Designing Upper-Body Gesture Interaction with and for People with Spinal Muscular Atrophy in VR

Jingze Tian*
Southeast University
The Hong Kong University of Science and Technology (Guangzhou)
tjz@seu.edu.cn

Yingna Wang*
The Hong Kong University of Science and Technology (Guangzhou)
ywang885@connect.hkust-gz.edu.cn

Keye Yu
University College London
The Hong Kong University of Science and Technology (Guangzhou)
zczlky4@ucl.ac.uk

Liyi Xu
Nanjing University of the Arts
The Hong Kong University of Science and Technology (Guangzhou)
xuliyi1@nua.edu.cn

Junan Xie
The Hong Kong University of Science and Technology (Guangzhou)
jxie622@connect.hkust-gz.edu.cn

Franklin Mingzhe Li
Carnegie Mellon University
mingzhe2@cs.cmu.edu

Yafeng Niu
Southeast University
nyf@seu.edu.cn

Mingming Fan†
The Hong Kong University of Science and Technology (Guangzhou)
The Hong Kong University of Science and Technology
mingmingfan@ust.hk

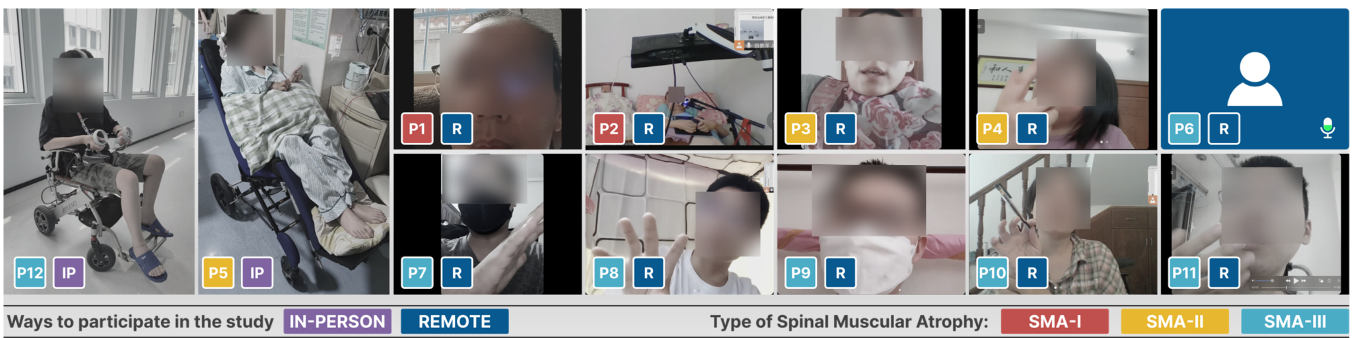


Figure 1: Twelve people with Spinal Muscular Atrophy designed upper-body gestures for VR commands. Two of them participated in person, while the other ten joined remotely. All participants actively took part in this research, and Participant 6 preferred to conduct the experiments without turning on the camera.

ABSTRACT

Recent research proposed gaze-assisted gestures to enhance interaction within virtual reality (VR), providing opportunities for people with motor impairments to experience VR. Compared to people with other motor impairments, those with Spinal Muscular Atrophy (SMA) exhibit enhanced distal limb mobility, providing them with more design space. However, it remains unknown what gaze-assisted upper-body gestures people with SMA would want and be able to perform. We conducted an elicitation study in which 12 VR-experienced people with SMA designed upper-body gestures for 26 VR commands, and collected 312 user-defined gestures. Participants predominantly favored creating gestures with their hands. The type of tasks and participants' abilities influence their choice of body parts for gesture design. Participants tended to enhance their body involvement and preferred gestures that required minimal physical effort, and were aesthetically pleasing. Our research will contribute to creating better gesture-based input methods for people with motor impairments to interact with VR.

*Both authors contributed equally to this research.
†Corresponding author

CHI '24, May 11–16, 2024, Honolulu, HI, USA

ACM ISBN 979-8-4007-0330-0/24/05…$15.00
https://doi.org/10.1145/3613904.3642884

CCS CONCEPTS

• **Human-computer interaction → Empirical studies in accessibility; Empirical studies in HCI**.

KEYWORDS

people with spinal muscular atrophy, virtual reality, upper-body gestures, user-defined gestures

ACM Reference Format:
Jingze Tian, Yingna Wang, Keye Yu, Liyi Xu, Junan Xie, Franklin Mingzhe Li, Yafeng Niu, and Mingming Fan. 2024. Designing Upper-Body Gesture Interaction with and for People with Spinal Muscular Atrophy in VR. In *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '24), May 11–16, 2024, Honolulu, HI, USA*. ACM, New York, NY, USA, 19 pages. https://doi.org/10.1145/3613904.3642884

1 INTRODUCTION

Spinal muscular atrophy (SMA) is the second most prevalent fatal autosomal recessive disorder following cystic fibrosis, afflicting approximately 1 in 6,000 to 1 in 10,000 live births [31, 38]. This disease is characterized by generalized muscle weakness and atrophy predominating in proximal limb muscles and classified into four phenotypes (SMA I, II, III, IV) based on onset age and motor function. Unlike other motor impairments such as limb loss, cerebral palsy, or amyotrophic lateral sclerosis (ALS), SMA primarily affects central body muscles, causing difficulties in breathing, swallowing, head control, and sitting [8]. People with SMA often exhibit better distal limb mobility, such as in the hands or feet, compared to proximal limb mobility. This leads to unique interaction patterns and provides them with opportunities to utilize technologies such as VR devices. However, research on SMA's impact on human-computer interaction (HCI) is still limited. The HCI community has not fully tackled the unique challenges faced by people with SMA.

Virtual reality (VR) can immerse users in a computer-generated environment that simulates reality both visually and interactively. VR enables people with limited mobility to engage in activities beyond their physical abilities and facilitates the exploration of inaccessible real-world experiences [12, 32]. Furthermore, VR promotes inclusivity and equality by granting equal mobility to all individuals [21].

However, VR devices, which are designed with implicit *Ability Assumption* [17], pose accessibility challenges for those with limited mobility, particularly due to inaccessible input methods like motion controllers and buttons [23, 26]. To utilize predominant input methods involving hand tracking or controller usage and interact with virtual objects, users need to elevate their hands to chest level, leading to prolonged muscle tension in the arms and resulting in significant fatigue. Moreover, individuals with upper-body motor impairments, particularly those with conditions like SMA, encounter greater challenges in manipulating intricate controllers and utilizing buttons that are not easily reachable. This arises from potential strength limitations that hinder their ability to access all buttons or simultaneously press and hold them.

Luckily, recent research has demonstrated the effectiveness of '*Gaze + Gestures*', the combination of eye-tracking and hand gestures, for VR operations [24, 33, 36, 45]. As eye movements are inherent parts of motor planning and precede actions [20], this approach leverages an interactive paradigm of *gaze selection* and *gesture confirmation*, capitalizing on the swift nature of eye movements with gestural dexterity. People with SMA have stronger distal upper limb mobility and can proficiently use their hands, making '*Gaze + Gestures*' interaction a viable option. '*Gaze + Gestures*', has utilized in commercial products like Apple Vision Pro [1], holds promise as an interactive modality of VR systems for people with SMA. However, it remains unclear how they would like to use this interaction paradigm for VR interaction and modify it using other upper-body parts.

Our research was motivated by the need for accessible VR input methods for people with SMA and these '*Gaze + Gestures*' works. Our user-defined gestures design method was inspired by the prior success of designing user-defined gestures for people with motor impairments in other contexts [63]. In this work, we engaged 12 people with SMA to design *upper-body gestures* for 26 VR common commands and extended the design that include *eye*, *mouth*, *face*, *head*, and *remained upper limb mobility* based on previous studies [13, 30, 41, 55]. During the study, participants watched video clips explaining each VR command and its effect (i.e., *referent*) and then designed and performed an upper-body gesture. Afterward, participants rated the effort required to design the gesture and the mental demand, physical demand, and satisfaction level of their created gestures using the 7-point Likert scale. Finally, we conducted semi-structured interviews to learn more about their considerations. In total, we collected 312 user-defined gestures and identified their preference and mental models to VR input methods.

Compared to the accessible gesture inputs for users with upper-body motor impairments for smartphone [10, 11, 62] and wearable devices [52], our study is unique in three aspects. Firstly, our gestures are designed for VR, which is beneficial [12] but inaccessible [26] for people with motor impairments. Secondly, our research specifically targets people with SMA. We use user-defined upper-body gestures to explore VR input accessible for people with motor impairments, grounded in their ability and creativity. Finally, compared to previous user-defined gestures research for people with motor impairments [10, 11], we expanded the gesture design space to the upper body, which provides more freedom for them to design.

Our paper makes the following contributions:

(1) We identified and described a set of 26 common commands to cover general VR interactions.

(2) We uncovered a taxonomy of user-defined upper-body gestures based on the gestures designed by people with SMA to complete the aforementioned set of commands.

(3) We derived the mental models and considerations of participants with SMA when designing upper-body gestures to increase the accessibility of VR interaction.

2 RELATED WORK

Our work is informed by prior work on *upper-body gestures input for people with upper-body motor impairments*, *gaze-assisted gestures interaction in VR*, and *user-defined gesture designs.*

[1] https://www.apple.com/apple-vision-pro/

2.1 Upper-body Gestures Input for People with Motor Impairments

Previous research has explored gesture-based interactions for individuals with upper-body motor impairments, utilizing readily available devices (e.g., smartphones [10, 11, 62] and wearables [52]) or custom technologies (e.g. [13, 15, 22, 41]). These approaches successfully transform movements into computing commands, demonstrating the feasibility and providing a diverse design space for inventing personalized gestures.

Hand-based gestures, such as stroke or motion gestures, are commonly used in computing device interactions. Empirical research indicates that individuals with motor impairments can accurately perform these gestures with devices worn on the wrist, finger, and head, although challenges related to repetition exist [51, 52]. Moreover, the performance of stroke gestures for individuals with motor impairments can be enhanced through computer modeling and synthesis [49].

Eye-based gestures, including eyelid and eye-gaze gestures, offer support to individuals with limited hand mobility. *Eye-gaze gestures* involve relative eye movement or gaze transitions on the screen [5, 9, 14, 62]. For example, Drewes et al. [9] translated various gaze directions into media control commands. Additionally, Zhang et al. [62] used gaze gestures (e.g., looking up or down) for character input. *Eyelid gestures* involve controlling the states and *duration* of eyelids [43]. Heikkila et al. used prolonged eye closure as a stop command for applications. Fan et al. [10, 11] introduced nine eyelid gestures designed by combining different sequences, frequencies, and *duration* of eye-opening and closing to control smartphones.

Gestures based on other body parts, such as the face or head, can also enhance gesture interaction for people with motor impairment. Facial features, like raising eyebrows or opening the mouth, can be utilized for customized computer control commands [41, 48] and to operate AR/VR headsets [55]. Head motions, such as turning or nodding, could be employed to control the movement of a dual-arm industrial robot [13] or adjust the volume of a sound system [30].

People with SMA, a subgroup of those with motor impairments, can benefit from gesture designs based on the above research. Unlike other motor impairments (e.g., ALS, cerebral palsy, limb deficiencies), individuals with SMA often have improved fine motor skills in distal limbs, specifically in hand dexterity and finger flexibility. This enhanced ability in distal limbs presents opportunities for diverse gesture designs. Despite limited research on their gestures design, we focus on exploring upper-body gestures created by people with SMA. Moreover, we extended the gestures design space to include all the upper-body parts, including eyes, mouth, face, head, limbs, etc., to allow them to better design a richer set of user-defined gestures.

2.2 Gaze-assisted Gestures Interaction in VR

In this work, we called the novel VR interaction methods that combine *gaze-assisted interaction* and *gesture-based interaction* as *gaze-assisted gestures interaction. Gaze-assisted interaction* involves using eye gaze to choose and visually indicate the object of interest, essentially serving as a virtual pointer in VR. The combination of gaze-based selection and gesture-based command input significantly reduces the physical effort compared to virtual hand and controller devices [33], which brings hope for people with SMA to use VR devices.

