



A Comparative Study of Pointing Techniques for Eyewear Using a Simulated Pedestrian Environment

Quentin Roy, Camelia Zakaria, Simon Perrault, Mathieu Nancel, Wonjung Kim, Archan Misra, Andy Cockburn

► To cite this version:

Quentin Roy, Camelia Zakaria, Simon Perrault, Mathieu Nancel, Wonjung Kim, et al.. A Comparative Study of Pointing Techniques for Eyewear Using a Simulated Pedestrian Environment. 17th IFIP Conference on Human-Computer Interaction (INTERACT), Sep 2019, Paphos, Cyprus. pp.625-646, 10.31219/osf.io/ktuy4 . hal-02272539

HAL Id: hal-02272539

<https://hal.science/hal-02272539>

Submitted on 27 Aug 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

A Comparative Study of Pointing Techniques for Eyewear Using a Simulated Pedestrian Environment

Quentin Roy^{1,2}, Camelia Zakaria¹, Simon Perrault³, Mathieu Nancel⁴,
Wonjung Kim¹, Archan Misra¹, and Andy Cockburn⁵

¹ Singapore Management University, Singapore

² School of Computer Science, University of Waterloo, Waterloo, Canada

³ Singapore University of Technology and Design (SUTD), Singapore

⁴ Inria, France & University of Lille, UMR 9189 - CRISTAL

⁵ University of Canterbury, Christchurch, New Zealand

quentin@quentinroy.fr, ncamelliaz.2014@phdis.smu.edu.sg,
perrault.simon@gmail.com, mathieu.nancel@inria.fr, kwj1189@gmail.com,
archanm@smu.edu.sg, andrew.cockburn@canterbury.ac.nz

Abstract. Eyewear displays allow users to interact with virtual content displayed over real-world vision, in active situations like standing and walking. Pointing techniques for eyewear displays have been proposed, but their social acceptability, efficiency, and situation awareness remain to be assessed. Using a novel street-walking simulator, we conducted an empirical study of target acquisition while standing and walking under different levels of street crowdedness. We evaluated three phone-based eyewear pointing techniques: indirect touch on a touchscreen, and two in-air techniques using relative device rotations around forward and a downward axes. Direct touch on a phone, without eyewear, was used as a control condition. Results showed that indirect touch was the most efficient and socially acceptable technique, and that in-air pointing was inefficient when walking. Interestingly, the eyewear displays did not improve situation awareness compared to the control condition. We discuss implications for eyewear interaction design.

1 Introduction

Although introduced in the 1960s [54], until recently see-through head-mounted displays (eyewear) were essentially dedicated to military and research environments [35]. Recently, however, there has been substantial commercial interest in developing eyewear technologies for public use, including Microsoft HoloLens or Epson Moverio. While current devices have limitations, like narrow field-of-view, hardware is quickly improving, creating new possibilities for interaction in the office and home as well as during daily activities such as commuting. Indeed, academic and industry experts have suggested that eyewear displays might underpin a foundational change for the next generation of mobile interaction [7,30,38,27].



Fig. 1. Illustrating the simulator: using a phone to point on eyewear while walking.

One potentially important advantage of eyewear displays is that they enable head-up interaction, which may enhance the user’s situational awareness while engaged in concurrent activities, such as walking along a busy sidewalk. Current phones, in contrast, encourage a posture in which the head is bent down rather than looking outwards at the environment, causing poor environmental focus and attention, and raising significant safety concerns [3,40,19,34,26].

Research on input methods for eyewear displays is in its infancy, especially when considering the interplay between the user’s *external* urban activity and their *internal* (eyewear-driven) task. We focus on two-dimensional pointing as a basic interaction for input on eyewear displays. While other interaction modalities are being explored (for example gesture and voice), pointing remains fundamental in most vision-based human-computer interfaces. However, the design and evaluation of pointing techniques for eyewear displays poses particular challenges: their design must be properly adapted to mobility (e.g. efficient interaction while walking), and evaluations should account for environmental factors (e.g. navigating through a crowd), including impact on situation awareness and social acceptability.

As a first step in exploring the design space of eyewear pointing techniques, we examined practical solutions that can be readily adapted and implemented using today’s off-the-shelf devices. We used a phone as an input device, a choice based on versatility (numerous sensors packaged in a small volume, allowing it to simulate handheld trackpads or in-air controllers for example), its ubiquitous ownership (smartglass users are likely to own and carry a phone) and its mobile pragmatics (for example bimanual techniques or bulky apparatus are impractical while walking).

We present an empirical study in which pointing techniques that have been proven useful in comparable contexts (like standing in front of an ultra-wall [39]) are adapted to phone input for eyewear displays—including variants of in-air pointing as well as using the phone as a hand-held trackpad. Direct touch on the phone, without eyewear, was used as control condition. We compare these techniques in three different environments: no simulator (while being stationary), simulated empty street and simulated busy street. Running such an experiment directly in a busy city street would put participants at risk, and is therefore

ethically undesirable. Therefore, inspired by the work of Schwebel [49], we developed a street simulator that enabled us to gain insights on the use of these techniques in pedestrian environments, while keeping our participants safe and preserving internal validity. Finally, we looked at three key metrics: perceived social acceptability, performance as a function of the simulated environment, and impact on situation awareness.

Our results demonstrate that (a) the trackpad technique was the most socially acceptable, most accurate, and fastest for eyewear, and (b) the in-air techniques (which are increasingly integrated in commercial AR products) tended to perform poorly and were subjectively unacceptable. Importantly, while a key expected benefit of eyewear is that their head-up view should improve situation awareness and safety while walking [38,30,27], our results indicate that this may not be true: results indicate that situation awareness was worse when using the candidate techniques (i.e., with eyewear) than in the control condition without eyewear.

We make three specific research contributions:

1. empirical evidence of the perceived social acceptability of pointing techniques for eyewear, collected through interviews and web survey;
2. empirical evidence of the relative performance of eyewear pointing techniques in terms of speed, accuracy and ability to maintain situation awareness (for example avoid simulated pedestrian hazards);
3. demonstration of a Virtual Reality method for safely evaluating interaction techniques in a simulated pedestrian environment.

