"Can It Be Customized According to My Motor Abilities?": Toward Designing User-Defined Head Gestures for People with Dystonia

Qin Sun
Yunqi Hu
iris119sun@sina.com
huyq@psych.ac.cn
CAS Key Laboratory of Behavioral Science, Institute of Psychology & Department of Psychology, University of the Chinese Academy of Sciences
Beijing, China

Mingming Fan
Computational Media, and Arts Thrust, The Hong Kong University of Science and Technology (Guangzhou), and Division of Integrative Systems and Design & Department of Computer Science and Engineering, The Hong Kong University of Science and Technology, Hong Kong SAR
China
mingmingfan@ust.hk

Jingting Li
Su-Jing Wang*
lijt@psych.ac.cn
wangsujing@psych.ac.cn
CAS Key Laboratory of Behavioral Science, Institute of Psychology & Department of Psychology, University of the Chinese Academy of Sciences
Beijing, China

ABSTRACT

Recent studies proposed above-the-neck gestures for people with upper-body motor impairments interacting with mobile devices without finger touch, resulting in an appropriate user-defined gesture set. However, many gestures involve sustaining eyelids in closed or open states for a period. This is challenging for people with dystonia, who have difficulty sustaining and intermitting muscle contractions. Meanwhile, other facial parts, such as the tongue and nose, can also be used to alleviate the sustained use of eyes in the interaction. Consequently, we conducted a user study inviting 16 individuals with dystonia to design gestures based on facial muscle movements for 26 common smartphone commands. We collected 416 user-defined head gestures involving facial features and shoulders. Finally, we obtained the preferred gestures set for individuals with dystonia. Participants preferred to make the gestures with their heads and use unnoticeable gestures. Our findings provide valuable references for the universal design of natural interaction technology.

CCS CONCEPTS

• **Human-centered computing → Empirical studies in HCI**.

KEYWORDS

Dystonia, Gesture interaction, Interaction technology, Interaction preferences, Human-computer interaction

ACM Reference Format:
Qin Sun, Yunqi Hu, Mingming Fan, Jingting Li, and Su-Jing Wang. 2024. "Can It Be Customized According to My Motor Abilities?": Toward Designing User-Defined Head Gestures for People with Dystonia. 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, 11 pages. https://doi.org/10.1145/3613904.3642378

*Corresponding author

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

ACM ISBN 979-8-4007-0330-0/24/05
https://doi.org/10.1145/3613904.3642378

1 INTRODUCTION

People have always presumed that mobile device design should be suited for hand operation. Smartphones with touchscreen input [30] mainly adopt multi-touch technology, requiring users to flexibly use hand gestures on the interface [21]. However, this approach poses challenges for people with upper body motor impairments, as the interactive experience is neither user-friendly nor convenient. Compared to people with motor impairments, only people without can benefit from the touchscreen with a lower error rate [8]. Many accessibility challenges regarding precision still exist for motor-impaired users [30]. As touchscreens have become mainstream, touchscreen-based interfaces should be usable for people with all abilities [11].

Researchers in human-computer interaction have made plenty of successful explorations about gesture interaction technology with the capabilities of people with disabilities [36]. As defined by Kurtenbach, G. & Hulteen, E. A., Gesture encompasses bodily movements that convey information [14]. In this context, nearly all body parts (e.g., mouth, head, arms, face, hand, shoulders, eyes, etc.) can engage in gesture interaction. This provides more possibilities for the motor impairment's interaction with mobile devices. Prior researchers conducted user studies that evaluated eyelid gestures and proposed body-based gestures for people with upper-body motor impairments to interact with smartphones without finger touch [7, 36, 42]. However, these user-designed gestures were from a wide range of upper-body motor impairments. It remains unknown whether the people with dystonia were included in the motor impairments with the same or different user-defined preferred gestures. As a kind of movement disorder, dystonia is characterized by sustained muscle contractions and abnormal postures of the trunk, neck, face, or arms and legs [28], especially facial dystonia, which present as blepharospasm (eye closing spasms), uncoordinated movements or poor eye-hand coordination causes poor fine motor skills such as eye movements [41]. Then, what preferences would they like to make for gesture interaction with smartphones?

In this paper, we extended the design space of body-based gestures. We investigated the accessibility of smartphone interaction for people with dystonia secondary to cerebral palsy by eliciting gestures on the touchscreen. We employed the guessability methodology of Wobbrock et al. [38] and the referents [1] followed the previous work [26, 36, 42] including 26 command operations on smartphones. Sixteen participants from the group of cerebral palsy with different levels of dystonia took part in the experiment. During the study, we presented participants with video clips demonstrating command actions and then requested them to define their preferred gestures for each referent. Using a think-aloud protocol, we collected qualitative data illuminating users' thinking. When the gesture was elicited, the participants were asked to rate their defined gestures on a 7-point Likert scale in terms of Goodness of Fit, Ease of Use, and Social Acceptance [5, 36, 42]. After completing all the tasks designed, we presented a short semi-structured interview to the participants for their experiment feedback.

A total of 416 user-defined gestures (26 referents x 16 participants) were elicited in the experiment. Building upon this dataset, we calculated the agreement score of each referent and analyzed insights learned from participants' feedback about user-defined gestures[2]. Subsequently, we assigned the best head-based gestures for 26 referents. The results show that the participants with dystonia had specific preferences for gestures to interact with smartphones without finger touch. Participants preferred to make the gestures with their heads and preferred the gestures with small movements that were less obvious in public. Besides that, we also compared the user-defined head-based gesture set obtained in our experiment with the ones from other studies for people with motor impairment.

In summary, the main contributions of this study are as follows:

(1) we took a first step to designing user-defined head gestures for people with dystonia, whose muscle ability poses challenges for the usage of user-defined gestures in prior work (e.g., [7, 18, 42]);
(2) we compared the user-defined gestures from this work with those from prior work and highlighted the commonalities and uniqueness of gestures for people with dystonia and possible rationales;
(3) we further provide design guidelines for making user-defined smartphone gestures more accessible for people with dystonia.

2 RELATED WORK

We build upon prior work on diverse mobility, needs, and gestures designed for people with body motor impairments.

2.1 Diverse mobility and needs among people with upper body motor impairments

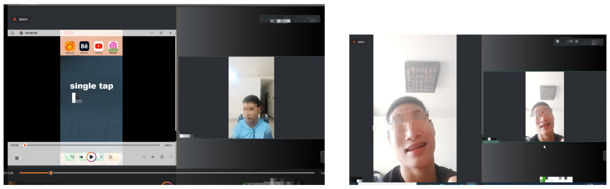
People with upper body motor impairments, such as tremors, muscular dystrophy, loss of arms, dystonia, or lack of sensation, have diverse mobility and different needs. Regarding the accessibility challenges for smartphone use, gesture interaction technology helps with the alternative. We can entirely focus on their remaining abilities. For example, some people who lose arms but with sound lower limbs can use their feet to interact with smartphones [10], while some only with arm tremors can use their eyelid gestures [7], or other body-based gestures [20, 24, 42] to interact with smartphones. Extending from hands to other body parts, incorporating gestures of other body parts will increase gesture variety [2], allowing any body part to convey interactive intentions, thereby encapsulating ample interactive information. From this perspective, this convenience makes available that users with various abilities have equal access to mobile intelligent products [2, 29], which motivated our research. In summary, regarding the accessibility challenges for smartphone use, gesture interaction technology can help with the alternative. Although people with dystonia find it hard to do much about fine motor tasks, they can utilize the remaining capabilities with corresponding body-based gestures interacting with smartphones. Inspired by such needs and to find more accessibility to smartphone interaction, we conducted a user study to find what user-defined gestures people with dystonia would want to use.

