Performer vs. Observer: Whose Comfort Level Should We Consider when Examining the Social Acceptability of Input Modalities for Head-Worn Display?

Fouad Alallah
University of Manitoba
Winnipeg, Manitoba, Canada
umalallf@cs.umanitoba.ca

Ali Neshati
University of Manitoba
Winnipeg, Manitoba, Canada
neshatia@cs.umanitoba.ca

Yumiko Sakamoto
University of Manitoba
Winnipeg, Manitoba, Canada
umsakamo@cs.umanitoba.ca

Khalad Hasan
University of British Columbia
Okanagan, Kelowna, Canada
khalad.hasan@ubc.ca

Edward Lank
University of Waterloo
Waterloo, Canada
lank@uwaterloo.ca

Andrea Bunt
University of Manitoba
Winnipeg, Manitoba, Canada
bunt@cs.umanitoba.ca

Pourang Irani
University of Manitoba
Winnipeg, Manitoba, Canada
pourang.irani@cs.umanitoba.ca

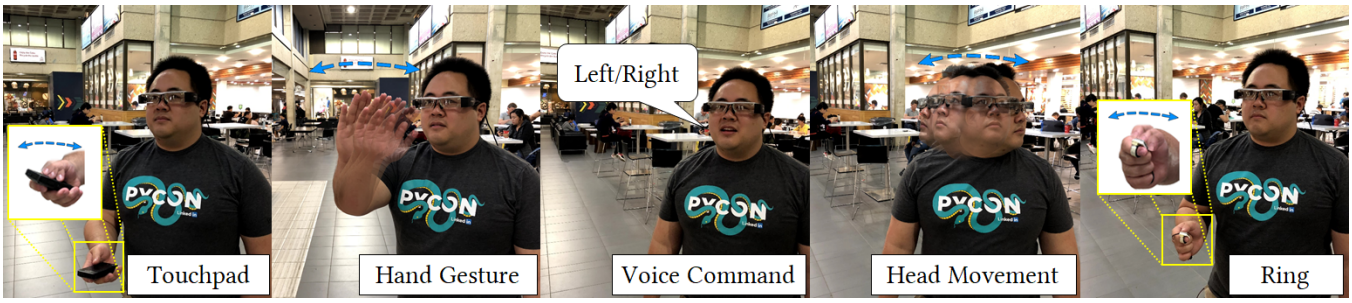

Figure 1: Head-Worn Display (HWD) Inputs that we examined in our studies. We evaluated the social acceptability of these interaction modalities from the perspective of observers as well as performers.

ABSTRACT

The popularity of head-worn displays (HWD) technologies such as Virtual Reality (VR) and Augmented Reality (AR) headsets is growing rapidly. To predict their commercial success, it is essential to understand the acceptability of these new technologies, along with new methods to interact with them. In this vein, the evaluation of *social acceptability* of interactions with these technologies has received significant attention, particularly from the performer's (i.e., user's) viewpoint. However, little work has considered social acceptability concerns from observers' (i.e., spectators') perspective. Although HWDs are designed to be personal devices, interacting with their interfaces are often quite noticeable, making them an ideal platform to contrast performer and observer perspectives on social acceptability. Through two studies, this paper contrasts performers' and observers' perspectives of social acceptability interactions with HWDs under different social contexts. Results indicate similarities as well as differences, in acceptability, and advocate for the importance of including both perspectives when exploring social acceptability of emerging technologies. We provide guidelines for understanding social acceptability specifically from the observers' perspective, thus complementing our current practices used for understanding the acceptability of interacting with these devices.

VRST'18, November 28-December 1, 2018, Tokyo, Japan

ACM ISBN 978-1-4503-6086-9/18/11…$15.00
https://doi.org/10.1145/3281505.3281541

CCS CONCEPTS

• **Human-centered computing → Human-computer interaction (HCI)**; *HCI design and evaluation methods*; User studies;

KEYWORDS

Social Acceptance, HWDs, Augmented Reality, Input Modalities.

ACM Reference Format:
Fouad Alallah, Ali Neshati, Yumiko Sakamoto, Khalad Hasan, Edward Lank, Andrea Bunt, and Pourang Irani. 2018. Performer vs. Observer: Whose Comfort Level Should We Consider when Examining the Social Acceptability of Input Modalities for Head-Worn Display?. In *Proceedings of ACM Woodstock conference (VRST'18).* ACM, New York, NY, USA, Article 4, 9 pages. https://doi.org/10.1145/3281505.3281541

1 INTRODUCTION

Virtual Reality (VR) and Augmented Reality (AR) technologies are gaining popularity. The social acceptability of these new technologies has significant implications because product success largely hinges on whether technology is accepted in society or not [39]. Although this "acceptability" or "one's psychological comfort level about using a technology" might appear to be a simple concept, various factors could be underlying it intricately.

To study technological acceptance in our society, prior studies have explored "Social Acceptability" or "Social Acceptance" of technology, defined as the level of comfort or discomfort from the performers' perspective (i.e., the users' perception) while using new technologies in different social contexts [1, 33]. However, this approach might not allow us to capture the construct of social acceptability fully: indeed, social acceptability from the user's perspective is the user's own perception of how socially comfortable they feel when they are using technology.

An additional measure of how accepting observers (or bystanders) are to new technology, or a new interaction modality might be important in understanding social acceptability further. Understanding this might allow us to, eventually, ease the awkwardness experienced by users in new or habitual social contexts. Consequently, this research step might lead us to the facilitation of the observers' adaptation of these new technologies. Although prior studies have investigated the observers' perceptions of social acceptability while witnessing a user operating a new technology [1, 11, 30–32], to our knowledge, there are no direct comparisons of the two perspectives, nor clear guidelines on how these perspectives need to be considered, if they differ.

In this paper, we examine the social acceptability of Head-Worn Displays (HWDs), as they are becoming available to users, and are beginning to infringe on traditional ways to perceive and interact with digital information. On both commercial and research platforms, HWDs have proven to be effective in numerous contexts [3, 4, 10, 21, 27, 35]. A variety of input techniques, ranging from strong overtones (e.g., hand gesture [7], head movement [13, 18], and voice command [18]) to relatively covert ones (e.g., touchpad [24] and ring [9]), have been demonstrated as a mean to interact with HWDs. These input techniques could have several, often contextual, limitations. For instance, voice commands may not be appropriate during a business meeting, whereas head movements might trigger others' unwanted attention, making performers feel awkward or uncomfortable. Naturally, a clearer understanding of social acceptability for HWD input methods is essential to promote technology adoption smoothly.