Previous studies have explored the application of *gaze-assisted gestures interaction* in various VR tasks (e.g., *3D object-related interaction* [33, 44, 45, 61], and *menu-related interaction* [24, 34, 39]) by decomposing complex VR tasks into multiple smaller sub-tasks, and build a larger design space with different integration, coordination, and transition between gaze and gestures [61]. *3D object-related interaction* tasks, mainly selection and manipulation could be simplified as "*gaze for selection, gesture for confirming and manipulation*" [45], like using *'Gaze + Pinch'* interaction [33]. Additionally, for *menu-related interaction*, Reiter et al. [39] used gaze to indicate menu selection, and rotational turn of the wrist to navigate the menu and manipulate continuous parameters. Lu et al. [24] proposed a Gaze-Pinch menu with continuously performs multiple gestures on the gazing object concurrently. However, these studies were designed by people without motor impairments, it remains unknown about the *gaze-assisted gestures* designed by individuals with motor challenges.

We aimed to fill the existing gap by focusing on tailoring this interaction paradigm to meet the unique needs of individuals with motor impairments (i.e., people with SMA in this work) in VR. To prevent constraints on the design and encourage them to fully leverage mobility capabilities, we used the method of user-defined gestures but did not fully develop the gaze-assisted gestures, to gain insights from their preferences and challenges they may face in VR.

2.3 User-defined Gestures Designs

User-defined gestures have been widely used as an elicitation study to discover and identify gesture vocabularies. It has been proved that user-defined gestures are easier to remember and learn than those defined by researchers [27]. Wobbrock et al. [58] started user-defined gesture for multi-touch surface computing, which was the first to employ users, rather than principles, in the development of a gesture set. They first recruited non-technical participants without prior experience using touch screens and presented the referents, or effects of an action to them, and then elicited the set of gestures meant to invoke them by using a think-aloud protocol and video analysis. This process for gesture design has been applied in a variety of domains, for example, keyboards [3], public displays [19], tangible systems [18, 50], smartwatches [2], in-car user interfaces [56], and augmented reality [35]. As for VR, Wu et al. [59] reported a research project on user-defined gestures for VR shopping applications which derived two gestures from each participant in the prior stage and selected the top-two gestures among all of them. Besides, Moran-Ledesma et al. [25] presented an elicitation study to manipulative gestures for 20 CAD-like and open-world game-like referents (the effect of an action in VR). Nanjappan et al. [28] presented a similar user-elicitation study for manipulating 3D objects in virtual reality environments. Our research was motivated by these user-defined methods. Specifically, our research adopts a similar user-centered approach by investigating what *gaze-assisted upper-body gestures* people with Spinal muscular atrophy would like to create and how they would want to use such gestures to accomplish tasks of the VR system.

Table 1: Participants' demographic information. Type indicates the severity of SMA, with I being the most severe type, followed by II and III. IP/R indicates whether the participant participated in the experiment online remotely (R) or offline in person (IP). All participants have experience using either smartphone-based, PC, or all-in-one VR headsets.

Participant	Age	Sex	SMA type	IP/R	Prior VR Experience		
					VR Devices	Time Length	Experience Content
P1	28	M	I	R	Smartphone-based VR headsets	>30min	360° movies
P2	25	M	I	R	Smartphone-based VR headsets	>30min	Interactive 360° videos
P3	26	F	II	R	Smartphone-based VR headsets	>5h	Interactive 360° videos
P4	26	F	II	R	Smartphone-based VR headsets	>30min	Interactive 360° videos
P5	20	F	II	IP	All-in-one VR headsets	>30min	360° movies
P6	42	F	III	R	All-in-one VR headsets	>15min	Games
P7	25	M	III	R	PC VR headsets, Smartphone-based VR headsets	>15min	Games, 360° movies
P8	28	M	III	R	All-in-one VR headsets	>30min/day, >2 year	Games
P9	20	M	III	R	Smartphone-based VR headsets	>30min	Interactive 360° videos
P10	26	F	III	R	All-in-one VR headsets, Smartphone-based VR headsets	>20min	360° movies
P11	37	M	III	R	All-in-one VR headsets, Smartphone-based VR headsets	>7h	360° movies, VR Chat
P12	37	M	III	IP	All-in-one VR headsets, Smartphone-based VR headsets	>30min	Interactive 360° videos

3 METHOD

3.1 Participants

We recruited twelve participants (7 male, 5 female, average age of 28.3 years, SD=6.8) through contact with a disability organization. Table 1 shows the demographic information. We conducted interviews in person (n=2) and remotely (n=10) with participants who were not convenient offline. To ensure participants can correctly understand the VR video content, we recruited participants with prior VR experience.

Two participants with SMA-I exhibited restricted bodily mobility, limited exclusively to few fingers on each hand and facial features such as eye, nose, and mouth movement . Three participants with SMA-II displayed unsteady hand movements and possessed partial control over their forearms, limited head, and trunk mobility, as well as constrained facial gestures. The remaining seven participants with SMA-III exhibited shaky hands and weakened hand and arm muscles, and they had difficulty lifting their hands above the chest level. Except for P8, all participants experienced difficulties in standing and walking. They relied on wheelchairs or bed for daily mobility. Notably, all participants exhibited clear and fluent communication ability. None of them had used upper-body gestures to control VR devices prior to the study. Participants received a compensation of $15 for their participation.

3.2 Commands Gathering

To gather the VR commands and present the effects to participants, we reviewed previous papers for VR commands but found none that included complete commands as a reference in our work. While there are three papers on user-defined VR gestures, each focuses on interactions within specific VR applications (e.g., CAD and open-world games [25], VR shopping applications [59], and manipulating 3D objects [29]) without addressing common VR interactions.

Prior work of user-defined gestures in Section 2.3 (e.g., designed for smartphone [63], smartwatch [2] and AR [35]) follows this two-step workflow to gather commands: 1) *Gathering commands with high level of commonality across various applications*; 2) *Joining the commands into several categories*. Inspired by this, we planned to follow a similar approach for gathering common VR commands. However, three key questions must be addressed before starting the process:

- Q1: How to select a diverse variety of VR applications?
- Q2: How to distill commonly used VR commands from various applications?
- Q3: How to objectively categorize VR commands?

The three questions were answered through *VR App Selection*, *Video Analysis*, and *Category Definition*, respectively.

In the *Category Definition* phase, we consulted official documents (e.g., Oculus and Magic Leap) and previous research, ensuring objective references. However, in the other stages of *VR App Selection* and *Video Analysis*, all materials were chosen by the authors, introducing a potential subjective bias. To mitigate this bias, we restructured the workflow by prioritizing *Category Definition*, where VR categories are established first, before proceeding to *VR App Selection* and *Video Analysis*.

In summary, as shown in Figure 2, we conducted a three-step analysis to obtain the commands used in our study.

1) ***Category Definition***. To answer Q3, we referred to materials from both academia and industries to obtain VR command categories. Initially, we looked into the recent work of Spittle et al [47]. This study conducted a systematic review of the state-of-the-art studies in immersive environments based on papers published between 2013 and 2020. It categorized seven immersive interaction tasks: pointing, selection, translation, rotation, scale, viewport, menu-based, and abstract. Then, we closely examined official documents from Oculus [2] and Magic Leap [3] to further identify these interaction task categories from an industrial perspective. Specifically, we identified commonalities in the VR command categories mentioned in these materials and evaluated whether each category was suitable for VR gesture design to derive the final VR command categories. The complete recording of this process is presented in Appendix A. Finally, we got four VR command categories shown in Table 2.

[2] https://developer.oculus.com/resources/hands-design-interactions/
[3] https://ml1-developer.magicleap.com/en-us/learn/guides/design-interaction-overview

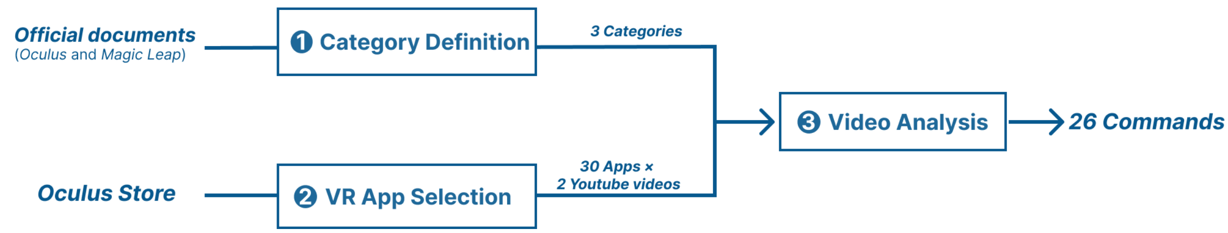


Figure 2: The process for gathering the common VR commands: 1) *Category Definition* established four categories of VR tasks, named *Selection*, *Manipulation*, *Menu-based Interaction* and *Locomotion*; 2) *Video Collection* specified VR tasks from popular VR application videos on YouTube; 3) *Video Analysis* categorized the VR tasks into the four defined categories, following the established coding rules.

Table 2: Four Fundamental Categories of VR Commands

Categories of VR Commands	Description
Selection	Initiating or confirming an action after pointing, such as grabbing an object up close or from a distance.
Manipulation	Moving, rotating, or resizing interactive elements, as well as altering their properties.
Menu-Based Interaction	Presenting a structured set of tabs, commands, or utilities for users to engage with.
Locomotion	Moving or changing the direction of an avatar's position within a virtual space.

Furthermore, we made two modifications to the four VR categories candidates. Firstly, we incorporated *distance*, a significant factor mentioned in the Oculus document [4], as a subcategory to further divide each VR command category into *near* and *far*. This distinction is made because VR commands differ within or beyond the user's arm's reach (e.g., a near button can be selected by poking directly, but a far one can be selected by ray casting). Secondly, as *Selection* and *Manipulation* often occur together while using VR (e.g., users often grasp an object first and then manipulate it), we grouped them into one category named *3D Object-related Commands*.

2) ***VR App Selection.*** To address Q1, we selected the Oculus Store as the app selection resource because its homepage categorizes VR applications into three tabs: Games [5], Applications [6], and Entertainment [7], facilitating the choice of diverse applications. Each tab further contains several subtabs, such as Games for strategists, Kinetic sports, Music games, etc., with each subtab encompassing various VR applications. Focusing on the diversity of VR applications, we meticulously examined each subtab and identified the "Most Popular" subtab as the representative category for our study. We reasoned that "Popular," influenced by player preferences, is more likely to include diverse applications compared to other subtabs that concentrate on specific content (e.g., Games for strategists, which exclusively features applications related to strategies).

We selected the top 10 most popular applications from Games, Applications, and Entertainment on August 13, 2023. Thus, we obtained $10 \times 3 = 30$ VR applications in final, as shown in Figure 3.

3) ***Video Analysis.*** To answer Q2, we conducted a video analysis to collect common VR commands from various applications.

[4]https://developer.oculus.com/resources/hands-design-interactions/
[5]https://www.oculus.com/experiences/quest/section/891919991406810/#/?_k=bok987
[6]https://www.oculus.com/experiences/quest/section/1453026811734318/#/?_k=8cv5gu
[7]https://www.oculus.com/experiences/quest/section/841434313157491/#/?_k=jxqawz

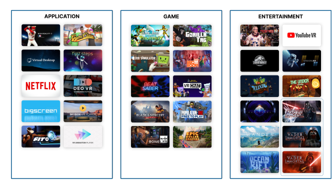


Figure 3: 30 VR applications in final. There are 32 VR apps in total, but in the Entertainment, *Vader Immortal: Episode I, II, and III* were regarded as one item because they had the same VR task but different contents.

To begin with, we searched for each application's name on YouTube, filtering the results by relevance. We selected two videos for each application, with a duration of over seven minutes, and from different publishers, resulting in a total of $30 \times 2 = 60$ videos. And then we collected VR tasks from those videos. Tasks refer to specific user actions in the context of each video. Based on these tasks (e.g., pulling a lever and pulling a drawer), commonly used VR commands will be distilled by summarizing and combining similar actions observed across various videos.

