International Journal of Human–Computer Interaction

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/hihc20

User-Defined Foot Gestures for Eyes-Free Interaction in Smart Shower Rooms

Zhanming Chen, Huawei Tu & Huiyue Wu

To cite this article: Zhanming Chen, Huawei Tu & Huiyue Wu (2023) User-Defined Foot Gestures for Eyes-Free Interaction in Smart Shower Rooms, International Journal of Human–Computer Interaction, 39:20, 4139-4161, DOI: 10.1080/10447318.2022.2109260

To link to this article: https://doi.org/10.1080/10447318.2022.2109260

INTERNATIONAL JOURNAL OF HUMAN-COMPUTER INTERACTION
2023, VOL. 39, NO. 20, 4139–4161
https://doi.org/10.1080/10447318.2022.2109260

User-Defined Foot Gestures for Eyes-Free Interaction in Smart Shower Rooms

Zhanming Chen[a], Huawei Tu[b], and Huiyue Wu[a,c]

[a]The School of Communication and Design, Sun Yat-Sen University, Guangzhou, China; [b]Department of Computer Science and Information Technology, La Trobe University, Melbourne, Australia; [c]Guangdong Key Laboratory for Big Data Analysis and Simulation of Public Opinion, Guangzhou, China

ABSTRACT
With the rapid development of natural human-computer interaction technologies, gesture-based interfaces have become popular. Although gesture interaction has received extensive attention from both academia and industry, most existing studies focus on hand gesture input, leaving foot-gesture-based interfaces underexplored, especially in scenarios where the user's hands are occupied for other interaction tasks such as washing the hair in smart shower rooms. In such scenarios, users often have to perform interactive tasks (e.g., controlling water volume) with their eyes closed when water and shampoo liquid flow along with their head to eyes area. One possible way to address this problem is to use eyes-free (rather than eyes-engaged), foot-gesture-based interactive techniques that allow users to interact with the smart shower system without visual involvement. Through our online survey, 71.60% of the participants (58/81) have the requirements of using foot-gesture-based eyes-free interactions during showers. To this end, we conducted a three-phase study to explore foot-gesture-based interaction to achieve eyes-free interaction in smart shower rooms. We first derived a set of user-defined foot gestures for eyes-free interaction in smart shower rooms. Then, we proposed a taxonomy for foot gesture interaction. Our findings indicated that end-users preferred single-foot (76.1%), atomic (73.3%), deictic (65.0%), and dynamic (76.1%) foot gestures, which markedly differs from the results reported by previous studies on user-defined hand gestures. In addition, most of the user-defined dynamic foot gestures involve atomic movements perpendicular to the ground (40.1%) or parallel to the ground (27.7%). We finally distilled a set of concrete guidelines for foot gesture interfaces based on observing end-users' mental model and behaviors when interacting with foot gestures. Our research can inform the design and development of foot-gesture-based interaction techniques for applications such as smart homes, intelligent vehicles, VR games, and accessibility design.

KEYWORDS
foot gesture; shower; eyes-free interaction; smart home; elicitation study

1. Introduction

In recent years, gesture-based interaction techniques have attracted significant interest from research communities and commercial sectors worldwide. However, most existing gesture-based studies focus on hand gesture input methods, leaving foot-gesture-based interfaces underexplored. Foot gestures can be very useful, especially in scenarios where the users need to perform concurrent tasks while their hands are being occupied for other interaction tasks. A concurrent task is one that combines two or more tasks in such a manner that each component task is performed independently and parallelly (Wu et al., 2021). Another advantage of foot-gesture-based interaction is that it does not require the engagement of the user's visual system into the interaction. A typical example is the shower scenario, in which users' hands are being used to wash their hair but they, at the same time, need to perform other interactive tasks (e.g., controlling water volume and/or temperature) with their eyes closed when water and shampoo liquid flow along with their head to eye area. One possible solution to address this problem is to use eyes-free (Findlater et al., 2011; Wu et al., 2021; Yan et al., 2018) (rather than eyes-engaged) foot-gesture-based interaction techniques that allow the users to interact with the smart shower system without any visual involvement.

The rapid development of novel technologies in ubiquitous computing and natural human-computer interaction (HCI), such as Ambient Assisted Living (AAL) (Maskeliūnas et al., 2019) and Wi-fi sensing (Zhang et al., 2019), has brought great potential into design more natural and intelligent foot-gesture-based interaction systems (Kim et al., 2019) and subsequently enhances user experience. However, previous studies on foot gesture systems primarily focused on providing basic interactive tasks such as pervasive entertaining and media access (Funk et al., 2015; Hoshino et al., 2015), few studies have focused on how to design a user-friendly foot-gesture-based system for eyes-free interaction for complex concurrent tasks such as the smart shower

CONTACT Huiyue Wu wuhuiyue@mail.sysu.edu.cn The School of Communication and Design, Sun Yat-sen University, Guangzhou, China
Supplemental data for this article is available online at https://doi.org/10.1080/10447318.2022.2109260

Figure 1. An illustrative scenario of eyes-free interaction in a smart shower room. The user is shampooing with eyes closed, and he performs foot gestures to achieve eyes-free interaction with smart shower devices naturally and conveniently.

system, VR games, and intelligent vehicles required by modern smart systems.

In this study, we would like to take advantage of users' eyes-free input ability to improve interaction in smart homes, by allowing users to interacting with the smart system using foot gestures. Figure 1 shows a promising use case of such an interaction. As shown, the user is shampooing with his eyes closed and hands occupied. His potential requirements include turning on/off the shower spray, increasing/decreasing the water volume, changing the direction of the shower spray, etc. In such an eyes-free condition, he can perform simple foot gestures to control the shower without interrupting his cleansing process. This makes the whole interaction process more natural, fluid, and immersive and subsequently significantly affects user experience when taking a shower.

To explore its feasibility and usability, we conducted a three-stage study on exploring user-defined foot gestures for eyes-free interaction in smart shower rooms. In Study 1, we certified the requirements with an online survey, generating a set of core tasks during showers. In the survey, 71.0% of the participants had interaction requirements in eyes-free conditions. In Study 2, we conducted a gesture elicitation study to generate a set of user-defined foot gesture proposals. Next, the user-defined foot gestures were evaluated in a benchmark test in Study 3. The results show that users preferred single-foot (76.1%), atomic (73.3%), deictic (65.0%), and dynamic (76.1%) foot gestures. In general, the main contributions of our work include:

- We established a taxonomy of foot gestures based on analyzing end-users' mental model and interaction behaviors, and described the design space of foot gestures for interacting with smart shower rooms.
- We generated a set of user-defined foot gestures, which were developed in light of the agreement participants exhibited in using foot gestures for each target task.
- We distilled several concrete design guidelines for foot gesture interaction in smart shower rooms.
- Our method and findings can inform the design of other foot-gesture-based applications such as smart homes, intelligent vehicles, VR games, and accessibility design.

2. Related work

In this section, we reviewed research efforts related to eyes-free interaction during showers. In the first part, we presented previous research on interaction techniques in smart shower rooms. After introducing applications of foot interaction in previous research, we presented existing methods of user-defined gesture design.

2.1. Interaction techniques in smart shower rooms

Previous research is mainly devoted to designing shower room entertainment systems to complement the whole-home entertainment experience. Jorro Beat, a shower spray providing tactile feedback based on music (Hoshino et al., 2015), is a typical example. A projection-based system can recognize uses' finger touching input by a thermal camera from the backside of the shower curtain (Funk et al., 2015). The system was designed to use applications such as music players and newsreaders during showers. Previous studies focused on the design of auxiliary shower environments for the elderly (Schlömer et al., 2017) and people with disabilities (Ferati et al., 2018).

Similar research focuses can be observed when it comes to the bathtub. Aquatop allowed users to manipulate objects or play games with hand gestures through the water surface and an interactive project (Koike et al., 2013; Takahashi et al., 2012). RapTapBath recognized hand-tapped tones and patterns on a bathtub's edge for interactions with the menu (Sumida et al., 2017). Some studies introduced rubbing sound perception, allowing users to perform tasks of entertain applications during bathing (Hirai et al., 2012; Kawakatsu & Hirai, 2018).

Studies mentioned above primarily focused on entertaining systems. To the best of our knowledge, solutions of eyes-free interaction issues during showers were rarely explored. Our research applied foot gestures to improve user experience for eyes-free interaction during showers.

2.2. Applications of foot-gesture-based interaction

Due to the differences of motor control abilities between hands and feet, it's essential to understand the design space of foot interaction. Each foot can be divided into three parts: the hindfoot, the midfoot, and the forefoot. Humans can only perform lower limb motions with three joints: ankle, knee, and hip. Specifically, there are 12 distinctive foot movements (Roaas & Andersson, 1982; Velloso et al., 2015), most of which are in pairs. Movements of ankles include dorsiflexion and plantar flexion, inversion, and eversion. Flexion and extension are the two main locomotion of the knees. Motions of hips include flexion and extension, abduction and adduction, medial rotation, and lateral rotation. Movements of each foot gesture are all constrained to a specific angle range. Some complex foot gestures are combinations of single ones.

Foot interaction has been proved to be an efficient approach to interacting with mobile phones when both hands are occupied (Fan et al., 2017). In addition, researchers have explored foot interaction in a variety of other applications, such as tabletop devices, VR, and AR. For tabletop devices, foot interaction was applied for spatial manipulation tasks through trackball (Pakkanen & Raisamo, 2004), fan-shaped menu interaction through heel rotation (Zhong et al., 2011), and pre-designed Tap-Kick-Click vocabulary in standing pose (Saunders & Vogel, 2016). For 3D object manipulation, FEETICHE was developed, which was a set of multimodal interactions combining hand and foot input (Lopes et al., 2019). Similar interaction methods were designed for touchless medical image systems (Paulo et al., 2019). Some studies introduced foot input to the head-mounting displays (HMDs) environment. One study applied foot tapings as an input modality for interacting with AR menus projected on the ground (Müller et al., 2019). Locomotion of the foot was tracked and mapped in the virtual space to improve users' embodied cognition (Pan & Steed, 2019). Another study applied foot position and pressure to change users' perspectives in the VR environment (Willich et al., 2020).

Foot gestures introduced in previous studies were mainly pre-defined by designers without consideration of using them in smart homes. User-defined gestures were more popular than pre-defined ones (Morris et al., 2010) because they were more in line with the user's mental model. In this work, we conducted an elicitation study to explore foot-gesture-based interaction under the eyes-free interaction context during showers.

2.3. User-defined gesture design

Compared with the pre-designed gesture interaction method, elicitation study is an effective way to improve the performance of gesture interaction (Nacenta et al., 2013). Researchers have proposed various elicitation methods to derive gestures with better usability.

Guessability study was commonly adopted by previous researchers (Villarreal-Narvaez et al., 2020). It includes three phases: (1) participants are firstly presented referents (i.e., the effect produced by the gesture proposals), then (2) they are asked to propose one gesture proposal for each referent, and (3) a subjective usability test is conducted to evaluate the performance of those gesture proposals (Wobbrock et al., 2005, 2009).

The choice-based elicitation method modifies the previous process to reduce negative impacts of gesture disagreement problem (Wu et al., 2019). It contains two revised phases: (1) conducting a traditional guessability study to collect gesture proposals from participants, and (2) inviting a new group of participants to re-evaluate and refine the gesture proposals and then finally generate a gesture vocabulary.

The intuitive and ergonomic gesture interface design method (Nielsen et al., 2004) introduced a benchmark test for evaluation during actual usage. It improves the deficiencies of the subjective evaluation process during elicitation study. The benchmark test includes different metrics (Ali et al., 2021; Wu et al., 2016), including matching degree, memorability, comfort, and learnability.