In this paper, we explore the social acceptability of HWD inputs from both the performers' (Study 1) and observers' (Study 2) perspectives. More specifically, we explore user and observer perceptions toward using five input modalities commonly used for commercial HWDs and for research prototypes: head movement, hand gesture, touchpad, voice command, and ring input. Based on the approaches used in [1], we also brought participants in both studies to public spaces, the interactions took place publically, and their reactions to the interactions were assessed. This approach allows us to explore participants' social acceptability level while they are actually "performing" the HWDs input modalities, or "observing" other people performing in real-world contexts.

Our exploration offers the following contributions: (i) we evaluate the social acceptability of five input methods for HWDs in different social contexts; (ii) we contrast performers' and observers' perspectives on the social acceptability of different input modalities for HWDs; and (iii) we develop guidelines for considering observers' perspectives in the assessment of social acceptability of new technologies.

2 RELATED WORK

Social acceptance of technology plays an essential role in the adoption of new technologies. Prior research explored social acceptance of new technologies from both performers' and observers' viewpoints, which we discuss below.

2.1 Performers' Perspective

To investigate users' opinion and attitude towards using a new technology, several studies looked at the social acceptability of emerging technologies from the users' perspective [1, 11, 30–32]. For instance, Ronkainen et al. [32] explored the idea of examining social acceptance of gestures to interact with mobile devices. The authors found that performers' social context (e.g., location and surrounding audience) is linked to their decision of accepting or rejecting a gesture. Rico and Brewster [31] extended this idea by exploring body-based (e.g., tapping the nose, belt, shoulder with the devices) and device-based (e.g., shaking and squeezing the device, and tapping the shoulder with the device) gestures. The authors provided participants with an online survey, which included a set of video clips of an actor performing these gestures. Participants were asked to specify the social contexts (i.e., the locations and audiences) in which they would feel comfortable performing the gestures as a performer. They found large variability in participants' perceptions of gesture acceptability. Moreover, researchers [30, 31] explored social acceptance of various multimodal interactions such as body-based gestures (e.g., shoulder rotation), device-based gestures (e.g., device shaking), arbitrary gestures (e.g., upright fist), and speech to interact with mobile devices. They found that performers recognized device-based gestures as more socially acceptable than body-based gestures. Their reasoning was that the visibility of the device in the user's hand will justify the users' somewhat awkward gestures in front of an audience. In contrast, arbitrary gestures were the least acceptable, possibly because the true motives for the gestures are not clearly conveyed to the audience.

Teng et al. [37] looked at input modality and user interaction for common game tasks on commercial HWDs (e.g., Google glasses and Epson Moverio) in public space. The study explored three types of input modalities: handheld, touch, and both non-handheld and non-touch. In their study, participants were asked to provide their feedback regarding their input modalities preferences (palm, ring,

finger, leg, back of the hand, and watch). Their results showed that more than 50% of participants preferred palm interaction over other types of interactions. Their results further indicated that participants' social acceptability perspectives might be connected to their preferences of the input modalities. Specifically, they favored less noticeable modalities.

Ahlström et al. [1] investigated performers' social comfort level while using hand gestures around mobile devices. More specifically, they examined different gesture properties such as the gesture position around the user, gesture size, and gesture length (i.e., duration). They found that quick and less noticeable gestures (i.e., small) are more socially acceptable to the performers. Alallah et al. [2] recently contrasted the feasibility of conducting in-laboratory style as well as crowdsourced social acceptability studies of commonly used HWD input modalities [2]. Their results showed that crowdsourcing platforms could be a potential alternative to the laboratory study style for examining social acceptability of HWD inputs. In contrast, in this paper, we are interested in understanding the social acceptability of HWDs inputs from performers' and observers' perspectives.

Studies explored performers' perspective of new technologies using two major methods. With the first method, researchers provide users a firsthand experience using the new technology, then collect users' feedback on the technology [2, 31]. With the second method, users watch a set of video clips, then they are asked to imagine themselves as a user of the technology. Finally, their feedback is collected [2, 16, 30, 31]. In our studies, we employed the first approach as it provides users a rather authentic experience of using the input modalities in public spaces.

2.2 Spectators' Perspective

While many social acceptability studies are conducted from the performers' perspective, the spectators' perspective has received much less attention. Profita et al. [28] examined spectators' social perception watching videos of actors interacting with a wearable e-textile interface. Results revealed that culture and gender of the actors have effects on spectators' attitudes toward acceptance of the wearable interface. Denning et al. [8] investigated spectators' reaction on watching other people using Augmented Reality (AR) in public space. The spectators showed either no interest or had negative reactions towards watching the person using the device. They further discovered that the bystanders were concerned about being secretly recorded by the onboard camera, deceitfully. Reeves et al. [29] explored the spectators' perceptions while watching performers interacting with public interfaces (e.g., in art, public performance act, or exhibition design). They found that spectators' perception is important in designing new technologies for public spaces. The authors classified the spectators' experiences of interfaces, based on performer' visibility (e.g., hidden, partially revealed, and fully revealed). Montero et al. [25] further investigated other factors such as interaction types (i.e., secretive, expressive, magical, or suspenseful gestures), and users' role (i.e., performer or observer) that influences the social acceptability of gestural interfaces. The author noted that the relationship between the gesture visibility and their effects on observers is an essential element on how spectators evaluate the social acceptability of gestural interfaces. For example, when the performer gesture is visible to the observer and the gesture has no meaning to it from the observer's viewpoint, it will be challenging for the observer to interpret the motivation behind the gesture and consequently lead to the observer's negative impression.

Observers' perspective is an important aspect to be considered when designing interaction gestures to be used in public spaces. The interaction with devices in such contexts might naturally attract observers' conscious or unconscious attention, and hence reactions. However, when observers recognize the performers' gestures (e.g., head-movement) as well as the intention behind this gesture (i.e., interaction modality), both performers and spectators could experience similar levels of social acceptability.