Secondly, we set a coding rule for better organizing VR tasks into several VR commands. Two researchers independently collected VR tasks by watching the videos. However, we faced a challenge where a single command could correspond to different VR tasks depending on the context. For instance, the action of pulling a lever upward and pulling a drawer backward, while distinct tasks share a similar motion involving flexing the forearm at the elbow joint. To address this ambiguity, we drew inspiration from Vuletic et al.'s methodology in their systematic literature review of hand gesture types [53]. They encountered similar ambiguity in coding hand gestures from different applied contexts while performing the same motion. Their rule for coding the gestures depends on their role or aim in the application. We established a coding rule inspired by

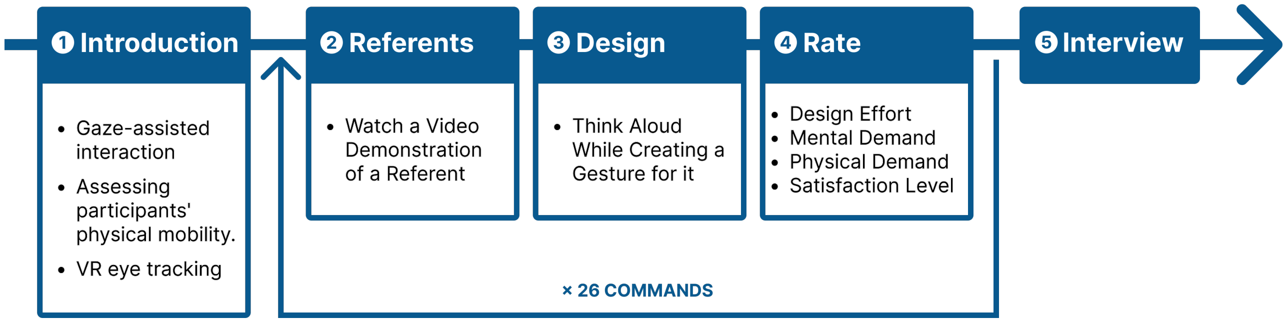


Figure 4: The Study Procedure.

Table 3: The list of the 26 VR commands, grouped into three categories of Menu-related, 3D object-related and locomotion-related commands.

Commands Category		No.	Commands
Menu-related (*Menu-based Interaction*)	Near	1	Access the Shortcut Menu
		2	Confirm a Nearby Selection
	Far	3	Access the Home Menu
		4	Confirm a Far Selection
		5	Scroll Up and Down
		6	Scroll Left and Right
		7	Zoom In or Zoom Out
		8	Drag
3D Object-related (*Selection & Manipulation*)	Near	9	Grab a Nearby Object
		10	Pinch a Nearby Object
		11	Drop an Object
		12	Throw an Object
		13	Hit or Attack the Target Object
		14	Chop the Target Object
		15	Move the Target Object
		16	Pull the Target Object
		17	Rotate the Target Object
		18	Shake or Swing the Target Object
		19	Wave Towards the Target Object
		20	Accumulate Force to Hit the Target Object
	Far	21	Grab a Distant Object
Locomotion-related (*Locomotion*)	Near	22	Artificial Locomotion
		23	Jump
		24	Turn
		25	Lean or Bend
	Far	26	Teleportation

their work: *VR tasks that have similar effects on the target object will share the same command code.*

Finally, two researchers used an open coding method [7] to group VR tasks and iteratively resolve conflicts through ongoing discussions. As a result, we identified 26 commands, as shown in Table 3.

3.3 Procedure

We used Tencent Meeting (a video conference application) to conduct and record all study sessions both in-person and remotely. All participants used personal computers, mobile phones, or tabletops to access the app and participate in our study. All participants, except P6, were positioned with one camera in front of them. The whole study procedure is shown in Figure 4

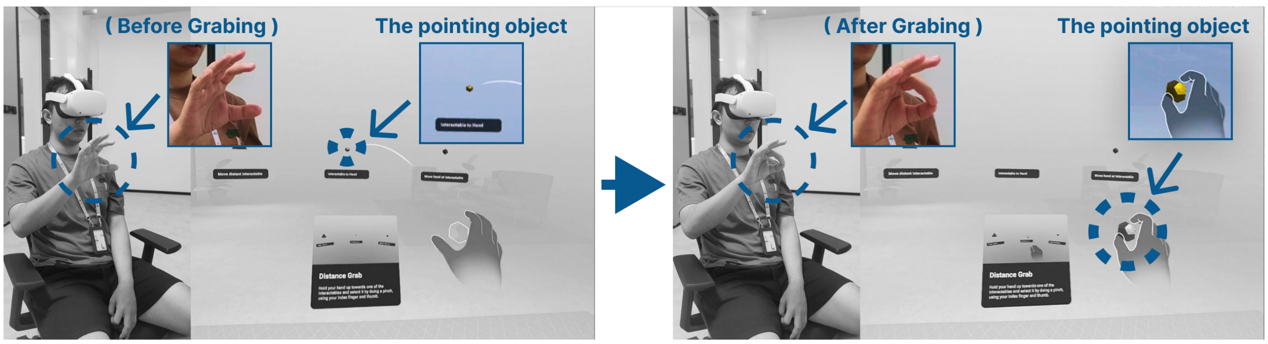
1) ***Introduction***. In the introduction phase, we first briefly introduced our project and then asked participants to self-report their motor abilities, including their daily activities and challenges using various devices, to better understand their capabilities. During the participant's demonstration of their physical state, we observed the maximum and minimum range of their movements and guided the participant to either modify the camera's angle or adjust their distance from the camera. This ensured that all of the participant's gestures were comprehensively displayed in the video. Then, participants were presented with materials including videos [8], images, and text to illustrate the methods of gaze-assisted interaction in VR. After viewing these materials, we confirmed the participants' comprehension.

2) ***Referent Watching***. Participants were first exposed to brief video clips illustrating the effects of each VR command (i.e., *referent*, which is firstly called by Wobbrock et al. [58]), prior to designing gestures. In Figure 5, an example video clip frame for the "Grab Distant Object" command is shown (the video is available in the supplementary materials). Each video clip consisted of two parts, as depicted in Figure 5 (a). On the right side of the screen, a researcher wearing the Quest2 headset recorded first-person views of executing the command. On the left side, the researcher's movements while performing the VR command in reality were presented to aid understanding. These two perspectives were edited to be in time synchronization. It's important to note that the able-bodied movements shown in the video were not gestures but merely a component of VR command instruction. We emphasized this to participants, encouraging them to design gestures based on their own capabilities and preferences.

[8]Links of the videos:

- https://www.bilibili.com/video/BV1jk4y1p7ax/?share_source=copy_web&vd_source=f1455f0b9f86a68ed04b8784b03662bc
- https://youtu.be/NzLrZSF8aDM?si=79Q4iNuJ45JL5NBc
- https://youtu.be/5GTOU6e8--I?si=KL6v6tJRv0DLJPGv

(a) The video clip frame of *Grab Distant Object*



(b) The video clip frame before and after *Grab Distant Object*

Figure 5: The effect of *Grab Distant Object* command: (a) demonstrates the video clip frame of *Grab Distant Object.* One researcher wearing Quest2 recorded first-person views of grabbing a distant object in VR (on the right side of the screen), while another researcher captured his actions in reality (on the left side of the screen). These two videos were edited in time synchronization to assist participants in comprehending VR tasks; (b) depicts the state before and after grabbing, showcasing an author's interaction with the VR system to execute a grabbing motion.

3) ***Gesture Design.*** After each watching, participants were asked to create an upper-body gesture for the VR command and perform the gesture to the moderators. During this process, we asked participants to think aloud to verbalize their thoughts during the design process. Researchers inquired about participants' thought processes as shared during the design phase. For example, if a participant expressed the desire to create a cool gesture, moderators delved deeper, asking why they wanted to design gestures with cool physical appearances and what characteristics they considered cool. This approach aimed to gain a more profound and comprehensive understanding of the mental models of individuals with SMA in the VR gesture design process.

To reduce gesture conflicts, we asked participants to design different gestures for each command within the same category. For commands that were in the different categories, participants were allowed to perform the same gesture. However, due to a large number of commands, some participants might have forgotten their previous designs. Thus, a moderator monitored the gestures already created, and if she found a conflict in design, she would remind the participants to change either the current or previously designed gesture to a different one. Participants were allowed to change their previous gestures at any point during the process.

4) ***Rate the Gesture.*** Upon completing each gesture design, participants were asked to rate four aspects of their gestures on a 7-point Likert scale, including design effort (the difficulty involved in designing gesture for the current referent), mental demand (the cognitive load required to execute the proposed gesture, such as memory), physical demand (the physical workload to execute the proposed gesture), and the overall satisfaction for the proposed gesture. Researchers followed up on these ratings to gain a deeper understanding of the participants' perspectives. For instance, if a gesture was rated high in physical demand but also high in overall satisfaction, researchers would inquire about the reasons behind these ratings.

5) ***Semi-structured Interview.*** After the completion of all gesture designs, we conducted semi-structured interviews with each participant. These interviews were tailored based on the participant's responses during designing gestures, allowing us to ask follow-up questions for more in-depth information. The interview focused on the participants' experiences with accessibility issues

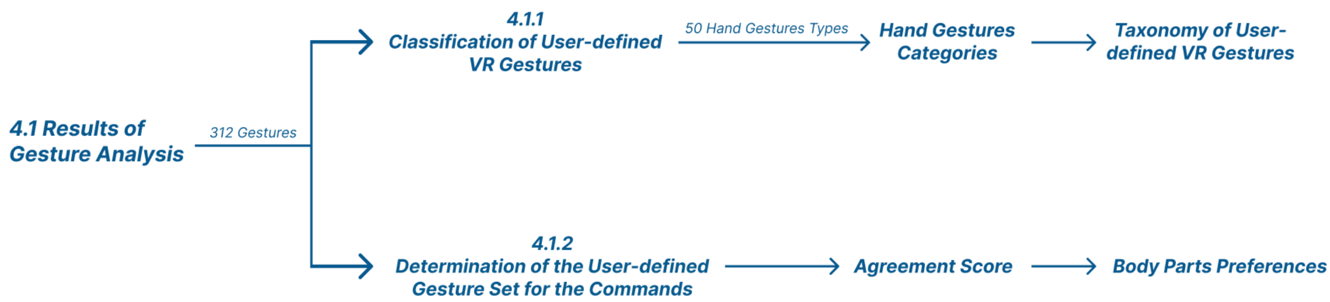


Figure 6: The Processing of User-defined Gesture Analysis

in current VR device usage, their preferences, and main considerations when designing gestures, their expectations for VR interaction methods, and any additional content they wished to add. For example, some participants suggested the desire for a universal gesture set integrated with eye movement and UI components, due to concerns about progressive muscle atrophy in the future.

3.4 Data Analysis

Our data analysis was divided into two parts. The first part involved categorizing and organizing user-defined gestures. The second part focused on a qualitative analysis of participants' mental models.

3.4.1 Gesture Analysis Method. The original analysis methods for user-defined gestures [58] involved two main steps: *First, classifying each gesture along four dimensions (form, nature, binding, and flow)* into a taxonomy to describe the gesture design space. *Second, grouping identical gestures and selecting the group with the largest consensus as the representative gesture for each referent* for future design references. Given our focus on SMA and VR, we made two significant changes, taking into account two factors: *body parts and similar patterns. Drawing inspiration from research on user-centered gesture design for individuals with motor impairments [63], which categorizes gesture taxonomy based on the involved body parts, we replaced four dimensions (form, nature, binding, and flow) with the involved body parts. We also referred to research on user-centered gesture design for VR [25, 59]. This prompted us to shift the criteria from "gestures must be identical" to "gestures with a similar pattern," acknowledging the diversity in VR gesture design.* The user-defined gestures were identified based on participants' descriptions and performed either through the camera or in person.

3.4.2 Mental Model Analysis Method. The data consisted of audio recordings from online meetings, including participants' expressions during the think-aloud design phase, their explanations for scoring the design effort, mental demand, physical demand, and satisfaction level of their designed gestures, as well as their overall concerns and expectations expressed in the final semi-structured interview. These recordings were transcribed into text scripts. Two researchers initially read through these scripts several times to gain an overall understanding of the participants' mental models in gesture design. Subsequently, the researchers independently coded the scripts using an open-coding approach [7]. Themes, subthemes, and specific contents were inductively constructed by assigning keywords to participants' responses. Repeating or similar keywords were grouped into higher-level categories. For example, the subtheme "Concerns about the Accuracy of Recognition of Micro Gestures" emerged when phrases like "wrong recognition," "worried about recognition," and "unrecognition" frequently appeared. The coders regularly discussed and reconciled any coding discrepancies. Further meetings with other co-authors were conducted to finalize agreements based on the preliminary coding. Ultimately, we identified four mental models of participants with SMA, which are detailed in Section 5: Mental Model Observations. A codebook of mental models is in the supplementary materials.

4 RESULTS

Our results were divided into two parts. The first part shows the classification results of user-defined gestures. The second shows four mental models of people with SMA when designing VR gestures.

4.1 Results of Gesture Analysis

Due to the diverse personalized preferences of user-defined gestures, we were unable to finalize a specific set of gestures. However, we developed a taxonomy and gained insights into participants' usage of different body parts by analyzing the distribution of gestures. The structure of this section is depicted in Figure 6.