2.2 Gestures Designed for People with Motor Impairments

Numerous research endeavors have been directed toward formulating gesture interaction. Recently research on stroke-gesture input for wearables (e.g., head-mounted displays, TouchRing, smartwatches, etc.) could increase accessibility for users with upper body motor impairments [18, 19, 31, 33]. However, these required the users' specific motor abilities of steady movements of their hand or finger to produce a straight, uninterrupted path on the touchscreen and lift off the finger to finalize input. Hu et al. [10] explored foot-based interaction with smartphones for people with upper body motor impairments and gained foot gesture sets, with the usage scenario limitation of reclining on the bed.

Fan et al. and Zhao et al. [7, 42] explored the potential of eyelid gestures and proposed above-the-neck gestures involving eye, mouth, and head movements for people with motor impairments to interact with mobile devices without finger touch, concluding that the participants prefer to make the gestures with their eyes. However, eye-based gestures include many fine motor tasks like blinking, winking, gazing, controlling eye movements, as well as opening and closing the eyes, along with possible combinations thereof [7, 9, 15, 36] which is difficult for people with dystonia due to poor coordination. In addition, like the participant with upper body motor impairments in the prior study suggested using his nose to create gestures [18], other upper body parts, especially ones involving facial gestures, tongue, and nose, can also be used to alleviate the sustained use of eyes in the original proposed user-defined gestures. Motivated by previous works, we conducted in-depth research on gesture interaction with smartphones in the cerebral palsy group with dystonia. We aimed to explore any unique interaction preferences and how they will use head gestures to interact with mobile devices.

[1]The word referent is following the terminology of Wobbrock et al. that used to denote the effect of a gesture command [38]. Jacob O. Wobbrock, Htet Htet Aung, Brandon Rothrock, and Brad A. Myers. 2005. Maximizing the guessability of symbolic input. In CHI '05 Extended Abstracts on Human Factors in Computing Systems. Portland, OR, USA, pp. 1869–1872. https://doi.org/10.1145/1056808.1057043.
[2]Source data is available on:
https://github.com/MELABIPCAS/Head_Gestures_Dystonia.git.



(a) (b)

Figure 1: Examples of the experimental setup for head gestures elicitation. Fig. 1(a) shows overview setup; Fig. 1(b) shows a participant in response to a referent.

3 METHOD

This section describes how people with dystonia design the head gestures for the referents selected from the smartphone. Our primary goal was to collect head gestures from people with dystonia and identify their preferred gestures set.

3.1 Participants

The number of participants in the experiment was based on the definition of user-defined gestures in previous studies. In previous studies, over 70% of the number of subjects ranged from 10 to 30. Moreover, small sample sizes are also reasonable in special situations, such as for people with disabilities [36]. Sixteen voluntary participants (N=16), comprising nine males and seven females, with an average age of 25 (SD = 6), were recruited online for the study. Table 1 shows the demographic information of the participants. All participants were diagnosed with dystonia and motor impairment. Five of the participants had a language barrier with pronunciation difficulties in fluency and clarity. Eleven individuals had a secondary or higher education background, while five had a primary or lower education. Before the study, they all had smartphone experience and never using similar gesture control devices. After the study, all the participants were compensated for their time.

3.2 Referents selection and experimental setup

The target device in the experiment is a smartphone. Our study was a within-subjects design with one independent variable: referents. We surveyed common referents in previous research e.g. [1, 7, 21, 26, 39, 42] so that the gesture set extends to a broad range of smartphone applications. And finally, twenty-six referents were settled for inducing material, which are 12 General commands, 10 App-related commands, and 4 Button-related commands [7, 42].

An overview of the experimental setup is shown in Fig. 1. The participants can pick a quiet and private time seated in front of their personal device running the video conference platform. The entire process was recorded by webcam. One experimenter was responsible for monitoring the experimental process, the gestural video recordings, and writing down participants’ oral descriptions. Participants need to keep their upper body displayed in the video during the experiment as presented in Fig. 1(a). Besides, Fig. 1(b) shows a participant responding to a referent.

3.3 Procedure

Fig. 2 shows the procedure of the study. Before the experiment, participants were informed of the details (e.g., the purpose, tasks, requirements, etc.) in an online questionnaire, which included general demographic information and a simple investigation of smartphone use. In addition, participants had to sign a consent form.

Once ready, participants watched the video clips. After clarity on the effect of tasks, the participant began to design the head gesture for each referent in a think-aloud manner. During the process, participants were encouraged to create more than one head gesture for each referent, perform them three times, and then select the best gesture from them. To reduce the sequential effect of task presentation, if participants ask to know about all tasks first, we will give participants one minute to glance over 26 tasks. Following the completion of a gesture design, participants were asked to evaluate the three aspects on a 7-point Likert scale. For 26 tasks, participants need to repeat 26 times of the above process.

After completing the elicitation process, we presented a post-experiment questionnaire to participants to collect their general feedback for the experiment. The study lasted approximately 1.5 to 2 hours. In the end, we obtained 416 (16 participants × 26 referents) user-defined preference head gestures.

3.4 Data collection and analysis

The data include recorded videos, interview transcriptions, gesture proposals, and participant subjective ratings. After the experiment, we transcribed the video records. Two researchers (the first and the second authors) encoded and classified these data based on a strategy similar to previous gesture papers [2, 5, 20, 25, 27, 39, 42]. We classified the gestures according to the different body parts involved while combining some gestures based on the similarities of facial action units. Any divergence in coding between them was resolved by talking together and asking a third expert (the fourth author) until reached a consensus. To compare our gesture set to previous studies, we also adopted the mathematical formula of Wobbrock et al. [38, 42] computing the agreement score Ac, which a great number of elicitation studies have used [4, 23, 27, 32, 39, 42]. where:

$$A_c = \sum_{P_i} \left(\frac{P_i}{P_c} \right)^2$$

Dystonia is characterized by abnormal involuntary movements with five levels based on clinical symptoms severity. After talking to the rehabilitation doctor, we separated the participants into two groups according to their behavioral expressions: severe and mild. The people in the severe group were the ones whose range of motion and position were oversized even when doing eye-based (e.g., blink, close the eyes…) gestures, the entire facial muscle is tense. In our study, 2 out of 16 participants were decided in the severe group while we counted their gesture proposals into data referring to both their actions and verbal descriptions.