2.3 Performers' and Observers' Perspective

A few studies have taken both performers' and observers' viewpoints into account in studying social acceptability of new technologies. For instance, Montero et al. [25] investigated both perspectives to understand the social acceptance of a set of interaction gestures. They found that the user's perception of others is an important factor, which helps the observers to clearly see the reason why the users are applying these gestures (i.e., to interact with a device). Koelle et al. [17] investigated a set of HWD and mobile usage scenarios to see how these are perceived by users (i.e., performers) and observers. The authors reported that HWD usage is perceived critically, but more positively from the performers' than from the spectators' perspective due to the unfamiliarity of the device. Ahlström et al. [1] also examined around-device gestures from both perspectives, and found that performers and observers had similar acceptance rating of such gestures. Lucero and Veteck [22] ran an in-situ evaluation to examine the effect of social context on interaction with a HWD from both performers' and observers' perspectives. In their study, participants were asked to walk on a street while wearing and interacting with a HWD. They reported that the majority of participants felt self-conscious about wearing and interacting with a HWD, while observers were confused, curious, and reacted carelessly around the participants.

We are unaware of any previous work that directly compared observers' and spectators' social acceptability. The majority of studies investigated the social acceptability of performers' and observers' perspective by asking participants to provide their reactions if they took the place of either a performer or/and observer [17, 25] or by analyzing video recording of observers' reactions [22].

3 INPUT MODALITIES

Researchers have investigated different input modalities to improve interactions with HWDs. These input modalities range from explicit or noticeable ones (e.g., hand gesture, head movement, and voice command) to subtle or less noticeable ones (e.g., touchpad and ring) [2, 9, 18]. Using noticeable input modalities in public spaces could naturally capture observers' unwanted attention due to their heightened visibility, and thus, make users feel socially awkward or uncomfortable. In contrast, subtle input modalities, which require less physical movement, should capture less social attention. Therefore, on the one hand, using subtle input modalities might be

perceived to be more socially acceptable by both performers and observers. On the other hand, however, when interactions are explicit, the observers might readily interpret these clear gestures as a mere method to interact with a device, making explicit gestures more socially acceptable. To study these opposing potential confounds, we chose to explore the following five noticeable as well as subtle input modalities to explore social acceptability.

3.1 Touchpad

Touchpads are often used with commercial HWDs to enable indirect interactions with virtual contents that are displayed on HWDs. For instance, Google glass [23] includes a touchpad on the device near the right temple. Epson Moverio [38] supports users input with a handheld touchpad. The touchpad input method generally requires users to perform gestural interactions such as tapping or swiping on the touchpad which is shown to be more intuitive than using another input modality such as head movements [20]. Touchpads appear to be quite useful for individuals with motor impairments as well [24].

3.2 Hand Gesture

Although hand gestures are primarily used during social communications (e.g., one wave at a friend to say hello), they have been used to interact with HWDs. Hand gestures are commonly captured with 2D or depth cameras which are mounted on the device. To recognize such gestures, advanced computer vision algorithms are often required. Hand gestures can be used for a variety of purposes; pointing, clicking [7], zoom-in/out [14], object manipulation [20], and menu item selection [19]. Interestingly, researchers showed that hand gestures are the preferred gestural technique for HWD inputs than voice command or touchpad [5]. However, understandably, large hand movements especially for a prolonged time period could trigger users' retraction on using this input style in public spaces. Thus, social acceptability of hand gesture needs to be explored.

3.3 Head Movement

Several studies examined the benefits and drawbacks of using head movement for HWDs. They focused on accessing content on multi-display environment in particular. Jay et al. [13], found that head movements could be used to overcome the limitations imposed by the devices narrow field of view. Kollee et al. [18] showed that participants preferred using head movement input method to hand gestures, presumably because head movement is more intuitive and hence requires less effort than hand gestures. Somewhat contradictory, Jakobsen et al. [39] suggested that this input could increase mental demand due to the need for switching between multiple windows.

What we must remember is that head movement is a rather explicit input modality technique, which might result in heightened awkwardness during the interaction. Thus, studying social acceptability associated with head movement is crucial. However, despite its potential contribution in the field, to our knowledge, only [2] examined the acceptance of head movement for HWDs with a crowdsourced and a laboratory study. We further explored the social acceptability related to head movement, from observers' as well as spactators' perspectives.

3.4 Voice Command

Voice command to interact with smart devices (e.g., smartphones, HWDs) is becoming popular. Google Glass and Microsoft HoloLens are two examples of HWD devices that allow users to use voice command as an input [12, 34]. Researchers suggest voice command as one of the input modalities with great potential to interact with HWDs [18]. However, this input method offers only limited functionalities and might raise users' privacy concerns especially when they are using it in public spaces.

3.5 Ring

Prior studies have demonstrated the potential of using a digital ring to interact with HWDs. For instance, Ens et al. [9] demonstrated a ring as a prominent input method for HWDs as it provides precise selection, with low-fatigue level for spatial interactions. Moreover, Kienzle and Hinckley [15] showed that this input modality could be used with on-device applications on-the-go (i.e., without requiring a mouse or touchpad). Additionally, this input offers subtle operations which might be more appropriate to be used in public spaces, relative to noticeable input methods such as hand gestures.

Several other input methods such as palm-based interaction [26], gaze interaction [36], wearable touchpad [24], and hand-to-face gestures [33] have been explored as a mean to interact with HWDs. In our studies, however, we chose to focus on the five inputs listed above as these are the most common input modalities used in commercial devices and explored in HCI.

4 STUDIES

We conducted two studies: Social acceptability from the performer's perspective was investigated in Study 1, while social acceptability from the observers' perspective was investigated in Study 2. For each study, participants experienced firsthand usage of the five input modalities. More specifically, in Study 1, participants performed the gestures themselves. In Study 2, participants were present while a confederate performed the gestures. In both cases, participants were asked to provide their feedback on the use of different input modalities in public spaces. This step was designed based on [1], so the users can get a real experience for using the input device, which helps them to readily imagine their feelings of using these inputs in different contexts. Additionally, we bring the participants in a public place to provide them a perception of using the inputs in a real usage context. However, since our studies solely focus on the perspectives (i.e., performers vs. observers), details of social contexts (i.e., locations or audience types) were not investigated.

4.1 Study 1 - Performers' Perspective

In this study, we investigate the social acceptability of HWD inputs from the performers' perspective. We conducted the study in a busy public place of a local university. This environment was selected to ensure that the participants experience using these input modalities in a suitable context, see Figure 1, where he or she is being watched by strangers.

