

Ubiquitous Interactions for Heads-Up Computing: Understanding Users' Preferences for Subtle Interaction Techniques in Everyday Settings

Shardul Sapkota*
shardul@u.yale-nus.edu.sg
NUS-HCI Lab, Department of
Computer Science, National
University of Singapore
Singapore
Yale-NUS College
Singapore

Ashwin Ram*
ashwinram@u.nus.edu
NUS-HCI Lab, Department of
Computer Science, National
University of Singapore
Singapore

Shengdong Zhao
zhaosd@comp.nus.edu.sg
NUS-HCI Lab, Department of
Computer Science, National
University of Singapore
Singapore

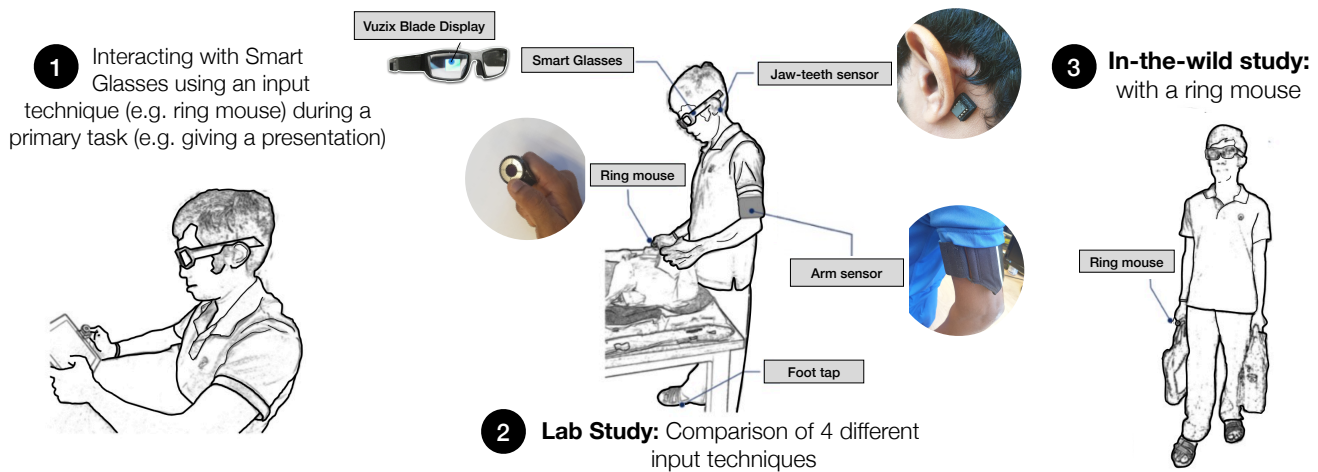


Figure 1: We evaluated four subtle interaction techniques (foot, arm, thumb-index-finger, and jaw-teeth) for smart glasses through a wizard-of-oz study under three hands-busy everyday scenarios: presenting in a sitting posture while holding a mobile device, standing while folding clothes, and walking while carrying two bags. To get an accurate understanding regarding the comfort of wearing the sensors and how the sensors would, in turn, affect the performance of the primary task, we mimicked the physical sensation of having a wearable by placing non-functioning artifacts on the respective positions as shown in (2) (no artifact was placed on the foot (ref Section 3.3)). Results indicate that while each technique has its niche, thumb-index-finger interaction using a ring mouse has the best overall performance and is preferred as a cross-scenario subtle interaction technique for heads-up computing.

ABSTRACT

In order to satisfy users' information needs while incurring minimum interference to their ongoing activities, previous studies have proposed using Optical Head-mounted Displays (OHMDs) with

*Both authors contributed equally to this research.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

MobileHCI '21, September 27-October 1, 2021, Toulouse & Virtual, France

© 2021 Association for Computing Machinery.

ACM ISBN 978-1-4503-8328-8/21/09...\$15.00

<https://doi.org/10.1145/3447526.3472035>

different input techniques. However, it is unclear how these techniques compare against one another in terms of being comfortable and non-intrusive to a user's everyday tasks. Through a wizard-of-oz study, we thus compared four subtle interaction techniques (feet, arms, thumb-index-fingers, and teeth) in three daily hands-busy tasks under different settings (giving a presentation-sitting, carrying bags-walking, and folding clothes-standing). We found that while each interaction technique has its niche, thumb-index-finger interaction has the best overall balance and is most preferred as a cross-scenario subtle interaction technique for smart glasses. We provide further evaluation of thumb-index-finger interaction with an in-the-wild study with 8 users. Our results contribute to an enhanced understanding of user preferences for subtle interaction techniques with smart glasses for everyday use.

CCS CONCEPTS

• **Human-centered computing** → **Interaction techniques**; **Empirical studies in HCI**; **Empirical studies in ubiquitous and mobile computing**.

KEYWORDS

Subtle Interactions; Wearable Computing; Smart Glasses; Heads-up Computing; Thumb-index-finger Interaction

ACM Reference Format:

Shardul Sapkota, Ashwin Ram, and Shengdong Zhao. 2021. Ubiquitous Interactions for Heads-Up Computing: Understanding Users' Preferences for Subtle Interaction Techniques in Everyday Settings. In *23rd International Conference on Mobile Human-Computer Interaction (MobileHCI '21)*, September 27–October 1, 2021, Toulouse & Virtual, France. ACM, New York, NY, USA, 15 pages. <https://doi.org/10.1145/3447526.3472035>

1 INTRODUCTION

Today's smartphone users have been referred to as the 'heads-down generation' as they stay glued to their phones with their necks flexed at an angle [43]. The heads-down style of interaction brings a number of adverse consequences to the users, demanding high attentional and physical constraints [38], and leading to social isolation [35] and musculoskeletal symptoms [21, 27]. Yet, the need for accessing information and processing it on-the-go remains a critical part of people's lives [18, 36].

To overcome some of the negative aspects of heads-down interaction with mobile devices, we envision a 'heads-up' style of computing by leveraging smart glasses or Optical Head-mounted Displays (OHMDs). To serve users' mobile information needs, OHMDs can provide just-in-time digital assistance to the users while they engage in a variety of daily activities under their natural postures. Such heads-up style of computing focuses on satisfying users' information needs with minimum interference to their ongoing activities so that users do not need to stop what they are doing to receive or interact with information. This leads to an important requirement of designing interactions that are synergistic to users' current movements and activities

One category of interaction technique that's more synergistic to users' ongoing activities is subtle interactions [1, 8, 42]. While subtle can refer to a broad facet of design and technology—being deceptive, hidden, non-intrusive, socially acceptable, requiring low effort, easy to perform, etc.—it has been more formally defined by Pohl et al [42] to operate on two levels: 1) on users, it must allow fine movements, have small space requirements, be non-intrusive and minimally disruptive, and 2) on viewers, it must be socially acceptable, hidden, and also minimally disruptive. Examples of subtle interaction techniques for hands-busy contexts, in particular, include interactions through the tapping of feet [17], thumb-fingers [2, 47], teeth [3, 54], and muscle contractions [9]. Each proposed technique has its advantages and is designed to be used in a number of eyes and hands-busy scenarios.

However, to date, there is no systematic evaluation of such interaction techniques across a variety of daily activities [42]. Particularly for interactions with OHMDs, this absence makes it harder for designers to make an informed decision on which interaction technique would users prefer in different situations. For instance, would

users prefer to use their thumb-index-finger or teeth to accept an incoming call while carrying grocery bags or cooking? Understanding user preferences for these interaction techniques would serve as an important step in increasing their user acceptance [12, 25] and supporting the vision of heads-up computing.

To that end, we focus on concretely evaluating subtle interaction techniques for OHMDs under possible heads-up usage scenarios from the user's perspective and providing an approach to systematically compare them. Since the experience of a subtle interaction technique is largely influenced by its implementation and the maturity of the sensing and AI algorithm used, different technique's experience can vary dramatically due to the challenges with the implementation. Thus, to find out the ceiling performance of each interaction technique and use that result as the basis for comparison, we combined Wizard of Oz with a technical probe to perform an investigation on 16 participants. To evaluate the subtlety of the user's own interaction, we (1) employed the NASA-TLX [22] evaluation that captures the physical and mental effort exerted by a user; to quantify the disruptiveness of the interaction to a user's ongoing task, we (2) computed Percentage of Interaction Overhead (PIO), a ratio of the total time taken (time to complete the ongoing task + time to perform the interaction) to the time taken to complete only the ongoing task; and, to gain insights into users' relative subjective preferences, we (3) analyzed users' ranking scores of the interaction techniques for the given task.

We evaluated four types of subtle-interaction techniques for OHMDs designed for hands-busy contexts under three representative everyday scenarios. Our results show that while each interaction technique has its niche, thumb-index-finger interaction has the best overall balance and is most preferred as a cross-scenario subtle interaction technique. This result leads to a possibility for thumb-index-finger interaction to be used as the default subtle interaction technique for heads-up computing with OHMDs.

As interactions in everyday life – outside of the lab's experimental setting – could be marked by more fluid transitions between hands-busy and hands-free scenarios and may be affected by social contexts, interactions with other people, etc, we wanted to understand whether the results of our comparative approach would still hold in more realistic scenarios. Thus, to validate our conclusion from the comparative study, we conducted an in-the-wild deployment of the thumb-index-finger interaction with 8 participants. Participants used a commercial ring mouse to input commands while wearing a pair of smart glasses as they engaged in daily indoor and outdoor activities in a 2-hour period. Results showed that while thumb-index-finger interaction demonstrates strong potential to be used with many everyday tasks for heads-up computing, a number of issues need to be addressed before it can be adopted for long-term use.

The contribution of this paper is threefold: (1) an evaluation method to quantify and compare the subtlety of different interaction techniques for smart glasses from the user's perspective; (2) a revisitation of the tradeoffs and appropriate use-cases associated with different interaction techniques in everyday hands-busy contexts; and (3) an insight into the attitudes and usage behaviors when using thumb-index-finger interaction in the wild. Based on these, we provide design recommendations for thumb-index-finger

interactions with smart glasses to make the technique more suitable for mass adoption.