2 Background and Related Work

Two categories of previous research are briefly reviewed in the following subsections: first, general background research on pointing methods that might be adapted to eyewear displays, and second, research that focuses on interaction while engaged in other activities such as walking a busy street.

2.1 Pointing with Eyewear

Pointing to targets is an elemental component of interaction with graphically displayed content, and it is therefore important that efficient and acceptable pointing methods are developed for eyewear displays. While alternative selection methods that negate the need for pointing have been proposed—such as speech (e.g., [31]) or hand-gestures (e.g., [36])—pointing-based methods offer substantial advantages due to their familiarity and learnability (‘see and point versus learn and remember’ [52]).

Abundant novel or improved pointing techniques are represented in the HCI literature, and many could be adapted for eyewear. A key requirement for pedestrian pointing, however, is that the method can be operated while standing or walking. When eyewear displays are explicitly considered, the most commonly suggested pointing techniques are based on in-air pointing, in which the

movement of the hand or a hand-held object is mapped to cursor movement (e.g., [9,17,22]); a similar in-air pointing method is used with the Microsoft Hololens. Another approach is to use trackpad-like interaction, with a dedicated device held in the hand [39,33], on the body [15,6,5,47,16] or the environment [14,56]. Finally, the use of eye-tracking or head-tracking is also possible [21,39].

All these pointing techniques are promising and most are valid candidates for eyewear pointing (providing the sensing mechanism can be made mobile). Research on large display interaction is also of interest as it often considers the users' need to stand or walk near the display [39,55,28]. Only a few previous works specifically investigated pointing on eyewear. The work of Jalaliniya et al. on head and eye tracking [21], or Hsieh et al. on gloves [17] are examples. However, as far as we know, none of these works comprehensively tackle social acceptability and formally investigate pointing performance under realistic urban movement, or used exotic, unrealistically bulky hardware. As a result, it is still unclear what is currently the best solution.

We focused on one-handed pointing techniques rather than bi-manual techniques because users often carry objects while walking. We used phones as input devices because they are readily available without requiring users to acquire a specialised input device; they also embed sensors that provide both touch (trackpad) and movement sensing (in-air controllers).

2.2 Interaction in Pedestrian Environments

The design of interaction techniques for use in pedestrian environments raises special challenges, including the need for the user to maintain situation awareness during interaction (to reduce safety concerns such as collisions with people or vehicles) and the need for the movements or actions required for interaction to be socially acceptable. In addition, there are also challenges for researchers in evaluating new technologies for pedestrian environments.

Situation awareness Several recent studies have highlighted evidence that the use of mobile phones in urban areas is elevating the risk of personal injury [3,40]. Rather than looking upwards and outwards at the environment, when interacting with a phone, the user's posture has the head bent down, causing poor environmental focus and divided attention [42,48,31]. This leads to the emergence of the "phone zombies" phenomenon: pedestrians who pay insufficient attention to their environment while looking at their phones, sometime walking into other people or traffic [3,40,19,34]. In an attempt to ease these problems, cities such as Singapore and Melbourne have started to install LED strips on pavements at pedestrian crossings [26].

Rather than altering the environment, another approach to improving situation awareness is to alter the interface [31], and to explore interaction mechanisms that are more fit for the challenges of pedestrian environments [30,57,31,43], such as eyes-free interaction [57,20] or the use of eyewear [30,17]. Researchers

have argued that eyewear displays allow more seamless integration between the display of information and the surrounding environment, and as a result, improve situation awareness [38,30,27]. However, there is a lack of empirical study testing this assumption, possibly due to the risks associated with placing experimental participants in congested urban settings.

Social Acceptability Montero et al. define social acceptability as the combination of the *user’s social acceptance*, which defines how comfortable a user is in executing a particular action, as well as *spectators’ social acceptance*, which refers to the impression it makes on witnesses of such action [37]. Interacting with eyewear displays need to be socially acceptable for public performance—this is especially relevant for eyewear given the current scepticism toward such devices [25]. While the social acceptability of actions may change as technologies become widespread [37], the likelihood of technology adoption is greatly improved if its interaction requirements are socially acceptable [12,45]. Several factors are known to influence the social acceptability of actions, including movement duration (the shorter the better) [8], and movement amplitude (small, discreet movements are better) [37,46].

Evaluating interaction techniques for pedestrian environments There are well known trade-offs between lab and field studies, with lab studies facilitating internal validity at the cost of external validity, and field studies the inverse [24,18].

Beyond concerns of internal validity, there are additional and important safety concerns that complicate the potential conduct of field studies in urban pedestrian environments [49]. Consequently, researchers have examined the use of simulations to reproduce some of the realistic interaction context in safe settings. When the research focus is on the act of walking (e.g., to understand motor perturbations to interaction caused by pacing) treadmills have been used [4,2,41]. And when the research focus is on environmental artefacts, video projection [32] and virtual reality [18,23,50,49] can be used. We chose the later approach. Notably, we were inspired by the work of Schwebel et al. who used a simulated street environment to investigate child safety at road crossings [50]. Using a simulated environment not only replicates common pedestrian constraints, but also provides better control of parameters and allows us to put participants in simulated risky situations without physical risks.

3 Perceived Social Acceptability

To reiterate, social acceptability is a key issue in the design of interaction techniques for use in public settings. We structured our investigation of social acceptability in two parts: (1) semi-structured interviews, (2) a large-scale web survey. The goal of the study was to seek participants’ perception of the social acceptability of the investigated interaction techniques. We focused on the user’s social acceptance [37].

