](#)]
- Osking, H.; Doucette, J.A. Enhancing Emotional Effectiveness of Virtual-Reality Experiences with Voice Control Interfaces. In Proceedings of the 5th International Conference on Immersive Learning (iLRN 2019), London, UK, 23–27 June 2019; pp. 199–209. [[CrossRef](#)]
- Wu, C.M.; Hsu, C.W.; Lee, T.K.; Smith, S. A virtual reality keyboard with realistic haptic feedback in a fully immersive virtual environment. *Virtual Real.* **2016**, *21*, 19–29. [[CrossRef](#)]
- Lin, J.; Han, P.H.; Lee, J.Y.; Chen, Y.; Chang, T.; Chen, K.; Hung, Y.A. Visualizing the Keyboard in Virtual Reality for Enhancing Immersive Experience. In Proceedings of the ACM SIGGRAPH 2017 Posters (SIGGRAPH 2017), Los Angeles, CA, USA, 30 July–3 August 2017; pp. 1–2. [[CrossRef](#)]
- Saredakis, D.; Szpak, A.; Birckhead, B.; Keage, H.A.D.; Rizzo, A.; Loetscher, T. Factors Associated with Virtual Reality Sickness in Head-Mounted Displays: A Systematic Review and Meta-Analysis. *Front. Hum. Neurosci.* **2020**, *14*, 1–17. [[CrossRef](#)]
- Tiferes, J.; Hussein, A.A.; Bisantz, A.; Higginbotham, D.J.; Sharif, M.; Kozlowski, J.; Ahmad, B.; O’Hara, R.; Wawrzyniak, N.; Guru, K. Are gestures worth a thousand words? Verbal and nonverbal communication during robot-assisted surgery. *Appl. Ergon.* **2019**, *78*, 251–262. [[CrossRef](#)]
- Vogel, D.; Balakrishnan, R. Distant Freehand Pointing and Clicking on Very Large, High Resolution Displays. In Proceedings of the 18th Annual ACM Symposium on User Interface Software and Technology (UIST 2005), Seattle, WA, USA, 23–26 October 2005; pp. 33–42. [[CrossRef](#)]
- Tao, D.; Diao, X.; Wang, T.; Guo, J.; Qu, X. Freehand interaction with large displays: Effects of body posture, interaction distance and target size on task performance, perceived usability and workload. *Appl. Ergon.* **2021**, *93*, 103370–103380. [[CrossRef](#)]
- Cohen, C.J.; Beach, G.; Foulk, G. A basic hand gesture control system for PC applications. In Proceedings of the 30th Applied Imagery Pattern Recognition Workshop (AIPR 2001), Washington, DC, USA, 10–12 October 2001; pp. 74–79. [[CrossRef](#)]
- Alkemade, R.; Verbeek, F.J.; Lukosch, S.G. On the Efficiency of a VR Hand Gesture-Based Interface for 3D Object Manipulations in Conceptual Design. *Int. J. Human-Comput. Interact.* **2017**, *33*, 882–901. [[CrossRef](#)]
- Williams, A.S.; Garcia, J.; Ortega, F. Understanding Multimodal User Gesture and Speech Behavior for Object Manipulation in Augmented Reality Using Elicitation. *IEEE Trans. Vis. Comput. Graph.* **2020**, *26*, 3479–3489. [[CrossRef](#)]

24. Lee, T.; Hollerer, R. Handy AR: Markerless Inspection of Augmented Reality Objects Using Fingertip Tracking. In Proceedings of the 11th IEEE International Symposium on Wearable Computers (ISWC 2007), Boston, MA, USA, 11–13 October 2007; pp. 23–90. [[CrossRef](#)]
25. Nai, W.; Rempel, D.; Liu, Y.; Barr, A.; Harris-Adamson, C.; Wang, Y. Performance and User Preference of Various Functions for Mapping Hand Position to Movement Velocity in a Virtual Environment. In Proceedings of the International Conference on Virtual, Augmented and Mixed Reality (VAMR 2017), Vancouver, BC, Canada, 9–14 July 2017; pp. 141–152. [[CrossRef](#)]
26. Huang, R.; Harris-Adamson, C.; Odell, D.; Rempel, D. Design of finger gestures for locomotion in virtual reality. *VRIH* **2019**, *35*, 1729–1735. [[CrossRef](#)]