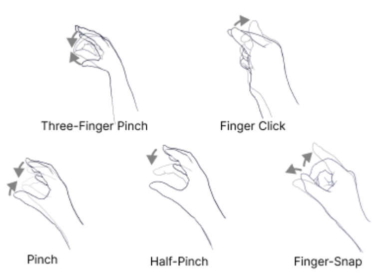
4.1.1 Classification of User-defined VR Gestures. We collected a total of 312 gestures (12 Participants × 26 Commands).

Hand Gestures Categories. While constructing the taxonomy based on body parts and movement patterns, we found a vast diversity in hand gestures. Our collection of 126 hand gestures, making up 40.3% of the total 312 gestures, included 50 different types. To categorize them into taxonomy breakdowns, we further refined the diverse hand gestures based on specific hand parts (e.g., thumb, index finger, palm, etc.) and similar movement patterns (e.g., swipe, slide, grip, strike, etc.). Figure 7 illustrates an example of this process. We grouped five gestures, including the three-finger pinch, thumb-to-index pinch, finger click, half-pinch, and finger snap, into the thumb-to-finger gesture type.

Table 4: The identification of five hand gesture types.

Hand Gestures Type	Description	Example
Thumb-to-Finger Gesture	Hand Gestures that involve the thumb touching the fingers signify confirmation or precise control.	Thumb-to-index Pinching, Three-Finger Pinch, Thumb-to-index Clicking, Thumb-to-index Half-pinching, Thumb-to-index Snaping quickly, etc.
Grasping or Closure Gestures	Hand Gestures that involve the proximity or closure of the palm or fingers convey a grasp or closing motion.	Five-finger Grasping, Clenching Fist, Four-finger Picking, index-finger hooking, etc.
Swinging or Extension Gestures	Hand Gestures that involve the spreading or flicking of the palm or fingers indicate a direction or signal a return to the initial state.	Index finger sliding to the left or right, Thumb sliding to the left or right, Hand palm opening, Index and thumb spreading apart, Hand palm swinging, etc.
Tapping or Striking Gestures	Hand Gestures that involve tapping or lightly tapping another surface in the air with the surface of the hand convey a clicking or tapping motion.	Both hands clapping together, Index finger tapping or pressing, One hand's fist striking the other hand, etc.
Symbolic Hand Signs	Hand Gestures that utilize specific hand or finger positions convey specific meanings or information.	Forming the "OK" sign with fingers, Single finger moving in a circular motion, Fingers forming the shape of the number "6", Fingers simulating a gun shape, Fingers outlining the shape of a person,etc.

(a) Thumb-to-finger Gesture type.



(b) Five specific user-defined gestures of Thumb-to-finger Gesture type.

Figure 7: Example of coding Thumb-to-finger Gesture Type: (a) illustrates the representative motion pattern of the Thumb-to-finger Gesture type, where the thumb touching the fingers indicates confirmation or precise control. (b) shows five user-defined gestures. Furthermore, these five gestures are grouped as the Thumb-to-finger Gesture, as shown in (a).

Finally, we identified five basic hand gesture patterns, as shown in Table 4, including thumb-to-finger gestures(TFG), grasping and closure gestures (GCG), swinging and extension gestures (SEG), tapping and striking gestures (TSG), shaping and symbolizing gestures (SSG). The complete visual illustration can be seen in Appendix B Figure 11.

Taxonomy of User-defined VR Gestures. We identified a user-defined gestures taxonomy, as shown in Table 5. We grouped the gestures into 15 categories based on the body parts involved. These 15 categories included *a single body part* and *the combinations of different body parts*: only hand, wrist, forearm, only shoulder, torso or chest, only head, only eyes, only mouth, hand && wrist, hand && forearm, head && mouth, eyes && hand, eyes && wrist, eyes && mouth and others.

categories The *Only Hand* category includes *Thumb-to-Finger Gestures (TFG), Grasping or Closure Gestures (GCG), Swinging or Extension Gestures (SEG), Tapping or Striking Gestures (TSG), Shaping or Symbolizing Gestures (SSG)* and the combinations of them. The combination of two basic hand gestures refers to the execution of two hand movements either subsequently or simultaneously. For instance, the *thumb-to-index pinch and horizontal slide* can be a sequential combination in TFG+SEG, where the thumb-to-index pinch belongs to TFG and the horizontal slide belongs to SEG.

The *Wrist* category includes two basic types: *Wrist Hook and Swing* and *Wrist Rotate. Wrist Hook and Swing* refers to movements of the wrist around the wrist joint in various directions. Additionally, *Wrist Rotate* typically includes the twisting of the hand around its axis. The *Forearm* category includes *Forearm Swipe* and *Forearm Swing*, with variations in the extent of forearm rotation centered around the elbow joint.

The *Only Shoulder* category includes *Shoulder Shrug* and *Shoulder Alternation Shake. Shoulder Alternation Shake* indicates the sequential and alternating forward movement of each shoulder, generating

Table 5: The fifteen categories of the user-defined upper-body gestures and the gestures types within each category.

Taxonomy	Breakdown	Taxonomy	Breakdown
Only Hand	Thumb-to-Finger Gesture (TFG)	Only Eyes	Eye Movement
	Grasping or Closure Gestures (GCG)		Wink
	Swinging or Extension Gestures (SEG)		Blink
	Tapping or Striking Gestures (TSG)		Eye Size
	Shaping or Symbolizing Gestures (SSG)		Gaze Dwell
	GCG/SEG + TSG		Eyebrows
	GCG/TFG + SEG		Eyebrows + Eye Movement
Wrist	Wrist Hook and Swing		Wink/ Eye Movement/Gaze Dwell + Eye Size
	Wrist Rotate	Only Mouth	Lip Movement
Forearm	Forearm Swipe		Mouth Open then Close
	Forearm Swing	Hand & Wrist	SEG/TSG/SSG + Wrist Hook and Swing
Only Shoulder	Shoulder Shrug		TFG/GCG/SEG/SSG + Wrist Rotate
	Shoulder Alternation Shake	Hand & Forearm	TFG/GCG/SEG/TSG/SSG + Forearm Swing
Torso or Chest	Torso Rotation and Tilt		GCG/SEG/TSG/SSG + Forearm Swipe
	Chest Lifting	Head & Mouth	Head Tilt + Lip Movement
Only Head	Head Turn	Eyes & Hand	GCG/SEG + Eye Movement
	Head Nod	Eyes & Wrist	Eye Size + Wrist Hook and Swing
	Head Tilt	Eyes & Mouth	Eyebrows + Mouth Open then Close
	Head Turn + Nod	Others	Eyes/Hand/Mouth + UI

a shaking motion. The *Torso or Chest* category includes *Torso Rotation and Tilt* and *Chest Lifting*.

The *Only Head* category includes *Head Turn, Head Nod, Head Tilt*, as well as *Head Turn and Nod*.

The *Only Eyes* category includes eight basic types of eye gestures: *Eye Movement*, *Wink*, *Blink*, *Eye Size*, *Gaze Dwell*, *Eyebrows*, and the combinations of these basic eye gestures. *Eye Movement* includes moving the eyes up and down, left and right, or eye rotating. *Wink* and *Blink* respectively indicate the rapid closure and reopening of one eye or both eyes. *Eye Size* includes wide opening, closing, and squinting of the eyes. *Gaze Dwell* is the act of maintaining prolonged eye contact with the target object. *Eyebrow* includes movements associated with raising or lowering the eyebrows.

The *Only Mouth* category includes *Lip Movement* and *Mouth Open then Close*. *Lip Movement* refers to various actions or changes primarily occurring in the area of the lips, such as *Wry Mouth* and *Pucker Lips*.

The *Others* category, introduced by P1 with SMA-I, diverges from the approach of designing unique gestures for each command. P1 advocates placing commands on menus or icon-based buttons and using available movements for selection.

4.1.2 Determination of the User-defined Gesture Set for the Commands. To derive the final user-defined gesture set from all gestures proposed by all participants, we first collated the gestures included in each command and counted the number of participants performing the same gesture. The number was also used to calculate the agreement score of the commands.

Agreement Score. The agreement score was initially proposed by Wobbrock et al. [58] and later widely used in studies uncovering user-defined gestures. It intuitively characterizes differences in agreement between target users for assigning a gesture to a given command. In general, the higher the agreement score of a command, the better the participants are in agreement with the gesture assigned to the command. We used the following equation to calculate the agreement score from prior user-defined gesture research [42, 58]:

$$A_c = \sum_{P_i} (\frac{P_i}{P_c})^2 \tag{1}$$

In Equation 1, c is one of the commands, A_c represents its agreement score based on participants' proposed gestures for this command. P_c is the total number of gestures proposed for c, which is the number of participants in our case (N=12). i represents a unique gesture. P_i represents the number of participants who propose the unique gesture i. Take the *Confirm a Far Selection* command as an example, 12 participants proposed 12 gestures in total, P_c equals 12. Among these gestures, there were 9 unique gestures: 3 (Blink Twice), 2 (Head Nod Once), 1 (Gaze for 5-10s), 1 (Furrowed Brow + Wink the Right Eye), 1 (Pout), 1 (Fingers simulating a Gun Shape), 1 (Pinch), 1 (Index Finger Tapping), and 1 (Thumb Swing). As a result, the agreement score of the *Select Far Selection Button* command was calculated as follows:

$$\left(\frac{3}{12}\right)^2 + \left(\frac{2}{12}\right)^2 + 7\left(\frac{1}{12}\right)^2 = 0.14 \tag{2}$$

Figure 9 shows the agreement score of the gestures proposed for each command. For most commands, the agreement score is low, which indicates that the participants proposed diverse gestures for most commands and less agreed on which gesture should be allocated to them. The agreement score of *Teleportation* is lowest since every participant proposed different gestures in this command and

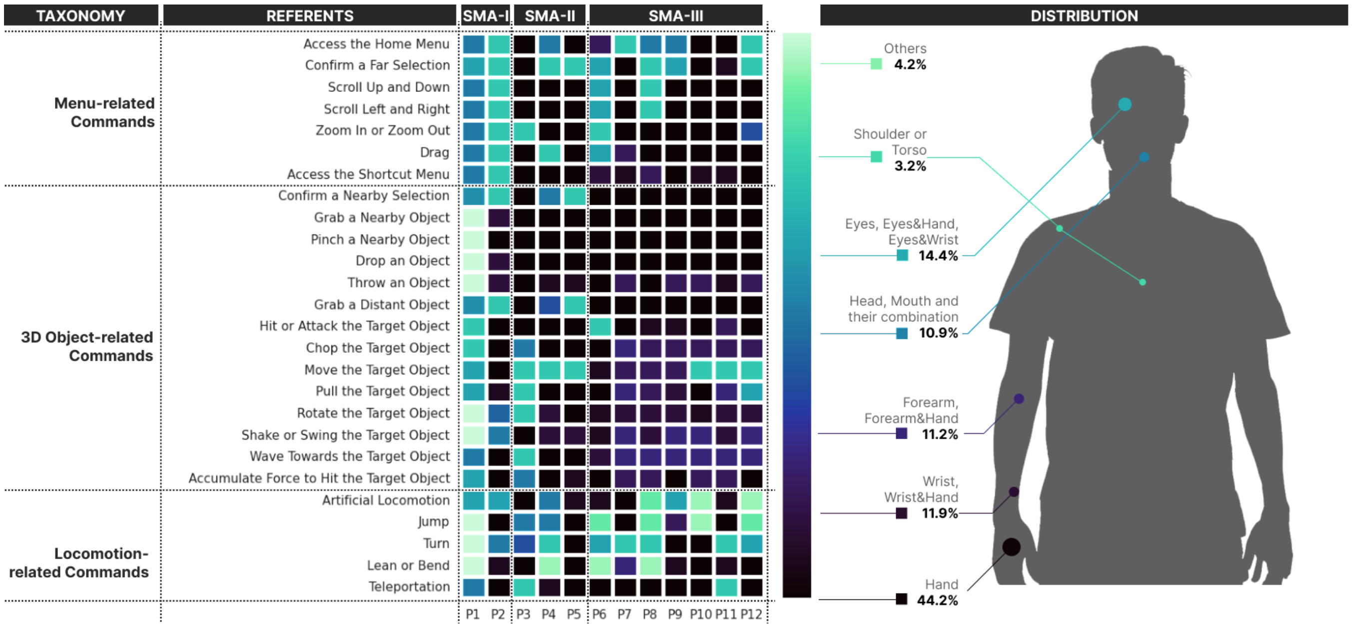


Figure 8: Participant-Designed Gesture Categories for VR Commands and the Distribution of Gesture Groups.