Table 1. Techniques included in our social acceptability study

In-Air Techniques	
Front Translation	Translation of the hand on a plane facing the user [39,55]
Down Translation	Translation of the hand on a plane parallel to the ground [22]
Front Rotation	Rotation of the wrist and forearm as if laser-pointing on a plane in front of the user (Fig. 2 left) [39,55]
Down Rotation	Rotation of the wrist and forearm as if laser-pointing on the ground [22]
Front Taps	in-air taps on a plane parallel to the user (back-and-forth movements of the forearm and the index finger) [9]
On-Body Touch Techniques	
Finger Touch	The tip of the index is used as a trackpad controlled with the thumb [5]
Palm Touch	The joined area of the four long fingers is used as a trackpad controlled with the thumb [6]
Pocket Touch	The pocket area is used as a trackpad [47]
On-Device Touch Technique	
Trackpad	A hand-held device is used as a touch surface (Fig. 2 right)

Our interview sessions were inspired by Rico and Brewster’s methodology [45]: participants were asked to perform different gestures as if they were interacting with the device, and we gathered their feedback on the gestures’ social acceptability (for public and private use). All interviews took place in a public setting within a local university campus.

For the web survey, participants watched online videos of the techniques, and were asked to rate their social acceptability. Videos are often used as a way to assess social acceptability [1,45,6,51]. The web survey was included to broaden participation in the study.

In our interview sessions, we examined a set of nine pointing techniques selected from previous literature (c.f. Table 1). Five of them were in-air gestures and one used a hand-held device as a trackpad. Though not the main focus of this work, we also included three body-touch techniques to learn about people’s perception of less-common input methods.

After the interview sessions, we discarded the techniques with very poor rankings and kept Front Rotation, Down Rotation, Trackpad, Finger Touch and Pocket Touch for the web survey. Palm Touch was excluded because it was comparable to Trackpad. Front Rotation was slightly modified to allow movements from both the elbow and the wrist. Pocket Touch, which received polarized feedback in our interviews, was also modified so that the control area was shifted to the side of the thigh, further away from the genitals.

3.1 Participants and Procedure

For the interview sessions, we recruited eight participants (5 female), aged 22 to 45 years old ($M = 32.1$, $SD = 6.5$) from our students and university staff. Seven participants lived in Singapore and one in France. No compensation was offered. Each session lasted 45 minutes.

For each of the nine pointing techniques, we carried out the following procedure: (1) demonstrated the pointing movements for that technique and made



Fig. 2. Pointing techniques explored during our experiment: Front Rotation (left), Relaxed Rotation (middle) and trackpad (right).

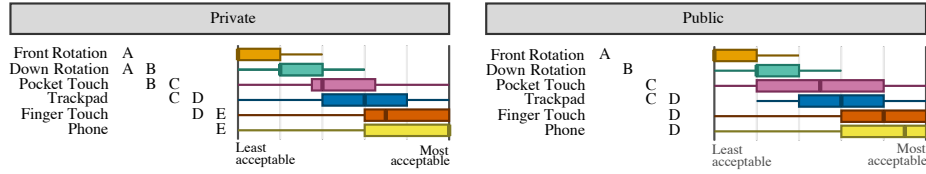


Fig. 3. Web survey: mean postulated social acceptability ranking for each technique. Levels not connected by the same letter are significantly different.

sure its principles were understood, (2) asked the participant to perform the movements for approximately 30 seconds in a busy public area of our university campuses, (3) conducted a semi-structured interview focusing on their perception of the social acceptability of the techniques, at home or in the street. We finished the session by asking our participants to rank all techniques by order of social acceptability for usage in a private setting (e.g. home) and public setting (e.g. a street).

From our web survey, we gathered 56 responses (25 female, 1 preferred not to disclose) from 18 to 57 years old ($M = 27.7$, $SD = 8.2$, 7 preferred not to disclose). The web survey was advertised using our university’s mailing lists. 50% of the participants were students (undergraduate and post-grads), 26% were IT Professionals and 5% in Academia and Research. They were mostly from South-East Asia ($n = 41$) and Europe ($n = 8$). The survey was divided into six parts, one dedicated to each TECHNIQUE, and a summary. In each survey part, participants were shown a short video of an actor walking in a street and demonstrating the use of the technique. Then they were asked to provide feedback on the perceived social acceptability of these techniques. Finally, they ranked techniques in order of perceived acceptability both for public and private contexts (1, most acceptable; 6, least acceptable).

3.2 Results and discussion

During the interview sessions and in the web survey, we asked participants about their perception of the social acceptability of the techniques in private and in public. We did not observe any statistically significant effect of the participants’ continent of origin on the recorded answers.

Private Use In the interviews, Finger Touch was ranked as being the most socially acceptable technique ($M = 1.8/9$), followed by Phone Touch ($M = 3.3/9$), Pocket Touch and Palm Touch (both $M = 3.8/9$). Among the in-air technique, Front Taps was ranked as the least socially acceptable ($M = 4.6/9$). A Friedman test showed a significant effect of TECHNIQUE on average ranking ($\chi^2(8) = 37.3$, $p < .00001$), although Bonferroni corrected analysis showed no pairwise differences.

In general, the interview results suggested that for private use, smaller on-body or on-device actions were perceived as more socially acceptable than larger in-air movements of the device. On-body movements were perceived to be less tiring (2 interview participants) and as a result easier to use in private, and on-device actions were reported as being familiar (5). The in-air techniques were considered as “tiresome” (4) and “intrusive” (5). Our web survey results tended to confirm this trend: we found a significant main effect of TECHNIQUE on the average ranking for private use ($\chi^2(5) = 108.3$, $p < .0001$, see Fig. 3 for post-hoc comparisons).

Public Use Results for use in public spaces reflected those for private use. Among the on-body techniques, Finger Touch was ranked as being the most socially acceptable ($M = 1.4/9$) followed by Trackpad ($M = 2.9/9$), Palm Touch ($M = 3.4/9$) and Pocket Touch ($M = 5.0/9$). Among in-air techniques Down Rotation was ranked best ($M = 4.6/9$) followed by Front Rotation ($M = 4.6/9$), Down Translation ($M = 6.5/9$), Front Taps ($M = 7.5/9$) and Front Translation ($M = 8.4/9$). A Friedman test showed a significant main effect of TECHNIQUE on average ranking ($\chi^2(8) = 43.2$, $p < .00001$).