Gesture elicitation methods have been applied by HCI researchers to derive user-defined hand gestures for surface computing (Wobbrock et al., 2009), mobile devices (Ruiz et al., 2011; Tu et al., 2020), wearable devices (Gheran et al., 2018), AR (Piumsomboon et al., 2013), and smart homes (Dong et al., 2015; Vogiatzidakis & Koutsabasis, 2019). In addition to the above-mentioned methods applied for hand-gesture-based interactions, researchers have also used the guessability method for the design of foot-gesture-based interaction applications. For example, Felberbaum and Lanir (2016, 2018) derived a set of user-defined foot gestures for typical GUI actions and the avatar controls. They also proposed a new metric – *specification score* to measure the degree to which a foot gesture is specific, preferable, and intuitive to a referent. In addition, user-defined foot gestures were also explored in other applications, such as performing navigation tasks for mobile devices (Alexander et al., 2012; Fukahori et al., 2015), locomotion tasks for VR environments (Kim & Xiong, 2021), manipulation tasks for intense video games (Silpasuwanchai & Ren, 2015), and large immersive AR maps (Austin et al., 2020).

Existing elicitation studies mostly focused on exploring upper-body gestures while ignoring the use of foot gestures (Villarreal-Narvaez et al., 2020). In contrast, we conducted three studies to explore the user's mental model of foot gesture interaction during showers, the design space of user-defined foot gestures, and the usability of the user-defined foot gesture set. Our work contributes to providing useful guidelines for the design of foot-gesture-based interfaces.

3. Study 1: Requirement analysis

It is essential to gather core requirements from end-users before developing a practical system. Therefore, we conducted an online survey to collect information from end-users about their interaction experience during showers.

Table 1. The originally derived task set for eyes-free conditions in smart shower rooms (sorted in descending order of frequency)

No.	Tasks	Frequency
1	**Turn on the shower (T_1)**	**39/58**
2	**Turn off the shower (T_2)**	**39/58**
3	**Increase water pressure in shower (T_3)**	**33/58**
4	**Decrease water pressure in shower (T_4)**	**33/58**
5	**Adjust the angle of the shower head (T_5)**	**22/58**
6	Raise water temperature in shower	21/58
7	Lower water temperature in shower	21/58
8	Switch the water outlet mode of the shower head	9/58
9	Turn on the bathroom heater	8/58
10	Turn off the bathroom heater	8/58
11	Raise the temperature of the bathroom heater	7/58
12	Lower the temperature of the bathroom heater	7/58
13	Turn on the exhaust fan	6/58
14	Turn off the exhaust fan	6/58
15	Increase the power of the exhaust fan	4/58
16	Decrease the power of the exhaust fan	4/58

Note: The **tasks in bold** are involved in the core task set.

3.1. Online survey

In the online survey, we collected participants' shower habits and interaction requirements during showers, for example, "What's your daily bathing routine? Shower only, bath only, or in both ways (Q1)." We further validated the generality of eyes-free interaction conditions and collected related requirements during showers. As an example, Question 5 is "Which activity(s) was (were) you used to do when your eyes were passively closed during shower?" Thirty tasks were involved in our questionnaire. These tasks were collected from previous studies about interactions during showers (Ferati et al., 2018; Funk et al., 2015; Kawakatsu & Hirai, 2018; Sumida et al., 2017). We also referred to interaction tasks provided by popular smart home platforms and commercial products such as the smart shower devices "Moen TS3302TB"[1] and "KOHLER K-527-1SN"[2] and the smart home control systems "Echo Dot"[3] and the "HomePod mini".[4] Participants were encouraged to provide other feedback if they wanted.

There were 113 valid samples (58 males and 55 females) left after screening out two invalid ones who were accustomed to soak in a tub rather than using a shower. The age of the participants ranged from 13 to 58 years old ($M = 29.42$, $SD = 12.587$). They were engaged in multiple occupations, including academic personnel, students, corporate employees, etc.

As a result, most participants (81/113, 71.68%) have experienced eyes-free conditions during showers at least once, most of which (58/81, 71.60%) had interaction requirements in eyes-free states. In summary, it is necessary to pay more attention to eye-free interaction during showers. Feedback from end-users laid a foundation for the latter two studies.

Table 1 shows all interaction requirements in eyes-free conditions during showers. As shown in Table 1, the task set included 16 tasks, which were derived from the analysis results of Question 5 in the questionnaire (Appendix 1).

3.2. Refinement of the task set

According to the recommendations by (Wobbrock et al., 2009; Wu et al., 2016), it is necessary to refine the task set derived from the survey. Thus, we conducted a brainstorming session with seven experts, who have 5–10 years of related experience in HCI. Our aim was to validate and refine the task set and streamline the number of tasks to a more reasonable level for foot-gesture interaction. The session lasted for 1.5 hours in a usability lab in our university.

All the experts believed that the 16 tasks summarized in Study 1 were too many to remember for ordinary users in conditions involving multiple concurrent tasks. After discussion, the first five tasks in Table 1 were retained in the core task set. For Tasks 6, 7, 9–12, the experts believed that these tasks were generally performed before the start or after the end of showers. Application scenarios for Tasks 8 and 13–16 were also rarely seen in general household appliances.

During the brainstorming, the experts proposed a new delimiter task —turn on the gesture control mode, to reduce the "Midas Touch" problem in vision-based gesture interfaces, which means the users' actions could potentially always be active (Wu & Wang, 2016). The "turn on the gesture control mode" task was used to activate the foot-gesture recognition system before the other five tasks and it would stop automatically when the user turns off the water outlet switch.

According to the results of the brainstorming section, the first five tasks in Table 1 and the newly added "turn on the gesture control mode" task proposed by experts were retained in the core task set, while the other 11 tasks were excluded. This setting is also consistent with the "7 ± 2" rule (Miller, 1956) about the short-term memory capacity of the human brain.

4. Study 2: Foot gesture elicitation

In Study 1, we collected requirements from end-users and generated a task set with six core tasks for eyes-free interaction in the smart shower room. In this section, we conducted an elicitation study, in which end-users were involved to design foot gesture proposals consistent with their mental models.

4.1. Participants

Thirty paid participants (15 males and 15 females) were recruited from a university. Their age ranged from 19 to 24 ($M = 22.90$, $SD = 1.012$). All of them were students with backgrounds in interaction design, engineering, and journalism. Eight of them had little experience in foot gesture interaction (e.g., using a dance machine). We chose these participants for two main reasons: our limit access to subject pools beyond this age group within a university and our expectation for the participants who can accept and learn new HCI technologies such as foot-gesture-based interaction with little difficulty.

4.2. Apparatus

We set up an environment that mimicked a smart shower room (Figure 2).

We adopted the "Wizard of Oz" approach (Cohen et al., 2008; Höysniemi et al., 2005) to simulate the task completion effects in response to participants' gestures. A Bluetooth stereo speaker was used to provide auditory feedback of foot gestures input with pre-recorded sound effects (T_1–T_4) and voice broadcast (T_5 and T_6) (Table 2). The corresponding auditory feedback was played by the experimenter as soon as the participants finished their foot gesture proposal.

The shower products, wet floor, and eye-patch were used to mimic the eyes-free conditions during showers. We also put an RGB camera in front of participants' lower limbs and a voice recorder clipped to their collar, with the aim to gather information of participants' actions and verbalization during the experiment.

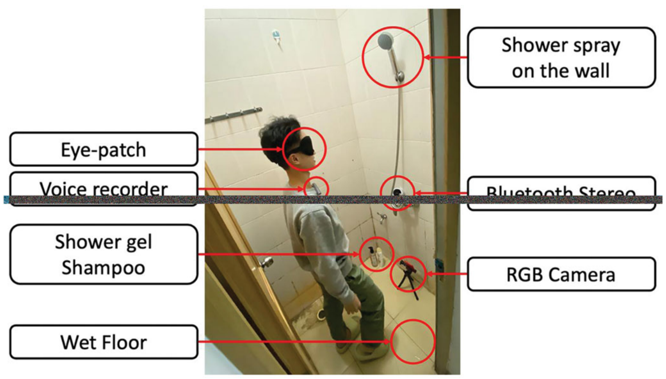


Figure 2. Experimental environment.

4.3. Procedure

Participants were briefed about the concept of foot gesture interaction and the process of elicitation study, and then finished the consent process. The core tasks were presented one at a time by slides on a laptop randomly. Each task was described on the slide by text, static images, or GIF images.

Wu et al. (Wu et al., 2022) and Cafaro et al. (Cafaro et al., 2018) suggested that, priming participants with a frame, or a scenario, could significantly reduce potential negative effects of legacy bias (Morris et al., 2014), i.e., reduce the number of legacy-inspired gestures and result in superior gesture vocabulary in gesture elicitation studies. Therefore, we used the *priming* technique to optimize our target set of user-defined foot gestures. As shown in Figure 3, participants were asked to follow the instructions of a video clip of fitness exercise as a warm-up *priming* technique in our study. The video was chosen from a fitness app called Keep (https://www.gotokeep.com/). It includes three warm-up exercises: (1) standing toe calf raises, (2) alternating quad stretch, and (3) standing hip rotation (left and right) (Check the supplement material for details). We selected the Keep app for two main reasons: (1) we expected to elicit a majority of 3D foot movements (foot gestures) by priming the participants with a scenario, in which the participants used their foot movements just as we wanted them to use in the final smart shower scenario; and (2) the fitness exercises tasks were disconnected from the target application

Table 2. The design of the pre-recorded sound effects and voice broadcast.

Task name	Sound effects/Voice instructions
T1: Turn on the shower	Play the sound of water coming out of the shower head.
T2: Turn off the shower	Stop the sound of water coming out of the shower head.
T3: Increase water pressure in shower	Increase the volume of the shower sound.
T4: Decrease water pressure in shower	Decrease the volume of the shower sound.
T5: Adjust the angle of the shower head	Voice instruction: "Please adjust water flow to the right/left of your body."
T6: Turn on the gesture control mode	Voice instruction: "Please turn on the gesture control mode."

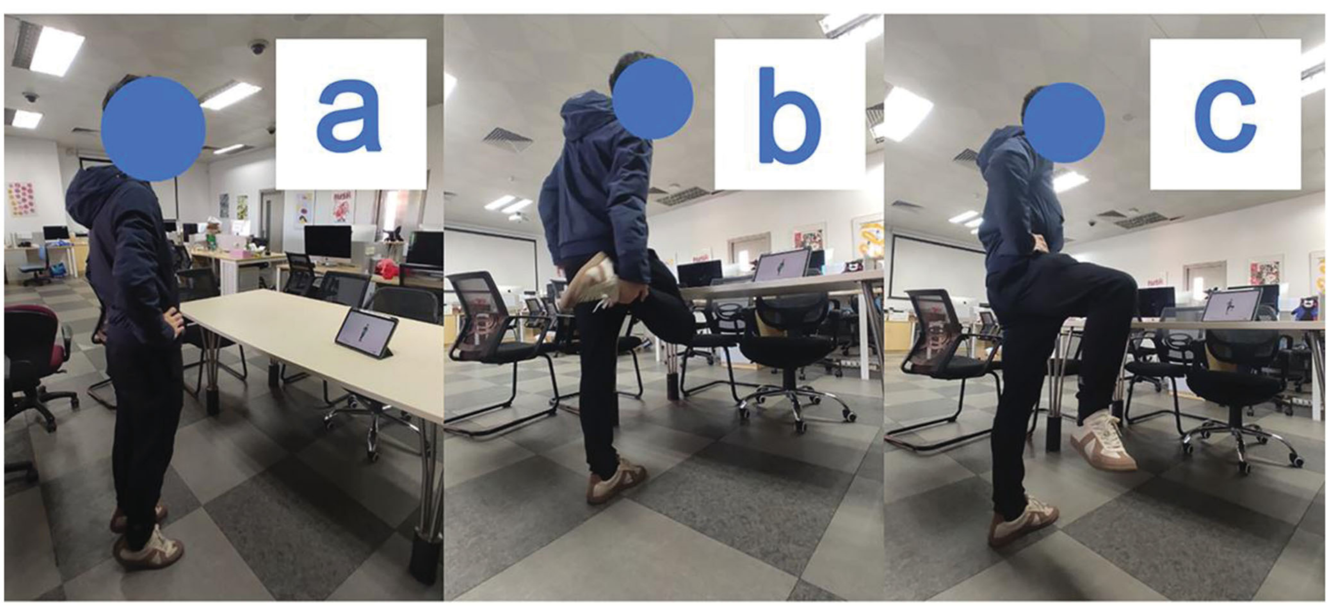


Figure 3. Experimental setup for the *priming* phase of the elicitations study. The *priming* exercises include: (a) standing toe calf raises; (b) alternating quad stretch, and (c) standing hip rotation (left and right).