4.1.1 Participants. We recruited 24 participants (14 males, 10 females), aged between 18 and 34 years (8 participants aged between 18-24, and 16 participants aged between 25-34), from a local university.

4.1.2 Method. We first showed participants five video clips where each video depicted a co-author performing image browsing tasks in a photo album. We used 1) right-to-left and left-to-right head movement, 2) hand movement, 3) finger swipe on touchpad, 4) finger slide gestures on the ring surface, and 5) "move next" and "move previous" voice commands to move an image from left to right and right to left. We then let the participants use the device so they can become familiar with the input modalities and the tasks. Subsequently, we asked participants to perform the image navigation task 10 times (5 times from left-to-right, and 5 times from right-to-left) for each input modality. This step provided them the real HWD usage experience of operating the input modalities in a public space.

Once they had completed 10 trials with one input modality, participants completed a questionnaire to assess social acceptance. We would like to note that the participants were specifically directed not to consider physical comfort while answering the questions. Rather we asked them to reflect on how they felt (e.g., embarrassment or discomfort) while performing these gestures in situations where they are observed by other people. The study lasted for roughly 30 minutes and the participants received $15 for their participation.

4.1.3 Questionnaire. The questionnaire mainly explored two questions: (Q1) How would participants feel performing the inputs in five different locations? (Q2) How would participants feel performing these interactions in front of five different audience groups?

Q13: On a scale of 1 to 5 (with **1** being **very socially uncomfortable**, and **5** being **very socially comfortable**), how do you feel performing **Ring** input in front of the following individual, please rate the following audience?

	1 Very socially uncomfortable	2	3	4	5 Very socially comfortable
In front of colleagues	○	○	○	○	○
In front of family	○	○	○	○	○
In front of strangers	○	○	○	○	○
In front of friends	○	○	○	○	○
When Alone	○	○	○	○	○

Figure 2: Questions used in this study.

There are three sections in the questionnaire: the first section contains information on the study, the second section includes questions on participants' basic demographic information, and the last section contains videos showing a co-author using the inputs with HWDs, and a set of questions asking participants' opinion on using the inputs. The videos are implemented to assist users' in recalling the gestures that they performed when interacting with the HWDs. Figure 2 shows an example question that we used in this study. The questions are used to collect participants' feedback on how comfortable they would feel using the inputs in front of five audiences (Colleagues, Family, Strangers, Friends, and Alone) and five locations (Sidewalk, At home, Public Transportation, Workplace, and Shopping Mall). We used five comfort levels, "Very socially comfortable", "Comfortable", "Neutral", "Uncomfortable" and "Very socially uncomfortable", to collect their opinion. We adopted these questions from [1, 31] and made adjustments for our study.

4.1.4 Result. First, Kolmogorov-Smirnov test showed that data was not normally distributed for both location and audience type. Thus, we conducted non-parametric analyses throughout, and thus, median values are reported.

4.1.4.1 Overall social acceptability across locations: We created an aggregate for each input modality to indicate participants' modality specific social acceptability across all the locations. Figure 3 summarizes our results.

The results of Friedman test showed that there were statistically significant differences among participants' social acceptability between modalities; X^2 (4, N = 24) = 69.16, p < .001 (See Figure 3, Left-hand side). We further conducted a series of Wilcoxon Signed Ranks Tests to identify where the differences were located (See Table 1).

We found that the performers' social acceptability of touchpad was different from hand gesture (z = -4.29, p < .001); head movement (z = -4.11, p < .001), and voice (z = -4.18, p < .001). Further analyses yielded that ring was different from voice (z = -4.12, p < .001); hand gesture (z = -4.29, p < .001); and head movement (z = -4.29, p < .001) (Please see Table 1 for the effect sizes, and Figure 3 for the medians). Please note that we interpreted the effect size (r) based on Cohen's convention (i.e., r = .10 as small; r = .30 as medium; and r = .50 as large) [6].

4.1.4.2 Overall social acceptability across audiences: Next, parallel to locations, aggregates were created for each input modality in order to indicate participants' modality specific sensitivity to various audiences. Again, Friedman test yielded statistically significant difference(s) among participants' social acceptability between modalities X^2 (4, N = 24) = 57.97, p < .001 (See Figure 3, Right-hand side).

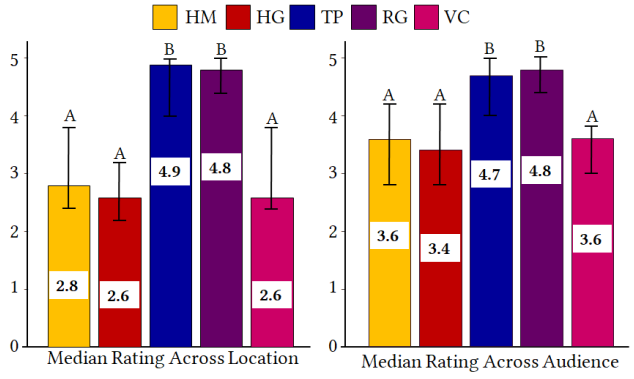

Figure 3: Medians of social acceptability for each modality across locations (left) and audience (right) in Study 1. HM = Head Movement; HG = Hand Gesture; TP = Touchpad; RG = Ring; VC = Voice Command. Error bars represent 95% CI. Matching alphabets indicate statistically ns differences (p > .05) while different alphabets indicate statistically significant differences (p < .05).

Table 1: Effect Sizes (r) from Wilcoxon Signed Ranks Tests in Study 1 (N = 24; p = .05). HM = Head Movement; HG = Hand Gesture; TP = Touchpad; RG = Ring; VC = Voice Command. *indicates that p <.05.

		HM	HG	TP	RG	VC
Locations	HM	x	.20	.84*	.88*	.22
	HG		x	.88*	.88*	.10
	TP			x	.20	.85*
	RG				x	.85*
	VC					x
Audiences	HM	x	.17	.83*	.84*	.03
	HG		x	.84*	.80*	.16
	TP			x	.48	.81*
	RG				x	.85*
	VC					x

Further analyses were conducted to locate the differences. The Wilcoxon Signed Ranks Tests yielded statistically significant results between voice and ring (z = -4.18, p < .001), head movement and ring (z = -4.12, p < .001), hand gesture and ring (z = -3.93, p <.001), voice and touchpad (z = -3.96, p < .001), head movement and touchpad (z = -4.07, p < .001), and hand gesture and touchpad (z = -4.10, p < .001). No other significant results were found (See Table 1 for the effect sizes, and Figure 3 for the medians).