Consistent with previous work [37,46], participants expressed concerns with high amplitude movements. In particular, five interview participants reported that Front Translation exceeded their “personal space”, and got “in the way of others”. Participants also expressed strong concerns on Front Taps [9] that made them appear as though they were pointing at others, explaining the large ranking difference compared to a private setting. This is potentially important, as contemporary implementations such as Microsoft Hololens use this modality as a primary means for interaction. Finally, Pocket Touch was polarizing in our interviews: half of our participants expressed little social concern, the other half strongly opposed to what they perceived as sexually suggestive (2 participants, 1 male and 1 female, even entirely refused to perform the gesture in public as per protocol).

As suggested by one of the interview participants, in the web survey videos we moved the control area for Pocket Touch further to the outside thigh region. This improved the ranking of the technique compared to our interviews. The rest of our web survey results tend to confirm the trend observed during the interviews: we found a main effect of TECHNIQUE on the average ranking for public use ($\chi^2(5) = 119.7$, $p < .0001$, see Fig. 3).

4 Performance and situation awareness

We explored pointing techniques enabled by everyday devices and usable by pedestrians. We compared the three techniques presented in Fig. 2: Front Rotation, Down Rotation and Trackpad. Front Rotation requires positioning the phone flat (screen up), then pressing and holding the screen while rotating the wrist and forearm left, right, up, or down to move the cursor⁶, not unlike tilt techniques [44]. Trackpad requires sliding the thumb on the screen to move the cursor. Relaxed Rotation requires holding the phone sideways while keeping the arm down in a relaxed position; the cursor is then moved by pressing and holding on screen while rotating the wrist left, right, up, or down. We designed Relaxed Rotation to require movements comparable in amplitude to Down Rotation. As a result, we believe that it should be perceived to have similar social acceptability (Down Rotation was perceived as the most socially acceptable in-air technique in the previous study). In all three techniques, target acquisition could be performed either by tapping on the screen or pressing one of the volume buttons. In practice, and due to the different grasps, the volume buttons were only used with Relaxed Rotation.

We included direct touch pointing on the phone display as a control condition, with participants instructed to hold and interact with the phone using one hand. Eyewear was disabled and removed in this condition, so participants had to look down while acquiring the targets.

Except for the control condition, all techniques made use of a mobile phone as an indirect, eyes-free controller. The controller was always manipulated with only one hand because pedestrians often need their other hand for activities such as opening doors, carrying bags, etc. In all techniques but direct touch (control), the visual feedback was exclusively displayed on the smartglasses.

Our social acceptability study included several on-body techniques that we did not include in this experiment because we wanted to focus on currently pragmatic phone-based techniques. Furthermore, our pre-tests and pilots indicated that the Down Rotation technique was excessively hard to control, so we eliminated it. In-air translation-based techniques and Front Taps were also excluded due to their poor social-acceptability findings in the previous study.

We compared the remaining techniques under three different environments: No Simulator in which participants stood while performing pointing task; Empty Street, where participants walked in an empty street simulation with no red lights or pedestrians; and Crowded Street, where participants walked in a street simulation including traffic lights and pedestrians (Fig. 1).

We formulated the following hypotheses:

H₁: Users achieve *the fastest pointing with Phone* because of their familiarity with traditional direct-touch pointing,

⁶ Our initial design used press-and-hold for clutching: the cursor was moving by default, but users could hold the screen to freeze it and reposition themselves. However, our pre-tests quickly revealed that it was counter-intuitive.

- H₂:** Users achieve *the lowest walking speed and poorest situational awareness with Phone* because they are required to look down (at the phone),
- H₃:** In the two street environments, users achieve *faster pointing with Trackpad than with Front Rotation and Relaxed Rotation*, because they are accustomed to trackpads and because the technique’s input is arguably less sensitive to walking movements,
- H₄:** Users achieve *the highest walking speed and best situational awareness using Trackpad* because they are not required to look down, and Trackpad’s input is arguably less sensitive to walking movements.

4.1 Street Simulation

Exploring safety or situation awareness in the wild implies putting participants at risk (e.g., within close vicinity to vehicle traffic), which is not ethically acceptable. Instead, inspired by previous works in social science [49], we rely on a street simulator (see Fig. 1) to investigate the ability of users to maintain situational awareness while interacting with the eyewear device. Participants stood in front of a wide display, and their body movements were tracked using fiducial markers. Walking on the spot caused the camera to move forward at a speed that the participant could control (treadmills, often used in previous works [4,2,41], do not allow pace control). Participants had to step sideways to avoid incoming pedestrians in the Crowded environment, and stop at red lights.

As realistic as it is, a simulation cannot be as externally valid as an in-the-wild experiment. The generalizability of our findings to real street scenarios remains for further work. Nevertheless, the method does require participants to remain aware of the situation and as a result provides actionable insights on situation awareness.

Street Elements Several factors influence a pedestrian’s walking behavior, such as street layout, illumination, and other pedestrians. Previous work in social sciences have focused on distracted behavior in road-crossings [49,50,3]. However, Oulasvirta et al. observed that the most attention-taxing situations encountered by pedestrians are when they walk in busy streets [42]. After discussion and further observations of our own behavior in the street, we included incoming pedestrians and changing traffic lights.

We used a simple street layout: a series of blocks with the same length and walkway width, not unlike some North-American cities. We designed these blocks to appear shorter (in length) than usual, to increase the number of intersections encountered by the participants.

Layout and Traffic Lights Each street block was separated by a crosswalk and a traffic light. Traffic lights could have four different behaviors: Fixed Green, Fixed Red, Changing Green and Changing Red. Fixed Green remained green, Changing Green and Changing Red changed from one to the other when participants were 0.016 to 0.039 blocks away. Changing Green and Fixed Red switched



Fig. 4. Adaptation of the vanishing point. No adaptation (left) vs adaptation (right).

to green after a wait time of 1 to 2.5 seconds. The ordering of Light behaviors were randomized, but we ensured that each behavior appeared at least once every four lights. Audio feedback of a car honk was played if participants jaywalked.