Table 3. Taxonomy of foot gestures.

Dimension		Category	Description	Frequency
Form	Static	Static posture	Users' foot is held in one location.	43
	Dynamic	Movement along the x-axis	Gesture consists of foot movement(s) along the x-axis.	16
		Movement along the y-axis	Gesture consists of foot movement(s) along the y-axis.	11
		Movement along the z-axis	Gesture consists of foot movement(s) along the z-axis.	55
		Movement in the x-y plane	Gesture consists of foot movement(s) in the x-y plane.	38
		Movement in the x-z plane	Gesture consists of foot movement in the x-z plane.	3
		Movement in the y-z plane	Gesture consists of foot movement in the y-z plane.	7
		Movement in 3D space	Gesture consists of foot movement in 3D space.	7
Num. of feet		Single foot	Gesture is performed by a single foot.	137
		Both feet	Gesture is performed by both feet.	43
Complexity		Atomic	Gesture consists of a single gesture.	132
		Compound	Gesture can be decomposed into simple gestures.	48
Flow		Continuous	Response occurs while the user performs a gesture.	84
		Discrete	Response occurs after a gesture is finished.	96
Nature		Symbolic	Foot gesture visually depicts a symbol.	1
		Deictic	Foot gesture indicates a position or direction.	117
		Abstract	Mapping between a foot gesture and a referent is arbitrary.	62
Symmetry		Symmetric	Both feet perform the same movement (yet not necessarily at the same time).	31
		Asymmetric	Two feet perform different movements.	149

context (taking a shower) but relevant in terms of technology (foot gesture recognition). Therefore, we could guarantee the fairness of the experimental conditions and also try to avoid introducing its own biases by the *priming* technique into the elicitation study.

After finishing the *priming* exercises, participants went into the simulated smart shower room to design foot gestures freely. For each task, they would hear a pre-recorded audio of the task description. Then, they were asked to design a foot gesture for each task after listening to the audio of the task description. The classic concurrent think-aloud protocol was applied during the designing process. As soon as the participants designed a gesture, they would hear the corresponding pre-recorded auditory feedbacks (e.g., the sound of the water flow).

After finishing the design process, participants were asked to fill in a questionnaire on their demographic data. A brief post hoc interview for extra information was also conducted. The elicitation study section for each participant lasted 35–45 minutes.

After the foot-gesture elicitation process, we invited five experts with more than seven years of experience in gesture interaction to group all gesture proposals in a brainstorming session. Proposals with the same pattern were merged into a single group. Then, the five experts reviewed the videos and audios recorded during the elicitation sessions, and discussed whether foot gestures with similar patterns could be merged into a single group. The session lasted for two hours in a usability lab.

4.4. Results of study 2

In this section, we presented the results and grouped our collected foot gesture proposals, and then calculated the agreement rates among gestures.

4.4.1. Gesture taxonomy

We collected 180 (30 participants × 6 tasks) user-defined foot gesture proposals in total. As shown in Table 2, we established a taxonomy for foot gestures based on these 180 proposals, by considering existing taxonomies in previous works (Karam & schraefel, 2005; Ruiz et al., 2011; Velloso et al., 2015; Wobbrock et al., 2009).

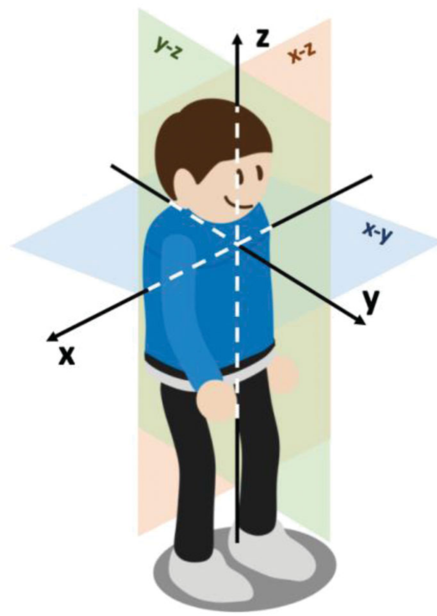


Figure 4. The world coordinate system used in this study. The coordinate system is consistent for both feet.

Our taxonomy for user-defined foot gestures includes five dimensions: *Form*, *Number of the foot (feet)*, *Complexity*, *Flow*, and *Nature*. Table 3 shows multiple categories within each dimension.

The *Form* dimension describes kinematic features of foot gestures. A gesture can be static or dynamic. Dynamic gestures occur along a single axis, in a plane, or in a 3D space. Figure 4 illustrates the world coordinate system we used in our study.

In the *Number of feet* dimension, we distinguish foot gestures performed by one foot and both feet.

In the *Complexity* dimension, we define atomic gesture as an independent foot gesture that cannot be further segmented and compound gesture as a combination of multiple

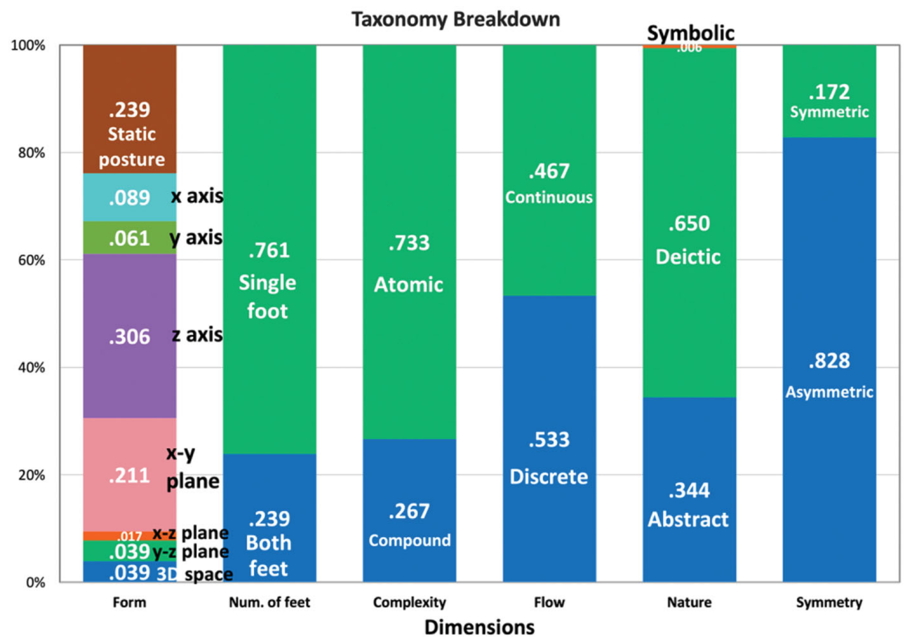


Figure 5. Distribution of foot gestures in each taxonomy category.

atomic gestures. For example, "stamping once" is an atomic gesture and "stamping twice" is a compound gesture.

A foot gesture's *Flow* is discrete if the gesture is recognized as an event and continuous if ongoing recognition is required.

In the *Nature* dimension, symbolic gestures are visual depictions, such as drawing a triangle by the toe. Deictic gestures usually indicate a position or direction, such as "sliding forwards or backwards." When the mapping between foot gestures and tasks is arbitrary, it comes to abstract gestures.

In the *Symmetry* dimension, symmetric gestures are those that both feet perform the same movement (e.g., pivot outwards on both heels). In contrast, asymmetric gestures refer to those that two feet perform different movements.

Our taxonomy of foot gestures covered all the distinguished user-defined foot gesture proposals. Figure 5 shows the percentage of foot gestures proposals in each category. Most of the foot gesture proposals fell into the category of single foot (76.1%), atomic (73.3%), and deictic (65.0%).

4.4.2. Grouping and merging

Based on the taxonomy, five experts grouped and merged similar foot gestures in a brainstorming session. The following example illustrates how the experts merge a group of foot gestures:

For the task "Turn on the shower (T_1)," one gesture was "the right foot stamps twice" and the other was "stamp the forefoot twice while keeping the heel on the ground." For these two gestures, participants explained that all they care about is whether the gesture movement involves stamping twice on the ground. It does not matter whether they touch the ground with the forefoot or the whole foot. Given the fact that the mental models of the participants about these two proposals are similar, the five experts grouped them into a single gesture. Table 4 shows a full list of grouped gesture proposals with frequency.

As shown in Figure 6, we finally generated a gesture set containing six foot gestures with the highest frequency by adopting the "Winner-take-all" strategy (Wobbrock et al., 2009).

For "Decrease water pressure in shower (T_4)," both "Slide backwards with one foot" and "Pivot on one heel" had the same frequency (proposed by 7 participants). Since "Slide backwards with one foot" has already been assigned to T_3, the experts decided to assign "Pivot on one heel" to T_4 to avoid conflict.

For "Turn on the gesture control mode (T_6)," the top frequency gesture was "Stamp twice with one foot," the same as the top gesture for T_1 – Turn on the shower and T_2 – Turn off the shower. Considering the popularity of the gesture among the participants and the potential conflict between different tasks, the experts finally chose the second top gesture, "Stand on tiptoe," for $T_{6.}$

4.4.3. Agreement rate

We calculated the degree of consensus among participants for the six core tasks based on the Agreement Rate (*AR*) formula (Vatavu & Wobbrock, 2015). *AR* is defined in Equation (1). The higher the value of *AR* was, the more likely participants proposed the same foot gesture for a target task.

$$AR(r) = \frac{|P|}{|P|-1}\sum_{P_i \subseteq P}\left(\frac{|P_i|}{|P|}\right)^2 - \frac{1}{|P|-1} \qquad (1)$$

where P is the set of all proposed gestures for Task r, $|P|$ the size of the set, and P_i subsets of identical gestures from P.

Figure 7 shows the *AR* of the six core tasks in descending order. According to (Vatavu & Wobbrock, 2015), the level of agreement is divided into low ($AR(r)$ $\leq$0.100), medium

Table 4. A full list of grouped gesture proposals with frequency.