4.1.5 Discussion. Overall, roughly two data clusters emerged; the higher social acceptability cluster and the lower social acceptability cluster: ring and touchpad belong to the higher acceptability cluster while the rest belong to the lower acceptability cluster. Altogether, the results were consistent with the notion of noticeability as discussed in the introduction. That is, less noticeable input modalities (touchpad and ring) were more socially acceptable than highly noticeable modalities from the performers' perspective.

4.2 Study 2 : Observers' Perspective

In Study 2, we explored observers' social acceptability regarding watching a stranger using the HWD inputs. We examined observers' perspective, based on [1], to deepen our understanding of social acceptability. Indeed, understanding observers' perspectives could potentially allow us to envision individuals' future use of the device, even when they have never actually used the device. Moreover, understanding not only the users' but also the observers' perspective, should offer a more holistic view of social acceptability, which could lead us to better predict the success of integrating new technologies in society.

4.2.1 Participants. We recruited 16 participants (9 males, 7 females), aged between 18 and 54 years from a local university.

4.2.2 Method. The study was conducted in the same public space as was used in Study 1. A co-author was using the HWD inputs to browse 50 images (i.e., 10 for each input) in the public space. This was included so the participants could easily imagine observing a user in five different locations and in front of five different audience groups. Once the co-author had finished all the tasks with one input, another co-author asked the participants to answer the questions that were parallel to the ones used in Study 1. Specifically, we asked participants about their feelings if they see a stranger performing each input modality, like the ones they saw in the public space, in different locations and in front of different audiences. After answering these questions, the co-author started using the next input, in a random order, to browse the images. The session lasted approximately 30 minutes and the participants received $15 as a compensation.

4.2.3 Result. For the consistency across studies, and after the assumptions were checked, we ran similar analyses to Study 1.

4.2.3.1 Overall social acceptability across locations: A Friedman test was conducted, and we found statistically significant differences among participants' social acceptability across modalities X^2 (4, N = 16) = 9.79, p < .05.

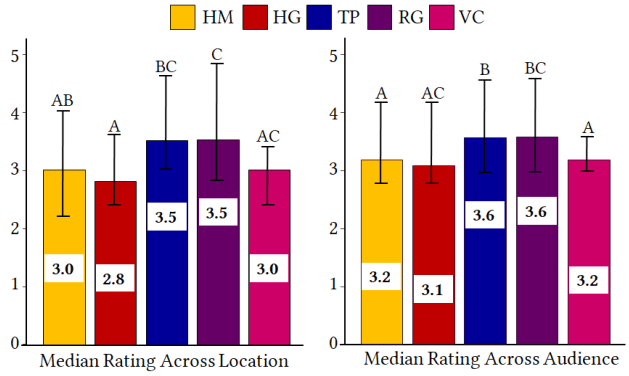

Figure 4: Medians of social acceptability for each modality across locations (left) and audience (right) in Study 2. Error bars represent 95% CI. HM = Head Movement; HG = Hand Gesture; TP = Touchpad; RG = Ring; VC = Voice Command. Matching alphabets indicate statistically ns differences (p > .05) while different alphabets indicate statistically significant differences at p < .05.

Further, Wilcoxon Signed Ranked Tests were conducted. We found multiple medium to large effects (See Table 2). Ring was significantly different from hand gesture (z = -2.63, p = .009), and head movement (z = -2.53, p = .011); hand gesture differed from touchpad (z = -2.13, p = .03) (See Figure 4 for the medians).

4.2.3.2 Overall social acceptability across audiences: Friedman test yielded another significant result among modalities X^2 (4, N = 16) = 19.54, p = .001). Further, Wilcoxon Signed Ranked Tests were conducted. We found ring was significantly different from voice (z = -2.55, p = .01), and head movement (z = -2.36, p = .02); touchpad was different from voice (z = -2.7, p = .007), hand gesture (z = -2.09, p = .04), and head movement (z = -2.56, p = .01). No other significant results were found (Please see Table 2 for the effect sizes, and Figure 4 for the medians).

4.2.4 Discussion. The analysis of the data indicated that participants' social acceptability differs between modalities across locations and audiences (p < 0.05). Importantly, we found medium to large effects in both location and audience results, and the results

Table 2: Effect Sizes (r) and p-values from Wilcoxon Signed Ranks Tests in Study 2 (N = 16). HM = Head Movement; HG = Hand Gesture; TP = Touchpad; RG = Ring; VC = Voice Command., *indicates that p <.05.

		HM	HG	TP	RG	VC
Locations	HM	x	.02	.47	.63*	.08
	HG		x	.53*	.66*	.07
	TP			x	.07	.39
	RG				x	.43
	VC					x
Audiences	HM	x	.03	.64*	.60*	.29
	HG		x	.52*	.46	.20
	TP			x	.47	.68*
	RG				x	.64*
	VC					x

pattern found in Study 2 was largely consistent with what we observed in Study 1 (See Figure 3 & 4): No large effects were found when touchpad and ring were compared (i.e., less noticeable) or hand gesture, head movement, and voice command were compared (i.e., highly noticeable), indicating the potential presence of two clusters.

4.3 Performers' vs. Observers' Perspective

Finally, we compared the data from two studies directly, in order to compare the two perspectives. We conducted non-parametric analyses to account for the sample size difference between Study 1 (n = 24) and Study 2 (n = 16).

4.3.1 Results.

4.3.1.1 Overall social acceptability across locations: We conducted Mann-Whitney U Tests to investigate the effect of perspective (i.e., performer vs. observer) when using five different modalities (Head Movement, Hand Gesture, Touchpad, Ring, and Voice) imagining five different locations (Sidewalk, At home, Public Transportation, Work place, and Shopping Mall). Two significant results emerged. There were statistically significant differences between performers' and observers' social acceptability regarding ring (U = 72.00, z = -3.37, p = 0.001, r = .53) and touchpad (U = 83.00, z = -3.08, p = 0.002, r = .49) with medium to large effects, where performers indicated higher social acceptability than observers did for both modalities. The effect of perspective did not have statistically significant impact on voice (U = 172.00, z = -.56, p = .58, r = .09), hand gesture (U = 160.00, z = -.89, p = .37, r = .14), and head movement (U = 180.50, z = -.32, p = .75, r = .05), and the effects were smaller (See Figure 5).