27. Lin, J.; Harris-Adamson, C.; Rempel, D. The Design of Hand Gestures for Selecting Virtual Objects. *Int. J. Human-Comput. Interact.* **2019**, *1*, 1–9. [[CrossRef](#)]
28. Lindeman, R.W.; Sibert, J.L.; Hahn, J.K. Towards Usable VR: An Empirical Study of User Interfaces for Immersive Virtual Environments. In Proceedings of the 1999 CHI Conference on Human Factors in Computing Systems (CHI 1999), Pittsburgh, PA, USA, 15–20 May 1999; pp. 64–71. [[CrossRef](#)]
29. Wang, Y.; MacKenzie, C.L. The Role of Contextual Haptic and Visual Constraints on Object Manipulation in Virtual Environments. In Proceedings of the 2000 CHI Conference on Human Factors in Computing Systems (CHI 2000), Hague, The Netherlands, 1–6 April 2000; pp. 532–539. [[CrossRef](#)]
30. Sodhi, R.; Poupyrev, I.; Glisson, M.; Israr, A. AIREAL: Interactive Tactile Experiences in Free Air. *ACM Trans. Graph.* **2013**, *32*, 1–11. [[CrossRef](#)]
31. Large, D.R.; Harrington, K.; Burnett, G.; Georgiou, O. Feel the noise: Mid-air ultrasound haptics as a novel human-vehicle interaction paradigm. *Appl. Ergon.* **2019**, *81*. [[CrossRef](#)]
32. Xiao, R.; Schwarz, J.; Throm, N.; Wilson, A.D.; Benko, H. MRTouch: Adding Touch Input to Head-Mounted Mixed Reality. *IEEE Trans. Vis. Comput. Graph.* **2018**, *24*, 1653–1660. [[CrossRef](#)] [[PubMed](#)]
33. Hincapié-Ramos, J.D.; Guo, X.; Moghadasiyan, P.; Irani, P. Consumed endurance: A metric to quantify arm fatigue of mid-air interactions. In Proceedings of the 2014 CHI Conference on Human Factors in Computing Systems (CHI 2014), Toronto, ON, Canada, 26 April–1 May 2014; pp. 1063–1072. [[CrossRef](#)]
34. Rempel, D.; Camilleri, M.J.; Lee, D.L. The design of hand gestures for human–computer interaction: Lessons from sign language interpreters. *Int. J. Hum. Comput. Stud.* **2014**, *72*, 728–735. [[CrossRef](#)] [[PubMed](#)]
35. Rico, J.; Brewster, S. Usable Gestures for Mobile Interfaces: Evaluating Social Acceptability. In Proceedings of the 2010 CHI Conference on Human Factors in Computing Systems (CHI 2010), Atlanta, GA, USA, 10–15 April 2010; pp. 887–896. [[CrossRef](#)]
36. Montero, C.S.; Alexander, J.; Marshall, M.T.; Subramanian, S. Would You Do That? Understanding Social Acceptance of Gestural Interfaces. In Proceedings of the 12th International Conference on Human Computer Interaction with Mobile Devices and Services (MobileHCI 2010), Lisbon, Portugal, 7–10 September 2010; pp. 275–278. [[CrossRef](#)]
37. Chan, E.; Seyed, T.; Stuerzlinger, W.; Yang, X.; Maurer, F. User Elicitation on Single-Hand Microgestures. In Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI 2016), Montréal, QC, Canada, 22–27 April 2016; pp. 3403–3414. [[CrossRef](#)]
38. Sharma, A.; Sol, J.R.; Steimle, J. Grasping Microgestures: Eliciting Single-Hand Microgestures for Handheld Objects. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI 2019), Glasgow, UK, 4–9 May 2019; pp. 1–13. [[CrossRef](#)]
39. Van Beurdena, M.H.P.H.; IJsselsteijna, W.A.; Hopfb, K. User centered design of gesture-based interaction technolog. In Proceedings of the 2011 3DTV Conference: The True Vision-Capture, Transmission and Display of 3D Video (3DTV-CON 2011), Antalya, TR, USA, 16–18 May 2011; pp. 1–4. [[CrossRef](#)]