Task	Gesture description	Gesture diagram	Freq.
Turn on the shower	**Stamp twice with one foot**	**(1)**	**11**
	Slide forwards with one foot	(2)	5
	Stamp once with one foot	(3)	4
	Step left/right with left/right foot	(4)	4
	Pivot outwards on both heels	(5)	2
	Stamp once with both feet in turn	(6)	1
	Stand on tiptoe	(7)	1
	Lift one heel	(8)	1
	Crouch	(9)	1
Turn off the shower	**Stamp twice with one foot**	**(1)**	**8**
	Stamp once with one foot	(3)	6
	Step right/left with left/right foot	(10)	5
	Crouch	(9)	4
	Slide forwards with one foot	(2)	2
	Slide backwards with one foot	(11)	1
	Stamp once with both feet in turn	(6)	1
	Stand on tiptoe	(7)	1
	Pivot inwards on both heels	(12)	1
	Draw a triangle with one foot	(13)	1
Increase water pressure in shower	**Slide forwards with one foot**	**(2)**	**7**
	Slide to the right with one foot	(14)	6
	Pivot on one heel clockwise	(15)	6
	Pivot on one toe counterclockwise	(16)	2
	Stand on tiptoe	(7)	2
	Lift one heel	(8)	2
	Raise one foot off the floor	(17)	2
	Slide backwards with one foot	(11)	1
	Tap on the floor with one toe in front of the right body	(18)	1
	Pivot outwards on both heels	(5)	1
Decrease water pressure in shower	**Slide backwards with one foot**	**(11)**	**7**
	Pivot counterclockwise on one heel	(19)	7
	Slide one foot to the left	(20)	5
	Pivot on one toe clockwise	(21)	3
	Crouch	(9)	3
	Lift one heel	(8)	2
	Slide forwards with one foot	(2)	1
	Tap on the floor with one toe in front of the left body	(22)	1
	Pivot inwards on both heels	(12)	1
Adjust the angle of the shower head	**Pivot on one heel**	**(23)**	**21**
	Slide to the left/right side of the body with one foot	(24)	7
	Stamp twice with one foot	(1)	1
	Stamp once with one foot	(3)	1
Turn on the gesture control mode	Stamp twice with one foot	(1)	7
	Stand on tiptoe	**(7)**	**5**
	Stamp once with one foot	(3)	4
	Pivot outwards on both heels	(5)	3
	Slide forwards with one foot	(2)	3
	Slide backwards with one foot	(11)	2
	Lift one heel	(8)	2
	Stamp once with both feet in turn	(6)	2
	Stand with both feet crossed	(25)	1
	Pivot on one heel clockwise	(15)	1

Note: Please refer to Appendix 2 to see the corresponding foot gesture diagram. The gestures in bold are involved in the final foot gesture set.

($0.100 < AR(r) \leq 0.300$), high ($0.300 < AR(r) \leq 0.500$) and very high ($AR(r) > 0.500$). Therefore, the average AR of all core tasks in our study was medium in magnitude ($M = 0.203$). The highest agreement rate appeared in T_5 (0.531).

5. Study 3: A benchmark test

We identified six core tasks for eyes-free interaction during showers and derived a set of user-defined foot gestures for these tasks in Study 1 and Study 2, respectively. In this section, we conducted a three-stage benchmark study to further validate the performance of these foot gesture proposals in practice.

5.1. Participants and apparatus

Twenty-four participants volunteered for the benchmark test with a reward of $2. All participants were students and faculty members from our university. Participants had different professional backgrounds, including interaction design, engineering, and journalism. Their age ranged from 18 to 54 ($M = 24.46$, $SD = 8.317$).

The benchmark test was conducted in a usability lab in our university. We used a laptop to show slides of tasks and foot gesture proposals to the participants. A stopwatch and a counter were also used in our experiment to record task recall time and assist participants in counting the times of their movements in Stage 3.

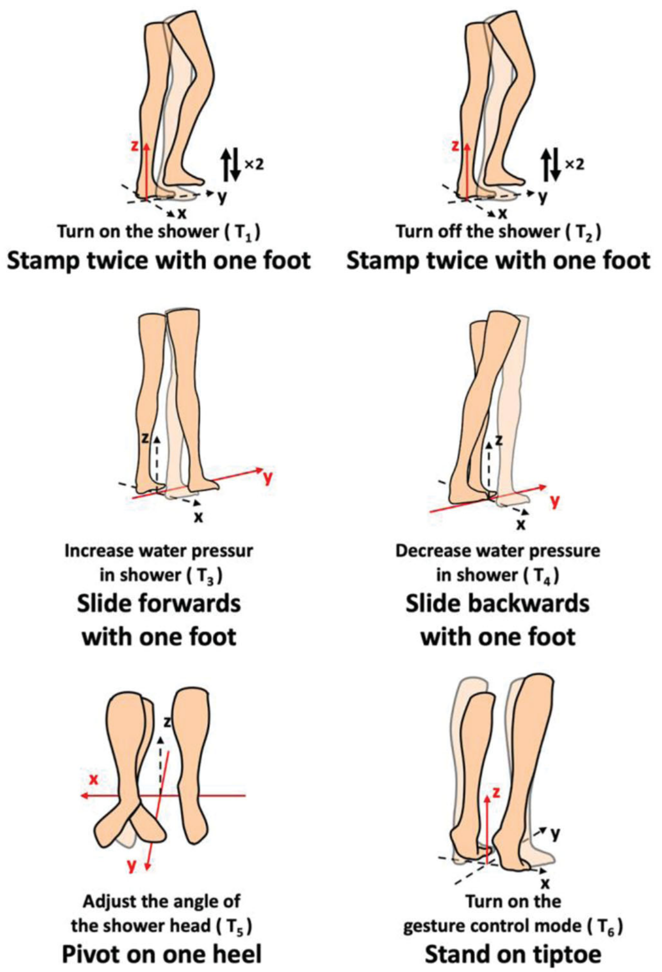


Figure 6. The user-defined gesture set for eyes-free interaction in smart shower rooms.

5.2. Procedure

Participants were briefed about the requirements of the benchmark test as well as the concept of foot gesture interaction, and then finished the consent process. After that, all participants went through a three-stage testing process (Figure 8).

The first stage tested the matching degree between the user-defined foot gestures and the corresponding target tasks. First, the names of the six tasks were shown to the participants on a piece of paper. Then, the participants would see the effects of the six top foot gestures on the slides presented with text descriptions and pre-recorded GIF images. The order of the six tasks was counterbalanced across participants. After that, participants were asked to choose the best gesture proposal for each task. An error was recorded if the participant mismatched a foot gesture for a given target task. We calculated the error rate of each task by Equation (2).

$$Error\ Rate = \frac{num.\ of\ error(s)}{num.\ of\ all\ tasks} \tag{2}$$

Before the second stage, participants were required to practice until they could remember all gesture proposals for the corresponding target tasks correctly. In the second stage, we examined the memorability of the foot gesture set, i.e., how easy it was for end-users to remember these proposals. Slides with the name of each task were presented to participants in random order. Then, participants were required to perform the corresponding foot gesture proposal as fast and accurate as possible. If the participants made a mistake in this process, they would be shown the slide of the correct gesture proposal again. The researchers recorded the errors and repeated the process until the participants finished all target tasks correctly. Three metrics were measured:

1. the number of error(s) occurred in the trials;

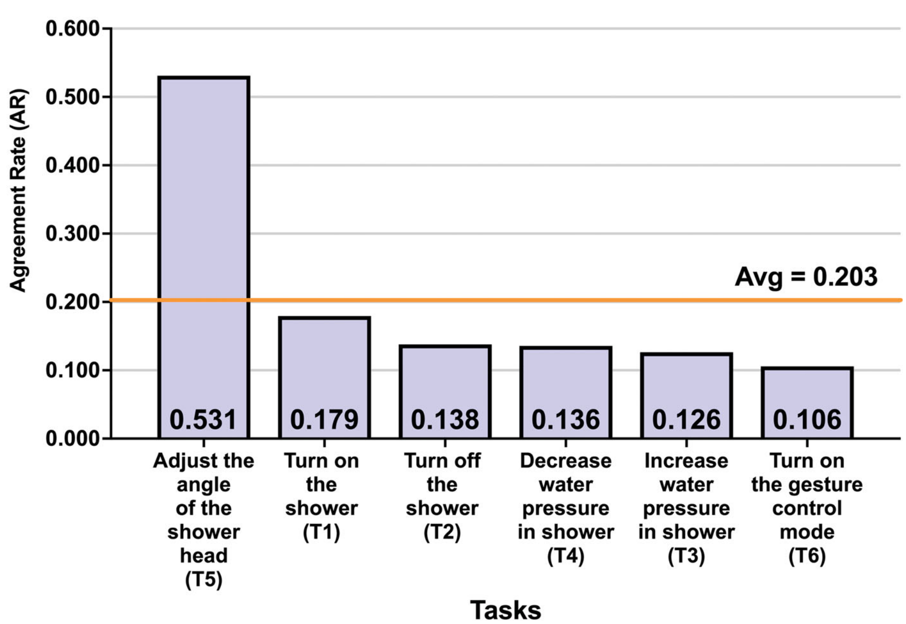


Figure 7. Agreement Rates (*AR*) of foot gesture proposals.

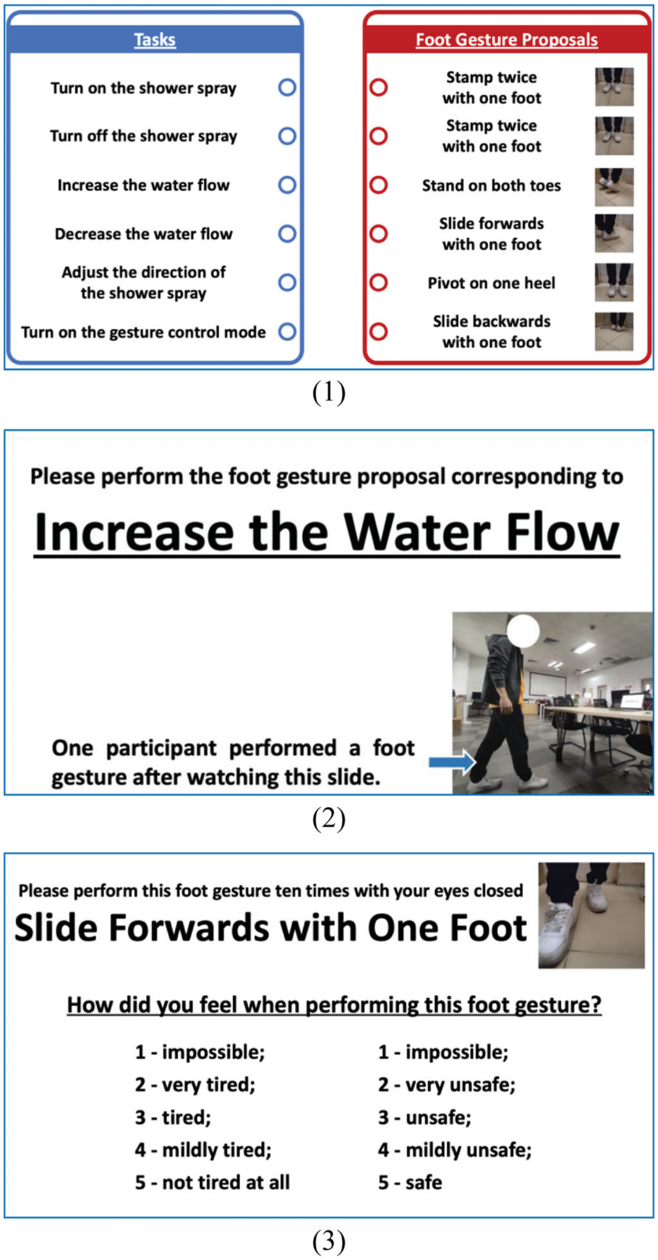


Figure 8. The three-stage benchmark test: (1) matching test, (2) memorability test, and (3) comfort and safety tests.