4.3.1.2 Overall social acceptability across audiences: Five parallel analyses were conducted to investigate the effect of perspectives when using five different modalities imagining five different audience types (Colleagues, Family, Strangers, Friends, and Alone). Analogous to locations, performers' and observers' perceived social acceptability regarding ring (U = 74.50, z = -3.31, p = .001, r = .52) and touchpad (U = 98.50, z = -2.63, p = 0.009, r = .42) differently at a statistically significant level, with medium to large effect sizes; again, performers indicated higher social acceptability than observers did for both modalities. The perspective effect did not have statistically significant impact on voice (U = 167.00, z = -.69, p = .49, r = .11), hand gesture (U = 164.00, z = -.78, p = .44, r = .12), and head movement (U = 181.50, z = -.29, p = .77, r = .05) and small effects were found for these (See Figure 5).

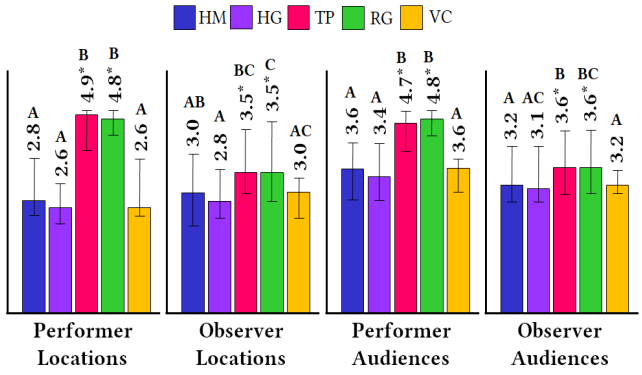

Figure 5: Medians of social acceptability. * indicates medium to large effects between performer vs. observer comparison, within locations or audiences at ps < .005. HM = Head Movement; HG = Hand Gesture; TP = Touchpad; RG = Ring; VC = Voice Command. Matching alphabets indicate statistically ns differences (p > .05) while different alphabets indicate statistically significant difference at p < .05.

4.3.2 Discussion. The statistical comparison of performers' and observers' perspective showed significant differences for ring and touchpad across both locations and audiences (ps < 0.005). Overall, performers' and observers' social acceptability levels were comparable in both social contexts (i.e., location and audience) when the participants were considering noticeable input modalities (e.g., hand gesture or head movement). However, when they were considering subtle or less noticeable input modalities, performers' social acceptability levels were higher than the observers' social acceptability level. That is, performers' and observers' social acceptability perception were not the same when it comes to subtle input modalities. This suggests that performers' and observers' perceptions of social acceptability differ in informing social acceptability and each can be a contributing factor. Further, since this perspective effect centered around the potential noticeability perception, we speculate that it could be possible that users perceive certain input modalities as less noticeable than the observers do, presumably due to the difference in the viewpoint. For instance, a person using a ring might be feeling that it is *very* subtle, while a person looking at the user might feel it is only *somewhat* subtle.

5 DESIGN GUIDELINES

Since understanding the social acceptability of new input modalities for HWDs is an important component for a product's success, we investigated social acceptability in a rather holistic manner. Based on our results indicating some differences between performers' and observers' perspectives, we would like to propose the following guidelines for the HWD input designers:

- Performers' social acceptability (i.e., how one feels socially comfortable during the usage of input modalities) does not always match the observer's perspectives. In our case, differences in perspectives depend on input modalities. Thus, designing HWDs input for different social contexts requires consideration of performers' and observers' acceptability.
- Both researchers and designers need to take the context into account when studying or designing input modalities for HWDs. If the technology is designed to be used in public spaces, researchers should consider subtle input modalities, which capture less unwanted attention.
- Input modalities that require less noticeable gestures to operate are generally more acceptable among the participants.

6 LIMITATION AND FUTURE WORK

Our studies were designed as a preliminary step toward understanding the social acceptability of input techniques for a HWD from performers' and observers' perspectives. Due to its exploratory nature, there are some limitations that we discuss. First, our sample sizes were small, and we acknowledge that larger sample sizes (i.e., more statistical power) should have allowed us to make stronger conclusions, especially with non-significant results (i.e., Study 2). Secondly, most participants were university students attending a local university, and this limits the generalizability of our data. The lack of participants' ethnicity data did not allow us to explore this issue further. Third, in our studies, we only considered a subset of HWDs, which are see-through glasses. Such see-through glasses allow users to see other people around them during the interactions. A further investigation is needed to explore any differences or similarities between see-through glasses and no see-through glasses (e.g., HTC Vive) where users cannot see others watching them while interacting with the device. Somewhat related to this point, our observers were led to be aware of the fact that the performer was using a HWD prior to their observation, and this relatively elevated awareness might have influenced our results: All the gestures were attributed to the HWD, automatically. Concerning the audiences and locations, because we focused our investigation on perspectives, we did not explore the effect of each location or audience type.

A future study should investigate the effect of each type, in conjunction with the perspectives, to understand the concept of social acceptability concerning new input modalities further. Moreover, although we did not explore technology adoption directly, social acceptability might be only one of the many factors influencing the users' decisions about adopting new input modalities. For example, the noticeability of the input modality might also be a factor linked to technology adoption. Finally, and related to this, although noticeability was not studied directly in our study, our data pattern implicated the potential connection between social acceptability and gesture noticeability. Furthermore, the noticeability of the HWD used in our study could have been lower because our HWD looks quite similar to regular glasses (See Figure 1) and relatively less bulky compared to other types of HWDs (e.g., Microsoft Hololens). To clearly investigate the function of noticeability, future studies should control noticeability for both gestures and the devices.

7 CONCLUSIONS

In this paper, two studies examined the social acceptability of five HWD input modalities from performers' and observers' perspectives. Our results indicated that HWDs' inputs that are less noticeable are more socially acceptable from both perspectives, relative to the highly noticeable ones. Further and more importantly, some differences between performers' and observers' perspectives on social acceptability were identified only around less noticeable modalities. This implies that social acceptability should be examined from both performers' and observers' perspectives at least for less noticeable input modalities. Finally, we presented a set of guidelines for HWDs designers and made suggestions on how to advance social acceptability research methodology.

ACKNOWLEDGMENTS

We thank all the participants for their time and feedback. This research was partially funded by a Natural Sciences and Engineering Research Council (NSERC) grant.