4 RESULTS

In the experiment, a total of 594 gestures were proposed for 26 referents. Based on participants’ subjective rate, we collected 416

Table 1: Participants' demographic information

ID	Age	Gender	Type	Level	Dystonia	Education
P1	21	Female	Speech/limb movement disorder	I	Ataxia type	Bachelor above
P2	27	Female	Hearing/ speech/ limb impaired	I	Spastic	Bachelor above
P3	25	Male	Speech/limb movement disorder	I	Mixed	senior school
P4	24	Female	Limb movement disorder	II	Spastic	Bachelor above
P5	42	Male	Limb movement disorder	II	Mixed	Bachelor above
P6	20	Male	Limb movement disorder	II	Spastic	primary school
P7	19	Male	Visual/ limb impaired	I	Spastic	Illiteracy
P8	23	Male	Limb movement disorder	III	Ataxia type	middle school
P9	20	Male	Limb movement disorder	II	Dyskinetic	primary school
P10	30	Male	Speech/limb movement disorder	II	Dyskinetic	middle school
P11	21	Female	Limb movement disorder	II	Spastic	Illiteracy
P12	22	Male	Limb movement disorder	II	Spastic	primary school
P13	31	Female	Limb movement disorder	III	Dyskinetic	senior school
P14	21	Female	Limb movement disorder	II	Spastic	middle school
P15	20	Female	Visual/ limb impaired	I	Spastic	middle school
P16	30	Male	Speech/limb movement disorder	II	Mixed	senior school

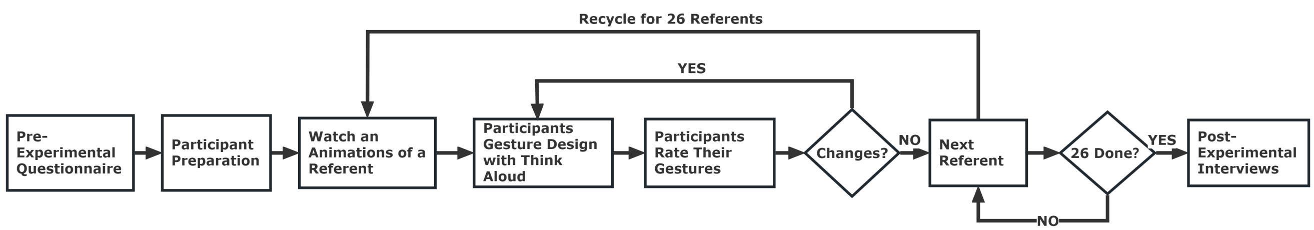


Figure 2: The diagram shows an overview of the experimental process.

gestures (26 referents x 16 participants). The results contained gesture taxonomy, agreement score, subjective rates, the preferred gesture set, and feedback observation through the experiment process.

4.1 Gesture taxonomy

There are certain differences in participants' perception of body-based gestures, while the range of motion and position for gestural interaction vary among individuals. Therefore, before analyzing, we classified the 416 gestures collected from the participant's behavior and descriptions of the gestures in the experiment and merged the same or similar body part movements.

4.1.1 Gesture Categories. According to the gesture taxonomy methods in previous studies [12, 23, 42], we classified the gestures according to the different body parts involved. Table 2 shows the details of the taxonomy strategy of head gestures. The body parts proposed by the participants included eyes, mouth, nose, tongue, teeth, and shoulders. We considered the parts of the nose, tongue, and teeth into the mouth category based on the similarities of facial action units. Consequently, four body parts were grouped into the gestures: the head, the mouth, the eyes, and the shoulder. Including the combinations (action sequence and frequency) of different body parts, we elicited 12 different dimensions of gesture.

4.1.2 Findings from the classification. We found 244 unique head gestures after merging the overlaps. Figure 3 shows the composition of head gestures. Three parts of the head, eyes, mouth, and their combination were used most frequently, accounting for a total of 77 percent. Among them, the proportion of head-based gestures was 34.8% (appeared 85 times), eyes-based gestures were 17.6% (appeared 43 times), and mouth-based gestures were 18% (appeared 44 times). From Fig. 3, it can be found that the gestural interaction based on head action is almost twice that based on eye action. This finding indicates that although eye-based gestures are more selective than head-based gestures, participants were more willing to choose head-based gestures, which differs from the conclusion mentioned in the paper [7, 42]. The possible reasons can be analyzed from the following two aspects: (i) gestural interaction frequently using eye movements and blinking can easily lead to a sense of tension in the eye muscles and nerves [6]. For people with dystonia, higher facial muscle tension may cause posture disorder when interacting with devices. In contrast, they prefer to use head-based gross movements. (ii) from the perspective of developmental psychology, human motor development follows a sequence from beginning to end, while the head motor is the earliest form of human development. In our daily conversation, we often use head movements such as nodding or shaking for feedback (such

Table 2: Head gesture classification based on facial muscle movements

Category	Dimension	Details
Single body part	Pose or motion	only head (direction and distance); only mouth (lips, cheek, nose, tongue); only eyes; only shoulder.
Parts Combinations	Action sequence	more than one part action at the same time, the other part action after one move.
	Action frequency	instant action, continuous action, once, twice, three times, repeat…

as nodding to agree or shaking to deny) [3]. From an evolutionary psychological perspective, compared to eye actions, people are more accustomed to conscious head movements [13].

4.2 Preferred head gesture set

We calculated the Agreement Score of the referents based on the frequency of the same head gesture. After addressing the issue of gesture conflicts, we assigned the preferred gesture set.

4.2.1 Agreement Score. We evaluated the degree of gestural consensus among our participants to quantify agreement. Fig. 4 illustrates the agreement of the gestures for each referent. There are 5 commands with high consensus (sliding left/right/up/down, phone lock), 15 commands with moderate consensus, and 6 commands with low consensus. The overall average agreement score for 26 referents was 0.201 (SD=0.132) obtained in the study, which implied a medium agreement according to the interpretations for the magnitudes of agreement score by Vatavu and Wobbrock [34]. These results are similar to the others reported in the literature [26, 42].

The maximum agreement was reached in Scroll up/down commands (Ac = 0.492). These are the paired commands, and in the experiment, we presented them together to the participants. Participants designed four interactive gestures based on head movement in different directions. The mapping gesture for “Scroll up” was “head up and look upward” chosen by 11 participants, and the corresponding gesture for “Scroll down” was “lower head and look downward” chosen by 11 participants. A high agreement score indicates that participants are more likely to use these gestures for commands. In contrast, the minimum agreement score was reached in referents of “Open Previous/Next App in the Background” (Ac = 0.07), both of which also had symmetry. Participants proposed 13 different head gestures, while the highest score was “head forward, head turn right/left, nod,” with only 2 participants choosing them. The result for these two commands with the lowest agreement score is consistent with the conclusion in [39], indicating that more complex commands would result in lower consensus.

The complexity of “Phone Lock” is greater than that of “Single Tap” or “Double Tap”. According to previous studies, the gesture agreement for conceptually complex commands should be relatively low, but there is a high level of agreement here. This is related to the psychological assumption that users tend to simplify commands. As P1 said after watching the video of the lock screen command, “Oh, it’s the same as locking and closing...”