Pedestrian Behavior Simulated pedestrians walked towards the participants at a speed randomly assigned between 2.46 and 3.78 blocks per minutes. They walked in straight lines, stopped to avoid “bumping” into participants, and respected traffic lights. Audio feedback of a pedestrian shouting “hey!” was also played if participants collided into them. In the Crowded Street condition, the street contained approximately 8 pedestrians per block (see companion video).

Steps and Position Tracking Fiducial markers [10] were attached on the participants’ ankles to track stomping motions, as well as the participant’s lateral position in front of the display. Our tracking algorithm enabled us to map the participants’ simulated walking speed as a function of both their stomping pace and the vertical amplitude of their steps.

Vanishing Point Adaptation The vanishing point of the scene was kept aligned with the participant’s position in front of the display when they stepped sideways, as opposed to constantly fixed at the center of the display, to further support the realism and immersiveness of the simulation (see video figure).

4.2 Participants and Apparatus

Twelve right-handed participants were recruited from Singapore Management University’s students and staff (7 female) aged 20 to 30 years old. Remuneration was the equivalent of 7.4 USD. All but one reported that they had used their phone while walking in the street at least once in the two days before the experiment.

The experimental software was run on an Epson Moverio BT-300 smart-glass, a Samsung S7 Edge smart-phone and two computers (one for the simulation, one for the devices). The simulation was run on a large TV monitor (75 inches diagonal, 1.65×0.93 meters) positioned 1 meter from the ground. Participants stood 1 meter from the display, and could move left and right in front of it.

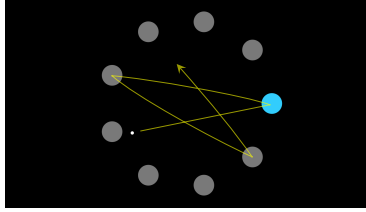


Fig. 5. Pointing task interface used during the experiment (black is transparent on the glasses). Each time a participant validates a target (currently the rightmost disk on the figure), a new one is highlighted until completion of the task. The superimposed arrow indicates the path to alternating targets following ISO 9241–9 standard procedure [53].

4.3 Task

Participants were instructed to perform an ISO 9241–9 standard Fitts multi-directional pointing task as established by Soukoreff et al. [53] (see Fig. 5) using one of the four techniques (see Fig. 2). We chose a Fitts’ Law task type for internal validity: controlling pointing distance and size simplifies comparison between techniques and with previous and future work. Except for the Phone condition, the display area on the eyewear appeared to be approximately 199.2×199.2 mm one meter away from the user (720×720 px), the radius of the targets layout (see Fig. 5) was 84.7 mm (306 px) and the radius of the targets was 10.5 mm (38 px). In the Phone condition, the total pointing area was 65.4×65.4 mm ($1,376 \times 1,376$ px), the radius of the target layout was 27.9 mm (585 px) and the radius of the targets was 3.5 mm (73 px). In both cases, the ratio of the layout radius on the target radius remains constant ($\frac{306}{38} = \frac{585}{73} = 8$), yielding the same Index of Difficulty [11]. The extra space around the targets discouraged the use of edge pointing.

All three eyewear techniques (Front Rotation, Relaxed Rotation, and Trackpad) were indirect and relative. We defined transfer functions to map participant’s input to cursor movements using Nancel et al.’s sigmoid function [39]:

$$G_t = G_{\min} + \frac{G_{\max} - G_{\min}}{1 + e^{-\lambda(v_t - V_{\inf})}}$$

with v_t and G_t respectively the input speed and gain at time t and λ a constant. Reporting generalizable (typically, physical) display units is complicated with smart glasses: the pixels can be perceived as if they were at any distance from the user’s eyes. Furthermore, since the virtual display is displayed as a flat surface facing the user, rather than a spherical one, angular units cannot be used consistently. For simplicity and generalizability, we report distance and speeds on the display as if the display was projected one meter away from the user’s eyes. The 1280×720 pixel map of the Moverio BT-300 corresponds to a 354×199 mm area one meter away, so one pixel is 0.277 mm wide⁷. We tuned its parameters separately for each technique (see Table 2).

⁷ Or 40 inches, 2.5 meters away according to the manufacturer.

Table 2. Parameters of the gain functions

Technique	$G_{\min}$	$G_{\max}$	λ	$V_{\inf}$
Trackpad	6.28 mm/mm	18.84 mm/mm	0.113 s/mm	26.46 mm/s
Front Rot.	0.48 mm/deg	2.9 mm/deg	10.47 ms/deg	143.24 deg/s
Relaxed Rot.	0.24 mm/deg	3.38 mm/deg	10.47 ms/deg	137.51 deg/s

In the No Simulator condition, participants executed the task while standing. In both street conditions, participants completed the task while navigating through the simulated street. Specifically in the Empty Street condition, participants were only required to walk on the spot. They were instructed to strive to maintain a natural walking speed. In the Crowded Street condition, they were asked to also avoid pedestrians and respect traffic rules, as in real world. If they failed to meet these rules, the simulation flashed red while the corresponding audio feedback was played (a man shouting “hey!” for pedestrians, and a car honk for the lights). We simplified this experiment by standardizing the walk to a straight path, under the hypothesis that it is reasonably similar to following a well-known path in term of cognitive load. This method was designed to simulate the most common external constraints encountered by pedestrians while enabling measures of pace, awareness, and interaction performance, as realistically as possible.

We finished the experiment with a short semi-structured interview during which participants provided subjective feedback. Before starting the experiment and as a training, participants were introduced to both conditions of the street simulator. During this training, we also asked participants to find what they thought was their usual walking speed and we recorded it for later comparison. Before each technique and under each environment, participants also had the opportunity to train themselves with the technique before starting.

4.4 Design

We used a 3×4 within-subjects design with the following factors and levels: ENVIRONMENT {No Simulator, Crowded Street, Empty Street} and TECHNIQUE {Front Rotation, Relaxed Rotation, Trackpad, Phone}. To ensure consistency of the simulation across all participants, we generated four predefined crowded street configurations (including pedestrian position and speed, lights, etc.).