REFERENCES

[1] David Ahlström, Khalad Hasan, and Pourang Irani. 2014. Are You Comfortable Doing That?: Acceptance Studies of Around-device Gestures in and for Public Settings. In *Proceedings of the 16th International Conference on Human-computer Interaction with Mobile Devices & Services (MobileHCI '14)*. ACM, New York, NY, USA, 193–202. https://doi.org/10.1145/2628363.2628381

[2] Fouad Alallah, Ali Neshati, Nima Sheibani, Yumiko Sakamoto, Andrea Bunt, Pourang Irani, and Khalad Hasan. 2018. Crowdsourcing vs Laboratory-Style Social Acceptability Studies?: Examining the Social Acceptability of Spatial User Interactions for Head-Worn Displays. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM, New York, NY, USA, Article 310, 7 pages. https://doi.org/10.1145/3173574.3173884

[3] Jarvis J. Arthur, Randall E. Bailey, Steven P. Williams, Lawrence J. Prinzel, Kevin J. Shelton, Denise R. Jones, and Vincent Houston. 2015. A review of head-worn display research at NASA Langley Research Center. In *Proceedings of Spie - the International Society For Optical Engineering*. 9470 – 9470. https://doi.org/10.1117/12.2180436

[4] Jarvis (Trey) J. Arthur III, Lawrence J. Prinzel III, Kevin J. Shelton, Lynda J. Kramer, Steven P. Williams, Randall E. Bailey, and Robert M. Norman. 2009. Synthetic Vision Enhanced Surface Operations With Head-Worn Display for Commercial Aircraft. *The International Journal of Aviation Psychology* 19, 2 (2009), 158–181. https://doi.org/10.1080/10508410902766507 arXiv:https://doi.org/10.1080/10508410902766507

[5] Florian Bemmann. 2015. User Preference for Smart Glass Interaction. (2015), 3–5.

[6] Jacob Cohen. 1988. A power primer. *Psychological Bulletin* 112(1) (1988), 155–159. http://dx.doi.org/10.1037/0033-2909.112.1.155

[7] Andrea Colaço, Ahmed Kirmani, Hye Soo Yang, Nan-Wei Gong, Chris Schmandt, and Vivek K. Goyal. 2013. Mime: Compact, Low Power 3D Gesture Sensing for Interaction with Head Mounted Displays. In *Proceedings of the 26th Annual ACM Symposium on User Interface Software and Technology (UIST '13)*. ACM, New York, NY, USA, 227–236. https://doi.org/10.1145/2501988.2502042

[8] Tamara Denning, Zakariya Dehlawi, and Tadayoshi Kohno. 2014. In Situ with Bystanders of Augmented Reality Glasses: Perspectives on Recording and Privacy-mediating Technologies. In *Proceedings of the 32Nd Annual ACM Conference on Human Factors in Computing Systems (CHI '14)*. ACM, New York, NY, USA, 2377–2386. https://doi.org/10.1145/2556288.2557352

[9] Barrett Ens, Ahmad Byagowi, Teng Han, Juan David Hincapié-Ramos, and Pourang Irani. 2016. Combining Ring Input with Hand Tracking for Precise, Natural Interaction with Spatial Analytic Interfaces. In *Proceedings of the 2016 Symposium on Spatial User Interaction (SUI '16)*. ACM, New York, NY, USA, 99–102. https://doi.org/10.1145/2983310.2985757

[10] Barrett M. Ens, Rory Finnegan, and Pourang P. Irani. 2014. The Personal Cockpit: A Spatial Interface for Effective Task Switching on Head-worn Displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*. ACM, New York, NY, USA, 3171–3180. https://doi.org/10.1145/2556288.2557058

[11] Euan Freeman, Stephen Brewster, and Vuokko Lantz. 2014. Towards Usable and Acceptable Above-device Interactions. In *Proceedings of the 16th International Conference on Human-computer Interaction with Mobile Devices & Services*

(MobileHCI '14). ACM, New York, NY, USA, 459–464. https://doi.org/10.1145/2628363.2634215

[12] G. P. Gowdham and N. Balasubramanian. 2015. A Review on the Google Glass Technology. *International Journal of Emerging Technology in Computer Science & Electronics (IJETCSE '15)* 12, 11 (Feb. 2015), 263–269. http://www.ijetcse.com/issues/?volume_id=22&paging=3

[13] Caroline Jay and Roger Hubbold. 2003. Amplifying Head Movements with Head-Mounted Displays. *Presence* 12, 3 (June 2003), 268–276. https://doi.org/10.1162/10547460376587952 1

[14] Brett Jones, Rajinder Sodhi, David Forsyth, Brian Bailey, and Giuliano Maciocci. 2012. Around Device Interaction for Multiscale Navigation. In *Proceedings of the 14th International Conference on Human-computer Interaction with Mobile Devices and Services (MobileHCI '12)*. ACM, New York, NY, USA, 83–92. https://doi.org/10.1145/2371574.2371589

[15] Wolf Kienzle and Ken Hinckley. 2014. LightRing: Always-available 2D Input on Any Surface. In *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology (UIST '14)*. ACM, New York, NY, USA, 157–160. https://doi.org/10.1145/2642918.2647376

[16] Marion Koelle, Abdallah El Ali, Vanessa Cobus, Wilko Heuten, and Susanne CJ Boll. 2017. All About Acceptability?: Identifying Factors for the Adoption of Data Glasses. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 295–300. https://doi.org/10.1145/3025453.3025749

[17] Marion Koelle, Matthias Kranz, and Andreas Möller. 2015. Don'T Look at Me That Way!: Understanding User Attitudes Towards Data Glasses Usage. In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '15)*. ACM, New York, NY, USA, 362–372. https://doi.org/10.1145/2785830.2785842

[18] Barry Kollee, Sven Kratz, and Anthony Dunnigan. 2014. Exploring Gestural Interaction in Smart Spaces Using Head Mounted Devices with Ego-centric Sensing. In *Proceedings of the 2Nd ACM Symposium on Spatial User Interaction (SUI '14)*. ACM, New York, NY, USA, 40–49. https://doi.org/10.1145/2659766.2659781

[19] Arun Kulshreshth and Joseph J. LaViola. 2014. Exploring the Usefulness of Finger-based 3D Gesture Menu Selection. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*. ACM, New York, NY, USA, 1093–1102. https://doi.org/10.1145/2556288.2557122