4.2.2 Conflict Strategy. According to the conclusion of the gesture distribution in section 4.1.2, the participants preferred head movements-based gestures. Therefore, we considered the head as the first choice in body gesture allocation, followed by the mouth and eyes. Take the command “Zoom out” (Table 3) for example, gesture 1 and gesture 2 had the same agreement score, and then, we assigned the head action as the preferred gesture to it. Participants’ cognitive patterns, past experiences, etc., have an impact on the choice of body-based gestural interaction [37]. We adjust specific gestures based on the analysis of the qualitative feedback during the elicitation process, such as following the body actions from simple to complex and pair command mapping, etc. As shown in table 3, the commands of “zoom in/out” are paired commands. Therefore, for the ”zoom in” command, although the gesture of “wide open mouth” proposed by the participants had a higher level of consensus, we assigned the “head forward” as the optimal gesture to it, considering the consistency of the paired commands. Besides that, if the same gesture was assigned to both a single command (such as rotation) and a command with symmetry (such as slide left/right), we followed a previous study [42] that prioritized the gesture to paired ones and the gesture with the second-highest score was allocated to the single command.

4.2.3 Final Head Gesture Set. Table 4 shows the optimal user-defined head gesture set for each command. SET 1 in the table is the user-defined head gestures mapping for 26 smartphone commands based on subjective rating, insight understanding, and the level of consensus. SET 2 is the alternative one second only to SET 1. And all gestures were proposed by more than one participant. Fig. 5 intuitively illustrates the final user-defined head gestures for the 26 referents. There were 10 were only head gestures, 4 were only mouth gestures, and 4 were only eye gestures.

We can find some relevance between the referent and the corresponding head gesture from the optimal gesture set. As to the paired commands (e.g., sliding up, down, left, and right), the allocated head gesture in high consensus proposed by the participants was head motion in the corresponding direction (up, down, left, and right). For the referents with the features of frequency (e.g., single tap and double tap), participants proposed head gestures based on action frequency (nodding/blinking once/twice), which was consistent with previous research findings. For the referent of “Zoom in/out,” the preferred gesture proposed by the participants is not based on changing the size of body parts but on real-life practice, which associated the distance causing the object size change from a visual perspective.

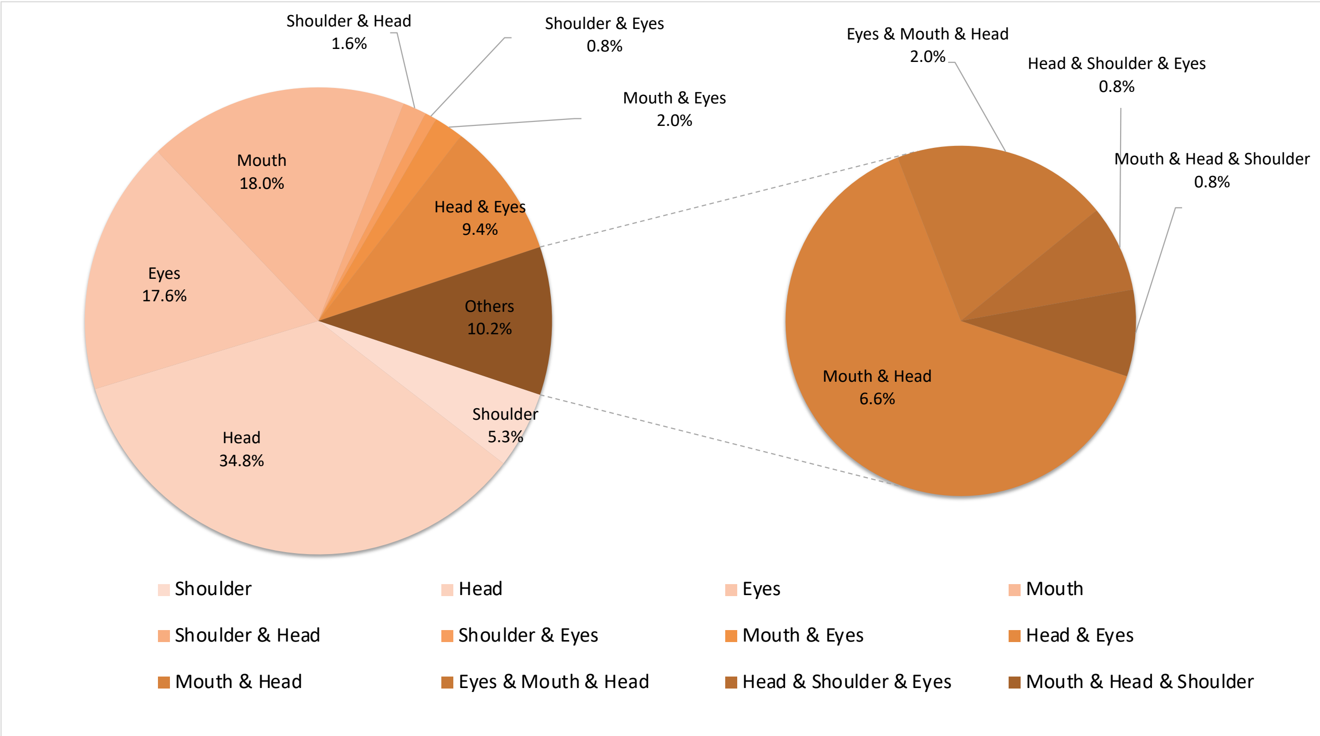


Figure 3: Distribution of gestures per category

Table 3: Strategy for Solving Gesture Conflicts

References	Gesture 1	Ac (gesture1)	Gesture 2	Ac (gesture2)	Gesture 3	Ac (gesture3)
Zoom out	head hypokinesis	0.035	squint eyes	0.035	Stick out the tongue	0.016
Zoom in	head forward	0.016	wide open eyes	0.035	Open the mouth	0.035

Conflict 1: For a command (e.g., Zoom out), have three types of gestures allocated, and gesture 1 conflicted with gesture 2 with the same highest agreement score. Then, we assigned the head action as the preferred gesture to it.

4.2.4 Subjective Ratings. Many recent investigations within the domain of gesture studies have incorporated subjective assessments as a pivotal metric for validating user-defined gesture sets [5, 17, 25, 40, 42]. In our experiment, after completion of the gesture elicited for a task, participants rated their gestures on three dimensions: goodness of fit, easiness of performing, and social acceptance. Subsequently, we divided the gestures of each task into two groups: the first group related to the user-defined best-preferred gestures, and the second group related to all others that do not belong to the best-preferred set. Then, we did the difference test for the participants' average goodness of fit (mean of large groups=5.83, mean of small groups=5.67), easiness of performing (mean of large groups=5.96, mean of small groups=5.96), and social acceptance (mean of large groups=5.86, mean of small groups=5.51) of the gestures in these two groups. The result showed that the three subjective dimensions were no significant difference in groups, which is consistent with the conclusion drawn in a previous paper [42]. The reason would be that participants always believed that the gestures they designed were the most suitable for them. Further, when we did the Pearson Correlation Test for the subjective ratings and agreement score, Pearson's correlation coefficient showed a positive correlation between social acceptance and the agreement score ($r=0.493$, $p=0.010 < 0.05$). This indicated that, in the interactive techniques, the participants' perception of social acceptance of the gestural interaction is an important factor. There was no significant correlation between the other two aspects.