2 RELATED WORK

Our work is broadly related to research on input techniques for smart glasses and subtle interaction techniques.

2.1 Input Techniques for Optical Head-mounted Displays (OHMDs)

Given the small display size and lack of a defined input space, interacting with OHMDs remains a challenge. Existing commercial solutions to interact with OHMDs include the use of the trackpad on the spectacle frame [56], an external handheld controller [49], or voice input [60]. However, these commercial solutions face their own challenges. First, with trackpads on the spectacle frame, users often experience muscle fatigue when raising their arms for extended periods [23, 32]. External handheld controllers, secondly, are cumbersome when the users' hands are occupied and are not always available [53]. Last, voice input can be inappropriate in noisy environments and often socially awkward to perform [30, 65]. These limitations have thus further challenged the wider adoption of OHMDs as an interaction paradigm for heads-up computing. In response to these limitations, researchers have introduced several other input modalities for OHMDs.

One of such input modalities makes use of the instrumental gloves. Hsieh et al. [24] proposed the use of a haptic glove for text-entry, scrolling, and point-and-select and evaluated the interaction in a public space for its social acceptance. The authors found that the interaction was unobtrusive and socially acceptable. In a similar study, Lee et al. [33] explored thumb-to-finger interaction using a glove for text entry with augmented reality (AR). The authors demonstrated increased average text entry rate compared to existing thumb-finger interactions and found that the system enhanced user mobility compared to other state-of-the-art solutions that required the use of both hands. Thumb-finger interaction had also been explored in Ghosh et al.'s [19] text-editing system with smart glasses where a ring based hand controller was used to complement voice interaction. The authors found that voice interaction with a ring for manual input improved text-editing better than typing on a smartphone for on-the-go contexts, until the user's attention span reached a certain limit. While these evaluations show the potential for finger-based interactions to be widely adopted, they are limited in the scope of evaluating social acceptance (from the perspective of the viewer) in contexts that did not specifically require hands-busy situations.

2.1.1 Hands-free Interactions with OHMDs. In a recent survey on interaction techniques for smart glasses, Lee et al. [32] categorized hands-free interaction evaluated in the literature to movements of the head, gaze, voice, and tongue. Voice interactions, available in Google Glasses and Microsoft HoloLens, were found to be less preferable to gestures or hand-held controller-based interactions [30, 32]. Head gestures [13, 57, 65]— which use the accelerometer and gyroscope inside the smart glasses to register input— had not been evaluated as a major input source, but rather for authentication [65] or gameplay [57], due to the restriction involving head movement [32]. Gaze movement [4, 48, 52, 58] was used to move

the mouse cursor on smart glasses but required obtrusive hardware, were greatly error-prone, and required frequent calibration [6, 32]. All of these proposed hands-free interactions, including tongue gestures [20, 66], have focused on evaluating the performance or the accuracy of the system but not user-preference in hands-busy situations. In addition, to the best of our knowledge, current literature also lacks an evaluation of other hands-free interaction techniques like feet, arms, and teeth as input techniques for smart glasses. This absence leaves the question open on whether hands-free interaction is even desirable during everyday interaction.

2.1.2 Comparative Evaluation of Interaction Techniques for OHMDs.

In addition to the exploration of finger based interactions, existing research has also looked into comparatively evaluating different interaction techniques. Esteves et al. [13] compared two hands-free (dwell, speech) and three hands-on (clicker, on-device, mid-air gesture) interactions to complement head-based input for VR and AR headsets. The authors reported the performance based on Fitts Law analysis and also report on perceived exertion and preference. They found that clicker and dwell worked best across those metrics. Tung et al. [53] explored user-defined game input and also evaluated the interaction in public settings. They compared (1) handheld trackpads, (2) gesture-based and wearable enabled (rings, watches) touch interactions, and (3) mid-air, head-body movement and voice enabled non-touch interactions and found that users preferred non-touch interactions over handheld interactions in gaming context. Our work differs from these comparative explorations in that we focus on evaluating hands-free interactions for subtlety from the user's perspective (comfort, disruptiveness, load) in everyday context rather than evaluating gesture-based interactions with existing interaction techniques for specific contexts like gaming. In addition, we decouple the technology from the interaction itself to understand user preference for only the interaction technique. As software and hardware capabilities continue to mature, our evaluation can thus provide guidance for understanding the ceiling performance of these interaction techniques.

2.2 Subtle Interactions for hands-free and eyes-free input

Recent efforts in developing systems for subtle interaction techniques have introduced multiple alternatives for users to interact with mobile devices. These include the interactions using arm-muscle contractions [8, 9], finger-based interactions [2], feet [17, 45], free-hand movement [59], jaw-teeth [3, 54], wrist [10, 44], and gaze [50, 51]-based interactions. However, each of these works claims to be subtle in a notion somewhat different from the other. These disparate interpretations raise more questions about the nature of subtle interactions than provide unifying accounts of what the interaction is actually supposed to be. The notion of 'subtle' in HCI thus requires some unpacking as it could refer to any of the following qualities: deceptive, hidden, non-intrusive, socially acceptable, requiring low effort, easy to perform, etc.

In a recent attempt to unify these different concepts, Pohl et al. [42] analyze the use of the term 'subtle' in the context of HCI and identify that an interaction is subtle when it operates on two levels— on users and the viewers. Pohl et al.'s analysis lends to a thorough review of existing works on subtle interaction systems and

highlights how these interaction properties are especially desirable during situational impairments [46], i.e., when the hands and/or eyes of the user are occupied, or when social etiquette and norms make the interaction undesirable to use a mobile phone.

2.2.1 Evaluating Subtle Interaction Techniques. As identified by Pohl et al., prior research has focused on evaluating subtle interaction techniques in concrete, measurable ways only from the perspective of the viewers. This evaluation is done by assessing and quantifying the social acceptance or deception of the interaction technique [1, 29, 41, 42]. Despite several works that have claimed to enable subtle interactions [2, 8, 17], there is a lack of a principled approach to evaluating subtle interaction techniques from the perspective of the users [42]. While this aspect of subtlety is commonly claimed, but has little evidence for it available, what benefit does the ‘subtle’ nature of the interaction bring to the user remains unclear. Our work thus aims to address this limitation by providing an empirical approach to first understand subtlety from the user’s perspective.

The evaluation of subtlety from the user’s perspective differs from the general usability studies in HCI in terms of the distinct characteristics defining subtlety—discreetness and non-intrusiveness. For instance, Bobeth et al. [5] evaluated the performance and acceptance of TV menu control using freehand gestures for older adults. The factors they considered were enjoyment, perceived ease of use, perception of control, and task completion time. Similarly, Verma et al. [55] evaluated gesture selection methods for large screen displays for both social acceptance and usability. The factors they considered were task completion time, user emotions, and preferred rankings of the gestures. In such studies, the focus is on whether users subjectively prefer or enjoy the interaction. However, for evaluating the subtlety of an interaction, the evaluation is more concretely defined by whether the interaction is disruptive to the user’s ongoing task and whether the gestures themselves are comfortable, and physically and mentally easier to perform.

3 STUDY 1: A COMPARISON OF SUBTLE INTERACTION TECHNIQUES IN EYES-AND-HANDS-BUSY SCENARIOS

While previous studies have proposed various subtle interaction techniques and the preferred use cases for their respective systems, it is unclear, how these different modes of input compare against one another in terms of being comfortable to the user and how they reduce the disruptiveness to the primary task. Particularly for OHMDs, having interactions that are subtle is desirable for realizing the vision of heads-up computing. Thus, to address this limitation, we evaluate the subtle nature of the interactions during hands-busy situations for their level of intrusiveness, comfort, and disruptiveness in quantifiable ways.

3.1 Subtle Interactions

In order to perform a more systematic comparison, we looked at the analysis performed by Pohl et al. [42] on subtle interactions and found that the gesture-based subtle interactions broadly fell under five categories based on the most prominent body-part engaged during the interaction: finger-, hand-, foot-, arm muscles-, and

teeth- based interactions. Finger-based interactions were put as a different category than hand-based to distinguish on-body finger-based interactions from those that required movements of the wrists and/or palms to interact with an external interface.

We ruled out hands-based interactions due to the constraints imposed by the unavailability of the hands. For finger-based interactions, given our scope of subtle interactions in hands-busy contexts, we focused on thumb-index-finger interactions which have the potential to be used in such contexts compared to other finger-based interactions (like on-object or on-air interactions) [47, 63]. We then proceeded with the following study using the four interaction techniques: arms, foot, thumb-index-finger, and jaw-teeth.

An ideal way to perform this comparison would be to implement the state of art interaction techniques for each of above mentioned body parts; however, we noticed one major practical issue of this approach— in order to perform a fair comparison, we need to achieve the theoretical ideal implementation for each technique, which is difficult - not to mention the theoretical ideal of each technique may yet to be discovered. In order to have a comparison that is close to the theoretical ideal situation of each technique, we combined the wizard-of-oz approach with the technical probes that were informed by prior works on gestural interactions specific to those body parts. This way we could ignore the issues from the technical implementation and focus on the user experience as the users performed the specific gestures for each mode of interaction.

3.2 Participants

16 participants (8 men and 8 women, age = 21.82 ± 3.39) (2 left-handed and 14 right-handed) were recruited from the university community via email and paper flyers. Participants were chosen if they reported having perfect vision or wore comfortable wearing contact lenses to avoid any viewing bias while wearing smart glasses. Since the implemented interface used color-coded cues, we ensured that the participants did not have color vision deficiencies by having them accurately describe all the visual cues they were able to see in the prompts during the practice phase. Two participants indicated that they had previously used foot-tap interactions while interacting with a music pedal whereas two other users were familiar with a ring mouse.