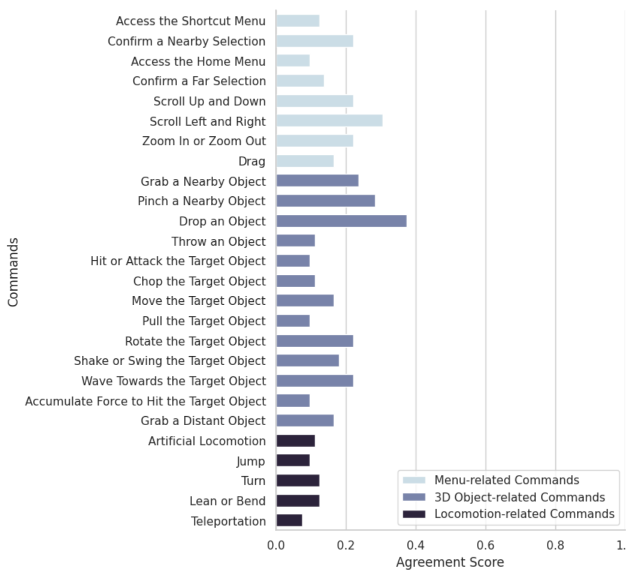


Figure 9: The agreement scores of the 26 commands. The higher the score, the higher the participants' consensus on which gesture type should be assigned.

this indicates no common gesture can be allocated to it. The diversity of gestures makes obtaining a common user-defined gesture set challenging. To better understand participants' preferences when designing VR gestures, we categorized the body parts used by each participant for each command, as shown in Figure 8.

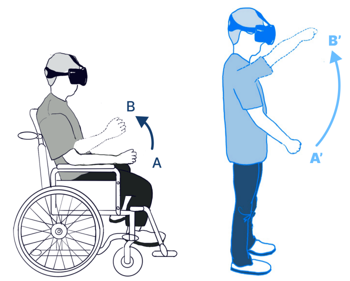
Body Parts Preferences. Figure 8 indicates that although the overall trend is that individuals with higher motor abilities use upper limb body parts more, there are differences observed among different participants and VR commands. For example, even participants with SMA-I, P1, and P2 preferred to use different body parts (refer to the first 2 columns in Figure 8). P1 relied more on body parts above the neck to design gestures. In cases where above-the-neck body parts were insufficient or appropriate , P1 used UI combined with body parts to solve the *"shortage of body parts to design"* (P1). Despite also having limited hand mobility, P2 still wanted to use his hands to design gestures for a better sense of body involvement in VR. He chose to use peripherals such as a mouse (which he also used in daily life for computer operations, as shown in Figure 1 P2) to support his hand muscles and perform small hand movements to design gestures.

The body parts used also varied for different categories of tasks. For *Menu-related Commands*, the proportion of using the eyes is higher compared to other categories. For *3D Object-related Commands*, the hands are being used more among all participants except P1. For *Locomotion-related Commands*, the body parts chosen were more diverse.

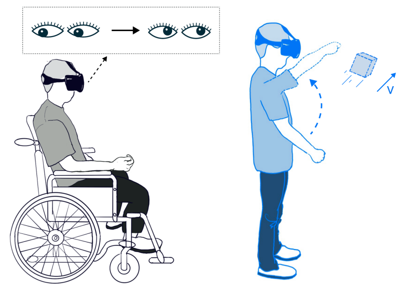
To gain deeper insights into the preferences and considerations of people with SMA when designing VR gestures, we analyzed the participants' feedback and identified four main mental models.

4.2 Mental Model Observations

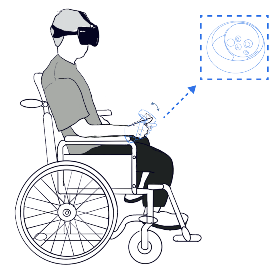
4.2.1 Use Local Movements to Map Unfeasible Large Range Motion. From the participants' self-reported abilities, we learned that they have significant difficulty performing large actions, particularly those requiring upper arm strength. These discrepancies between the limitations in upper arm muscle strength and the large motion-scale tasks in VR influenced their experience, leading to self-disappointment and aversion towards VR devices.



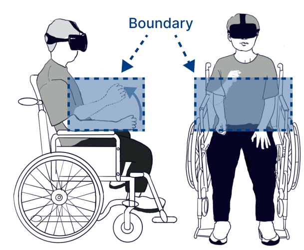
(a) Participants prefer to use their hand to make a small gesture to map a large movement in VR.



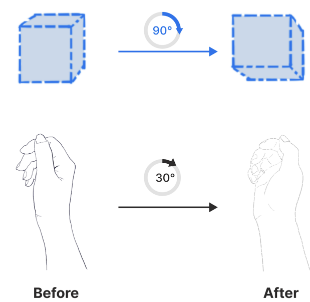
(b) Participants employ eye movements that mirror the hand gesture required to execute VR commands, such as 'throw 3D objects.'



(c) Participants mimic the use of a VR controller through their gestures.



(d) Participants adjust the height of their gestures to fall within a range from the waist to just below the chest.



(e) Participants rotate their wrists by a smaller degree, such as 30°, to manipulate the virtual object by a larger degree, such as 90°.

Figure 10: The example illustration of five strategies for conducting large actions in VR: (a) Utilize Retaining Hand Mobility to Mimic Weak Proximal Upper Limbs; (b) Substitute with above-the-neck body parts; (c) Imitating Existing Peripherals; (d) Restrict the performance boundary of gestures; (e) Enhance small movements through offset.

Strategies for Conducting Large Actions in VR. We identified five strategies that participants employed to complete the large-scale actions while also preserving their body involvement in VR.

1) *Utilize Retaining Hand Mobility to Mimic Weak Proximal Upper Limbs.* Participants (N=10) with remaining hand abilities initially opted for an approach that involved using their hands to mimic the weaker proximal upper limbs, as shown in Fig 10 (a). Although other distal body parts may exhibit superior mobility compared to the hands, participants demonstrated a preference for utilizing their remaining hand mobility to enhance body involvement in VR, as P9 noted, *"the sense of bodily involvement that arises from using the hands can compensate for the difficulty of performing gestures."* P6 incorporated the wrist and fingers as a substitute for the entire arm in gesture design and humorously described this method as *" recreating a miniature VR world with fingers."*

In addition to using their hands, participants also expanded their hand mobility with other tools. For example, when designing a chopping action for 3D objects, P10 picked up a pen as an aid to extend her hands, allowing the limbs below the wrist to become a complete arm.

2) *Substitute with above-the-neck body parts.* Participants with severe atrophy of hand muscles, like those with SMA-I, tend to rely on using body parts above the neck to substitute small movements for larger ones. Their choices of body parts are purposeful. They tend to select body parts with characteristics similar to VR task effects, such as those that can simulate the motion trajectory of 3D objects or the direction of hand force. For example, P3 chose to move her eyes from the bottom left corner to the upper right corner when designing a gesture for VR command *throw 3D objects*, as shown in Fig 10 (b).

3) *Imitating Existing Peripherals.* Participants used their body parts to imitate existing devices to accomplish large movements. For example, when designing locomotion tasks, participants found it challenging to design gestures for tasks with a sense of distance. When unable to think of a direct body part to use, they referred to existing devices capable of achieving locomotion. P6 and P11 used their wrists and thumbs, respectively, as analogies for the joystick, as shown in Fig 10 (c).

4) *Restrict the performance boundary of gestures.* All participants expressed their desire to restrict the performance boundary of VR operations. Participants with SMA Type I or Type II, who primarily adopt a lying position, wished to limit the range around the waist. For participants with SMA Type III, they hoped to control the height

within the range from below the chest to the waist, as shown in Fig 10 (d).

5) *Enhance small movements through offset.* After designing micro gestures, participants also sought to enhance the effects of small movements through offset, as shown in Fig 10 (e), including using discrete gestures to initiate continuous actions and using small gestures to large movements. For example, P10 used wrist rotation to rotate 3D objects, but she chose to *"rotating the wrist by 5 degrees in reality, which would correspond to rotating the virtual object by 25 degrees."* For the same task, P11 hoped to activate continuous rotation by slightly rotating her wrist.

Concerns about the Accuracy of Recognition of Micro Gestures. Although we had reminded the participants before the design phase that there was no need to consider whether the current technology or equipment could support their gesture design, participants still worried about whether the VR device could accurately recognize their micro gestures, as P2 said, *"My mouth can only open as wide as one finger can fit in, and I am not sure if the device can recognize it."* Besides concerns about micro-gesture non-recognition, participants also worried that limiting the range of gesture performance might cause gestures to fall outside the recognition area.

In addition to being unrecognizable, participants also expressed concerns about gesture misrecognition due to limb tremors, particularly when the movement range is relatively subtle. Hand tremors are a common condition among our participants. P8's physical ability is better than other participants, but his hand tremors still significantly impact his daily life, such as mistyping.

This concern about the accuracy of gesture recognition affects the participants' perception of the design process's difficulty and satisfaction with the designed gestures. However, participants still hope to rely on the improvement of VR performance to reduce their burden. In addition to expecting high-precision gesture recognition, participants also hope that the VR system recognizes the motion trend rather than the specific moving body parts. For example, when P7 designed the action of waving toward target objects, he mentioned,*"I hope VR can recognize the direction of the swing, regardless of whether it is performed with the arm or wrist."*

4.2.2 *Minimize Physical Efforts within Capabilities.*

Participants would minimize the physical efforts exerted in their designed gestures. This may be related to the generalized muscle weakness among people with SMA, specifically manifested in the weakness of force and the difficulty of maintaining it.

Strategies to Minimize Physical Efforts. Participants employed three strategies to reduce their physical efforts. Firstly, all participants would repeatedly measure the smallest body parts used in their gestures, such as the number of fingers or the extent of clenching a fist. Through this continuous testing, they sought to find the most effort-efficient combination of body parts. For example, when designing a gesture for grabbing a 3D object, P11 changed the gesture from using five fingers to two after repeatedly testing. Secondly, participants also considered the duration of the gesture use and tended to use gestures with shorter durations to reduce physical load. Finally, besides the physical effort required for the gesture itself, participants also considered the operation frequency of VR tasks. And they tend to prioritize using the most flexible body parts to complete more frequently used tasks. For example, participants (N=4) chose blinking as a gesture for high-frequency VR tasks, as P4 said, *"I prefer to use blinking because it is relatively easy for me, allowing me to complete the task with minimal effort."*

The Balance Between Energy Conservation and Consistency. Participants (N=7) are willing to bear the additional physical burden to maintain consistency between reality and the virtual world. For example, P11 was willing to use more fingers in the gesture for pulling a 3D object to simulate the force performed in the task. However, when designing gestures that require significant physical effort, they potentially sacrifice consistency. For example, when designing the gesture for pinching an object, P7 opted to abandon the more experiential loose grip in favor of a more effortless tight grip.

4.2.3 *Consider Social Encouragement and Acceptance.*

Participants prefer gestures with a cool physical appearance to convey to others that they are doing something cool and to encourage themselves to use these gestures more frequently. They also try to avoid gestures that may evoke negative associations.

Gestures with Cool Physical Appearances. Participants tried to design gestures with cool physical appearances based on the characteristics of VR. Participants mainly drew inspiration from popular culture, magical stories, and science fiction films to design appealing gestures. For example, when designing a gesture for hitting 3D objects, P7 chose classic moves from the martial arts world, a Chinese popular culture, because he thought *"both VR and martial arts can transcend human limitations to accomplish some incredible things"*, and he also believed that using popular ethnic culture could reduce the understanding barriers between VR users and viewers.

The cool and creative aspects of gestures can increase participants' satisfaction and enhance engagement. Hence, when designing gestures that cannot fully utilize the characteristics of VR, participants' satisfaction may decrease. P8 mentioned that he was not satisfied with his designed gestures and was not very willing to use them due to a lack of creativity. He explained, *"I think VR is a very technologically advanced product, and cool gestures also make VR more appealing."*

Social Acceptance. Participants were concerned about others' perceptions when designing gestures. They were reluctant to design gestures that were too unusual or would evoke negative associations. In addition to negative associations, P7 mentioned the unique understanding barriers caused by the isolation of VR users from others.

4.2.4 *Design Gestures across Time Span and Abilities.*

Participants design gestures not only based on their abilities but also consider their past experiences and potential future physical conditions.

On the one hand, participants would reflect on their previous better physical conditions or even refer to people without motor impairments, and they tend to design gestures that resemble those performed by individuals without motor impairments. This approach allows them to gain the perception that their motor abilities remain intact. As P4 expressed,

> "When designing a gesture, I think about how I would perform the tasks if I did not have a disability, and then I try to get as close to that state as possible. It is

> like an idealized state where I imagine myself being the same as before, or even without a disability. In this way, I can do anything in VR, just like able-bodied people."