4.3 Feedback and Observation

Three themes were summarized for the insights of elicited gestures.

Usage scenario Participants considered the contexts both private and public. When designing for the "Next Button" task, P4/P14 raised similar concerns: " Can I set based on context? The first body part on my mind is my tongue. It is quite comfortable to do

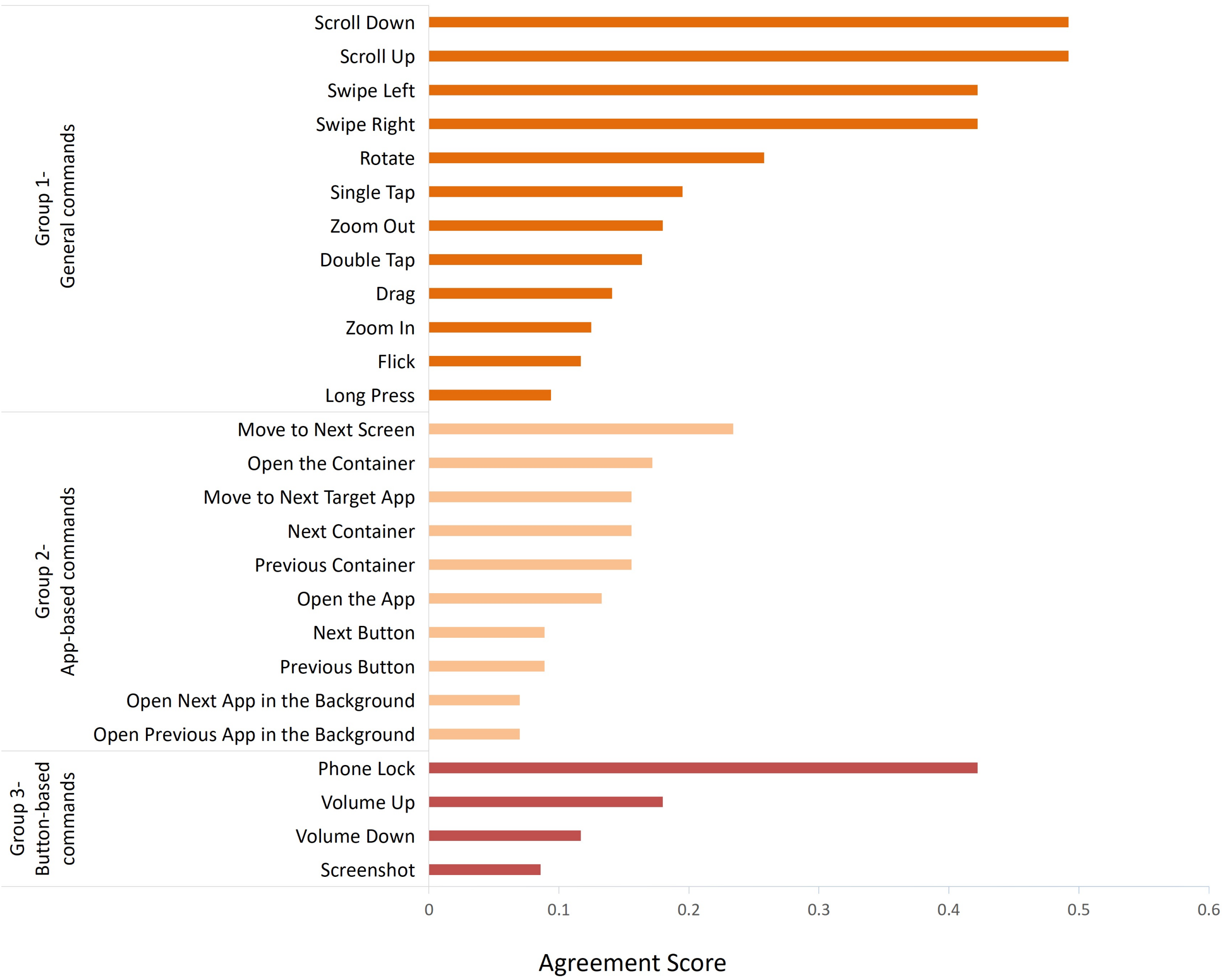


Figure 4: Agreement scores sorted by referents (the higher the better; low agreement: 0–0.1, medium agreement: 0.1–0.3, high agreement: 0.3–0.5, very high agreement: >0.5)

on my own, but I wouldn’t do it in public. I think that sticking out my tongue is a bit ugly.” In design processing, participants are often intuited to suggest comfortable gestures based on their first impression. Then, they present a willingness to use it in private but not in public, which is similar to the conclusion in the paper [22].

Structural metaphor The gestural interaction is not enough to reveal its metaphorical nature but is related to the user’s psychological model [4]. Taking the “Zoom In/Out” task as an example, P2 said, “It is much more intuitive to use the eyes or the head to infinitely close the phone. This felt like being unable to see clearly and moved closer for a better look.” Participants would associate it with scenes of life that they would attempt to get close enough to see clearly when the target looks blurry. This leads to the idea of designing a head gesture based on distance. When designing head gestures for paired commands, participants often selected body parts that can achieve symmetric actions, preferring to keep the consistency of task orientation. For example, the “Swipe/Flick” task includes up, down, left, and right in the dimension of surface and motion, and the participants selected the corresponding head to turn up, down, left, and right. The result of the study was that the level of consensus for these tasks was very high. Besides, some participants conduct the gestures for tasks based on the function of sensory organs, such as using the mouth for Volume up/down, using eyes-based gestures for select tasks, etc.

Influence of the severity of dystonia on gesture performance We grouped the participants into severe and mild in terms of dystonia severity. From behavior observation through the experiment process, the gesture actions in mild groups can almost be identified and comply with what they described. Participants