3.3 Apparatus and Sensor Placement

In order to get an accurate understanding regarding the comfort of wearing the sensors and how the sensors would, in turn, affect the performance of the primary task, we mimicked the physical sensation of having a wearable by placing non-functioning artifacts on the respective positions for the foot (no artifact was placed), arm (two plastic arm-bands; 400 x 60 x 20 mm, 28.3 g), thumb-index-finger (a Sanwa Supply ring-mouse; 28 x 36.7 x 34.7 mm, 9.6 g), and jaw-teeth (two microchips mounted on a foam, attached behind ear with skin-safe dressing tape; 15 x 13 x 6 mm, 5 g) as shown in Figure 1. The on-body placement and design choices of these artifacts were informed by the studies [3, 8, 17, 47] which had evaluated the gestural inputs using these modes of interactions. No artifact was placed for the foot-based interaction in the experiment as our pilot studies showed that users could not tell apart the presence or

absence of having artifacts which resembled the apparatus in [17] placed under the shoe sole while performing the foot taps.

In order to evaluate the interaction techniques in eyes-and-hands busy contexts like walking or standing up and looking around, we used the Vuzix Blade smart glasses as the platform to provide visual stimuli to participants. The Vuzix Blade has a 480x480 px display that is vertically centered on the right glass and has a web-server running on Android 5.1.

Since the artifacts that were used in place of the actual sensors had no sensing capabilities, the interaction with the smart glasses was done through a wizard-of-oz approach as described in section 3.5.

3.4 Delimiter

Since a single action is prone to accidental triggers, to avoid false positives, we created a compound gesture involving the repetition of the same basic action three times in quick successions, simulating a triple click. Such compound gestures are much less likely to trigger accidentally [31] and are thus more realistic as a viable gesture to be performed in real world scenarios. In addition, the repetition not only accounts for the additional time taken to perform the activation gesture but also reduces the mental load on the participants of having to remember a new wake up gesture for each mode of interaction.

3.5 Task and Interactions Performed

Four interaction techniques were used in the study, which included interactions based on foot, arm, thumb-index-finger, and teeth. These interaction techniques were chosen based on the rationale provided in the subtle interactions section above.

We used two tasks (Tasks A and B as shown in Table 1) which differed in the number of steps required to complete the interaction. Both tasks were received visually via the smart glasses without any audio cue. The first task involved a phone call (Call) where users had to either accept/reject the call depending on whether the caller was Family (accept) or Work (reject). This was a short task involving 4 interaction steps (3 for delimiter and 1 for accept/reject). The second task, menu selection, involved navigating a menu consisting of 4 options and selecting the option highlighted in red. This task required anywhere between 5 to 8 interaction steps (3 for delimiter, upto 4 for navigating the options and 1 for selection) to complete and was representative of tasks that require a longer duration of interaction.

For both tasks, the experimenter triggered the prompt on the smart glasses and removed the prompt after visually observing that the participant had completed the interaction. Although the display was not updated to reflect the interactions being performed (like updating the selections in the menu navigation task), the experimenter ensured that the participant performed the interaction as specified in Table 1. To minimise the possibility of any delay in visual feedback, especially for difficult-to-notice interactions such as jaw-teeth, timeouts were also set based on the interaction times from the practice sessions to ensure that the visual feedback remained for a pre-defined time.

3.6 Context

We picked three common, everyday hands-busy scenarios as contexts for performing the subtle gestural interactions. To avoid systematic biases towards certain input modes, we first borrowed the taxonomy and categorization of everyday activities from previous works that chart activities of daily living (ADLs). While ADLs are primarily used to evaluate the ability of older adults and rehabilitating patients to perform daily crucial tasks for unassisted living [34], they are representative of what the rest of the population engages in everyday for tasks that do not require assistance. We found the categorization in Aaron Dollar's [11] work to be particularly useful when deciding on everyday hands-busy tasks.

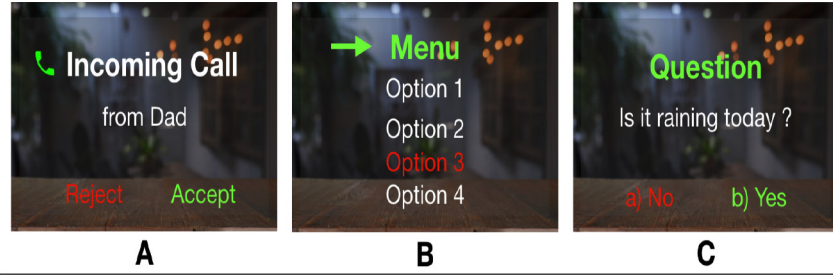
From these broad range of activities, we scoped down the particular contexts based on the constraints determined by the research questions. One set of the constraints included experimental requirements with activities which (1) involved the use of both hands and (2) could be consistently done within the same time so that the disruption to primary tasks could be quantified based on the task completion time (TCT).

The other set of these constraints included the limitations and preferences described in prior research that investigated foot, arm, thumb-index-finger, and teeth based interactions:

- (1) Fukahori et al.[17] proposed using foot based interactions while sitting or standing, and when both hands were busy like when "scrolling the web page of a recipe while cooking or browsing a slide during presentation with hand gestures."
- (2) Costanza et al.[8] proposed using the biceps to control mobile systems "while walking or standing using a subtle contraction with the arm relaxed and on the side."
- (3) Sharma et al. [47] and Wolf et al. [63] proposed that not all fingers were engaged in certain hands busy situations like grabbing an object by a hook (like a shopping bag) or a handle (of a bicycle), generally leaving the thumb free to perform thumb-to-index finger gestures.
- (4) Ashbrook et al. [3] proposed using teeth interactions in hands-busy situations and found better classification accuracy when avoiding eating or talking.

Thus, we picked three contexts from the ADL, at least one from each category as shown in Table 2, and where at least one context ensured the relative advantage of using that particular input mode:

- (1) Giving a presentation (Domestic Activity- Office task): Participants held an ipad as they were seated and read a script such that they were giving a presentation in an intimate setting with a colleague while holding a mobile device for reference. Task completion time was measured in terms of the time taken to read the script. A new script of the same length and complexity was generated at each trial using travel guide information [61] for different countries, with the Flesch reading ease scores [28] fixed between 45-55.
- (2) Carrying bags (Extra Domestic- Shopping) : Participants carried two identical shopping bags of dimensions (in cm) approximately 40 x 30 x 10 weighing 1.5 kg each, one in each hand, and walked once to and fro on a 16 m long path that was slightly inclined at the ends. This path was chosen as it was representative of the slopes in real world terrains. The



Interaction Technique	Delimiter	A			B			C		
		Gestures								
Foot	Triple tap (dominant foot)	Left tap/ Right tap as described by Alexander et al. [1]			Dominant foot tap to move the arrow down + Non-dominant foot tap to select the option			-		
Arm	Triple contraction of the biceps (dominant hand)	Left arm contraction/ Right arm contraction as described by Costanza et al. [9]			Dominant arm contraction to move the arrow down + Non-dominant arm contraction to select the option			-		
Thumb-index-finger	Triple press on the ring-mouse (worn on the dominant hand)	Left button tap/ Right button tap			Down button tap to move the arrow + Center button tap to select the option			Left button tap/ Right button tap		
Jaw-Teeth	Triple bite	Left click/Right click as described by Vega et al. [55]			Left click to move the arrow down + Right click to select the option			-		

Table 1: Interactions that participants performed for the 3 different tasks (A, B, and C) that were prompted on the smart glasses. Task A (call) and Task B (menu selection) were used in both Study 1 and 2 while Task C (SMS response) was used only in Study 2.

Context	Foot	Arm	Thumb-Index-Finger	Jaw-Teeth
Giving a presentation while holding an ipad	✓	✓	✓	
Folding clothes	✓			✓
Carrying bags		✓	✓	✓

Table 2: Table showing the relative advantage of each input technique based on prior research

weight of the bags was chosen based on a previous work [37] such that it replicated the effect of holding realistic objects while keeping the strain on the participants to a minimum. Task completion time was measured in terms of the time taken to complete one lap.

- (3) Folding clothes (Personal- Dressing): Participants stood up and folded the clothes placed on the table (as instructed in a sequential, consistent way [15]) with both hands. Since the previous two contexts required the participants to be either sitting or walking due to the nature of the tasks, we asked the participants to stand up in this context to understand the effect of performing the interaction techniques in a different posture. Participants were given one t-shirt, a shirt, and a pair of pants to fold. Task completion time was measured in terms of the time taken to finish folding the clothes. The order of the clothes was randomised after each trial.

3.7 Study Design

We compared the four different input modes against one another in three different contexts while the participants engaged in common phone application tasks. We conducted a 4 x 3 x 2 factorial within-subject study with *Interaction Technique*, *Context*, and *Task* as the three independent variables. The order of *Interaction Technique* was counterbalanced using a Latin square. Since we were not interested

in comparing the user experience across different *Contexts*, we sequentially ordered the contexts in increasing difficulty (verified through a pilot study with 4 users) following the study design of a previous work [67] to allow the participants to ease into the more complex tasks. The sequence of the activities was as follows: presenting, carrying bags, and folding clothes.

3.8 Procedure

Participants first familiarized themselves with all four input modes. The experimenter then introduced the smart glasses application to the participants and showed them the visual cue of the phone call and the menu selection task. A 3-minute practice phase for each combination of input mode and context was also included so that the participants were familiar with the gesture mappings and instructions for each context. Participants were asked to complete the tasks at their natural pace. The experimenter then recorded the time they took to complete the activity for the specified context without any gestures as the baseline.

Participants were asked to complete the task for each context as soon as possible after they saw the prompt. The experiment was video recorded to get an accurate estimate of the task completion times, which were used in computing the Percentage of Interaction Overhead (PIO). There was a one-minute break between

each task. After each setup, participants filled out the NASA-TLX questionnaire for the contexts.