The experiment was divided into three parts, one for each ENVIRONMENT condition. Each of these parts were divided into four blocks, each one dedicated to a technique. In accordance with ISO 9241–9 [53], each block started with the cursor centered and an initial unmeasured target selection, followed by four selections of each of the 9 targets in opposing order (see Fig. 5). ENVIRONMENT and TECHNIQUE orders were counterbalanced using a Latin Square. In the street conditions, the simulation ran uninterrupted during a block, and restarted afterwards. For all trials, we measured selection time, wrong selections (clicks outside of the target), “walking” speed, number of pedestrian collisions, and jaywalking. Participants were allowed to take breaks between each

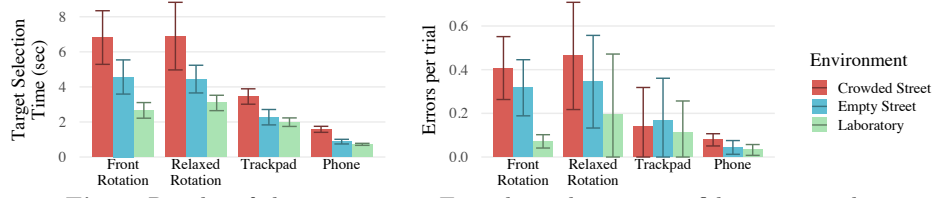


Fig. 6. Results of the experiment. Error bars show .95 confidence intervals.

block. We recorded 12 participants $\times$ 4 techniques $\times$ 3 environments $\times$ 9 targets $\times$ 4 repetitions = 5,184 trials. The experiment lasted one hour per participant.

4.5 Results

We ran two-way ANOVAs with two within-subjects factors (TECHNIQUE and ENVIRONMENT) on *Selection Time*, *Selection Error*, *Walking Speed*, pedestrian collisions, and jaywalking. We applied Greenhouse-Geiser sphericity correction when needed, with adjusted p -values and degrees of freedom. We ran pairwise t -tests with Bonferroni correction applied to the p -values of post-hoc tests.

Selection Time As per ISO 9241-9 [53], Selection Time was measured as the time between two target selections (first excluded). We observed a significant main effect of TECHNIQUE on *Selection Time* ($F_{2,07,22.8} = 61.7$, $p < .0001$). Post-hoc comparisons show significant differences between all pairs (all $p < .0001$), except Relaxed Rotation $\times$ Front Rotation (“ $\ll$ ” indicates significant differences):

$$\text{Phone } 1.06 \text{ s} \ll \text{Trackpad } 2.58 \text{ s} \ll \text{Front Rot. } 4.69 \text{ s}, \text{ Relaxed Rot. } 4.8 \text{ s}$$

ENVIRONMENT also had an effect on *Selection Time* ($F_{2,22} = 42.2$, $p < .0001$). We found significant differences between all pairs (all $p < .0001$):

$$\text{No Simulator } 2.11 \text{ s} \ll \text{Empty Street } 3.04 \text{ s} \ll \text{Crowded Street } 4.7 \text{ s}$$

We also found a significant TECHNIQUE $\times$ ENVIRONMENT interaction ($F_{2,38,26.18} = 6.5$, $p < .01$). Fig. 6-top summarizes the averaged times across conditions.

Selection Errors We observed significant main effects of both TECHNIQUE ($F_{2,0,22.09} = 10.5$, $p < .001$) and ENVIRONMENT ($F_{2,22} = 4.4$, $p = .024$) on *Selection Errors*. Pairwise t -tests showed significant differences between all pairs of Techniques except between Relaxed Rotation and Front Rotation:

$$\text{Phone } 0.05 \ll \text{Trackpad } 0.14 \ll \text{Front Rot. } 0.26, \text{ Relaxed Rot. } 0.33$$

We found a significant difference between the No Simulator ($M = 0.1 \text{ errors/trial}$) and Crowded Street (0.22) conditions ($p < .001$). Fig. 6 illustrates the results.

Walking Speed We draw comparisons of walking speed between techniques and between the two street conditions (Empty Street and Crowded Street). We measured the time it took each participant to walk through an entire block whilst

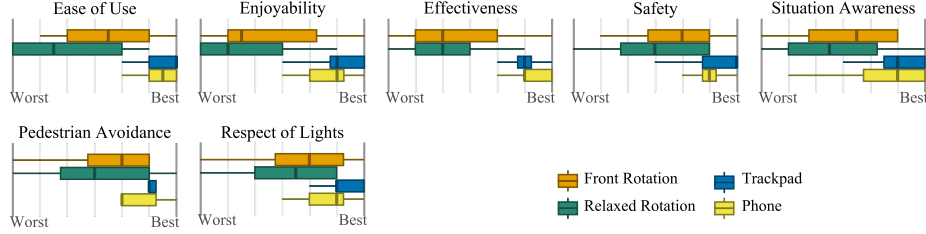


Fig. 7. Subjective Assessments of the techniques.

in the simulation. We found a significant main effect of Street Environment ($F_{1,11} = 6.69$, $p < .05$), TECHNIQUE ($F_{3,33} = 17.74$, $p < .0001$), as well a ENVIRONMENT $\times$ TECHNIQUE interaction ($F_{3,33} = 6.93$, $p < .001$) on the *Walking speed*. Post-hoc comparison showed significant differences between two groups of techniques ($p < .01$, $^{sec/block}$):

$$\begin{array}{ccccccc} \text{Phone} & \text{Trackpad} & & \text{Front Rot.} & \text{Relaxed Rot.} \\ 20.9 & 21.4 & \ll & 28.1 & 29.1 \end{array}$$

Jaywalking and Collisions We only considered the Crowded Street scenario when investigating jaywalking and pedestrian collisions. Note that because of the differences in selection time (Fig. 6-top), we computed the average number of jaywalkings and pedestrian-collisions per block (a block is a unit of distance in the simulator). We did not found a significant effect of TECHNIQUE on jaywalking ($p = .13$, with 0.1 jaywalk/block on average), nor on collisions ($p = .49$, 0.659 collisions/block on average).