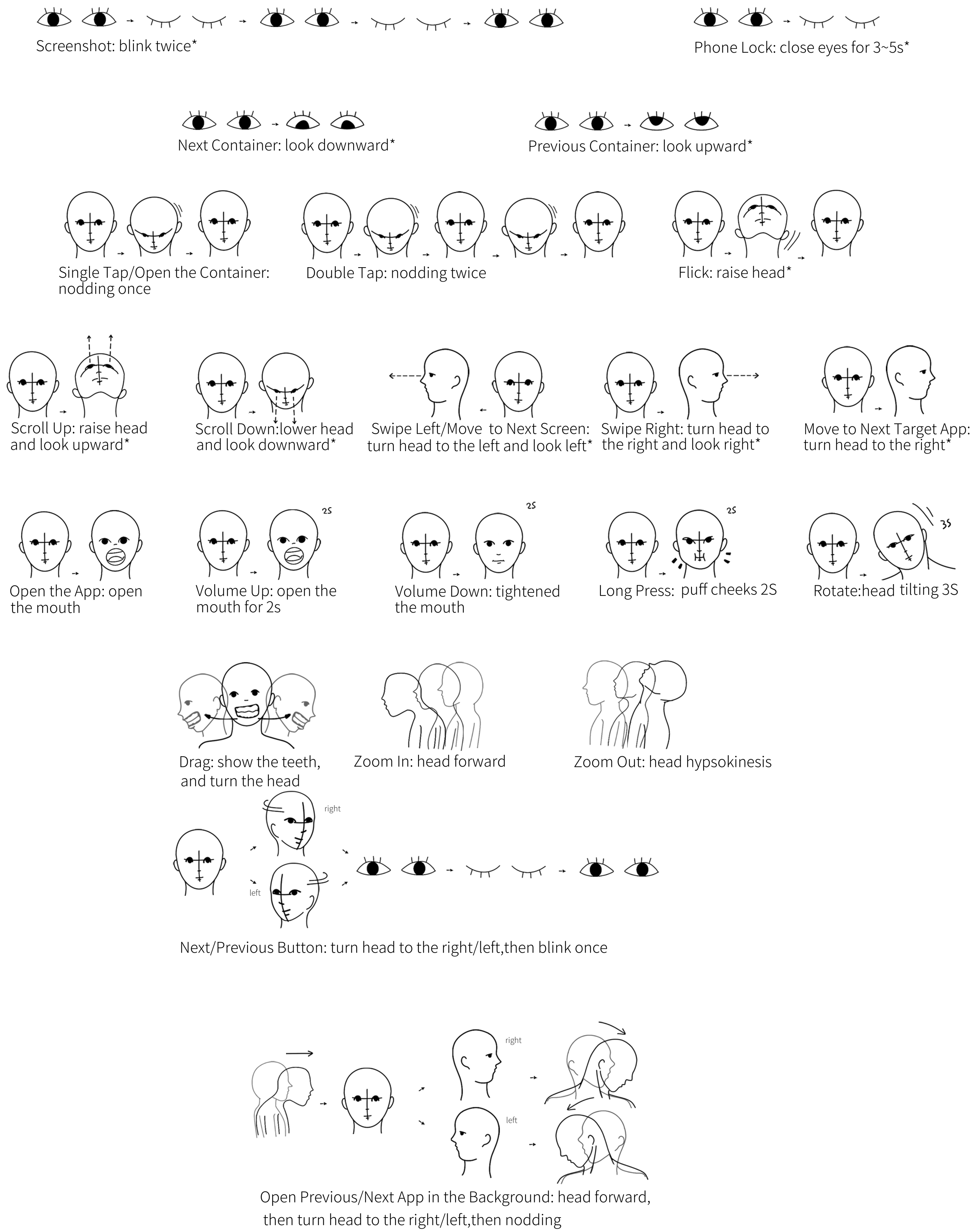


Figure 5: Visual illustrations of the final user-defined head gestures for 26 referents. (the gestures marked with * are similar to the prior paper [43])

Table 4: The preferred gesture set

Referents	SET 1	SET 2	SET 3
Single Tap	nodding once	blink once	turn head
Double Tap	nodding twice	blink twice	
Flick	raise head	open mouth	smile
Long Press	puff cheeks 2S	open mouth	Close eyes 2s
Scroll Up	raise head and look upward		
Scroll Down	lower head and look downward		
Swipe Left	turn head to the left and look left	head tilting to the left	
Swipe Right	turn head to the right and look right	head tilting to the right	
Zoom In	head forward	wide open eyes	Open the mouth
Zoom Out	head hypokinesis	squint eyes	Stick out the tongue
Drag	show the teeth, and turn the head	gaze, and stick out the tongue	Wry mouth
Rotate	head tilting 3S	Turn head to the left	shrug the left shoulder
Open the App	open the mouth	Head shake and blink twice	Eye movement, and blink
Move to Next Screen	turn head to the left and look left	Turn head to the left	
Open the Container	nodding once	eye movement and blink twice	Blink 3times
Next Button	turn head to the right, then blink once	shrug the right shoulder and blink twice	Head forward
Previous Button	turn head to the left, then blink once	shrug the left shoulder and blink twice	head hypokinesis
Next Container	look downward	slowly lower head	
Previous Container	look upward	slowly raise the head	
Move to Next Target App	turn head to the right	eyes moving	turn head to the left
Open Previous App in the Background	head forward, then turn head to the right, then nodding		
Open Next App in the Background	head forward, then turn head to the left, then nodding		
Phone Lock	close eyes for 3~5s	blink twice	
Volume Up	open the mouth for 2s	tongue out and upturned	raise head for 10s
Volume Down	tightened the mouth	lower the head	pout
Screenshot	blink twice	eyes blink 3 times	nodding 3times

preferred to perform gestures that were simple and easy to remember. When designing a gesture for the task of “double tap”, P2 said: “Considering this command using frequently, simple actions are not getting too tired.” However, only in terms of behaviors is it hard to see what kind of gestures are made for tasks by participants in severe groups. Taking P3 as an example, when he made the gesture of “closing the eyes” for the task, the whole facial muscle tensed (e.g., furrowing his brow, keeping the mouth tightly closed). Meanwhile, we found that his tongue is very flexible and can make many mouth-based gestures such as sticking out the tongue, turning the tongue left/right, bulging the cheek, etc.

5 DISCUSSION

In this section, we conducted a comparative study between previous works for motor impairment and discussed the commonalities and uniqueness of gestures for people with dystonia and possible rationales. Besides, we reflect on design implications.

5.1 Comparative study to prior papers

Our work through an elicitation study found that the people with dystonia exhibited head gesture preference, compared to the preference towards eye-based gestures for people with upper body motor impairment. [7, 42]. This difference would be linked to the specific motor characteristics of individuals with dystonia. We extend Zhao et al.’s conclusion that interactive gestures are the same for 10 out of 26 common tasks, while the other 16 common tasks are with different allocated gestures (marked * in Fig. 5). This difference indicates that groups with diverse abilities would have a personal-tailed gesture set. The participants in the previous are the people with a wide range of upper-body motor impairments, which include dystonia.

5.2 Reflect on design implications

Customized gestures based on motor abilities Our findings provide the guidelines for smartphones to be more accessible for people with dystonia. From gesture classification, we found that people with dystonia prefer the gross motion of head-based gestures followed by mouth-based gestures due to specific motor characteristics. Especially for severe dystonia, it's difficult to see what gestures they make, identifying by most facial actions, but they can make many tongue-based gestures smoothly. The high dexterity of the tongue should make it a good candidate for people with upper body impairment. In this study, about 5% of participants suggested making gestures with their tongues. However, tongue gesture interaction studies are almost for severe motor impairment such as full-body paralysis (e.g., [16, 22]). A possible solution is to investigate more system customization services adapting to users with diverse abilities, for example, designing different head gesture set packages for smartphone tasks in the same category.

Potential challenges Our study mainly explores interaction technology solutions from the user's perspective: What interactive gestures can they engage in, and are they willing to do without considering recognition accuracy? To a general camera, it's difficult to accurately detect some gestures with micro action, such as tightening the mouth. For dystonia, muscular spasms might result in high recognition error rates, generating frustration for the user. Furthermore, due to diverse motor characteristics, the participants proposed gestures based on what they could perform involving various facial action units (e.g., head, mouth, tongue, teeth, etc.). The gained final head gesture sets are typical and generalizable to people with mild dystonia. However, whether and to what extent such user-defined gestures can be generalized to people with other types of motor impairments remains unclear. Future work is warranted to answer such a question.