3.9 Results

3.9.1 Percentage of Interaction Overhead. Percentage of Interaction Overhead (PIO) measures how disruptive the interaction is to the primary task. We measure the time for participants to complete the primary task without any secondary tasks or disruptions as Task time, and the time to complete the primary task while performing the interactions as the Overall time. PIO is computed using the formula of $((\text{Overall time}/\text{Task time}) - 1) \times 100$. The value of PIO indicates the percentage of time to the task time caused by performing the interaction (i.e., 20 means the interaction causes 20% overhead to be used to perform the primary task).

A repeated measures analysis of variance (ANOVA) was performed on the Percentage of The Interaction Overhead.

There was a significant main effect for *Context* ($F_{2,30} = 6.47, p = 0.005$), *Task* ($F_{1,15} = 35.38, p < 0.001$) and *Technique* ($F_{3,45} = 25.49, p < 0.001$). Overall, participants were less impeded while using the thumb-index-finger ($25.42\% \pm 1.92\%$) than the arm ($29.10\% \pm 2.00\%$) ($p < 0.001$). Similarly, the foot interaction ($37.41\% \pm 1.99\%$) caused a significant interaction overhead as compared to using the arm ($p < 0.001$).

There were also significant interaction effects observed on the TCT: *Context x Task* ($F_{2,30} = 13.05, p < 0.001$) and *Context x Technique* ($F_{6,90} = 90.05, p < 0.001$). Post hoc multiple means comparison tests with Bonferroni correction revealed that in the carrying context, the foot took a significantly longer time to complete than all other forms of interactions ($p < 0.001$). However, for the presentation task, the foot ($30.09\% \pm 2.56\%$) had a significantly lesser overhead than using jaw-teeth ($39.40\% \pm 3.28\%$) ($p = 0.002$). Additionally, arm, thumb-index-finger and jaw-teeth interactions were fastest in the carrying context as compared to their use during presenting or folding contexts. For instance, arm interaction had a lower overhead in the carrying context ($12.50\% \pm 1.12\%$) than during presenting ($36.50\% \pm 2.69\%$) ($p < 0.001$) or folding ($p < 0.001$). For the other context and tasks, there were no significant differences between the interaction techniques.

3.9.2 Task Load. A factorial RM-ANOVA was conducted on the unweighted overall NASA-TLX scores after applying Aligned Rank Transformation [62]. There was a significant main effect for *Context* ($F_{2,30} = 3.32, p = 0.05$), *Task* ($F_{1,15} = 7.27, p = 0.017$) and *Technique* ($F_{3,45} = 15.27, p < 0.001$).

A significant interaction effect was observed between *Context x Technique* ($F_{6,90} = 31.26, p < 0.001$). Post hoc comparison using Wilcoxon signed rank tests and Bonferroni correction showed that in the presentation while holding a device context, the jaw-teeth (5.27 ± 0.39) and arm (5.09 ± 0.4) had a significantly higher task load than foot (2.72 ± 0.29) and thumb-index-finger (2.35 ± 0.35) ($p < 0.001$). While carrying bags, foot interaction (5.05 ± 0.37) had the highest task load as compared to thumb-index-finger (2.24 ± 0.37) ($p < 0.001$), arm (3.39 ± 0.31) ($p < 0.001$) and jaw-teeth (3.36 ± 0.35) ($p = 0.03$). In the case of folding clothes, thumb-index-finger interaction (2.63 ± 0.36) was found to have significantly lesser task load than arm (4.95 ± 0.43) ($p < 0.001$) and showed a similar statistical trend compared to the foot interaction (3.82 ± 0.46) ($p = 0.06$).

In addition to the above results we also found that the foot had significantly higher task load while folding (3.82 ± 0.46) than while presenting (2.72 ± 0.29) ($p < 0.001$). In contrast, arm interaction had a much lower task load while carrying bags (3.39 ± 0.31) than in the case of presenting (5.09 ± 0.4) ($p < 0.001$).

3.9.3 Subjective Preferences. Users also ranked the interaction techniques for each context-task pair based on their preference as shown in Figure 4. Overall, the thumb-index-finger interaction was found to be the most preferred form of interaction (14/16). Participants' opinion regarding jaw-teeth was polarised with users either preferring it the least (6/16) or placing it within the top two choices (6/16). Arm, on the other hand, was the somewhat less preferred technique with most participants (10/16) finding it a little more awkward and difficult to perform. However, the arm was preferred in the carrying context (4/16), when gripping the bags in a supinated position.

For the presentation context, foot and thumb-index-finger were equally preferred (8/16) for attending phone calls whereas for longer tasks like menu selection the thumb-index-finger is most preferred (10/16). In the context of carrying bags and walking, the foot was not preferred (15/16) for interaction. This is expected as there is a strong conflict between using the foot for interactions and walking. Surprisingly, the thumb-index-finger interaction was still preferred by many while carrying (7/16) despite the need to use hands for carrying bags, although several users also found the jaw-teeth (4/16) and arm (4/16) to be equally appealing. Arm, however, was the least preferred form of interaction when folding clothes, in which case users found using thumb-index-finger (9/16) or jaw-teeth (5/16) to be the most convenient.

3.10 Discussion

In the *presenting task*, the most preferred technique is thumb-index-finger and foot, followed by arm and jaw-teeth. While the thumb-index-finger was highly preferred due to its convenience and natural feel, accessing the ring buttons while simultaneously holding the iPad imposed some challenges to a few users. Hence, these users opted for the foot interaction while presenting as it did not disrupt the primary task and was more intuitive to perform than jaw-teeth. "*Tapping our feet in response to something is more common than biting (which is more commonly associated with eating or subconsciously done when experiencing certain emotions)*" (P10). Similarly, users felt strained using the arm interaction as it was difficult to clench the biceps while holding the iPad stationary.

While *carrying bags and walking*, the thumb-index-finger was once again the most preferred mode of interaction although to a lesser degree followed by jaw teeth, arm and foot. The decreased preference for thumb-index-finger interaction is due to difficulty ascertaining the accuracy of the button press when the ring mouse was not in their field of view. Users also found arm and jaw-teeth somewhat easy to use as it did not interrupt their primary task, with users inclined towards jaw-teeth as it was more comfortable and less awkward to perform than flexing the arm multiple times. Nevertheless, both techniques lacked a feedback mechanism which, irrespective of context, affected users' perception of the accuracy of their interactions using these modes. Apart from being disruptive, the foot was also physically demanding and frustrating to use during the slopes of the path: "*foot was even more difficult now*".

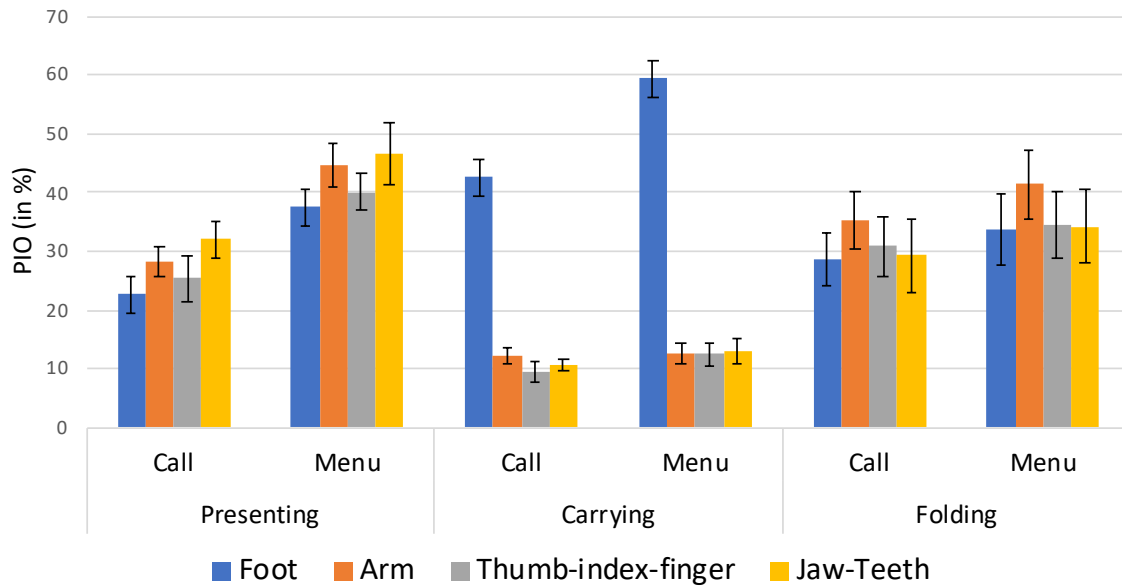


Figure 2: Comparison of the Percentage of Interaction Overhead (PIO) of the different interaction techniques across contexts and tasks.

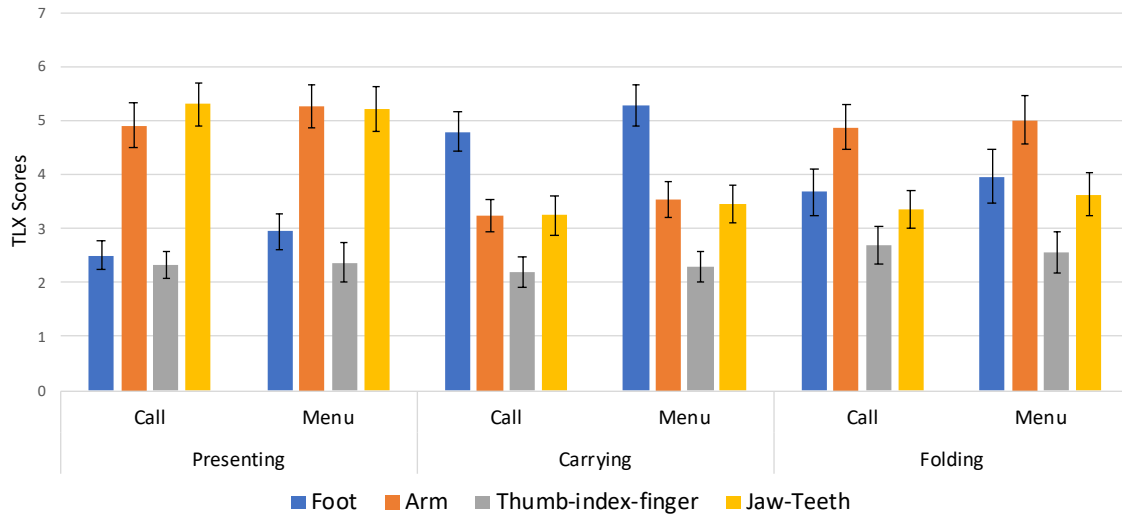


Figure 3: Comparison of the overall unweighted NASA TLX scores of the different interaction techniques across contexts and tasks for 16 participants

especially for [selecting] option 4 where I lose balance on the sloping areas of the path” (P4).