On the other hand, participants also expressed concerns about progressive muscle atrophy and would like to design a universal gesture set for the future. For example, at the end of the design process, P2 proposed to redesign all gestures by combining UI components with body parts, as he said,

> "If possible, I would list all operations on the side of the screen, and I could either look at it for a few seconds or tap my nose to confirm."

P1 believed that having a backup solution was necessary and could provide him with a sense of security. However,P1 also acknowledged the lack of immersion and body involvement with this method.

5 DISCUSSION

In this paper, we discuss the implications of the user-defined upper-body VR gestures designed by and for people with SMA. VR offers an opportunity for individuals with limited mobility to act beyond their physical capabilities, promoting inclusivity and equality [12, 21]. However, current VR devices with ability assumptions pose challenges in input methods for them [17, 26], and there are few attempts made at alternative accessible VR input methods for users with motor impairments [60]. This study used a video elicitation study to investigate what upper-body gestures people with SMA prefer in VR. This was pivotal for us to understand their design considerations and to provide some useful design suggestions.

5.1 Key Takeaway

By involving people with SMA in the design process, we identified a taxonomy of user-defined upper-body gestures and their mental models. All participants in the elicitation study expressed their desire to experience VR with gestures in the future. Gestures involving hands were the most diverse and preferred. The type of task and participants' abilities influence the choice of body parts for gesture design. We identified four mental models that people with SMA employed when designing gestures. They preferred to use local movements to map unfeasible extensive motion, intending to minimize physical effort in their gestures. They also focused on designing gestures with visually appealing appearances, and they aimed to create gestures adaptable to changes over time and their abilities.

5.2 Design Considerations for Accessible Gesture Input in VR

Our findings suggest the need for design approaches that capitalize on users' motor abilities and preferences. In the following, we present four practical implications informed by our empirical findings to facilitate the development of accessible VR input methods for people with motor impairments.

Design visually appealing gestures to encourage people with motor impairments to engage in VR. Our findings highlight the importance of creating VR input gestures that are visually appealing to encourage people with motor impairments to engage in VR. Prior studies also identified that social acceptance is a crucial consideration in the design of gesture-based interactions for people with motor impairments in other contexts [10, 63]. However, our study extends beyond the realm of social acceptance. Informed by our study, the aesthetics and social encouragement address a need that goes beyond mere social acceptance and can enhance the engagement of users with motor impairments. Therefore, when designing accessible VR input gestures for people with motor impairments, it is essential to fully utilize the characteristics of VR to create gestures that are not only functional but also appealing and engaging.

Improve recognition accuracy of micro gestures by people with motor impairments in VR. Participants expressed concerns about whether VR devices can accurately detect their micro-gestures. Efforts have been made to improve gesture recognition accuracy [4, 37, 46, 54]. However, these approaches primarily target input methods designed for able-bodied individuals, potentially leading to technological incompatibility. Our study's focus on this issue highlights the need for technology that accommodates the specific challenges faced by people with motor impairments, such as limb tremors in SMA, necessitating specialized gesture recognition technologies.

Designing more personalized user-defined gestures for people with motor impairments. While designing gestures for people with SMA, a standard gesture set might not be optimal for a particular user. The results of our gesture analysis indicate that they have different physical conditions and habitual perceptions. And the difference in VR commands may also influence their preference. Thus, it is important that an individual user with motor impairment can customize their VR gestures, similar to the work in the mobile phone context by Ahmetovic et al [1].

Combine alternative or adapted input devices with user-defined gestures for people with motor impairments in VR. Some participants expressed their interest in combining alternative input devices like joysticks and adapted keyboards with VR gestures, particularly for those with SMA who have weaker anti-gravity hand muscle capabilities. Although it remains unclear how to create an ecology that combines input devices with gestures for people with motor impairments in VR [57], our findings underscore this need to enhance the VR experience while preserving the unique preferences and abilities of users with motor impairments.

Using user-defined gestures for motor rehabilitation in VR. Participants expressed their expectations for VR to facilitate rehabilitation by incorporating their more severely affected body parts into gesture design, which could encourage more frequent use of these areas and potentially slow down muscle degeneration. Existing VR rehabilitation methods primarily involve designing games with task-specific training scenarios, which can be relatively simple and repetitive [6, 16, 40]. Using user-defined gestures for motor rehabilitation in VR suggests a more natural and personalized integration of rehabilitation exercise into VR interactions.

5.3 Reflections on Gesture Design with Able-bodied Movement References and Video

Participants were influenced by the able-bodied movements demonstrated in the video, despite our instruction to focus on their own abilities and use the video examples solely for understanding VR commands. Our observations and interviews revealed that all participants, regardless of their motor abilities, initially perceived the video actions as reference gestures and then adapted them based on their actual motor skills. For instance, when designing the *Pinch a Nearby Object* commands, individuals with relatively stronger motor abilities, such as P12 and P10, crafted hand gestures by incorporating the example movements of pinching with the thumb and forefinger. Conversely, participants with weaker motor abilities, like P1-5, primarily considered their capabilities in the example movements before making adjustments to their designs. According to their feedback, the use of able-bodied movement references not only reduces their memory load but also facilitates the creation of memorable gestures. As P3 mentioned,

> "I study VR gestures used by able-bodied individuals both in daily life and online videos. If there are 100 common gestures, and I create 100 new ones, the total to remember would be 200. However, with some overlap, it could be reduced to 150."

Therefore, as individuals with SMA commonly look to able-bodied movements before designing gestures, we suppose that even when using a VR headset to experience commands firsthand instead of relying on our examples, they still consider the actions of those without impairments as a reference for their gesture design. This can be confirmed in future research, potentially influencing the focus of future HCI designs for individuals with motor impairments.

In our study, we utilized 2D videos to demonstrate VR command effects within a 3D immersive environment. While this method may lack the complete immersive VR experience, it could potentially limit participants' comprehension of each VR command. For example, P11 and P12 both had queries about differentiating between *Confirm a Nearby Selection* and *Confirm a Far Selection* because they seemed similar in the 2D video, while we addressed these queries through video review and detailed explanations.

To enhance understanding, we engaged participants with prior VR experience and provided verbal descriptions as a supplement for each VR command. However, it remains unknown if participants would change their design in our work when they wore a VR headset and were shown the VR command effects in immersive scenarios.

Future research could explore a more hands-on approach, involving participants with motor impairments in designing gestures directly within a VR environment. Techniques like the 'Wizard of Oz' could be employed to effectively link gestures with VR command effects. Such an approach has the potential to offer a more authentic and immersive experience, leading to more intuitive and effective gesture design.

6 LIMITATION AND FUTURE WORK

Impact of Single Camera on Gesture Observation. Initially, we proposed using two cameras to provide a comprehensive view of participants' gesture design process—one in front and one on the side. However, due to mobility and device constraints, participants were limited to a single camera. This limitation may have hindered our ability to fully observe gesture articulations despite detailed inquiries. When comparing the gestures of two offline participants with those participating online, it appeared that the impact of the scope limitation was minimal. Future studies could enhance insights by integrating offline user studies for comprehensive comparison and validation of findings.

Limited Coverage of People with SMA. Another limitation of this study is the limited number of participants, with a majority being individuals with SMA type III. Due to geographical dispersion, mobility constraints, and privacy concerns of the participants, we chose to conduct the experiments via online video. To ensure participants can understand the VR referent, we tried to recruit participants with prior VR experience, though the current inaccessibility of VR input methods poses challenges for individuals with SMA types I and II to use VR. Considering the diversity in motor abilities across SMA types, our study ultimately involved 12 participants (2 SMA-I, 3 SMA-II, and 7 SMA-III). However, our participants exhibited the diversity in their gesture design process and results. Future research could include a broader spectrum of participants, particularly focusing on those with SMA types I and II, to gain deeper insights into their considerations.

Potential Impact of Non-exhaustive Command Selection. We identified 26 common commands from the popular applications from three VR application categories. This strategy was employed in an effort to capture a diverse and representative set of common VR commands. However, we acknowledge this approach may not fully capture the wide range of interaction possibilities within VR and tends to focus more on current, prevalent technologies. This limitation might restrict the breadth of our findings. For instance, the number of commands identified in different categories may vary, potentially affecting the distribution of body parts used in each command category. Future research could expand the range of command selection, exploring a wider array of VR applications, including those that are emerging. And the common commands we categorized can serve as a foundational reference for further research.

7 CONCLUSION

We have adopted a user-centered approach by involving participants with SMA to design user-defined upper-body gestures for VR interactions. We initially identified 26 common commands in VR through a content analysis of 60 videos and then derived a taxonomy of user-defined gestures by analyzing the 312 gestures. Gestures involving hands were the most diverse and preferred. The type of task and participants' abilities influence the choice of body parts for gesture design. The participants preferred using localized movements to map unfeasible extensive motions, aiming to engage their hands for better body involvement despite their limited hand mobility. They favored gestures that required minimal physical effort based on their abilities and had visually appealing physical appearances based on their perception of VR characteristics. Additionally, they aimed to create gestures adaptable to changes over time and their abilities. In light of these findings, we highlight design considerations and demonstrate future work to enhance the accessibility of VR for people with motor impairments.

ACKNOWLEDGMENTS

This work is partially supported by 1) 2024 Guangzhou Science and Technology Program City-University Joint Funding Project (PI: Mingming Fan); 2) 2023 Guangzhou Science and Technology Program City-University Joint Funding Project (Project No. 2023A03J0001); 3) Guangdong Provincial Key Lab of Integrated Communication, Sensing and Computation for Ubiquitous Internet of Things (No.2023B1212010007).

REFERENCES

[1] Dragan Ahmetovic, Daniele Riboli, Cristian Bernareggi, and Sergio Mascetti. 2021. RePlay: Touchscreen Interaction Substitution Method for Accessible Gaming. In *Proceedings of the 23rd International Conference on Mobile Human-Computer Interaction* (New York, NY, USA) *(MobileHCI '21)*. Association for Computing Machinery, 1–12. https://doi.org/10.1145/3447526.3472044

[2] Shaikh Shawon Arefin Shimon, Courtney Lutton, Zichun Xu, Sarah Morrison-Smith, Christina Boucher, and Jaime Ruiz. 2016. Exploring Non-touchscreen Gestures for Smartwatches. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, San Jose California USA, 3822–3833. https://doi.org/10.1145/2858036.2858385

[3] Gilles Bailly, Thomas Pietrzak, Jonathan Deber, and Daniel J. Wigdor. 2013. Métamorphe: augmenting hotkey usage with actuated keys. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, Paris France, 563–572. https://doi.org/10.1145/2470654.2470734

[4] Nanziba Basnin, Lutfun Nahar, and Mohammad Shahada Hossain. 2021. An Integrated CNN-LSTM Model for Micro Hand Gesture Recognition. In *Intelligent Computing and Optimization* (Cham) *(Advances in Intelligent Systems and Computing)*, Pandian Vasant, Ivan Zelinka, and Gerhard-Wilhelm Weber (Eds.). Springer International Publishing, 379–392. https://doi.org/10.1007/978-3-030-68154-8_35

[5] Jeffrey P Bigham, Mingzhe Li, Samuel C White, Xiaoyi Zhang, Qi Shan, and Carlos E Guestrin. 2021. On-the-fly calibration for improved on-device eye tracking. US Patent 11,106,280.

[6] Mónica S. Cameirão, Sergi Bermúdez I Badia, Esther Duarte Oller, and Paul F. M. J. Verschure. 2009. The Rehabilitation Gaming System: A Review. *Studies in Health Technology and Informatics* 145 (2009), 65–83. pmid:19592787

[7] John W. Creswell. 2007. *Qualitative Inquiry and Research Design: Choosing among Five Approaches, 2nd Ed.* Sage Publications, Inc. xvii, 395 pages.

[8] Basil T. Darras, Jennifer A. Markowitz, Umrao R. Monani, and Darryl C. De Vivo. 2015. Spinal Muscular Atrophies. In *Neuromuscular Disorders of Infancy, Childhood, and Adolescence*. Elsevier, 117–145. https://doi.org/10.1016/B978-0-12-417044-5.00008-1

[9] Heiko Drewes and Albrecht Schmidt. 2007. Interacting with the Computer Using Gaze Gestures. In *Human-Computer Interaction – INTERACT 2007*, Cécilia Baranauskas, Philippe Palanque, Julio Abascal, and Simone Diniz Junqueira Barbosa (Eds.). Vol. 4663. Springer Berlin Heidelberg, Berlin, Heidelberg, 475–488. https://doi.org/10.1007/978-3-540-74800-7_43 Series Title: Lecture Notes in Computer Science.