Subjective Assessments We asked participants to rate the techniques on a 7-level Likert-scale in term of (1) ease of use, (2) enjoyability, (3) effectiveness, (4) safety, (5) situation awareness, (6) ease to avoid pedestrians and (7) respect of traffic lights. A Friedman’s test showed a significant main effect of TECHNIQUE on all questions: ease of use ($\chi^2(3) = 23.4$, $p < .0001$), enjoyability ($\chi^2(3) = 18.5$, $p < .001$), effectiveness ($\chi^2(3) = 22.2$, $p < .001$), safety ($\chi^2(3) = 14.9$, $p < .01$), situation awareness ($\chi^2(3) = 12.2$, $p < .01$), ease to avoid pedestrians ($\chi^2(3) = 11.9$, $p < .01$) and respect of traffic lights ($\chi^2(3) = 9.9$, $p = .019$). Figure 7 shows these results.

At the end of the experiment, participants also ranked the techniques from most-preferred to least-preferred, specifically in a crowded environment. Five participants ranked Trackpad best, while five others ranked it second-best. Accordingly, five participants preferred the Phone first, one participant second-best, and another one third-best. One participant ranked Phone worst. Front Rotation was ranked second by three participants, while Relaxed Rotation was consistently ranked either third-best or least-preferred. Many participants felt that the Trackpad technique poses less danger on the streets, in comparison to frequent look-ups while interacting with the phone.

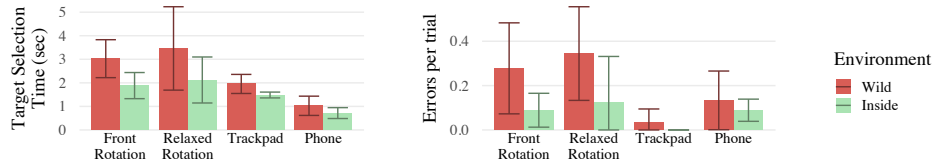


Fig. 8. Results of the ecological validity experiment. Error bars show .95 confidence intervals.

4.6 Ecological Validity Experiment

As an extra validation step, we ran an ecological validity experiment to challenge our simulation-bound results in a real street situation. For safety reasons, it was not ethically acceptable to use external participants. Four of the authors took part in the experiment to test the four pointing techniques in the (actual) wild. We used the same techniques and protocol, with two ENVIRONMENT conditions: Wild and Inside, and 5 repetitions of each pointing target instead of 4 during the controlled experiment. In the Wild condition, the authors walked along a busy underground concourse. We measured an average $22.9 \text{ pedestrians/minute}$ at this location and time, with a large variance. In the Inside condition, the authors performed the pointing tasks standing but without walking. We counterbalanced the order of the techniques using a Latin Square, and measured both selection times and selection errors.

Due to the small population, we only report descriptive statistics, shown in Fig 8. None of the four authors collided with a pedestrian. The results of this experiment show the same trend as our main experiment. Participants were generally faster and more accurate, which can be explained by a higher expertise with the techniques and by a lower pedestrian density.

4.7 Limitations

Techniques Some participants reported difficulties with Relaxed Rotation due to the width of the phone: it made it difficult to press-and-hold or click. Relaxed Rotation may perform differently with a more ergonomically adapted in-air controller. Two participants also reported visual fatigue. Hopefully, this issue can be resolved on future eyewear displays.

Generalizability Our street-walking simulator allows us to put participants in controlled situations resembling crowded streets. This novel methodology enables to gather preliminary insights on situation awareness without putting participants at risk. Concerns over generalizability to real-world situations are eased through our limited validity experiment, but further experimental validation is difficult due to risks of participant harm. When Schwebel et al. run into a similar problem, they argue that at least three indicators can still be considered: immersion, interactiveness, and realism [49].

Though not as immersive as a virtual-reality “cave”, we join Schwebel et al. in the argument that a large display area covering most of the participant’s field

of view provides sufficient immersion. At least four of our participants agreed and described our simulation as “immersive”.

We think interactiveness is a strength of our simulator as it included pace and position control (using feet tracking) instead of less interactive controller like a joystick, or pure absence of control as when using treadmills.

On the realism side, our participants’ opinions were more divided: three reported the simulation as “unrealistic” and four stated that it was “realistic”. Criticisms mostly concerned the in-place stepping mechanism, although two participants described the speed control as being “natural”. This is a trade-off for the interactiveness required by our experiment. When VR treadmills are commercialized, allowing pace control, they may provide a better alternative.

Allowing participants to change direction, like walking around a corner, would allow more complex paths to be walked and add interesting factors. Similarly, half of our participants observed that real-world pedestrians typically give way when a collision is about to occur (rather than stop in front of the participant in our simulation). This behavior can easily be added, though we needed participants to actively avoid the pedestrians.

Less trivially, the stakes of colliding into a pedestrian or jaywalking remain limited compared to real-life. Keeping participants safe is of course the main point of using a simulation, but recent approaches such as force-feedback [29] could be put to use to produce physical sensations at no cost of safety.

5 Discussion

Participants were able to point faster and with fewer errors using the Phone technique. We therefore find support for $\mathbf{H}_1$. Contrary to our expectations, the better performance of Phone did not come at the cost of slower walking speed or worse situation awareness. Therefore, we reject $\mathbf{H}_2$ and $\mathbf{H}_4$. This can be explained by a strong discrepancy in participants’ prior experience between the control and candidate techniques.

Trackpad emerged as the best pointing technique for eyewear in terms of speed and error rate, in both the No Simulator and our street simulations. Therefore, we find support for $\mathbf{H}_3$. Trackpad was also perceived as more socially acceptable, easier to use and more enjoyable than the in-air techniques. These results could be influenced by the smaller movements or lower cognitive burden associated with highly familiar touch interaction.