For the *folding* context, the arm was the least preferred mode of interaction due to its disruptive and physically demanding nature. Surprisingly, despite being minimally disruptive, most users disliked the foot mode of interaction since performing multiple foot taps from the standing posture was unpleasant. “it not only required me to stop my task but also required me to shift my balance in order to complete the interaction and thus was the most inconvenient” (P7).

The ring, although slightly disruptive, was still more convenient to use, allowing users to quickly finish the interaction and resume the primary task.

In summary, the results largely met the expectations we had before conducting the study— each interaction technique will have its advantages in certain contexts (foot in sitting and presentation, jaw-teeth in standing and folding, etc.) yet will face difficulties when the interaction and primary task require the same body parts (foot in walking and carrying, jaw-teeth in presentation, etc.). Thus, these

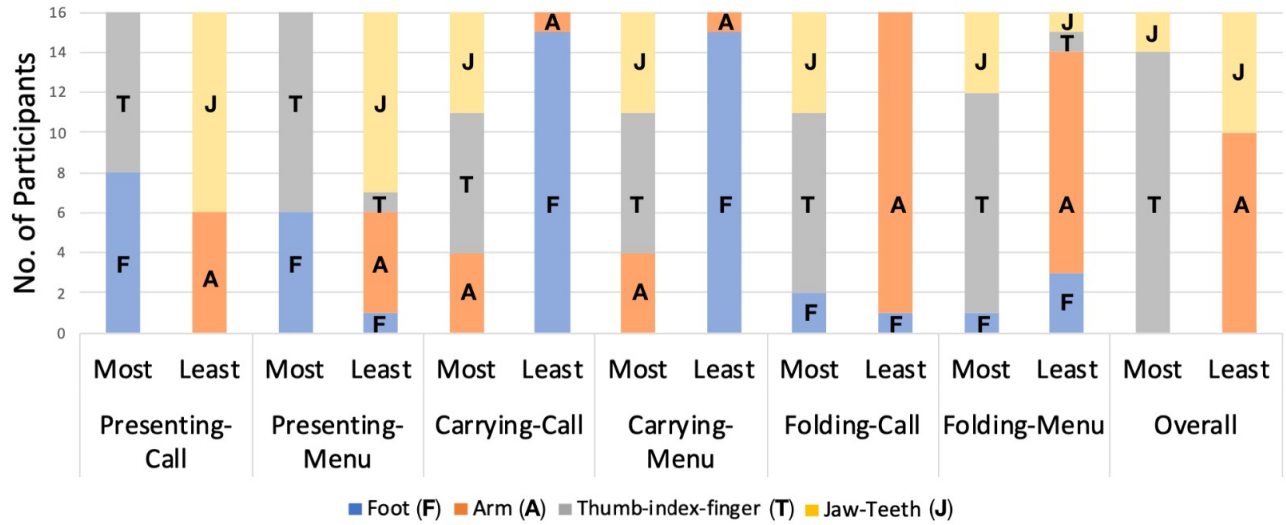


Figure 4: The Most and Least preferred interaction techniques for 16 participants in each context-task pair (Most=Most preferred, Least=Least Preferred)

techniques are more suitable to be deployed to specific contexts. However, one technique (thumb-index finger interaction) stood out as it was regarded to be the fastest and easiest form of interaction to perform across all contexts. It was either rated first or second in all three scenarios, even for the folding task, where we expected it to impose a strong conflict with the primary task. Further investigation indicated that although there is a conflict between the thumb-index-interaction and the folding task, its impact was minimalized due to the dexterity of our fingers, and familiarity of using fingers for interaction – “I found the thumb-index-finger interaction very similar to what I’m already used to using a mouse” (P5) “clicking the button was more ‘braindead’ and more convenient, allowing me to multitask better” (P9). We also found that the prominent tactile feedback offered by the ring was beneficial for users to self-evaluate the success of their interaction, especially for menu selection tasks that involved more interaction steps. Thus, we believe these conclusions could be generalized to any other form of ring device providing tactile feedback during interaction.

4 STUDY 2: IN-THE-WILD INVESTIGATION OF THE THUMB-INDEX-FINGER INTERACTION

From the wizard-of-oz study, we found that thumb-index-finger subtle interaction technique performed on a ring mouse was considered to have the best overall performance across three different scenarios. This raised an intriguing possibility that thumb-index-finger interaction can be used as “the” subtle interaction technique for heads-up computing in everyday use. However, the type of activities we encounter in everyday life is much more than the 3 scenarios considered in study 1. Thus, to test whether thumb-index-finger interaction can serve as the subtle interaction for heads-up computing in everyday use, we conducted an in-the-wild study.

To get a more comprehensive understanding of participants’ experience using the interaction technique in different scenarios, we systematically chose activities from the Activities of Daily Living (ADL) [11] as shown in Table 3. We then examined participants’ daily schedule, and picked 2-hr slots during which they were likely to perform activity instances (e.g. cooking, cleaning the house) from the ADLs. Note that while these activities represented the primary tasks that the participants engaged in during the 2-hr slots, the activities didn’t necessarily fill the entire time slot. Participants were free to do other tasks that they would normally do – like check their phone, watch TV, talk to a friend, etc – in the 2-hr period. They were also instructed to take off the ring-mouse and the smart glasses in any event that could get the devices wet. During the experiment, the display in the smart glasses prompted the participants with the interaction tasks which the participants responded to using thumb-finger-interaction.

4.1 Participants

Eight volunteers (3 men and 5 women, age = 24.6 ± 2.34) took part in our investigation. Four were students and four worked full-time office workers. They were asked to wear the Vuzix Blade smart glasses, through which they received the prompts for interactions, and a small ring-mouse (Sanwa Supply 400-MA077) on the index finger of their dominant hand to perform the interaction. Both of these devices were the same as the one used in study 1. Participants were checked for color vision deficiencies by having them accurately describe all the visual cues they were able to see in the prompts during the practice phase.

4.2 Implementation

The ring-mouse was programmed to interact with 8 tasks of two broad types: 1) four short tasks that accept binary input (yes/no) (Task A and C in Table 1) and 2) one long task (with four subtasks, Task B in Table 1) that accept repeated keystrokes for navigation

ADL Category	Activity Instances (no. of instances)
Domestic Activities	
Food Preparation	cutting vegetables (1), sauteing (1), transferring food between utensils (2), arranging dishes (1), making soup (1)
Housekeeping	replacing bedsheet (1), cleaning bathroom (2), sweeping and mopping (1), wiping desk (1), organising wardrobe (1), folding shirt (1), organising study desk (1)
Laundry	loading washing machine (2), putting cloth to dry (1)
Technology Use	watching TV (1), online shopping on laptop (2), browsing internet on laptop (3), watching videos on phone (1), browsing internet on phone (2), texting on phone while walking (1), texting on phone while sitting (1)
Hobby/sport	running (3), painting (5), playing guitar (1)
Extra-Domestic Activities	
Transportation	walking on the footpath (2), walking in a park (1), walking in mall (2), waiting to cross road (1), waiting for elevator (1)
Shopping	talking to a cashier (1), scanning groceries (1), receiving food delivery (1)
Employment related tasks	typing on laptop (3), fixing laptop (2), talking to colleague (1), Opening door (1)
Physical Self-Maintenance	
Feeding/Medication	eating with cutlery (3), eating with hands (1)
Toileting	brushing teeth (1)

Table 3: A summary of the Activity of Daily Living (ADL) Categories and the Instances that the 8 participants covered. A detailed participant-wise breakdown is given in Table 4

and selection. The short tasks prompted the participants to respond to a call (accept/reject) or SMS (choose between two available answers). The long tasks prompted the participants to navigate a menu with 4 options and select the highlighted option. While the SMS prompt was similar to the accept/reject call prompt in terms of the interactions, both tasks were included to mimic additional application scenarios that users could interact with in a realistic scenario.

As a shorter notification time interval is likely to annoy or frustrate the users [16], we designed the tasks to be triggered at random once every 10-20 minutes, with an average of 15 minute per interaction request, especially since we wanted the participants to fill up a survey after every interaction. We picked this time interval based on prior work [40] that had evaluated the effect of notification interruptions on users' cognitive load. Thus, in the 2 hour period, participants got a total of 8 prompts to interact with a different application each time.

We developed a host server (a Python Flask server running on a Mac-Book Air, 2018). The 2.4 Ghz wireless ring-mouse was connected to the Mac-Book using a USB dongle. The button presses of the ring-mouse were detected as key-press events using a javascript webpage (which was also hosted on the Python Flask server). The Python server subscribed to the button presses of the ring-mouse through a socket connection and relayed the commands to the Vuzix Blade Server through a POST request to update the display.

4.3 Procedure

Prior to the deployment, we met each participant in a 15-minute session where we passed the smart glasses, the ring-mouse, and the Mac-Book Air hosting the server and familiarized them with the setup. The participants were then left alone to their activities. In

situations where participants had to commute, they were told to put the Mac-Book Air inside a bag (the computer could run the server even when the lid was down). The Python server and the Vuzix Blade Server were connected to a mobile hotspot to send and receive the POST requests. The participants were prompted on their smart-glasses to respond to a short survey after each interaction task. Through the survey, we recorded the rating of the comfort, disruptiveness, and social perception of the interaction itself using Likert scales (1-10) and the immediate activity they were engaged in prior to performing the interaction using a form that could be accessed from their mobile phones. Each survey was designed to be finished within 1 minute. After finishing the session, we met the participants again for an interview and filled in a summary survey.