[10] Mingming Fan, Zhen Li, and Franklin Mingzhe Li. 2020. Eyelid Gestures on Mobile Devices for People with Motor Impairments. In *The 22nd International ACM SIGACCESS Conference on Computers and Accessibility*. ACM, Virtual Event Greece, 1–8. https://doi.org/10.1145/3373625.3416987

[11] Mingming Fan, Zhen Li, and Franklin Mingzhe Li. 2022. Eyelid gestures for people with motor impairments. *Commun. ACM* 65, 1 (Jan. 2022), 108–115. https://doi.org/10.1145/3498367

[12] Kathrin Gerling, Patrick Dickinson, Kieran Hicks, Liam Mason, Adalberto L. Simeone, and Katta Spiel. 2020. Virtual Reality Games for People Using Wheelchairs. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM, Honolulu HI USA, 1–11. https://doi.org/10.1145/3313831.3376265

[13] Muhammad Abdul Haseeb, Maria Kyrarini, Shuo Jiang, Danijela Ristic-Durrant, and Axel Gräser. 2018. Head Gesture-based Control for Assistive Robots. In *Proceedings of the 11th PErvasive Technologies Related to Assistive Environments Conference*. ACM, Corfu Greece, 379–383. https://doi.org/10.1145/3197768.3201574

[14] Henna Heikkilä and Kari-Jouko Räihä. 2012. Simple gaze gestures and the closure of the eyes as an interaction technique. In *Proceedings of the Symposium on Eye Tracking Research and Applications*. ACM, Santa Barbara California, 147–154. https://doi.org/10.1145/2168556.2168579

[15] Shoya Ishimaru, Kai Kunze, Koichi Kise, Jens Weppner, Andreas Dengel, Paul Lukowicz, and Andreas Bulling. 2014. In the blink of an eye: combining head motion and eye blink frequency for activity recognition with Google Glass. In *Proceedings of the 5th Augmented Human International Conference*. ACM, Kobe Japan, 1–4. https://doi.org/10.1145/2582051.2582066

[16] Won-Seok Kim, Sungmin Cho, Jeonghun Ku, Yuhee Kim, Kiwon Lee, Han-Jeong Hwang, and Nam-Jong Paik. 2020-10. Clinical Application of Virtual Reality for Upper Limb Motor Rehabilitation in Stroke: Review of Technologies and Clinical Evidence. *Journal of Clinical Medicine* 9, 10 (2020-10), 3369. Issue 10. https://doi.org/10.3390/jcm9103369

[17] Georgios Kouroupetroglou and Priti Das (Eds.). 2014. *Assistive Technologies and Computer Access for Motor Disabilities:*. IGI Global. https://doi.org/10.4018/978-1-4666-4438-0

[18] Christian Kray, Daniel Nesbitt, John Dawson, and Michael Rohs. 2010. User-defined gestures for connecting mobile phones, public displays, and tabletops. In *Proceedings of the 12th international conference on Human computer interaction with mobile devices and services*. ACM, Lisbon Portugal, 239–248. https://doi.org/10.1145/1851600.1851640

[19] Ekaterina Kurdyukova, Matthias Redlin, and Elisabeth André. 2012. Studying user-defined iPad gestures for interaction in multi-display environment. In *Proceedings of the 2012 ACM international conference on Intelligent User Interfaces*. ACM, Lisbon Portugal, 93–96. https://doi.org/10.1145/2166966.2166984

[20] Michael F. Land and Mary Hayhoe. 2001. In what ways do eye movements contribute to everyday activities? *Vision Research* 41, 25-26 (Nov. 2001), 3559–3565. https://doi.org/10.1016/S0042-6989(01)00102-X

[21] Rena B Lewis. 1998. Assistive Technology and Learning Disabilities: Today's Realities and Tomorrow's Promises. *JOURNAL OF LEARNING DISABILITIES* 31, 1 (1998).

[22] Franklin Mingzhe Li, Di Laura Chen, Mingming Fan, and Khai N Truong. 2021. "I Choose Assistive Devices That Save My Face" A Study on Perceptions of Accessibility and Assistive Technology Use Conducted in China. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–14.

[23] Franklin Mingzhe Li, Michael Xieyang Liu, Yang Zhang, and Patrick Carrington. 2022. Freedom to Choose: Understanding Input Modality Preferences of People with Upper-body Motor Impairments for Activities of Daily Living. In *Proceedings of the 24th International ACM SIGACCESS Conference on Computers and Accessibility*. 1–16.

[24] Lu Lu, Pengshuai Duan, Xukun Shen, Shijin Zhang, Huiyan Feng, and Yong Flu. 2021. Gaze-Pinch Menu: Performing Multiple Interactions Concurrently in Mixed Reality. In *2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. IEEE, Lisbon, Portugal, 536–537. https://doi.org/10.1109/VRW52623.2021.00150

[25] Marco Moran-Ledesma, Oliver Schneider, and Mark Hancock. 2021. User-Defined Gestures with Physical Props in Virtual Reality. *Proceedings of the ACM on Human-Computer Interaction* 5, ISS (Nov. 2021), 1–23. https://doi.org/10.1145/3486954

[26] Martez Mott, John Tang, Shaun Kane, Edward Cutrell, and Meredith Ringel Morris. 2020. "I just went into it assuming that I wouldn't be able to have the full experience": Understanding the Accessibility of Virtual Reality for People with Limited Mobility. In *The 22nd International ACM SIGACCESS Conference on Computers and Accessibility*. ACM, Virtual Event Greece, 1–13. https://doi.org/10.1145/3373625.3416998

[27] Miguel A. Nacenta, Yemliha Kamber, Yizhou Qiang, and Per Ola Kristensson. 2013. Memorability of pre-designed and user-defined gesture sets. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, Paris France, 1099–1108. https://doi.org/10.1145/2470654.2466142

[28] Vijayakumar Nanjappan, Hai-Ning Liang, Feiyu Lu, Konstantinos Papangelis, Yong Yue, and Ka Lok Man. 2018. User-elicited dual-hand interactions for manipulating 3D objects in virtual reality environments. *Human-centric Computing and Information Sciences* 8, 1 (Dec. 2018), 31. https://doi.org/10.1186/s13673-018-0154-5

[29] Vijayakumar Nanjappan, Hai-Ning Liang, Feiyu Lu, Konstantinos Papangelis, Yong Yue, and Ka Lok Man. 2018. User-elicited dual-hand interactions for manipulating 3D objects in virtual reality environments. *Human-centric Computing and Information Sciences* 8, 1 (Dec. 2018), 31. https://doi.org/10.1186/s13673-018-0154-5

[30] Tomi Nukarinen, Jari Kangas, Oleg Špakov, Poika Isokoski, Deepak Akkil, Jussi Rantala, and Roope Raisamo. 2016. Evaluation of HeadTurn: An Interaction Technique Using the Gaze and Head Turns. In *Proceedings of the 9th Nordic Conference on Human-Computer Interaction*. ACM, Gothenburg Sweden, 1–8. https://doi.org/10.1145/2971485.2971490

[31] Shuji Ogino, Debra G.B. Leonard, Hanna Rennert, Warren J. Ewens, and Robert B. Wilson. 2002. Genetic risk assessment in carrier testing for spinal muscular atrophy. *American Journal of Medical Genetics* 110, 4 (July 2002), 301–307. https://doi.org/10.1002/ajmg.10425

[32] Siyou Pei, Alexander Chen, Chen Chen, Franklin Mingzhe Li, Megan Fozzard, Hao-Yun Chi, Nadir Weibel, Patrick Carrington, and Yang Zhang. 2023. Embodied Exploration: Facilitating Remote Accessibility Assessment for Wheelchair Users with Virtual Reality. In *Proceedings of the 25th International ACM SIGACCESS Conference on Computers and Accessibility*. 1–17.

[33] Ken Pfeuffer, Benedikt Mayer, Diako Mardanbegi, and Hans Gellersen. 2017. Gaze + pinch interaction in virtual reality. In *Proceedings of the 5th Symposium*

on Spatial User Interaction. ACM, Brighton United Kingdom, 99–108. https://doi.org/10.1145/3131277.3132180

[34] Ken Pfeuffer, Lukas Mecke, Sarah Delgado Rodriguez, Mariam Hassib, Hannah Maier, and Florian Alt. 2020. Empirical Evaluation of Gaze-enhanced Menus in Virtual Reality. In *26th ACM Symposium on Virtual Reality Software and Technology*. ACM, Virtual Event Canada, 1–11. https://doi.org/10.1145/3385956.3418962

[35] Thammathip Piumsomboon, Adrian Clark, Mark Billinghurst, and Andy Cockburn. 2013. User-defined gestures for augmented reality. *Augmented Reality* (2013).

[36] Matti Pouke, Antti Karhu, Seamus Hickey, and Leena Arhippainen. 2012. Gaze tracking and non-touch gesture based interaction method for mobile 3D virtual spaces. In *Proceedings of the 24th Australian Computer-Human Interaction Conference*. ACM, Melbourne Australia, 505–512. https://doi.org/10.1145/2414536.2414614

[37] Manuel Prätorius, Dimitar Valkov, Ulrich Burgbacher, and Klaus Hinrichs. 2014. DigiTap: An Eyes-Free VR/AR Symbolic Input Device. In *Proceedings of the 20th ACM Symposium on Virtual Reality Software and Technology* (Edinburgh, Scotland) *(VRST '14)*. Association for Computing Machinery, New York, NY, USA, 9–18. https://doi.org/10.1145/2671015.2671029

[38] Thomas W. Prior, Pamela J. Snyder, Britton D. Rink, Dennis K. Pearl, Robert E. Pyatt, David C. Mihal, Todd Conlan, Betsy Schmalz, Laura Montgomery, Katie Ziegler, Carolee Noonan, Sayaka Hashimoto, and Shannon Garner. 2010. Newborn and carrier screening for spinal muscular atrophy. *American Journal of Medical Genetics Part A* 152A, 7 (May 2010), 1608–1616. https://doi.org/10.1002/ajmg.a.33474

[39] Katharina Reiter, Ken Pfeuffer, Augusto Esteves, Tim Mittermeier, and Florian Alt. 2022. Look & Turn: One-handed and Expressive Menu Interaction by Gaze and Arm Turns in VR. In *2022 Symposium on Eye Tracking Research and Applications*. ACM, Seattle WA USA, 1–7. https://doi.org/10.1145/3517031.3529233

[40] Tyler Rose, Chang S. Nam, and Karen B. Chen. 2018. Immersion of Virtual Reality for Rehabilitation - Review. *Applied Ergonomics* 69 (2018), 153–161. https://doi.org/10.1016/j.apergo.2018.01.009

[41] David Rozado, Jason Niu, and Andreas Duenser. 2015. FaceSwitch - Low-Cost Accessibility Software for Computer Control Combining Gaze Interaction and Face Gestures. In *Proceedings of the Annual Meeting of the Australian Special Interest Group for Computer Human Interaction*. ACM, Parkville VIC Australia, 197–201. https://doi.org/10.1145/2838739.2838809

[42] Jaime Ruiz, Yang Li, and Edward Lank. 2011. User-defined motion gestures for mobile interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, Vancouver BC Canada, 197–206. https://doi.org/10.1145/1978942.1978971

[43] R. Shaw, E. Crisman, A. Loomis, and Z. Laszewski. 1990. The eye wink control interface: using the computer to provide the severely disabled with increased flexibility and comfort. In *[1990] Proceedings. Third Annual IEEE Symposium on Computer-Based Medical Systems*. IEEE Comput. Soc. Press, Chapel Hill, NC, USA, 105–111. https://doi.org/10.1109/CBMSYS.1990.109386

[44] Ludwig Sidenmark, Christopher Clarke, Xuesong Zhang, Jenny Phu, and Hans Gellersen. 2020. Outline Pursuits: Gaze-assisted Selection of Occluded Objects in Virtual Reality. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM, Honolulu HI USA, 1–13. https://doi.org/10.1145/3313831.3376438

[45] Dana Slambekova, Reynold Bailey, and Joe Geigel. 2012. Gaze and gesture based object manipulation in virtual worlds. In *Proceedings of the 18th ACM symposium on Virtual reality software and technology*. ACM, Toronto Ontario Canada, 203–204. https://doi.org/10.1145/2407336.2407380

[46] Mohamed Soliman, Franziska Mueller, Lena Hegemann, Joan Sol Roo, Christian Theobalt, and Jürgen Steimle. 2018. FingerInput: Capturing Expressive Single-Hand Thumb-to-Finger Microgestures. In *Proceedings of the 2018 ACM International Conference on Interactive Surfaces and Spaces* (Tokyo, Japan) *(ISS '18)*. Association for Computing Machinery, New York, NY, USA, 177–187. https://doi.org/10.1145/3279778.3279799

[47] Becky Spittle, Maite Frutos-Pascual, Chris Creed, and Ian Williams. 2023. A Review of Interaction Techniques for Immersive Environments. *IEEE Transactions on Visualization and Computer Graphics* 29, 9 (Sept. 2023), 3900–3921. https://doi.org/10.1109/TVCG.2022.3174805

[48] Wei Sun, Franklin Mingzhe Li, Benjamin Steeper, Songlin Xu, Feng Tian, and Cheng Zhang. 2021. Teethtap: Recognizing discrete teeth gestures using motion and acoustic sensing on an earpiece. In *26th International Conference on Intelligent User Interfaces*. 161–169.