2. the number of rounds that the participants needed to finish the required tasks correctly;
3. task recall time, i.e., the time between the participants saw a task on the slide and correctly finished the corresponding gesture.

In the third stage, we investigated the degree of comfort and safety, i.e., how comfortable and safe it would be when performing the gesture proposals in a smart shower room. Participants were randomly shown the task names and the corresponding GIF images of foot gesture proposals on the slides. Then participants were asked to perform each gesture ten times with their eyes closed. After that, participants were required to rate the degree of comfort (1 = impossible, 2 = very tired, 3 = tired, 4 = mildly tired, 5 = not tired at all) and safety (1 = impossible, 2 = very unsafe, 3 = unsafe, 4 = mildly unsafe, 5 = safe) by using a 5-point Likert scale. The benchmark test session for each participant lasted 20-25 minutes.

5.3. Results of study 3

In Stage 1, the average error rate was .04 ($SD = 0.156$). One participant (P_6) mistakenly assigned the "Stamp twice with one foot" to "Increase and Decrease water flow" instead of "Turn on and off the shower spray." Another one (P_7) mismatched the two gestures for the "Increase and Decrease water flow" tasks.

The average number of errors in the first round was 0.04 ($SD = 0.200$) in Stage 2, and the average number of rounds required to perform all proposals correctly was 1.04 ($SD = 0.200$). Most participants (15/24) skipped the practice stage and performed all gesture proposals correctly. The average task recall time among all tasks (Figure 9) was 1.85 seconds ($SD = 0.157$). A Friedman test found no statistical significance among the tasks ($\chi^2 = 8.74$, $p = .11 > .05$).

Next, we analyzed the participants' subjective ratings on the derived gestures on Stage 3. The average scores of comfort and safety are 4.28 ($SD_{(comfort)} = 0.83$) and 4.42 ($SD_{(safety)} = 0.90$), respectively. Figures 10 and 11 show the average scores of comfort and safety for the five gestures. A Friedman test found no statistical significance among scores of comfort ($\chi^2 = 4.28$, $p = .37 > .05$) and safety ($\chi^2 = 7.59$, $p = .11 > .05$).

5.4. Discussion of study 3

The results of Study 3 indicate that the user-defined foot gestures are generally intuitive and easy to memorize, as well as comfortable and safe for eyes-free interaction during showers.

The majority of participants (22/24) correctly matched gestures with the corresponding tasks. This result validates the intuitiveness of the user-defined gesture set. Through a post hoc interview, we found that the errors participants made were mainly attributed to random factors (e.g., carelessness). No evidence was found to be related to the unreasonable mapping relationship between the gestures and the tasks. The very low error rate (0.04) can also prove the intuitiveness of use for our gesture set, as one participant stated:

> "As for me, it is quite easy to identify the corresponding gesture for a given target task." (P_1)

In addition, the user-defined gesture set was also easy to recall since most participants (23/24) performed all gestures correctly in the first round on Stage 2. Only one participant made a mistake accidentally, and she quickly realized her mistake when the next slide was presented to her. Participants stated that:

> "These gestures are intuitive and match to the corresponding tasks closely. I can remember all of them without too much practice." (P_{14})

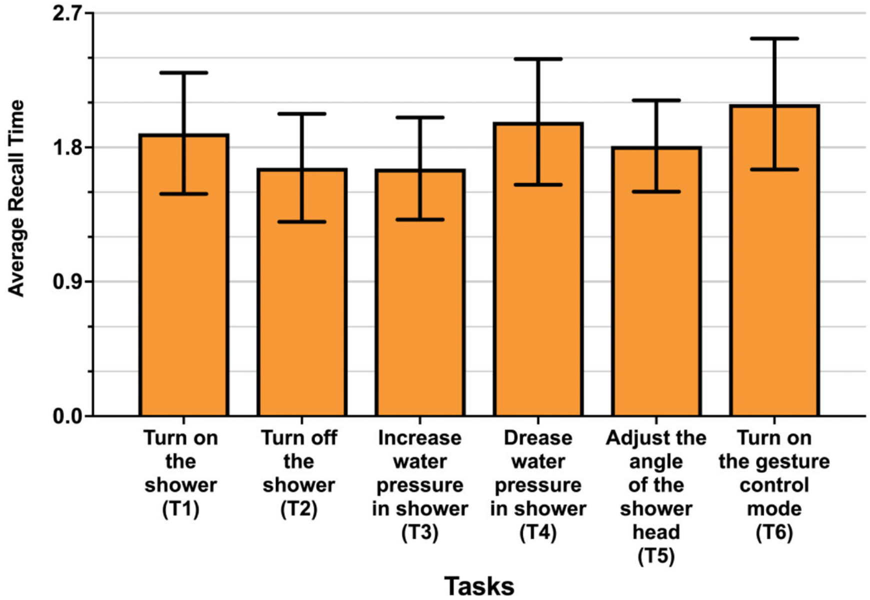


Figure 9. Average recall time among tasks, with 95% confidence intervals.

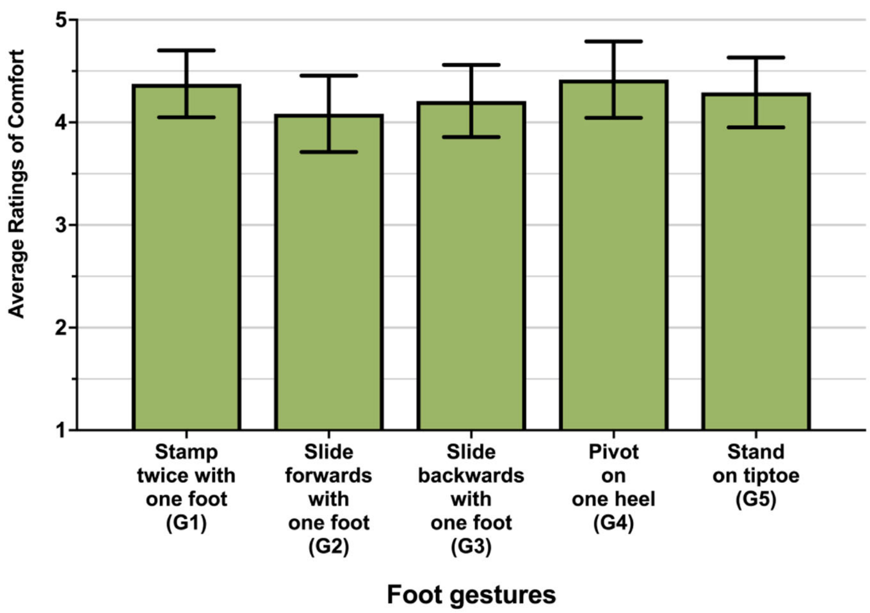


Figure 10. Average ratings of comfort among 5 top foot gestures, with 95% confidence intervals.

High subjective ratings of comfort and safety were received with an average score of 4.28 and 4.42, respectively. Some participants expressed their satisfaction with our design:

> "These foot gestures do not need too much physical effort to perform. Most importantly, they are so cool! I just feel like I am dancing in the bathroom. I personally can't wait to use such a system" (P_{20})

6. Implications for foot gesture design

Based on observations of the user's mental model and behaviors, we proposed the following guidelines for foot gesture design.

6.1. Simplified mental models for foot gesture interactions under the context of eyes-free interaction during shower

Compared to the human hand, the human foot has lower DOF (Degree-Of-Freedom). Therefore, designers should be aware that foot gestures should be simple, iconic, and easy to perform and remember. As shown in Figure 4, end-users preferred single (76.1%), atomic (73.3%), deictic (65.0%), dynamic (76.1%), and asymmetric (82.8%) foot gestures. In addition, many dynamic foot gestures involve movements along the z-axis (40.1%) and within the x–y plane (27.7%).

Compared to previous studies on freehand gesture design, our study reported a higher percentage of participants who preferred to use singulative gestures, e.g., 76.1% single-foot

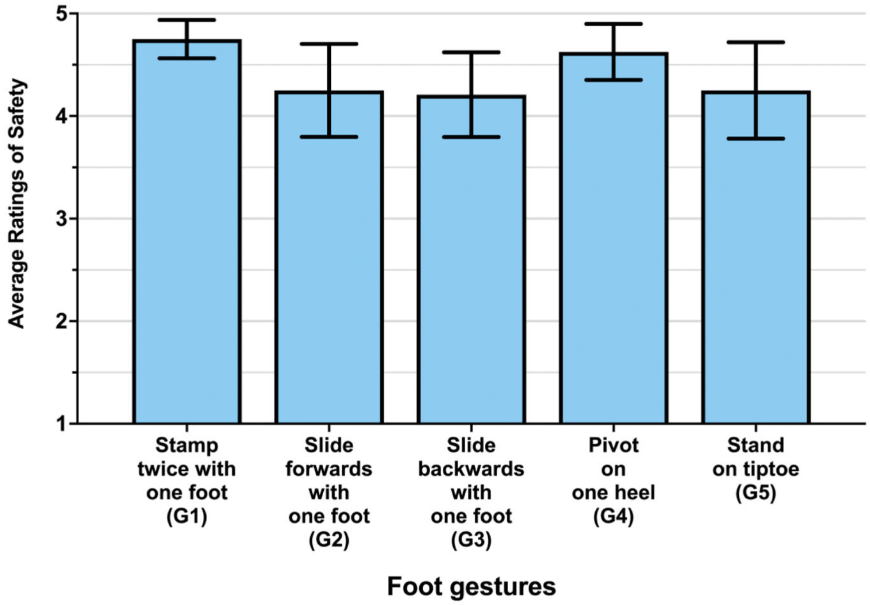


Figure 11. Average ratings of safety among 5 top foot gestures, with 95% confidence intervals.

gestures in our study vs. 64.6% single-hand gestures in Wobbrock et al. (2009) research. Moreover, a higher percentage of asymmetric foot gestures was found in our study compared to the result reported by Felberbaum and Lanir (2018) (82.8 vs. 67.4%). This was not surprising since the participants had to keep one foot standing during the showering process. In addition, we found a much higher proportion of gestures movements along a single axis relative to the results reported by Wobbrock et al. (45.6 vs. 17.7%). For gesture classification in the *Nature* dimension, participants in our study proposed a lower proportion of symbolic foot gestures (0.6%) and metaphorical foot gestures (0%). In contrast, more participants preferred symbolic (10.0%) and metaphorical gestures (18.9%) reported by (Wobbrock et al., 2009).

In summary, interaction designers should be aware of users' simplified mental models and design simple foot gestures that can be easily performed and remembered by end-users, and consequently correctly recognized by the system.

6.2. Number of foot gestures preferred by end-users

Typically, the number of gestures used in a hand-gesture-based interaction system ranged from 1 to 35 (Vuletic et al., 2019). Unlike hand-gesture-based applications, in which the users only need to center on performing a single primary task by using hand gestures, for example, performing a "view product details" task in an immersive VR shopping environment using freehand gestures (Wu et al., 2022), foot gesture interaction in this study was used for controlling a complex concurrent task, in which the user is washing their hair using their hands with their eyes closed while performing foot gestures for controlling the shower spray. Such a complex concurrent task requires the involvement and coordination of the user's hands, feet, and mind. Therefore, it is impractical to design too many gestures at once in a foot-gesture-based system. Otherwise, it will exert too much cognitive burden on users.

In total, five foot gestures were designed for six core objective tasks (note: the "Stamp twice with one foot" gesture was reused for a binary pair task—Switching On/Off the Shower Spray). Experimental results verified our gesture design scheme. As shown in Figures 7 and 8, the average number of recall errors in the first round was 0.04, i.e., the users need only an average of 1.04 times to correctly perform all required gesture proposals in the recall experiment. Therefore, system designers should carefully reduce foot gestures that the target system can support to a considerably limited number in practice to maintain an acceptable level of memorability for end-users.