4.4 Results and Discussion

In total, we obtained 59 valid instances of users reacting to prompts gathered during a diverse set of contexts which could be broadly classified as shown in Table 4. We made this classification based on the overall score for disruptiveness, comfortability, and awkwardness and by verifying whether the overall score matched with the participant's attitude towards the interaction during those activities. In cases where multiple participants performed the same activity, the scores were averaged. Given the comfortability (1=least comfortable, 10=most comfortable), disruptiveness (1=least disruptive, 10=most disruptive), and awkwardness (1=least awkward, 10=most awkward) ratings, *Activities suitable for thumb-index-finger interaction* had comfortability score range from 7-10 AND disruptiveness score range from 1-4 AND awkwardness score range from 1-4. *Activities that may/may not be suitable for thumb-index-finger interaction* had at least one of scores out of the three in the range 5-6. *Activities that were unsuitable for thumb-index-finger interaction*

had either comfortability score range from 1-4 OR disruptiveness score range from 7-10 OR awkwardness score range from 7-10.

4.4.1 Comfortability. Interacting with the ring mouse was in general found to be comfortable (7.27 ± 2.29), except in certain both-hands-busy and ring-hand-busy situations. For example, in the context of eating, the interaction was more comfortable when the participant was on a table (e.g. eating soup on a table with a spoon) compared to the same interaction which was less convenient when eating while holding the plate with another hand—*“I felt uncomfortable pressing the buttons as I was holding the plate with one hand and eating with the other”* (P1). Similarly, the ring mouse was found to be uncomfortable for interaction while holding long objects in the ring bearing hand—*“The ring was slightly bulky and while painting I had to adapt around the protruding aspect of the ring.”* (P7).

4.4.2 Disruptiveness. As expected, the ring interaction was found to be minimally disruptive in both-hands-free and stationary contexts such as when watching TV or waiting to cross a road. On the other hand, the ring was disruptive during ring-hand-busy contexts where the mobility of participants’ fingers on the ring bearing hand was compromised. For example, P1 contrasted the experience interacting with the ring while cutting vegetables and pouring oil: *“[While pouring oil] I had to keep the bottle down and then interact as I was afraid I would lose grip if I move my fingers... [whereas while cutting onions] I could interact when holding the knife as my thumb was much closer to the buttons”*. This was also valid in outdoor contexts like shopping where users frequently pick and examine objects: *“[As I was] holding on to some items, it was difficult to interact with the ring-mouse as all my fingers were wrapped around the object...and [the thumb] was unreachable”*(P4). Apart from this, interacting with the ring was also found to be disruptive in social conversation settings because: *“I had to mentally detach from the conversation, and after interaction I had to recollect where I left the conversation before resuming”* (P3).

4.4.3 Comparison to Smartphones. Overall, participants favoured the ring over the smartphone for two main reasons. First, operating the ring only required a single free finger as opposed to phones where both hands were required: *“I liked the fact that I only need the thumb [for interaction] which is especially helpful for cooking where my other hand is occupied”* (P1). Second, the combination of ring mouse with smart glasses was more convenient and faster to use than phones as the interaction was direct thereby skipping the overhead of taking and unlocking the phone. Interestingly, the interview also revealed that participants would prefer to wear the ring even if their phone was around, with the ring being compared to having a smartwatch. *“People use smartwatches even when they have their phones on them as it’s much easier to access and I found the ring-mouse to be similarly useful”* (P2).

4.4.4 Issues for Long Term Adoption. While our analysis suggests that the ring can provide versatile subtle interaction support, users were skeptical about wearing the ring mouse for long duration and incorporating it into their daily routine. This was due to several issues in its current form factor: (1) Despite being small, wearing the ring for two hours was still somewhat fatiguing due to its bulky and slightly heavy nature: *“I felt it was bulky, it was in the way of many things such as painting [...] Also it becomes sweaty while*

engaging in some active tasks making it a bit irritating” (P6), (2) The non-waterproof nature of the ring introduced safety concerns regarding exposing it to water, especially in the cooking context where there is a frequent need to wash objects: *“I liked the ring for the most part, but still I wouldn’t use it because I can’t wash things wearing it and I need to do that a lot while cooking”* (P1), (3) The grip provided by the ring was insufficient for certain users, requiring them to be careful not to let it fall. (4) The act of pressing the button itself was found to be unnatural and less comfortable compared to flicking or sliding motions: *“Doing a flick feels more comfortable even from a social perspective when someone is watching”* (P3)

4.4.5 Design Considerations for Long Term Adoption. The above concerns highlight the need to design a form factor that can offer improved thumb-index-finger interaction experience. In particular, users were inclined towards a lightweight waterproof device having a smooth surface and personalised grip. One way to achieve this is to redesign the ring-mouse as an ordinary ring that completely encircles the finger. This form-factor was previously explored in iRing [39] in which the gestures were detected only based on the movements of the index finger. Future research could build on this work to allow a richer set of thumb-index-finger interactions like swipes or flicks on the ring. However, when the hands are busy and the ring is positioned in the base of the index finger, the interaction might require extending the thumb which could be potentially fatiguing. This observation also warrants an evaluation of the placement of the ring in the index finger. Previous investigations— for example, Tiptap [26] or Magic Finger [64]— have proposed using the tip and/or the middle region of the index finger. These system designs could potentially help overcome the fatigue from having to extend the thumb to the base of the index finger.

4.4.6 Thumb-index finger as a cross-scenario interaction technique for heads-up computing. The primary goal of this study was to understand whether the thumb-index finger style of interaction is a suitable natural pairing for heads-up computing on smart glasses. Our findings suggest that it shows great promise with users finding the interaction style to be comfortable for cross-scenario use— be it indoors or outdoors and in both-hands-busy or ring-hand-busy scenarios with minimal disruption to their primary task: *“It was convenient during busy times like when I need to get some housework done but attend some calls or see some updates.”* (P7). While some discomfort was experienced with the current form factor of the device used to achieve this interaction, the interaction in itself was found to be easy to perform. Moreover, the interaction overall was felt to be very discreet (2.5 ± 1.83), and that this feeling holds across different social contexts (e.g. getting food delivery, talking with other people) makes thumb-index finger technique a viable candidate for heads-up computing using OHMDs.

5 OVERALL DISCUSSION AND HOW TO INTERPRET OUR RESULTS

In this paper, we first investigated how different subtle interactions compare against one another in different hands-busy contexts. To do so, we conducted a wizard-of-oz comparison with technical probes to gain insights into users’ own evaluations of the subtlety of the interactions. We acknowledge that the user experience with

Activities Suitable For Thumb-Index-Finger Interaction		Activities That May/May Not Be Suitable For Thumb-Index-Finger Interaction		Activities Unsuitable For Thumb-Index-Finger Interaction	
Both-hands-busy		Both-hands-busy		Both-hands-busy	
Activity instance	Overall Rating	Activity instance	Overall Rating	Activity instance	Overall Rating
Organising wardrobe	9,1,1	Fixing laptop	9,5,5,6	Folding shirt	8,9,1
Browsing internet on phone	9,1,5, 2	Making soup	7,4,5	Texting on phone while walking	8,7,2
Playing guitar	9,2,3	Cleaning bathroom	6,6,2	Sweeping/mopping	4,7,1
Texting on phone while sitting	9,4,3	Eating with cutlery	5,4,3,2,7	Cutting vegetables	3,7,1
Replacing bedsheet	10,4,1	Ring-hand-busy		Eating with hands	2,9,1
Putting cloth to dry	8,3,1	Scanning groceries	9,6,3	Ring-hand-busy	
Loading washing machine	7,5, 3,1	Brushing teeth	8,6,1	Watching videos on phone	8,7,1
Typing on laptop	8,7,2,7,3	Online shopping with laptop	5,5,3,5	Browsing on laptop	8,7,2
Ring-hand-busy		Painting	5,5,4,7,3,5	Organising study desk	5,8,1
Arranging dishes	10,1,1	Both-hands-free		Opening door	6,7,7
Transferring food between utensils	9,3,5,1	Walking in a park	9,5,2	Wiping desk	4,7,1
Receiving food delivery	10,3,4	Talking to a cashier	8, 6, 2	Both-hands-free	
Sauteing	8,3,1	Walking in a mall	6, 4,5, 2,5	Talking to colleague	9,7,4
Both-hands-free		Running	7,7,6,2		
Watching TV	9,1,1				
Waiting to cross road	9,2,2				
Waiting for elevator	7,2,1				

Table 4: Thumb-index-finger interaction in different contexts covered by the 8 participants. The activity instance are categorised based on the suitability for thumb-index finger interaction using the overall rating of the interaction’s comfortability (1=least comfortable, 10=most comfortable), disruptiveness (1=least disruptive, 10=most disruptive), and awkwardness (1=least awkward, 10=most awkward). E.g., Watching TV (9, 1, 1) indicates that performing interaction in the context of watching TV is very comfortable (9), least disruptive (1), and least awkward (1).