[49] Ovidiu-Ciprian Ungurean, Radu-Daniel Vatavu, Luis A. Leiva, and Réjean Plamondon. 2018. Gesture Input for Users with Motor Impairments on Touchscreens: Empirical Results based on the Kinematic Theory. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, Montreal QC Canada, 1–6. https://doi.org/10.1145/3170427.3188619

[50] Consuelo Valdes, Diana Eastman, Casey Grote, Shantanu Thatte, Orit Shaer, Ali Mazalek, Brygg Ullmer, and Miriam K. Konkel. 2014. Exploring the design space of gestural interaction with active tokens through user-defined gestures. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, Toronto Ontario Canada, 4107–4116. https://doi.org/10.1145/2556288.2557373

[51] Radu-Daniel Vatavu and Ovidiu-Ciprian Ungurean. 2019. Stroke-Gesture Input for People with Motor Impairments: Empirical Results & Research Roadmap. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM, Glasgow Scotland Uk, 1–14. https://doi.org/10.1145/3290605.3300445

[52] Radu-Daniel Vatavu and Ovidiu-Ciprian Ungurean. 2022. Understanding Gesture Input Articulation with Upper-Body Wearables for Users with Upper-Body Motor Impairments. In *CHI Conference on Human Factors in Computing Systems*. ACM, New Orleans LA USA, 1–16. https://doi.org/10.1145/3491102.3501964

[53] Tijana Vuletic, Alex Duffy, Laura Hay, Chris McTeague, Gerard Campbell, and Madeleine Grealy. 2019. Systematic literature review of hand gestures used in human computer interaction interfaces. *International Journal of Human-Computer Studies* 129 (Sept. 2019), 74–94. https://doi.org/10.1016/j.ijhcs.2019.03.011

[54] Jérémy Wambecke, Alix Goguey, Laurence Nigay, Lauren Dargent, Daniel Hauret, Stéphanie Lafon, and Jean-Samuel Louis de Visme. 2021. M[Eye]Cro: Eye-Gaze+Microgestures for Multitasking and Interruptions. *Proc. ACM Hum.-Comput. Interact.* 5, EICS, Article 210 (may 2021), 22 pages. https://doi.org/10.1145/3461732

[55] Ker-Jiun Wang, Quanbo Liu, Yifan Zhao, Caroline Yan Zheng, Soumya Vhasure, Quanfeng Liu, Prakash Thakur, Mingui Sun, and Zhi-Hong Mao. 2018. Intelligent Wearable Virtual Reality (VR) Gaming Controller for People with Motor Disabilities. In *2018 IEEE International Conference on Artificial Intelligence and Virtual Reality (AIVR)*. IEEE, Taichung, Taiwan, 161–164. https://doi.org/10.1109/AIVR.2018.00034

[56] Florian Weidner and Wolfgang Broll. 2019. Interact with your car: a user-elicited gesture set to inform future in-car user interfaces. In *Proceedings of the 18th International Conference on Mobile and Ubiquitous Multimedia*. ACM, Pisa Italy, 1–12. https://doi.org/10.1145/3365610.3365625

[57] Johann Wentzel, Sasa Junuzovic, James Devine, John Porter, and Martez Mott. 2022. Understanding How People with Limited Mobility Use Multi-Modal Input. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 4, 17 pages. https://doi.org/10.1145/3491102.3517458

[58] Jacob O. Wobbrock, Meredith Ringel Morris, and Andrew D. Wilson. 2009. User-defined gestures for surface computing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, Boston MA USA, 1083–1092. https://doi.org/10.1145/1518701.1518866

[59] Huiyue Wu, Yu Wang, Jiali Qiu, Jiayi Liu, and Xiaolong (Luke) Zhang. 2019. User-defined gesture interaction for immersive VR shopping applications. *Behaviour & Information Technology* 38, 7 (July 2019), 726–741. https://doi.org/10.1080/0144929X.2018.1552313

[60] Momona Yamagami, Sasa Junuzovic, Mar Gonzalez-Franco, Eyal Ofek, Edward Cutrell, John R. Porter, Andrew D. Wilson, and Martez E. Mott. 2022. Two-In-One: A Design Space for Mapping Unimanual Input into Bimanual Interactions in VR for Users with Limited Movement. *ACM Trans. Access. Comput.* 15, 3, Article 23 (jul 2022), 25 pages. https://doi.org/10.1145/3510463

[61] Difeng Yu, Xueshi Lu, Rongkai Shi, Hai-Ning Liang, Tilman Dingler, Eduardo Velloso, and Jorge Goncalves. 2021. Gaze-Supported 3D Object Manipulation in Virtual Reality. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. ACM, Yokohama Japan, 1–13. https://doi.org/10.1145/3411764.3445343

[62] Xiaoyi Zhang, Harish Kulkarni, and Meredith Ringel Morris. 2017. Smartphone-Based Gaze Gesture Communication for People with Motor Disabilities. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, Denver Colorado USA, 2878–2889. https://doi.org/10.1145/3025453.3025790

[63] Xuan Zhao, Mingming Fan, and Teng Han. 2022. "I Don't Want People to Look At Me Differently": Designing User-Defined Above-the-Neck Gestures for People with Upper Body Motor Impairments. In *CHI Conference on Human Factors in Computing Systems*. ACM, New Orleans LA USA, 1–15. https://doi.org/10.1145/3491102.3517552

A THE REFERENCE FOR THE SIX VR CATEGORIES

In Section 3.2, We selected six VR categories candidates from the referenced paper [47], or official documents of Oculus or HTC Vive, as shown in Table 6. We initially finalized six VR categories candidates, as shown in Table 7. It is noted that *Pointing* and *Viewport Control* need to receive continuous real-time signals from the VR device, they are not well-suited for transforming into commands for user-defined gestures. Therefore, we finally gathered four VR Commands Category presented in Table 2.

Table 6: Categories of VR Commands and their References. Each VR Commands category is supported by at least one reference, sourced from the referenced paper, or official documents of Oculus or HTC Vive.

Our Work		Reference		
VR Categories Candidate	**Description**	**A review of interaction techniques for immersive environments [47]**	**Oculus Document**	**HTC Vive Document**
Pointing	The act of locating interactive elements through methods such as virtual hand hovering or ray casting.	Pointing	Pointer Cursor	Content Targeting
Selection	Initiating or confirming an action after pointing, such as grabbing an object up close or from a distance.	Selection	Select Something	Selection: Grab and Place
Manipulation	Moving, rotating, or resizing interactive elements, as well as altering their properties.	Translation Rotation Scaling	Move Something Rotate Something Resize Something	Manipulation
Viewport Control	Zooming and panning within an environment using dedicated functions.	Viewport Control		Targeting
Menu-Based Interaction	Presenting a structured set of tabs, commands, or utilities for users to engage with.	Menu-based	Buttons Pinch-and-Pull Components	Context Menu
Locomotion	Moving or changing the direction of an avatar's position within a virtual space.		Locomotion	

Table 7: Six VR Categories candidates and their descriptions.

VR Categories Candidates	**Description**	**Whether Selected as Gestures Design**
Pointing	The act of locating interactive elements through methods such as virtual hand hovering or ray casting.	No
Selection	Initiating or confirming an action after pointing, such as grabbing an object up close or from a distance.	Yes
Manipulation	Moving, rotating, or resizing interactive elements, as well as altering their properties.	Yes
Viewport Control	Zooming and panning within an environment using dedicated functions.	No
Menu-Based Interaction	Presenting a structured set of tabs, commands, or utilities for users to engage with.	Yes
Locomotion	Moving or changing the direction of an avatar's position within a virtual space.	Yes

B THE COMPLETE VISUAL ILLUSTRATION OF HAND GESTURES CATEGORIES

The complete visual illustration of five basic hand gesture patterns, as shown in Figure 11

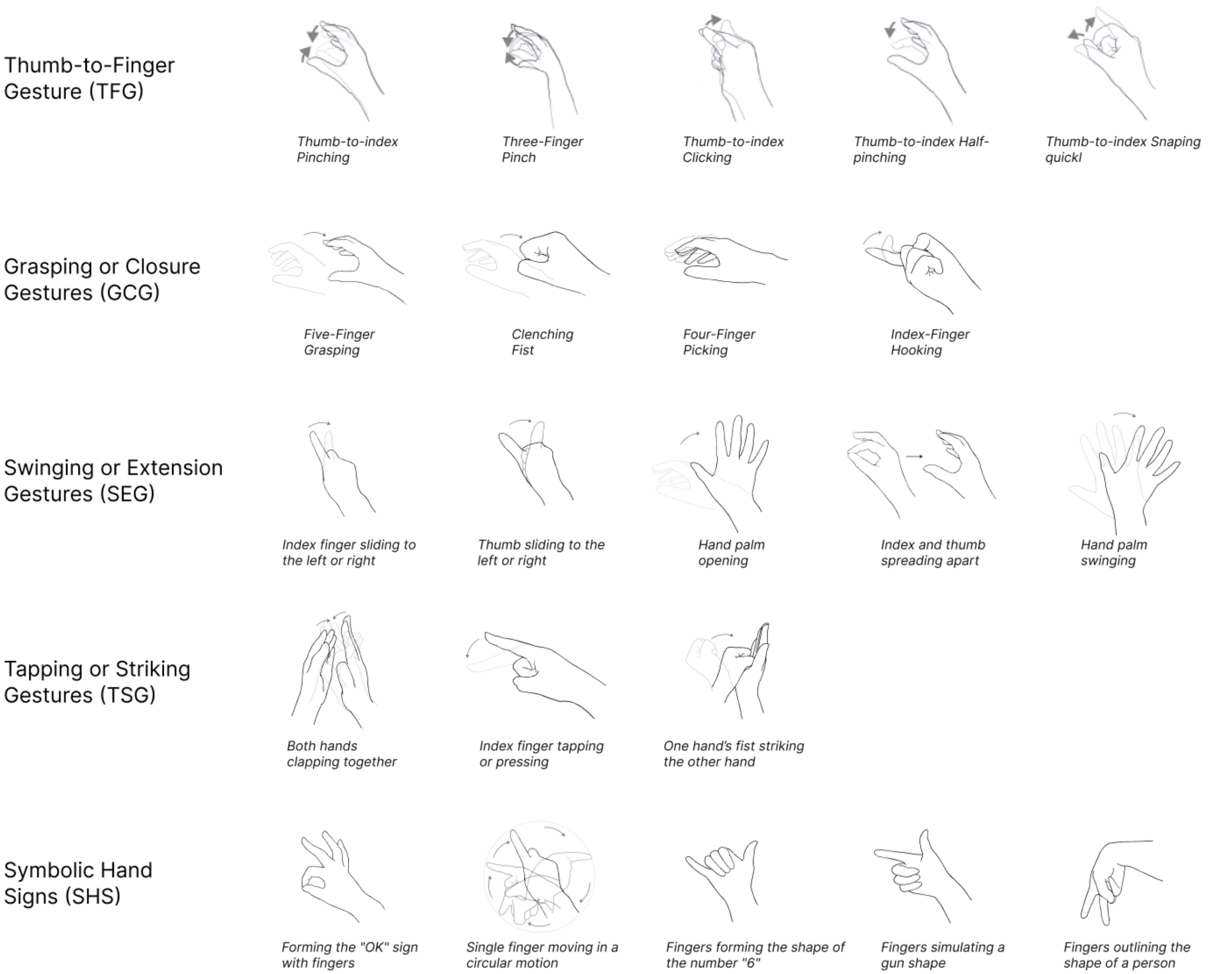


Figure 11: The complete visual illustration of Hand Gestures Categories.