40. Kela, J.; Korpiä, P.; Mäntyjärvi, J.; Kallio, S.; Savino, G.; Jozzo, L.; Marca, S.D. Accelerometer-based gesture control for a design environment. *Pers. Ubiquitous Comput.* **2006**, *10*, 285–299. [[CrossRef](#)]
41. Wobbrock, J.O.; Morris, M.R.; Wilson, A.D. User-defined gestures for surface computing. In Proceedings of the 2009 CHI Conference on Human Factors in Computing Systems (CHI 2009), Boston, MA, USA, 4–9 April 2009; pp. 1083–1092. [[CrossRef](#)]
42. Pereira, A.; Wachs, J.P.; Park, K.; Rempel, D. A User-Developed 3D Hand Gesture Set for Human–Computer Interaction. *Hum. Factors* **2015**, *4*, 607–621. [[CrossRef](#)]
43. Ruiz, J.; Li, Y.; Lank, E. User-Defined Motion Gestures for Mobile Interaction. In Proceedings of the 2011 CHI Conference on Human Factors in Computing Systems (CHI 2011), Vancouver, BC, Canada, 7–12 May 2011; pp. 197–206. [[CrossRef](#)]
44. Wu, H.; Zhang, S.; Liu, J.; Qiu, J.; Zhang, X.L. The Gesture Disagreement Problem in Free-hand Gesture Interaction. *Int. J. Human-Comput. Interact.* **2019**, *35*, 1102–1114. [[CrossRef](#)]
45. Voida, S.; Podlaseck, M.; Kjeldsen, R.; Pinhanez, C. A Study on the Manipulation of 2D Objects in a Projector/Camera-Based Augmented Reality Environment. In Proceedings of the 2005 CHI Conference on Human Factors in Computing Systems (CHI 2005), Portland, OR, USA, 2–7 April 2005; pp. 611–620. [[CrossRef](#)]
46. Hinckley, K.; Baudisch, P.; Ramos, G.; Guimbretiere, F. Design and analysis of delimiters for selection-action pen gesture phrases in scriboli. In Proceedings of the 2005 CHI Conference on Human Factors in Computing Systems (CHI 2005), Portland, OR, USA, 2–7 April 2005; pp. 451–460. [[CrossRef](#)]

47. Vatavu, R.; Wobbrock, J.O. Formalizing Agreement Analysis for Elicitation Studies: New Measures, Significance Test, and Toolkit. In Proceedings of the 2015 CHI Conference on Human Factors in Computing Systems (CHI 2015), Seoul, Korea, 18–23 April 2015; pp. 1325–1334. [[CrossRef](#)]
48. Visser, B.; Korte, E.D.; Van der Kraan, I.; Kuijer, P. The effect of arm and wrist supports on the load of the upper extremity during VDU work. *Clin Biomech* **2000**, *15*, S34–S38. [[CrossRef](#)]
49. Bashabsheh, A.H.; Alzoubi, H.H.; Ali, M.Z. The application of virtual reality technology in architectural pedagogy for building constructions. *Alex. Eng. J.* **2019**, *58*, 713–723. [[CrossRef](#)]
50. Aigner, R.; Wigdor, D.J.; Benko, H.; Haller, M.; Lindlbauer, D.; Ion, A.; Zhao, S.; Tzu, J.; Valino, K.; Electronica, A. Understanding Mid-Air Hand Gestures: A Study of Human Preferences in Usage of Gesture Types for HCI. In *Microsoft Research TechReport MSR-TR-2012-111*; 2012. Available online: <https://www.microsoft.com/en-us/research/publication/understanding-mid-air-hand-gestures-a-study-of-human-preferences-in-usage-of-gesture-types-for-hci/> (accessed on 9 June 2021)
51. Alpern, M.; Minardo, K. Developing a Car Gesture Interface for Use as a Secondary Task. In Proceedings of the 2003 Extended Abstracts on Human Factors in Computing Systems (CHI 2003), Ft. Lauderdale, FL, USA, 5–10 April 2003; pp. 932–933. [[CrossRef](#)]
52. May, K.R.; Gable, T.M.; Walker, B.N. A Multimodal Air Gesture Interface for In Vehicle Menu Navigation. In Proceedings of the 6th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI 2014), Seattle, WA, USA, 17–19 September 2014; pp. 1–6. [[CrossRef](#)]