6.3. Design space of foot gestures

As an essential part of our body, our feet play an irreplaceable role in supporting our body standing. Although this study explored the feasibility of foot gestures as an alternative interaction manner under the eyes-free context in a smart shower room, we found that the participants used foot gestures with great caution for safety. One proof is that it's rare to see participants devise large-scale gestures (i.e., a gesture that requires the involvement of all body parts) that may cause the user to lose balance or even fall on the slippery floor (e.g., jump with both feet off the ground). In contrast, large-scale hand gestures are commonly seen in previous gesture elicitation studies (Cafaro et al., 2018; Chen et al., 2018). Large-scale foot gestures were also derived in previous elicitation studies for computer games (Felberbaum & Lanir, 2018; Kim & Xiong, 2021; Silpasuwanchai & Ren, 2015), such as jumping, jogging, and swinging legs. In addition, we found that all participants kept at least one foot or at least the toes or the heel on the ground while performing foot gestures. Based on our observation, the maximum sliding distance of the user's non-dominant foot on the floor (i.e., dynamic gestures along the x-axis, y-axis, and in the x–y plane) should not exceed two feet. In addition, the maximum lifting height of the user's non-dominant foot along the z-axis should not exceed knee height.

Another interesting finding was that the deictic foot gestures were more preferred by the participants (65.0%), compared to the symbolic (0.6%) and the metaphorical ones (0%). The latter two categories were very popular in hand gesture interfaces (Wobbrock et al., 2009). The majority of these deictic foot gestures were simple dynamic gestures, which involve two atomic movements: (1) linear movements (e.g., "Slide forwards/backwards with one foot" presented in Figure 5) and (2) arc movements (e.g., "Pivot on one heel" in Figure 5). The percentage of atomic gesture is also founded to be higher than that reported by Felberbaum and Lanir (2018) (76 vs. 57.3%). Such information can help system designers better understand the design space of foot gestures.

6.4. Implications for eyes-free interaction

This study can also inform the design of eyes-free interaction. In recent ten years, software-based plug-ins (Kane et al., 2011) and tactile interfaces (Kane et al., 2013) have been developed for blind users to improve their ability to manipulate objects on tabletop screens rapidly and accurately. Beyond such applications on screens, some researchers also developed eyes-free interaction techniques with hand gestures for the users to accurately select virtual objects out of view in VR environments without any eye engagement (Wu et al., 2021; Yan et al., 2018). In this study, we explored the feasibility of eyes-free foot gesture interaction techniques when performing complex concurrent tasks in a smart shower room. We believe that the observation of the user's mental model and behaviors, the foot gesture taxonomy, and the foot gesture design guidelines distilled from the elicitation study provide helpful insights for system designers to create better eyes-free interaction interfaces.

7. Conclusion

To facilitate users to perform more natural interaction during shower when they have to close their eyes (e.g., washing their hair in the shower with their eyes closed), we conducted a three-phase study to (1) collect end-users' actual requirements through online questionnaires, (2) derive instinctive foot-gesture proposals from end-users by leveraging an extended gesture elicitation study method, and (3) validate the user-elicited foot-gestures in a benchmark test. As a result, we present a complete user-defined foot-gesture set based on the observations and analysis of end-users' mental models and behaviors. We also establish a taxonomy of foot-gestures and explore the foot-gesture design space. Finally, we propose some concrete guidelines for foot-gesture-based interaction. Our method and findings can inform the design of foot-gesture-based techniques for eyes-free interaction in other related application scenarios such as smart homes, intelligent vehicles, VR games, and accessibility design (Ferati et al., 2018).

There are some limitations in our study. First, similar to traditional gesture elicitation studies (Ali et al., 2021; Ruiz et al., 2011; Vatavu, 2012), this paper is not intended to go into too many technical details, because our primary focus here is the in-depth understanding of end-users' mental models and the exploration of the design space for foot-gesture-based user interface design. An interesting next step is to develop robust machine learning and pattern recognition algorithms for foot gesture detection and recognition in the eyes-free interaction context. Second, most participants in our gesture elicitation study were students from a university aged between 19 and 24. To further generalize our findings, additional research is needed to investigate the usability of the user-defined foot gesture set by involving a more general population of actual users. Third, we investigated the use context of the foot gestures in the smart shower scenario. Future work can be extended by modifying and/or adapting the foot-gesture design space we proposed to more general scenarios. Fourth, we are interested in putting the user-defined foot gestures into a digital repository for future research.

Notes

1. https://www.amazon.com/dp/B01MY07CZG/?tag=homeawes-20&th=1
2. https://www.amazon.com/dp/B005ECLU2Q/?tag=fhmeag-20
3. https://www.amazon.com/Echo-Dot/dp/B07XJ8C8F5/ref=hsx_sh_dp_dp_bdg2_dsk
4. https://www.apple.com/homepod-mini/

Acknowledgments

The authors would like to thank the editor and the anonymous reviewers for their insightful comments.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the National Natural Science Foundation of China under [Grant No. 61772564], the Guangdong Basic and Applied Basic Research Foundation under [Grant No. 2021A1515011990], and Guangdong Key Laboratory for Big Data Analysis and Simulation of Public Opinion under [Grant No. 2017B030301003].

ORCID

Zhanming Chen http://orcid.org/0000-0002-9913-7239
Huiyue Wu http://orcid.org/0000-0001-7027-518X

References

Alexander, J., Han, T., Judd, W., Irani, P., & Subramanian, S. (2012). *Putting your best foot forward: Investigating real-world mappings for* *foot-based gestures* [Paper presentation]. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Austin, Texas, USA.

Ali, A., Morris, M. R., & Wobbrock, J. O. (2021). *"I Am Iron Man": Priming improves the learnability and memorability of user-elicited gestures* [Paper presentation]. Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems, Yokohama, Japan.

Austin, C. R., Ens, B., Satriadi, K. A., & Jenny, B. (2020). Elicitation study investigating hand and foot gesture interaction for immersive maps in augmented reality. *Cartography and Geographic Information Science*, *47*(3), 214–228. https://doi.org/10.1080/15230406.2019.1696232.

Cafaro, F., Lyons, L., & Antle, A. N. (2018). *Framed guessability: Improving the discoverability of gestures and body movements for full-body interaction* [Paper presentation]. Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, Association for Computing Machinery, Montreal, QC, Canada.

Chen, Z., Ma, X., Peng, Z., Zhou, Y., Yao, M., Ma, Z., Wang, C., Gao, Z., & Shen, M. (2018). User-defined gestures for gestural interaction: Extending from hands to other body parts. *International Journal of Human–Computer Interaction*, *34*(3), 238–250. https://doi.org/10.1080/10447318.2017.1342943.

Cohen, P., Swindells, C., Oviatt, S., & Arthur, A. (2008). *A high-performance dual-wizard infrastructure for designing speech, pen, and multimodal interfaces* [Paper presentation]. Proceedings of the 10th international conference on Multimodal interfaces, Chania, Crete, Greece.

Dong, H., Danesh, A., Figueroa, N., & Saddik, A. E. (2015). An elicitation study on gesture preferences and memorability toward a practical hand-gesture vocabulary for smart televisions. *IEEE Access*. *3*, 543–555. https://doi.org/10.1109/ACCESS.2015.2432679.

Fan, M., Ding, Y., Shen, F., You, Y., & Yu, Z. (2017). *An empirical study of foot gestures for hands-occupied mobile interaction* [Paper presentation]. Proceedings of the 2017 ACM International Symposium on Wearable Computers, Maui, Hawaii.

Felberbaum, Y., & Lanir, J. (2016). *Step by STEP: Investigating foot gesture interaction* [Paper presentation]. Proceedings of the International Working Conference on Advanced Visual Interfaces, Bari, Italy.

Felberbaum, Y., & Lanir, J. (2018). *Better understanding of foot gestures: An elicitation study* [Paper presentation]. Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, Montreal, QC, Canada.

Ferati, M., Babar, A., Carine, K., Hamidi, A., & Mörtberg, C. (2018). Participatory design approach to internet of things: Co-designing a smart shower for and with people with disabilities. In M. Antona & C. Stephanidis (Eds.), *Universal access in human-computer interaction. Virtual, augmented, and intelligent environments, International Conference on Universal Access in Human-Computer Interaction*. Cham.

Findlater, L., Wobbrock, J. O., & Wigdor, D. (2011). *Typing on flat glass: examining tenfinger expert typing patterns on touch surfaces* [Paper presentation]. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Vancouver, BC, Canada.

Fukahori, K., Sakamoto, D., & Igarashi, T. (2015). *Exploring subtle foot plantar-based gestures with sock-placed pressure sensors* [Paper presentation]. Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, Seoul, Republic of Korea.

Funk, M., Schneegass, S., Behringer, M., Henze, N., & Schmidt, A. (2015). *An interactive curtain for media usage in the shower* [Paper presentation]. Proceedings of the 4th International Symposium on Pervasive Displays, Saarbruecken, Germany.

Gheran, B.-F., Vanderdonckt, J., & Vatavu, R.-D. (2018). *Gestures for smart rings: Empirical results insights and design implications* [Paper presentation]. Proceedings of the 2018 Designing Interactive Systems Conference, Hong Kong, China.

Hirai, S., Sakakibara, Y., & Hayakawa, S. (2012). Bathcratch: Touch and sound-based DJ controller implemented on a bathtub. In A. Nijholt, T. Romão, & D. Reidsma (Eds.), *Advances in computer entertainment, International Conference on Advances in Computer Entertainment Technology*, Berlin, Heidelberg.

Hoshino, K., Koge, M., Hachisu, T., Kodama, R., & Kajimoto, H. (2015). *Jorro Beat: Shower tactile stimulation device in the bathroom* [Paper presentation]. Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems, Seoul, Republic of Korea.

Höysniemi, J., Hämäläinen, P., Turkki, L., & Rouvi, T. (2005). Children's intuitive gestures in vision-based action games. *Communications of the ACM*, *48*(1), 44–50. https://doi.org/10.1145/1039539.1039568.

Kane, S. K., Morris, M. R., Perkins, A. Z., Wigdor, D., Ladner, R. E., & Wobbrock, J. O. (2011). *Access overlays: Improving non-visual access to large touch screens for blind users* [Paper presentation]. Proceedings of the 24th annual ACM Symposium on User Interface Software and Technology, Santa Barbara, California, USA.

Kane, S. K., Morris, M. R., & Wobbrock, J. O. (2013). *Touchplates: Low-cost tactile overlays for visually impaired touch screen users* [Paper presentation]. Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility, Bellevue, Washington.

Karam, M., & schraefel, m. c. (2005). *A taxonomy of gestures in human computer interactions* (project report). A. T. o. C.-H. Interactions. https://eprints.soton.ac.uk/261149/.

Kawakatsu, R., & Hirai, S. (2018, 19–23 March). *Rubbinput: An interaction technique for wet environments utilizing squeak sounds caused by finger-rubbing* [Paper presentation]. 2018 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops), Athens, Greece.

Kim, T., Blum, J. R., Alirezaee, P., Arnold, A. G., Fortin, P. E., & Cooperstock, J. R. (2019). Usability of foot-based interaction techniques for mobile solutions. In S. Paiva (Ed.), *Mobile solutions and their usefulness in everyday life* (pp. 309–329). Springer International Publishing. https://doi.org/10.1007/978-3-319-93491-4_16

Kim, W., & Xiong, S. (2021). User-defined walking-in-place gestures for VR locomotion. *International Journal of Human-Computer Studies*, *152*, 102648. https://doi.org/10.1016/j.ijhcs.2021.102648.