6 LIMITATIONS AND FUTURE WORK

We employed the same methodology [38]as the newest research [42]. However, the potential limitations of this traditional method should be considered. In our work, we defined the similarity criteria to identify the proposals for each referent and employed the A formula to calculate the agreement score, which depends on the number of proposals provided by participants. Although the gestural proposals for each referent were agreed upon by more than one participant in this paper, this approach would lead to an unstable zero-agreement level along with the sample size expanding [35]. Nevertheless, it is interesting to apply other recent measures of agreement for further investigations. In addition, although the participant's previous experience with other interfaces is desirable and helpful for exploring more accessible gestural interaction with smartphones for people with dystonia, legacy bias would be a potential concern for effectiveness in exploring new interactions.

Our study aimed to explore the gestural interaction with smartphones for people with dystonia. While using multiple methods across the study, due to the inherent varied abilities of the sample, our results are preliminary, and we believed that our finding of the gesture set would be a good framework for people with dystonia to interact with smartphones without touch. In the future, more investigations with expanding the number of participants are necessary to achieve more generalizable findings.

7 CONCLUSION

We took the first step to explore the user-defined head gesture interactive technology on smartphones in people with dystonia. Based on participants' agreement over 416 gesture samples, subjective rating, and understanding of participants' feedback, we obtained a user-defined gesture set. The findings notably illuminate that, although eye-based gestures are more diverse, participants were more willing to choose head-based gestures followed by mouth-based gestures. Furthermore, we compared our result with the conclusions in previous related gesture studies and expended their findings that although some overlap, people with dystonia have a preference for gesture choices based on their conditions. Finally, we highlight the reflection on design implications together with the limitations and future work.

ACKNOWLEDGMENTS

This paper is partially supported by 1) the National Natural Science Foundation of China (62276252, 62106256); 2) the Youth Innovation Promotion Association CAS; 3) Guangdong Provincial Key Lab of Integrated Communication, Sensing and Computation for Ubiquitous Internet of Things (No.2023B1212010007); 4) 2024 Guangzhou Science and Technology Program City-University Joint Funding Project (PI: Mingming Fan).

REFERENCES

[1] Muhammad Anshari and Yabit Alas. 2015. Smartphones habits, necessities, and big data challenges. *Journal of High Technology Management Research* 26, 2 (2015), 177–185. https://doi.org/10.1016/j.hitech.2015.09.005

[2] Zhen Chen, Xiaochi Ma, Zeya Peng, Ying Zhou, Mengge Yao, Zheng Ma, Ci Wang, Zaifeng Gao, and Mowei Shen. 2017. User-Defined Gestures for Gestural Interaction: Extending from Hands to Other Body Parts. *International Journal of Human-Computer Interaction* 24, 3 (2017), 238–250. https://doi.org/10.1080/10447318.2017.1342943

[3] Darwin and Charles. 1978. The Expression of the Emotions in Man and Animals. *Journal of Nervous & Mental Disease* 123, 1 (1978), 90. https://doi.org/10.7208/chicago/9780226220802.001.0001

[4] Haiwei Dong, Ali Danesh, Nadia Figueroa, and Abdulmotaleb El Saddik. 2015. An Elicitation Study on Gesture Preferences and Memorability Toward a Practical Hand-Gesture Vocabulary for Smart Televisions. *IEEE Access* 3 (2015), 543–555. https://doi.org/10.1109/ACCESS.2015.2432679

[5] Ze Dong, Thammathip Piumsomboon, Jingjing Zhang, Adrian Clark, Huidong Bai, and Rob Lindeman. 2020. A Comparison of Surface and Motion User-Defined Gestures for Mobile Augmented Reality. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI EA '20)*. Association for Computing Machinery, New York, NY, USA, 1–8. https://doi.org/10.1145/3334480.3382883

[6] Heiko Drewes. 2010. *Eye Gaze Tracking for Human Computer Interaction.* PhD. https://doi.org/10.5282/edoc.11591

[7] Mingming Fan, Zhen Li, and Franklin Mingzhe Li. 2021. Eyelid Gestures for People with Motor Impairments. *Commun. ACM* 65, 1 (dec 2021), 108–115. https://doi.org/10.1145/3498367

[8] Leah Findlater, Karyn Moffatt, Jon E. Froehlich, Meethu Malu, and Joan Zhang. 2017. Comparing Touchscreen and Mouse Input Performance by People With and Without Upper Body Motor Impairments. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI '17)*. Association for Computing Machinery, New York, NY, USA, 6056–6061. https://doi.org/10.1145/3025453.3025603

[9] Tiago Guerreiro. 2021. Technical perspective: Eyelid gestures enhance mobile interaction. *Commun. ACM* 65, 1 (2021), 107. https://doi.org/10.1145/3498365

[10] Xiaozhu Hu, Jiting Wang, Weiwei Gao, Chun Yu, and Yuanchun Shi. 2021. FootUI: Assisting People with Upper Body Motor Impairments to Use Smartphones with Foot Gestures on the Bed. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI EA '21)*. Association for Computing Machinery, New York, NY, USA, Article 436, 7 pages. https://doi.org/10.1145/3411763.3451782

[11] Shaun K. Kane, Jacob O. Wobbrock, and Richard E. Ladner. 2011. Usable Gestures for Blind People: Understanding Preference and Performance. In *International Conference on Human Factors in Computing Systems.*

[12] Maria Karam and M. C. Schraefel. 2005. A Taxonomy of Gestures in Human Computer Interactions. *zeitschrift fur naturforschung a* (2005). https://doi.org/10.1515/zna-1973-0512.

[13] R. Kjeldsen. 2001. Head gestures for computer control. In *Proceedings IEEE ICCV Workshop on Recognition, Analysis, and Tracking of Faces and Gestures in Real-Time Systems*. 61–67. https://doi.org/10.1109/RATFG.2001.938911

[14] Gordon Kurtenbach and Eric A Hulteen. 1990. *The Art of Human-Computer Interface Design*. Gestures in Human-Computer Communication.