[20] Lik-Hang Lee and Pan Hui. 2017. Interaction Methods for Smart Glasses. *CoRR* abs/1707.09728 (2017). http://arxiv.org/abs/1707.09728

[21] David Liu, Simon A. Jenkins, and Penelope M. Sanderson. 2009. Clinical Implementation of a Head-Mounted Display of Patient Vital Signs. In *Proceedings of the International Symposiun on Wearable Computers*. 47–54. https://doi.org/10.1109/ISWC.2009.36

[22] Andrés Lucero and Akos Vetek. 2014. NotifEye: Using Interactive Glasses to Deal with Notifications While Walking in Public. In *Proceedings of the 11th Conference on Advances in Computer Entertainment Technology (ACE '14)*. ACM, New York, NY, USA, Article 17, 10 pages. https://doi.org/10.1145/2663806.2663824

[23] Zhihan Lv, Liangbing Feng, Haibo Li, and Shengzhong Feng. 2014. Hand-free Motion Interaction on Google Glass. In *Proceedings of SIGGRAPH Asia 2014 Mobile Graphics and Interactive Applications (SA '14)*. ACM, New York, NY, USA, Article 21, 1 pages. https://doi.org/10.1145/2669062.2669066

[24] 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 (CHI '15)*. ACM, New York, NY, USA, 221–230. https://doi.org/10.1145/2702123.2702188

[25] Calkin S. Montero, Jason Alexander, Mark T. Marshall, and Sriram Subramanian. 2010. Would You Do That?: Understanding Social Acceptance of Gestural Interfaces. In *Proceedings of the 12th International Conference on Human Computer Interaction with Mobile Devices and Services (MobileHCI '10)*. ACM, New York, NY, USA, 275–278. https://doi.org/10.1145/1851600.1851647

[26] Florian Müller, Niloofar Dezfuli, Max Mühlhäuser, Martin Schmitz, and Mohammadreza Khalilbeigi. 2015. Palm-based Interaction with Head-mounted Displays. In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct (MobileHCI '15)*. ACM, New York, NY, USA, 963–965. https://doi.org/10.1145/2786567.2794314

[27] D. F. Ormerod, B. Ross, and A. Naluai-Cecchini. 2002. Use of a see-through head-worn display of patient monitoring data to enhance anesthesiologists' response to abnormal clinical events. In *Proceedings. Sixth International Symposium on Wearable Computers,*. 131–132. https://doi.org/10.1109/ISWC.2002.1167229

[28] Halley P. Profita, James Clawson, Scott Gilliland, Clint Zeagler, Thad Starner, Jim Budd, and Ellen Yi-Luen Do. 2013. Don'T Mind Me Touching My Wrist: A Case Study of Interacting with On-body Technology in Public. In *Proceedings of the 2013 International Symposium on Wearable Computers (ISWC '13)*. ACM, New York, NY, USA, 89–96. https://doi.org/10.1145/2493988.2494331

[29] Stuart Reeves, Steve Benford, Claire O'Malley, and Mike Fraser. 2005. Designing the Spectator Experience. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '05)*. ACM, New York, NY, USA, 741–750. https://doi.org/10.1145/1054972.1055074

[30] Julie Rico and Stephen Brewster. 2010. Gesture and Voice Prototyping for Early Evaluations of Social Acceptability in Multimodal Interfaces. In *Proceedings of the International Conference on Multimodal Interfaces and the Workshop on Machine Learning for Multimodal Interaction (ICMI-MLMI '10)*. ACM, New York, NY, USA, Article 16, 9 pages. https://doi.org/10.1145/1891903.1891925

[31] 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 (CHI '10)*. ACM, New York, NY, USA, 887–896. https://doi.org/10.1145/1753326.1753458

[32] Sami Ronkainen, Jonna Häkkilä, Saana Kaleva, Ashley Colley, and Jukka Linjama. 2007. Tap Input As an Embedded Interaction Method for Mobile Devices. In *Proceedings of the 1st International Conference on Tangible and Embedded Interaction (TEI '07)*. ACM, New York, NY, USA, 263–270. https://doi.org/10.1145/1226969.1227023

[33] Marcos Serrano, Barrett M. Ens, and Pourang P. Irani. 2014. Exploring the Use of Hand-to-face Input for Interacting with Head-worn Displays. In *Proceedings of the 32Nd Annual ACM Conference on Human Factors in Computing Systems (CHI '14)*. ACM, New York, NY, USA, 3181–3190. https://doi.org/10.1145/2556288.2556984

[34] Allen G. Taylor. 2016. *Develop Microsoft HoloLens Apps Now*. Apress, New York City, MA.

[35] Lisa C. Thomas. 2009. Evaluation of the Use of a Head Worn Display (HWD) for Flight Support in the Commercial Flight Deck. *SAE International Journal of Aerospace-V118-1EJ* 2 (11 2009), 109–114. https://doi.org/10.4271/2009-01-3144

[36] Takumi Toyama, Daniel Sonntag, Andreas Dengel, Takahiro Matsuda, Masakazu Iwamura, and Koichi Kise. 2014. A Mixed Reality Head-mounted Text Translation System Using Eye Gaze Input. In *Proceedings of the 19th International Conference on Intelligent User Interfaces (IUI '14)*. ACM, New York, NY, USA, 329–334. https://doi.org/10.1145/2557500.2557528

[37] Ying-Chao Tung, Chun-Yen Hsu, Han-Yu Wang, Silvia Chyou, Jhe-Wei Lin, Pei-Jung Wu, Andries Valstar, and Mike Y. Chen. 2015. User-Defined Game Input for Smart Glasses in Public Space. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. ACM, New York, NY, USA, 3327–3336. https://doi.org/10.1145/2702123.2702214

[38] Cheng-Yao Wang, Wei-Chen Chu, Po-Tsung Chiu, Min-Chieh Hsiu, Yih-Harn Chiang, and Mike Y. Chen. 2015. PalmType: Using Palms As Keyboards for Smart Glasses. In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '15)*. ACM, New York, NY, USA, 153–160. https://doi.org/10.1145/2785830.2785886

[39] Brian Winston. 1998. *Media Technology and Society: A History: from the Telegraph to the Internet*. Routledge. https://books.google.com.sa/books?id=TZOF_1GZRmYC