Koike, H., Matoba, Y., & Takahashi, Y. (2013). *AquaTop display: Interactive water surface for viewing and manipulating information in a bathroom* [Paper presentation]. Proceedings of the 2013 ACM international conference on Interactive tabletops and surfaces, St. Andrews, Scotland, United Kingdom.

Lopes, D., Relvas, F., Paulo, S., Rekik, Y., Grisoni, L., & Jorge, J. (2019). *FEETICHE: FEET Input for Contactless Hand gEsture Interaction* [Paper presentation]. The 17th International Conference on Virtual-Reality Continuum and its Applications in Industry, Brisbane, QLD, Australia.

Maskeliūnas, R., Damaševičius, R., & Segal, S. (2019). A review of internet of things technologies for ambient assisted living environments. *Future Internet*, *11*(12), 259. https://www.mdpi.com/1999-5903/11/12/259.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *63*(2), 81–97. https://doi.org/10.1037/h0043158.

Morris, M. R., Danielescu, A., Drucker, S., Fisher, D., Lee, B., schraefel, m c., & Wobbrock, J. O. (2014). Reducing legacy bias in gesture elicitation studies. *Interactions*, *21*(3), 40–45. https://doi.org/10.1145/2591689.

Morris, M. R., Wobbrock, J. O., & Wilson, A. D. (2010). *Understanding users' preferences for surface gestures* [Paper presentation]. Proceedings of Graphics Interface 2010, Ottawa, Ontario, Canada.

Müller, F., McManus, J., Günther, S., Schmitz, M., Mühlhäuser, M., & Funk, M. (2019). *Mind the tap: Assessing foot-taps for interacting with head-mounted displays* [Paper presentation]. Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, Glasgow, Scotland, UK.

Nacenta, M. A., Kamber, Y., Qiang, Y., & Kristensson, P. O. (2013). *Memorability of pre-designed and user-defined gesture sets* [Paper presentation]. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Paris, France.

Nielsen, M., Störring, M., Moeslund, T. B., & Granum, E. (2004). A procedure for developing intuitive and ergonomic gesture interfaces for HCI. In A. Camurri & G. Volpe (Eds.), *Gesture-based communication in human-computer interaction*. Springer.

Pakkanen, T., & Raisamo, R. (2004). *Appropriateness of foot interaction for non-accurate spatial tasks* [Paper presentation]. CHI '04 Extended Abstracts on Human Factors in Computing Systems, Vienna, Austria.

Pan, Y., & Steed, A. (2019). How foot tracking matters: The impact of an animated self-avatar on interaction, embodiment and presence in shared virtual environments. *Frontiers in Robotics and AI*, *6*, 104. https://doi.org/10.3389/frobt.2019.00104.

Paulo, S. F., Relvas, F., Nicolau, H., Rekik, Y., Machado, V., Botelho, J., Mendes, J. J., Grisoni, L., Jorge, J., & Lopes, D. S. (2019). Touchless interaction with medical images based on 3D hand cursors supported by single-foot input: A case study in dentistry. *Journal of Biomedical Informatics*, *100*, 103316. https://doi.org/10.1016/j.jbi.2019.103316.

Piumsomboon, T., Clark, A., Billinghurst, M., & Cockburn, A. (2013). User-defined gestures for augmented reality. In P. Kotzé, G. Marsden, G. Lindgaard, J. Wesson, & M. Winckler (Eds.), *Human-computer interaction – INTERACT 2013, IFIP Conference on Human-Computer Interaction, Berlin, Heidelberg*. Springer.

Roaas, A., & Andersson, G. B. J. (1982). Normal range of motion of the hip, knee and ankle joints in male subjects, 30–40 years of age. *Acta Orthopaedica Scandinavica*, *53*(2), 205–208. https://doi.org/10.3109/17453678208992202.

Ruiz, J., Li, Y., & Lank, E. (2011). *User-defined motion gestures for mobile interaction* [Paper presentation]. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Vancouver, BC, Canada.

Saunders, W., & Vogel, D. (2016). *Tap-Kick-Click: Foot interaction for a standing desk* [Paper presentation]. Proceedings of the 2016 ACM Conference on Designing Interactive Systems, Brisbane, QLD, Australia.

Schlömer, I., Klein, B., & Roßberg, H. (2017). A robotic shower system – evaluation of multimodal human-robot interaction for the elderly. *Gesellschaft Für Informatik e.V*, https://doi.org/10.18420/MUC2017-WS17-0415

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

Sumida, T., Hirai, S., Ito, D., & Kawakatsu, R. (2017). *RapTapBath: User interface system by tapping on a bathtub edge utilizing embedded acoustic sensors* [Paper presentation]. Proceedings of the 2017 ACM International Conference on Interactive Surfaces and Spaces, Brighton, United Kingdom.

Takahashi, Y., Matoba, Y., & Koike, H. (2012). *Fluid surface: Interactive water surface display for viewing information in a bathroom* [Paper presentation]. Proceedings of the 2012 ACM International Conference on Interactive Tabletops and Surfaces, Cambridge, Massachusetts, USA.

Tu, H., Huang, Q., Zhao, Y., & Gao, B. (2020). Effects of holding postures on user-defined touch gestures for tablet interaction. *International Journal of Human-Computer Studies*, *141*, 102451. https://doi.org/10.1016/j.ijhcs.2020.102451.

Vatavu, R.-D. (2012). *User-defined gestures for free-hand TV control* [Paper presentation]. Proceedings of the 10th European Conference on Interactive TV and Video, Berlin, Germany.

Vatavu, R.-D., & Wobbrock, J. O. (2015). *Formalizing agreement analysis for elicitation studies: New measures significance test and tool kit* [Paper presentation]. Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, Seoul, Republic of Korea.

Velloso, E., Schmidt, D., Alexander, J., Gellersen, H., & Bulling, A. (2015). The feet in human–computer interaction: A survey of foot-based interaction. *ACM Computing Surveys*, *48*(2), 1–35. https://doi.org/10.1145/2816455.

Villarreal-Narvaez, S., Vanderdonckt, J., Vatavu, R.-D., & Wobbrock, J. O. (2020). *A systematic review of gesture elicitation studies: What can we learn from 216 studies?* [Paper presentation]. Proceedings of the 2020 ACM Designing Interactive Systems Conference, Association for Computing Machinery, Eindhoven, Netherlands, 855–872.

Vogiatzidakis, P., & Koutsabasis, P. (2019). Frame-based elicitation of mid-air gestures for a smart home device ecosystem. *Informatics*, *6*(2), 23. https://www.mdpi.com/2227-9709/6/2/23.

Vuletic, T., Duffy, A., Hay, L., McTeague, C., Campbell, G., & Grealy, M. (2019). Systematic literature review of hand gestures used in human computer interaction interfaces. *International Journal of Human-Computer Studies*, *129*, 74–94. https://doi.org/10.1016/j.ijhcs.2019.03.011.

Willich, J. V., Schmitz, M., Müller, F., Schmitt, D., & Mühlhäuser, M. (2020). *Podoportation: Foot-based locomotion in virtual reality* [Paper presentation]. Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, Association for Computing Machinery, Honolulu, HI, 1–14.

Wobbrock, J. O., Aung, H. H., Rothrock, B., & Myers, B. A. (2005). *Maximizing the guessability of symbolic input* [Paper presentation]. CHI '05 Extended Abstracts on Human Factors in Computing Systems, Portland, OR, USA.

Wobbrock, J. O., Morris, M. R., & Wilson, A. D. (2009). *User-defined gestures for surface computing* [Paper presentation]. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Boston, MA, USA.

Wu, H., Fu, S., Yang, L., & Zhang, X. (2022). Exploring frame-based gesture design for immersive VR shopping environments. *Behaviour & Information Technology*, *41*(1), 96–117. https://doi.org/10.1080/0144929X.2020.1795261.

Wu, H., Huang, K., Deng, Y., & Tu, H. (2021). Exploring the design space of eyes-free target acquisition in virtual environments. *Virtual Reality*, *26*, 513–524. https://doi.org/10.1007/s10055-021-00591-6

Wu, H., & Wang, J. (2016). A visual attention-based method to address the Midas touch problem existing in gesture-based interaction. *The Visual Computer*, *32*(1), 123–136. https://doi.org/10.1007/s00371-014-1060-0.

Wu, H., Wang, J., & Zhang, X. (2016). User-centered gesture development in TV viewing environment. *Multimedia Tools and Applications*, *75*(2), 733–760. https://doi.org/10.1007/s11042-014-2323-5.

Wu, H., Zhang, S., Liu, J., Qiu, J., & Zhang, X. (2019). The gesture disagreement problem in free-hand gesture interaction. *International Journal of Human–Computer Interaction*, *35*(12), 1102–1114. https://doi.org/10.1080/10447318.2018.1510607.

Yan, Y., Yu, C., Ma, X., Huang, S., Iqbal, H., & Shi, Y. (2018). *Eyes-free target acquisition in interaction space around the body for virtual reality* [Paper presentation]. Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, Association for Computing Machinery, 42.

Zhang, T., Song, T., Chen, D., Zhang, T., & Zhuang, J. (2019). WiGrus: A wifi-based gesture recognition system using software-defined radio. *IEEE Access*. *7*, 131102–131113. https://doi.org/10.1109/ACCESS.2019.2940386.

Zhong, K., Tian, F., & Wang, H. (2011, 12–15 June). Foot menu: Using heel rotation information for menu selection [Paper presentation]. 2011 15th Annual International Symposium on Wearable Computers, San Francisco, CA.

About the Authors

Zhanming Chen is a graduate student at the School of Communication and Design, Sun Yat-Sen University, China. His research interests include human-computer interaction, elicitation study, and usability engineering. He obtained a Bachelor of Marketing from Sun Yat-Sen University, Guangzhou, China, in 2019.

Huawei Tu is an Assistant Professor at La Trobe University, Australia. His research area is Human-computer Interaction, with special interests in multimodal interaction and user interface design. He has published more than 30 research papers including top-tier HCI journal papers (e.g. ACM TOCHI) and conference papers such as ACM CHI.

Huiyue Wu is a Full Professor at Sun Yat-Sen University, Guangzhou, China, where he is also the director of the HCI Laboratory. He is the author of five books and more than 40 publications in the field of HCI (e.g., IJHCS, IJHCI). His research interests include human-computer interaction and virtual reality.

Appendix

Appendix 1. The online questionnaire in study 1

Note: Items marked with * are required.

1. What's your daily bathing routine? Shower only, bath only, or in both ways [Single Choice] *
   ○ Shower only
   ○ Bath only
   ○ In both ways
2. How often do you take showers? [Single Choice] *
   ○ More than once per day
   ○ Once per day
   ○ Once every two days
   ○ Once every three days
   ○ Once a week
   ○ Others ________________ *
3. Have you ever experienced the "passive eyes closure" status during shower? [Single Choice] * (The "passive eyes closure" status refers to the frequent or continuous temporary eyes closure due to water or shampoo liquid flowing along with the head to eyes. It often occurs during hair washing process.)
   ○ Yes
   ○ No (If participants chose this option, the survey was ended.)
4. How often do you experience "passive eyes closure" during shower? [Single Choice] *
   ○ Every time I take a shower
   ○ Often, but not every time I take a shower
   ○ Occasionally
   ○ Only experienced once
   ○ Others ________________ *
5. Which activity(s) was (were) you used to do when your eyes were passively closed during shower? [Multiple Choices] *
   □ A. Shower device(s) manipulation (Examples: turn on/off the shower spray, adjust the water volume and water temperature, etc.)
   □ B. Ventilation device(s) manipulation (Example: turn on/off exhaust fans, etc.)
   □ C. Room temperature control device(s) manipulation (Example: turn on/off the heater or cooler, adjust the setting temperature, etc.)
   □ D. Audio device(s) manipulation (Example: turn on/off the smart stereo or Bluetooth stereo, adjust the audio volume, etc.)
   □ E. Video device(s) manipulation (Example: turn on/off the video display, switch the currently playing video, adjust the video volume, etc.)
   □ Others ________________*

Participants are required to finish the following questions according to their selections in Question 5.