[15] Zhen Li, Mingming Fan, Ying Han, and Khai N. Truong. 2020. IWink: Exploring Eyelid Gestures on Mobile Devices. In *Proceedings of the 1st International Workshop on Human-Centric Multimedia Analysis* (Seattle, WA, USA) *(HuMA'20)*. Association for Computing Machinery, New York, NY, USA, 83–89. https://doi.org/10.1145/3422852.3423479

[16] Zheng Li, Ryan Robucci, Nilanjan Banerjee, and Chintan Patel. 2015. Tongue-n-Cheek: Non-Contact Tongue Gesture Recognition. In *Proceedings of the 14th International Conference on Information Processing in Sensor Networks* (Seattle, Washington) *(IPSN '15)*. Association for Computing Machinery, New York, NY, USA, 95–105. https://doi.org/10.1145/2737095.2737109

[17] Yiqin Lu, Bingjian Huang, Chun Yu, Guahong Liu, and Yuanchun Shi. 2020. Designing and Evaluating Hand-to-Hand Gestures with Dual Commodity Wrist-Worn Devices. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 4, 1 (2020), Article 20. https://doi.org/10.1145/3380984

[18] Meethu Malu, Pramod Chundury, and Leah Findlater. 2018. Exploring Accessible Smartwatch Interactions for People with Upper Body Motor Impairments. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3173574.3174062

[19] Meethu Malu and Leah Findlater. 2015. Personalized, Wearable Control of a Head-Mounted Display for Users with Upper Body Motor Impairments. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (Seoul, Republic of Korea) *(CHI '15)*. Association for Computing Machinery, New York, NY, USA, 221–230. https://doi.org/10.1145/2702123.2702188

[20] Katsutoshi Masai, Kai Kunze, Daisuke Sakamoto, Yuta Sugiura, and Maki Sugimoto. 2020. Face Commands - User-Defined Facial Gestures for Smart Glasses. In *2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. 374–386. https://doi.org/10.1109/ISMAR50242.2020.00064

[21] Fatih Nayebi, Jean-Marc Desharnais, and Alain Abran. 2012. The state of the art of mobile application usability evaluation. In *2012 25th IEEE Canadian Conference on Electrical and Computer Engineering (CCECE)*. IEEE, 1–4.

[22] Shuo Niu, Li Liu, and D. Scott Mccrickard. 2018. Tongue-able Interfaces: Prototyping and Evaluating Camera Based Tongue Gesture Input System. *Smart Health* 11 (2018), 16–28. https://doi.org/10.1016/j.smhl.2018.03.001

[23] Mohammad Obaid, Markus Häring, Felix Kistler, René Bühling, and Elisabeth André. 2012. User-Defined Body Gestures for Navigational Control of a Humanoid Robot. In *Social Robotics*, Shuzhi Sam Ge, Oussama Khatib, John-John Cabibihan, Reid Simmons, and Mary-Anne Williams (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 367–377.

[24] Jorge-Luis Pérez-Medina, Santiago Villarreal, and Jean Vanderdonckt. 2020. A Gesture Elicitation Study of Nose-Based Gestures. *Sensors* 20, 24 (2020), 7118. https://www.mdpi.com/1424-8220/20/24/7118

[25] Julie Rico and Stephen Brewster. 2010. Usable Gestures for Mobile Interfaces: Evaluating Social Acceptability. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Atlanta, Georgia, USA) *(CHI '10)*. Association for Computing Machinery, New York, NY, USA, 887–896. https://doi.org/10.1145/1753326.1753458

[26] 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* (Vancouver, BC, Canada) *(CHI '11)*. Association for Computing Machinery, New York, NY, USA, 197–206. https://doi.org/10.1145/1978942.1978971

[27] Chaklam Silpasuwanchai and Xiangshi Ren. 2015. Designing concurrent full-body gestures for intense gameplay. *International Journal of Human-Computer Studies* 80 (2015), 1–13. https://doi.org/10.1016/j.ijhcs.2015.02.010

[28] Daniel Tarsy and David K. Simon. 2006. Dystonia. *New England Journal of Medicine* 355, 8 (2006), 818–829. https://doi.org/10.1056/NEJMra055549

[29] Feng Tian, Fei Lyu, Xiaolong Luke Zhang, Xiangshi Ren, and Hongan Wang. 2017. An Empirical Study on the Interaction Capability of Arm Stretching. *International journal of human-computer interaction* 33, 7-9 (2017), 565–575. https://doi.org/10.1080/10447318.2016.1265782

[30] Shari Trewin, Cal Swart, and Donna Pettick. 2013. Physical accessibility of touchscreen smartphones. In *Proceedings of the 15th international ACM SIGACCESS conference on computers and accessibility*. 1–8.

[31] 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* (Montreal QC, Canada) *(CHI EA '18)*. Association for Computing Machinery, New York, NY, USA, 1–6. https://doi.org/10.1145/3170427.3188619

[32] Jean Vanderdonckt, Nathan Magrofuoco, Suzanne Kieffer, Jorge Pérez, Ysabelle Rase, Paolo Roselli, and Santiago Villarreal. 2019. Head and Shoulders Gestures: Exploring User-Defined Gestures with Upper Body *(Design, User Experience, and Usability. User Experience in Advanced Technological Environments)*. Springer International Publishing, 192–213. https://doi.org/10.1007/978-3-030-23541-3_15

[33] 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* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3290605.3300445

[34] Radu-Daniel Vatavu and Jacob O. Wobbrock. 2015. Formalizing Agreement Analysis for Elicitation Studies: New Measures, Significance Test, and Toolkit. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (Seoul, Republic of Korea) *(CHI '15)*. Association for Computing Machinery, New York, NY, USA, 1325–1334. https://doi.org/10.1145/2702123.2702223

[35] Radu-Daniel Vatavu and Jacob O. Wobbrock. 2022. Clarifying Agreement Calculations and Analysis for End-User Elicitation Studies. *ACM Trans. Comput.-Hum. Interact.* 29, 1 (2022), 1073–0516. https://doi.org/10.1145/3476101

[36] Santiago Villarreal-Narvaez, Jean Vanderdonckt, Radu-Daniel Vatavu, and Jacob O. Wobbrock. 2020. A Systematic Review of Gesture Elicitation Studies: What Can We Learn from 216 Studies?. In *Proceedings of the 2020 ACM Designing Interactive Systems Conference* (Eindhoven, Netherlands) *(DIS '20)*. Association for Computing Machinery, New York, NY, USA, 855–872. https://doi.org/10.1145/3357236.3395511

[37] Christopher D. Wickens, William S. Helton, Justin G. Hollands, and Simon Banbury. 2021. *Engineering Psychology and Human Performance* (5th edition ed.). New York. https://doi.org/10.4324/9781003177616

[38] Jacob O. Wobbrock, Htet Htet Aung, Brandon Rothrock, and Brad A. Myers. 2005. Maximizing the Guessability of Symbolic Input. In *CHI '05 Extended Abstracts on Human Factors in Computing Systems* (Portland, OR, USA) *(CHI EA '05)*. Association for Computing Machinery, New York, NY, USA, 1869–1872. https://doi.org/10.1145/1056808.1057043

[39] 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* (Boston, MA, USA) *(CHI '09)*. Association for Computing Machinery, New York, NY, USA, 1083–1092. https://doi.org/10.1145/1518701.1518866

[40] Huiyue Wu and Jianmin Wang. 2012. User-Defined Body Gestures for TV-based Applications. In *Fourth International Conference on Digital Home*. IEEE, 415–420. https://doi.org/10.1109/ICDH.2012.23.

[41] Wei Xiaomei, Y. U. Yong, and Jiang Li. 2017. Quantifying oromoter muscle tone using MyotonPro device:A pilot study. *Chinese Journal of Rehabilitation Medicine* 32, 7 (2017), 768–772.

[42] 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 *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 1, 15 pages. https://doi.org/10.1145/3491102.3517552