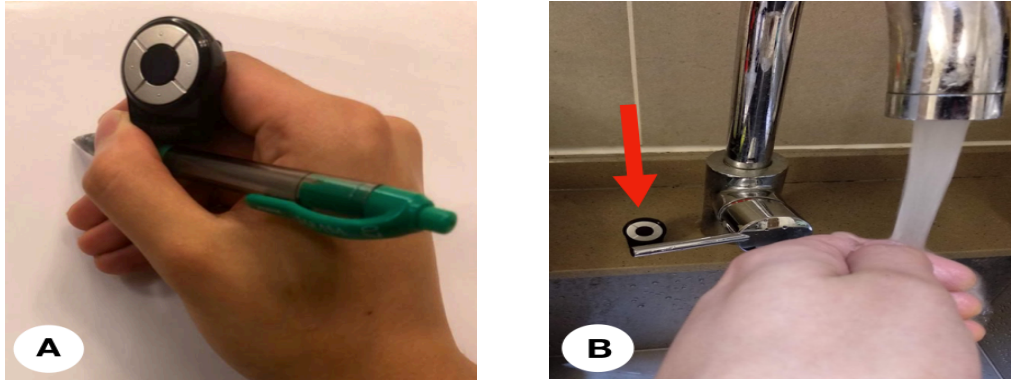


Figure 5: Scenarios where participants found thumb-index-finger interaction with the ring-mouse to be uncomfortable or disruptive. The ring-mouse was bulky, especially when holding an object (A), or had to be taken off when it could get wet (B).

an actual implementation of the system would be different from just having a physical artifact mimic the sensation of the device. However, given our emphasis on capturing users’ experiences while performing the interaction, the wizard-of-oz evaluation helped us to mimic an ideal system for each of those interaction techniques. Our results indicate that thumb-index-finger interaction using a ring mouse is the best cross-scenario subtle interaction technique among the tested candidates. This result is likely to hold even with real implementations. This is because the ring mouse is a relatively mature technology. So its real world usage experience will not be

much different from the Woz study. On the other hand, jaw-teeth, foot, and arm-based subtle interaction techniques are less mature, and likely to face more difficulties (such as lower accuracy) with a real implementation, making them less favored by users. Thus, the relative ranking will still put thumb-index-finger interaction above the other techniques.

Although we picked only three representative contexts each of which required different states of mobility, the same contexts could be carried out in multiple ways– for instance, carrying bags while sitting down or giving a presentation without the need of holding

a mobile device. As expected, the outcome of the preference of the interaction technique in those contexts may be different. However, given the analysis of the reasons for those preferences, our results were not meant to be prescriptive for different situational impairments but rather descriptive of why they may or may not be preferable. In addition, the gestures that the participants performed were restricted to only tapping movements. In light of several works that have explored user-elicited gestures (single-handed microgestures [7] and grasping microgestures [47] for thumb-finger interaction or microgestures for foot-based interaction [14]), the scope of our work limited our exploration of user-elicited gestures for subtle interaction techniques, especially when the primary task required the same body parts. This understanding could have lent new insights into user's relative preferences for subtle interaction techniques.

For both our experiments, participants were instructed to respond to the prompt as soon as they were able to see it. However, in realistic settings, the prompts may not always require an immediate response—for instance, with the SMS based task, the participants may choose to respond to the notification at a later time. This non-urgent interaction could easily give a user time to place down the bags they are holding or a plate of food they are eating in order to respond, lowering the requirement for the interaction to be subtle. Thus, investigating the effect of urgency of a task on the requirement for subtlety of an interaction could have provided further insights into the factors affecting subtle interactions.

To evaluate the potential of using thumb-index-finger interaction as a cross-scenario everyday interaction technique, we conducted an in-the-wild investigation with 8 volunteers. Although we carried the investigation in a wider set of contexts than the wizard-of-oz comparison, we could not extend the evaluation beyond 2 hours due to the battery life of the Vuzix Blade smart-glasses. Aware of such limitations, we caution our readers against generalizing our findings of the in-the-wild deployment. Nonetheless, it helped us to understand the important considerations that are necessary in answering the question of can thumb-index-finger serve as a cross-scenario subtle interaction technique for everyday use. That several participants raised similar concerns on issues of form-factor, comfort, and cross-scenario suitability highlights the challenges for having a thumb-index-finger interaction for mass adoption.

6 CONCLUSION AND FUTURE WORK

We present findings from the first study to evaluate the subtlety of the interaction for OHMDs from users' own perspective by systematically comparing four different subtle interaction techniques in three representative hands-busy contexts. Our results show that despite each interaction having its relative advantage, thumb-index-finger interaction—to the contrary of existing expectations on designing systems for hands-busy tasks—had the best overall cross-scenario preference as an interaction technique for smart glasses. The follow-up in-the-wild investigation helped us further understand the challenges with having a system that enables thumb-index-finger interaction with smart glasses for everyday use and thus enable. These findings provide design recommendations to inspire future subtle interaction techniques for heads-up computing that seamlessly weave into our lives. Future works can

further refine the design of both the hardware and software aspects of thumb-index-finger based subtle interaction techniques to support the everyday usage in heads-up computing for ubiquitous environments.

ACKNOWLEDGMENTS

This project is funded by the NUS Advanced Robotics Centre. We sincerely thank Debjyoti Ghosh for his help with the experiment design, Nuwan Janaka for his help with the Smart Glasses application server, Zhitao Wang and Sheryl Wee for their help with the experiment logistics, and Pallavi Mohan for her early input in the paper draft. We also thank the reviewers for their insightful comments that helped to improve the paper.

REFERENCES

- [1] Fraser Anderson, Tovi Grossman, Daniel Wigdor, and George Fitzmaurice. 2015. Supporting subtlety with deceptive devices and illusory interactions. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. 1489–1498.
- [2] Daniel Ashbrook, Patrick Baudisch, and Sean White. 2011. NENYA: subtle and eyes-free mobile input with a magnetically-tracked finger ring. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 2043–2046.
- [3] Daniel Ashbrook, Carlos Tejada, Dhwanit Mehta, Anthony Jimenez, Goudam Muralitharam, Sangeeta Gajendra, and Ross Tallents. 2016. Bitey: An exploration of tooth click gestures for hands-free user interface control. In *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services*. 158–169.
- [4] Mihai Băce, Teemu Leppänen, David Gil De Gomez, and Argenis Ramirez Gomez. 2016. ubiGaze: ubiquitous augmented reality messaging using gaze gestures. In *SIGGRAPH ASIA 2016 Mobile Graphics and Interactive Applications*. 1–5.
- [5] Jan Bobeth, Susanne Schmehl, Ernst Kruijff, Stephanie Deutsch, and Manfred Tscheligi. 2012. Evaluating performance and acceptance of older adults using freehand gestures for TV menu control. In *Proceedings of the 10th European conference on Interactive tv and video*. 35–44.
- [6] Andreas Bulling, Raimund Dachsel, Andrew Duchowski, Robert Jacob, Sophie Stellmach, and Veronica Sundstedt. 2012. Gaze interaction in the post-WIMP world. In *CHI'12 Extended Abstracts on Human Factors in Computing Systems*. 1221–1224.
- [7] Edwin Chan, Teddy Seyed, Wolfgang Stuerzlinger, Xing-Dong Yang, and Frank Maurer. 2016. User elicitation on single-hand microgestures. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. 3403–3414.
- [8] Enrico Costanza, Samuel A Inverso, and Rebecca Allen. 2005. Toward subtle intimate interfaces for mobile devices using an EMG controller. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 481–489.
- [9] Enrico Costanza, Samuel A Inverso, Rebecca Allen, and Pattie Maes. 2007. Intimate interfaces in action: Assessing the usability and subtlety of EMG-based motionless gestures. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 819–828.
- [10] Artem Dementyev and Joseph A Paradiso. 2014. WristFlex: low-power gesture input with wrist-worn pressure sensors. In *Proceedings of the 27th annual ACM symposium on User interface software and technology*. 161–166.
- [11] Aaron M Dollar. 2014. Classifying human hand use and the activities of daily living. In *The Human Hand as an Inspiration for Robot Hand Development*. Springer, 201–216.
- [12] Sébastien Duval and Hiromichi Hashizume. 2005. Perception of wearable computers for everyday life by the general public: impact of culture and gender on technology. In *International Conference on Embedded And Ubiquitous Computing*. Springer, 826–835.
- [13] Augusto Esteves, Yonghwan Shin, and Ian Oakley. 2020. Comparing selection mechanisms for gaze input techniques in head-mounted displays. *International Journal of Human-Computer Studies* 139 (2020), 102414.
- [14] Mingming Fan, Yizheng Ding, Fang Shen, Yuhui You, and Zhi Yu. 2017. An empirical study of foot gestures for hands-occupied mobile interaction. In *Proceedings of the 2017 ACM International Symposium on Wearable Computers*. 172–173.
- [15] Marie Kondo Folding. [n.d.]. *The Life-Changing Magic of Tidying Up: The Japanese Art of Decluttering and Organizing*. <https://goop.com/food/decorating-design/the-illustrated-guide-to-the-kondo-mari-method/>
- [16] Pascal E Fortin, Elisabeth Sulmont, and Jeremy Cooperstock. 2019. Detecting perception of smartphone notifications using skin conductance responses. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–9.