5-A. Which activity(s) of the shower device(s) manipulation was (were) you used to do when your eyes were passively closed during shower? [Multiple Choices] *
   □ Turn on the shower
   □ Turn off the shower
   □ Increase water pressure in shower
   □ Decrease water pressure in shower
   □ Raise water temperature in shower
   □ Lower water temperature in shower
   □ Adjust the angle of the shower head
   □ Switch the water outlet mode of the shower head
   □ Other activity(s): ________________*

5-B. Which activity(s) of the ventilation device(s) manipulation was (were) you used to do when your eyes were passively closed during shower? [Multiple Choices] *
   □ Turn on the exhaust fan
   □ Turn off the exhaust fan
   □ Increase the power of the exhaust fan
   □ Decrease the power of the exhaust fan
   □ Other activity(s): ________________*

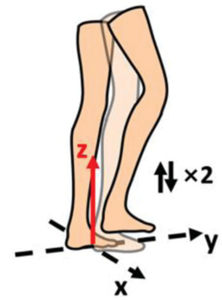
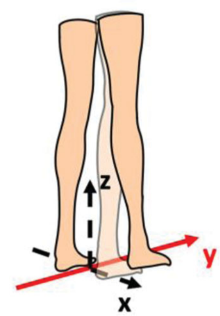
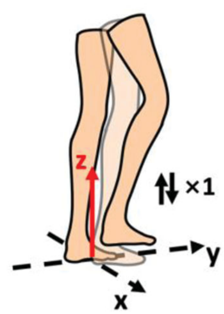
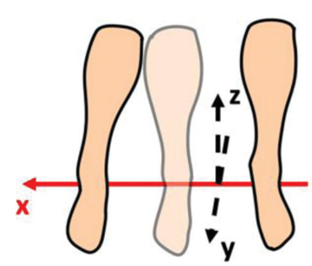
5-C. Which activity(s) of the room temperature control device(s) manipulation was (were) you used to do when your eyes were passively closed during shower? [Multiple Choices] *
   □ Turn on the bathroom heater
   □ Turn off the bathroom heater
   □ Raise the temperature of the bathroom heater
   □ Lower the temperature of the bathroom heater
   □ Other activity(s): ________________*

5-D. Which activity(s) of the audio device(s) manipulation was (were) you used to do when your eyes were passively closed during shower? [Multiple Choices] *
   □ Turn on the audio device(s)
   □ Turn off the audio device(s)
   □ Turn up the audio volume
   □ Turn down the audio volume
   □ Previous audio
   □ Next audio
   □ Other activity(s): ________________*

5-E. Which activity(s) of the video device(s) manipulation was (were) you used to do when your eyes were passively closed during shower? [Multiple Choices] *
   □ Turn on the video display(s)
   □ Turn off the video display(s)
   □ Turn up the video volume
   □ Turn down the video volume
   □ Previous video
   □ Next video
   □ Full screen the video
   □ Exit the full screen status
   □ Other activity(s): ________________*

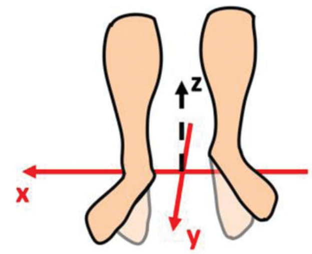
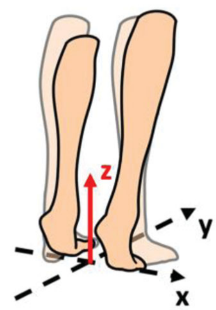
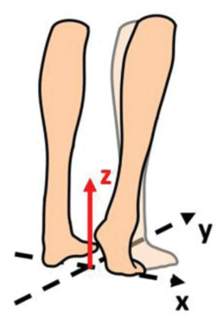
6. What is your gender: [Single Choice] *
   ○ Male
   ○ Female
   ○ N/A
7. Please fill in your age: * ___________
8. Please fill in your occupation, if you are a student, please fill in your major: * ________________

Appendix 2. The foot gesture diagram in study 2.

Number of foot gesture	Foot gesture description	Diagram
(1)	Stamp twice with one foot	**(1)** 
(2)	Slide forwards with one foot	**(2)** 
(3)	Stamp once with one foot	**(3)** 
(4)	Step left/right with left/right foot	**(4)** 

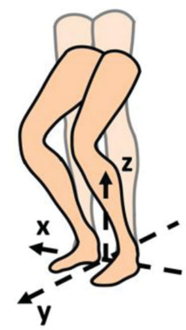
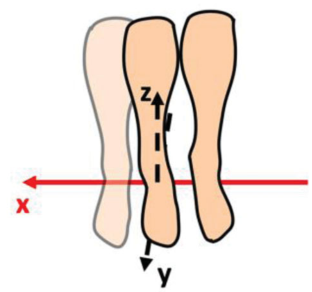
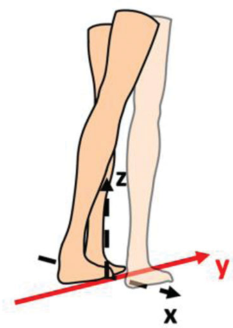
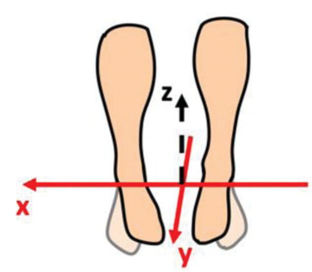
(*continued*)

Appendix 2. Continued.

Number of foot gesture	Foot gesture description	Diagram
(5)	Pivot outwards on both heels	**(5)** 
(6)	Stamp once with both feet in turn	**(6)** 
(7)	Stand on tiptoe	**(7)** 
(8)	Lift one heel	**(8)** 

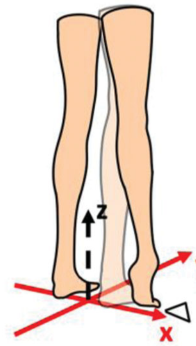
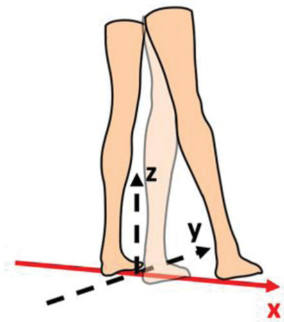
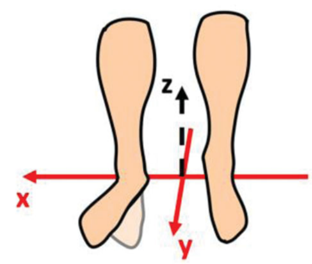
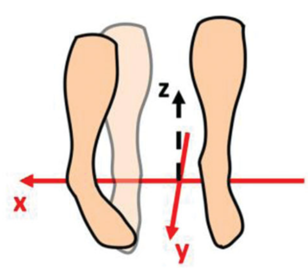
(*continued*)

Appendix 2. Continued.

Number of foot gesture	Foot gesture description	Diagram
(9)	Crouch	**(9)** 
(10)	Step right/left with left/right foot	**(10)** 
(11)	Slide backwards with one foot	**(11)** 
(12)	Pivot inwards on both heels	**(12)** 

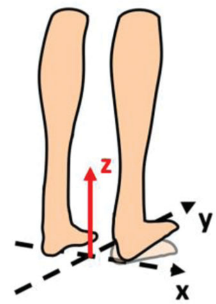
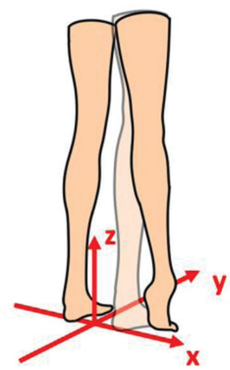
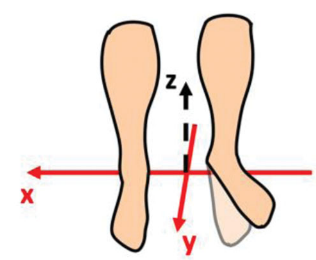
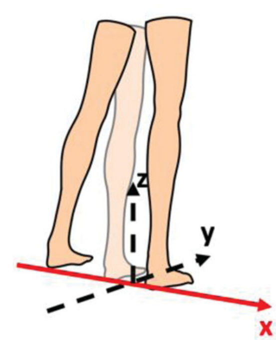
(*continued*)

Appendix 2. Continued.

Number of foot gesture	Foot gesture description	Diagram
(13)	Draw a triangle with one foot	**(13)** 
(14)	Slide to the right with one foot	**(14)** 
(15)	Pivot on one heel clockwise	**(15)** 
(16)	Pivot on one toe counterclockwise	**(16)** 

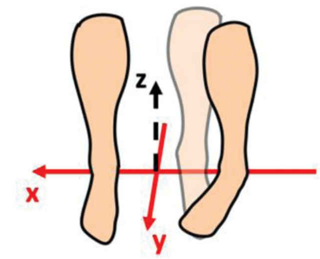
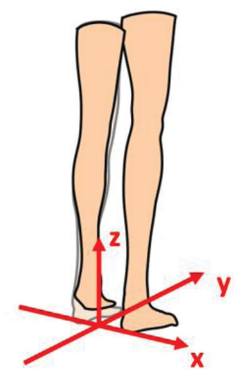
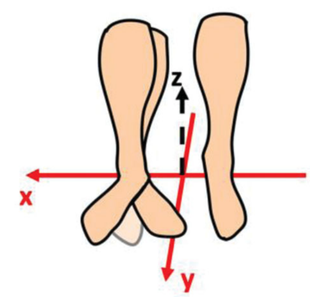
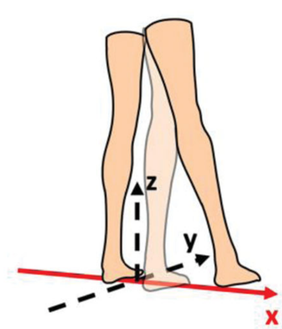
(*continued*)

Appendix 2. Continued.

Number of foot gesture	Foot gesture description	Diagram
(17)	Raise one foot off the floor	**(17)** 
(18)	Tap on the floor with one toe in front of the right body	**(18)** 
(19)	Pivot counterclockwise on one heel	**(19)** 
(20)	Slide one foot to the left	**(20)** 

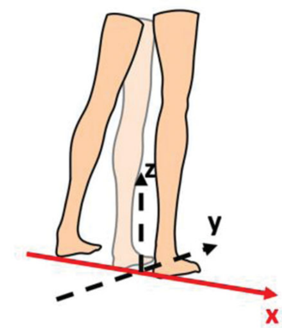
(*continued*)

Appendix 2. Continued.

Number of foot gesture	Foot gesture description	Diagram
(21)	Pivot on one toe clockwise	(21) 
(22)	Tap on the floor with one toe in front of the left body	(22) 
(23)	Pivot on one heel	(23) 
(24)	Slide to the left/right side of the body with one foot	(24-1) 

(*continued*)

Appendix 2. Continued.

Number of foot gesture	Foot gesture description	Diagram
		(24-2) 
(25)	Stand with both feet crossed	**(25)**