- [17] Koumei Fukahori, Daisuke Sakamoto, and Takeo Igarashi. 2015. Exploring subtle foot plantar-based gestures with sock-placed pressure sensors. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. 3019–3028.
- [18] Debjyoti Ghosh, Pin Sym Foong, Shengdong Zhao, Can Liu, Nuwan Janaka, and Vinitha Erusu. 2020. Eyeditor: Towards on-the-go heads-up text editing using voice and manual input. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–13.
- [19] Sarthak Ghosh, Hyeon Cheol Kim, Yang Cao, Arne Wessels, Simon T Perrault, and Shengdong Zhao. 2016. Ringinteraction: Coordinated Thumb-index Interaction Using a Ring. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*. 2640–2647.
- [20] Mayank Goel, Chen Zhao, Ruth Vinisha, and Shwetak N Patel. 2015. Tongue-in-cheek: Using wireless signals to enable non-intrusive and flexible facial gestures detection. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. 255–258.
- [21] Ewa Gustafsson, Sara Thomée, Anna Grimby-Ekman, and Mats Hagberg. 2017. Texting on mobile phones and musculoskeletal disorders in young adults: a five-year cohort study. *Applied ergonomics* 58 (2017), 208–214.
- [22] Sandra G Hart and Lowell E Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In *Advances in psychology*. Vol. 52. Elsevier, 139–183.
- [23] Juan David Hincapié-Ramos, Xiang Guo, and Pourang Irani. 2014. The consumed endurance workbook: A tool to assess arm fatigue during mid-air interactions. In *Proceedings of the 2014 companion publication on Designing interactive systems*. 109–112.
- [24] Yi-Ta Hsieh, Antti Jylhä, Valeria Orso, Luciano Gamberini, and Giulio Jacucci. 2016. Designing a willing-to-use-in-public hand gestural interaction technique for smart glasses. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. 4203–4215.
- [25] Armağan Karahanoglu and Çiğdem Erbuğ. 2011. Perceived qualities of smart wearables: determinants of user acceptance. In *Proceedings of the 2011 conference on designing pleasurable products and interfaces*. 1–8.
- [26] Keiko Katsuragawa, Ju Wang, Ziyang Shan, Ningshan Ouyang, Omid Abari, and Daniel Vogel. 2019. Tip-tap: battery-free discrete 2D fingertip input. In *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology*. 1045–1057.
- [27] David M Kietrys, Michael J Gerg, Jonathan Dropkin, and Judith E Gold. 2015. Mobile input device type, texting style and screen size influence upper extremity and trapezius muscle activity, and cervical posture while texting. *Applied ergonomics* 50 (2015), 98–104.
- [28] J Peter Kincaid, Robert P Fishburne Jr, Richard L Rogers, and Brad S Chissom. 1975. *Derivation of new readability formulas (automated readability index, fog count and flesch reading ease formula) for navy enlisted personnel*. Technical Report. Naval Technical Training Command Millington TN Research Branch.
- [29] Marion Koelle, Swamy Ananthanarayan, and Susanne Boll. 2020. Social acceptability in hci: A survey of methods, measures, and design strategies. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–19.
- [30] 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*. 40–49.
- [31] Juyoung Lee, Shaurye Aggarwal, Jason Wu, Thad Starner, and Woontack Woo. 2019. SelfSync: exploring self-synchronous body-based hotword gestures for initiating interaction. In *Proceedings of the 23rd International Symposium on Wearable Computers*. 123–128.
- [32] Lik-Hang Lee and Pan Hui. 2018. Interaction methods for smart glasses: A survey. *IEEE access* 6 (2018), 28712–28732.
- [33] Lik Hang Lee, Kit Yung Lam, Tong Li, Tristan Braud, Xiang Su, and Pan Hui. 2019. Quadmetric optimized thumb-to-finger interaction for force assisted one-handed text entry on mobile headsets. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 3, 3 (2019), 1–27.
- [34] Dawei Liang and Edison Thomaz. 2019. Audio-based activities of daily living (adl) recognition with large-scale acoustic embeddings from online videos. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 3, 1 (2019), 1–18.
- [35] Ming-I Brandon Lin and Yu-Ping Huang. 2017. The impact of walking while using a smartphone on pedestrians' awareness of roadside events. *Accident Analysis & Prevention* 101 (2017), 87–96.
- [36] Aliaksandr Malokin, Giovanni Circella, and Patricia L Mokhtarian. 2015. How do activities conducted while commuting influence mode choice? Testing public transportation advantage and autonomous vehicle scenarios. In *94th annual meeting of the transportation research board*. 11–15.
- [37] Alexander Ng, Stephen A. Brewster, and John H. Williamson. 2014. Investigating the Effects of Encumbrance on One- and Two- Handed Interactions with Mobile Devices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Toronto, Ontario, Canada) (CHI '14). Association for Computing Machinery, New York, NY, USA, 1981–1990. <https://doi.org/10.1145/2556288.2557312>
- [38] Hugo Nicolau and Joaquim Jorge. 2012. Touch typing using thumbs: understanding the effect of mobility and hand posture. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 2683–2686.
- [39] Masa Ogata, Yuta Sugiura, Hirotaka Osawa, and Michita Imai. 2012. iRing: intelligent ring using infrared reflection. In *Proceedings of the 25th annual ACM symposium on User interface software and technology*. 131–136.
- [40] Tadashi Okoshi, Julian Ramos, Hiroki Nozaki, Jin Nakazawa, Anind K Dey, and Hideyuki Tokuda. 2015. Attelia: Reducing user's cognitive load due to interruptive notifications on smart phones. In *2015 IEEE International Conference on Pervasive Computing and Communications (PerCom)*. IEEE, 96–104.
- [41] Jennifer Pearson, Simon Robinson, Matt Jones, Anirudha Joshi, Shashank Ahire, Deepak Sahoo, and Sriram Subramanian. 2017. Chameleon devices: investigating more secure and discreet mobile interactions via active camouflaging. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. 5184–5196.
- [42] Henning Pohl, Andreea Muresan, and Kasper Hornbæk. 2019. Charting subtle interaction in the hci literature. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–15.
- [43] Glaukus Regiani Bueno, Lucas França Garcia, Sonia Maria Marques Gomes Bertolini, and Tiago Franklin Rodrigues Lucena. 2019. The head down generation: musculoskeletal symptoms and the use of smartphones among young university students. *Telemedicine and e-Health* 25, 11 (2019), 1049–1056.
- [44] Jun Rekimoto. 2001. Gesturewrist and gesturepad: Unobtrusive wearable interaction devices. In *Proceedings Fifth International Symposium on Wearable Computers*. IEEE, 21–27.
- [45] Jeremy Scott, David Dearman, Koji Yatani, and Khai N Truong. 2010. Sensing foot gestures from the pocket. In *Proceedings of the 23rd annual ACM symposium on User interface software and technology*. 199–208.
- [46] Andrew Sears, Min Lin, Julie Jacko, and Yan Xiao. 2003. When computers fade: Pervasive computing and situationally-induced impairments and disabilities. In *HCI international*, Vol. 2. 1298–1302.
- [47] Adwait Sharma, Joan Sol Roo, and Jürgen Steimle. 2019. Grasping Microgestures: Eliciting Single-hand Microgestures for Handheld Objects. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–13.
- [48] Dana Slambekova, Reynold Bailey, and Joe Geigel. 2012. Gaze and gesture based object manipulation in virtual worlds. In *Proceedings of the 18th ACM symposium on Virtual reality software and technology*. 203–204.
- [49] Sony.Com. 2021. SmartEyeglass SED-E1. Retrieved February 4, 2021 from <https://developer.sony.com/develop/smarteyeglass-sed-e1/>
- [50] Srinivas Sridharan, Reynold Bailey, Ann McNamara, and Cindy Grimm. 2012. Subtle gaze manipulation for improved mammography training. In *Proceedings of the symposium on eye tracking research and applications*. 75–82.
- [51] Srinivas Sridharan, Brendan John, Darrel Pollard, and Reynold Bailey. 2016. Gaze guidance for improved password recollection. In *Proceedings of the Ninth Biennial ACM Symposium on Eye Tracking Research & Applications*. 237–240.
- [52] 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*. 329–334.
- [53] Ying-Chao Tung, Chun-Yen Hsu, Han-Yu Wang, Silvia Chyow, 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*. 3327–3336.
- [54] Tomás Vega Gálvez, Shardul Sapkota, Alexandru Dancu, and Pattie Maes. 2019. Byte. it: discreet teeth gestures for mobile device interaction. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–6.
- [55] Surbhit Verma, Himanshu Bansal, and Keyur Sorathia. 2015. A study for investigating suitable gesture based selection for gestural user interfaces. In *Proceedings of the 7th International Conference on HCI, IndiaHCI 2015*. 47–55.
- [56] Vuzix.Com. 2021. Vuzix Blade Upgraded Smart Glasses. Retrieved February 4, 2021 from <https://www.vuzix.com/products/blade-smart-glasses-upgraded>
- [57] Florian Wahl, Martin Freund, and Oliver Amft. 2015. Using smart eyeglasses as a wearable game controller. In *Adjunct Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2015 ACM International Symposium on Wearable Computers*. 377–380.
- [58] Colin Ware and Harutune H Mikaelian. 1986. An evaluation of an eye tracker as a device for computer input2. In *Proceedings of the SIGCHI/GI conference on Human factors in computing systems and graphics interface*. 183–188.
- [59] David Way and Joseph Paradiso. 2014. A usability user study concerning free-hand microgesture and wrist-worn sensors. In *2014 11th International Conference on Wearable and Implantable Body Sensor Networks*. IEEE, 138–142.
- [60] Wikipedia.org. 2021. Google Glass. Retrieved February 4, 2021 from https://en.wikipedia.org/wiki/Google_Glass
- [61] Wikitravel. [n.d.]. *Country Description*. <https://wikitravel.org/en/>
- [62] Jacob O. Wobbrock, Leah Findlater, Darren Gergle, and James J. Higgins. 2011. The Aligned Rank Transform for Nonparametric Factorial Analyses Using Only Anova Procedures (CHI '11). Association for Computing Machinery, New York, NY, USA, 143–146. <https://doi.org/10.1145/1978942.1978963>

- [63] Katrin Wolf, Sven Mayer, and Stephan Meyer. 2016. Microgesture detection for remote interaction with mobile devices. In *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct*. 783–790.
- [64] Xing-Dong Yang, Tovi Grossman, Daniel Wigdor, and George Fitzmaurice. 2012. Magic finger: always-available input through finger instrumentation. In *Proceedings of the 25th annual ACM symposium on User interface software and technology*. 147–156.
- [65] Shanhe Yi, Zhengrui Qin, Ed Novak, Yafeng Yin, and Qun Li. 2016. Glassgesture: Exploring head gesture interface of smart glasses. In *IEEE INFOCOM 2016-The 35th Annual IEEE International Conference on Computer Communications*. IEEE, 1–9.
- [66] Qiao Zhang, Shyamnath Gollakota, Ben Taskar, and Raj PN Rao. 2014. Non-intrusive tongue machine interface. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 2555–2558.
- [67] Shengdong Zhao and Ravin Balakrishnan. 2004. Simple vs. compound mark hierarchical marking menus. In *Proceedings of the 17th annual ACM symposium on User interface software and technology*. 33–42.