The in-air techniques, Front Rotation and Relaxed Rotation, performed worse in every condition. Performance with these techniques was also more adversely affected by street crowding than the other techniques (selection times increased dramatically in the Street conditions). Despite the higher movement amplitude and the need to keep the forearm up, Front Rotation was found easier to use than Relaxed Rotation. This might be due to the additional joint involved (wrist + elbow vs. wrist only) [13]. We observe significant differences in selection time for all environments, and increased errors in the Crowded Street compared to

the No Simulator condition. The participants were also able to walk faster in Empty Street than in Crowded Street.

Interestingly, Trackpad and Phone performance were not significantly affected by the simulated Empty Street condition and were close to the No Simulator's. The two techniques did not significantly differ in term of situation awareness with our simulator. The two in-air techniques suffered substantially more from the simulator.

Though these results could only be safely obtained using a simulation, we believe they provide valuable insights sufficient to reliably recommend the use of the Trackpad for eyewear pointing by pedestrians (provided that it is implemented with an efficient transfer function⁸).

Contrarily to previous assumptions [31,38,30,27] we were surprised that the use of eyewear did not improve situation awareness in our simulator; indeed, situation awareness with eyewear was worse than regular phone interaction with harder-to-use input techniques. User feedback was divided: five participants reported that it was easier to deal with divided attention using the smartglasses, while four others stated the opposite. One obvious caveat on this, however, is that people are highly familiar with current touchscreen interaction, and our participants' performance with eyewear conditions might improve with familiarization. In terms of pure input performance, as expected, Phone was superior.

6 Conclusion

This work contributed the first empirical study of eyewear pointing while mobile, taking into account environmental awareness and perceived social acceptability. In our street simulations or in a quiet building, participants were faster using a hand-held trackpad than with every investigated variant of in-air techniques. The trackpad was also perceived as the most socially acceptable technique.

Research on eyewear for pedestrians is still in its infancy. Our results indicate muted benefits regarding situation awareness. However, we remain confident that eyewear might expose situation awareness advantages, in particular for more passive tasks such as reading, and we plan to explore it as future work. Also, we would like to explore body-based techniques, like Finger- and Pocket Touch as they were deemed highly acceptable. But, there are also many other promising paths to investigate, as outlined in our related works.

We believe that our simulation-based method is promising, particularly given the ethical concerns associated with evaluation in-the-wild. While the generalizability of our simulation cannot be fully assessed, we argue it provides valuable insights on situation awareness. However, the generalizability of other aspects, such as path finding, might be easier to investigate in future work. Simulations will never be as externally valid as in-the-wild studies, but they require less resources and allows much greater control. Compared to other form of lab studies, we argue they do can provide opportune improvements of external validity.

⁸ During initial tests, the trackpad shipped with the Epson Moverio glasses proved to be particularly cumbersome to use (in contrast with our transfer function).

References

1. Bailly, G., Müller, J., Rohs, M., Wigdor, D., Kratz, S.: ShoeSense: A New Perspective on Gestural Interaction and Wearable Applications. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. pp. 1239–1248. ACM (2012). <https://doi.org/10.1145/2207676.2208576>
2. Barnard, L., Yi, J.S., Jacko, J.A., Sears, A.: An Empirical Comparison of Use-in-motion Evaluation Scenarios for Mobile Computing Devices. *Int. J. Hum.-Comput. Stud.* **62**(4), 487–520 (apr 2005). <https://doi.org/10.1016/j.ijhcs.2004.12.002>
3. Basch, C.H., Ethan, D., Zybert, P., Basch, C.E.: Pedestrian Behavior at Five Dangerous and Busy Manhattan Intersections. *Journal of Community Health* **40**(4), 789–792 (aug 2015). <https://doi.org/10.1007/s10900-015-0001-9>
4. Bergstrom-Lehtovirta, J., Oulasvirta, A., Brewster, S.: The Effects of Walking Speed on Target Acquisition on a Touchscreen Interface. In: Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services. pp. 143–146. ACM (2011). <https://doi.org/10.1145/2037373.2037396>
5. Chan, L., Liang, R.H., Tsai, M.C., Cheng, K.Y., Su, C.H., Chen, M.Y., Cheng, W.H., Chen, B.Y.: FingerPad: Private and Subtle Interaction Using Fingertips. In: Proceedings of the 26th Annual ACM Symposium on User Interface Software and Technology. pp. 255–260. ACM (2013). <https://doi.org/10.1145/2501988.2502016>
6. Chen, K.Y., Lyons, K., White, S., Patel, S.: uTrack: 3D Input Using Two Magnetic Sensors. In: Proceedings of the 26th Annual ACM Symposium on User Interface Software and Technology. pp. 237–244. ACM (2013). <https://doi.org/10.1145/2501988.2502035>
7. Costanza, E., Inverso, S.A., Allen, R.: Toward Subtle Intimate Interfaces for Mobile Devices Using an EMG Controller. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. pp. 481–489. ACM (2005). <https://doi.org/10.1145/1054972.1055039>
8. Dobbstein, D., Hock, P., Rukzio, E.: Belt: An Unobtrusive Touch Input Device for Head-worn Displays. In: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems. pp. 2135–2138. ACM (2015). <https://doi.org/10.1145/2702123.2702450>
9. Ens, B., Byagowi, A., Han, T., Hincapié-Ramos, J.D., Irani, P.: Combining Ring Input with Hand Tracking for Precise, Natural Interaction with Spatial Analytic Interfaces. In: Proceedings of the 2016 Symposium on Spatial User Interaction. pp. 99–102. ACM (2016). <https://doi.org/10.1145/2983310.2985757>
10. Fiala, M.: Designing Highly Reliable Fiducial Markers. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **32**(7), 1317–1324 (jul 2010). <https://doi.org/10.1109/TPAMI.2009.146>
11. Fitts, P.M.: The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology* **74**, 381–391 (1954). <https://doi.org/10.1037/h0055392>
12. Goffman, E.: The presentation of self in everyday life. 1959. Garden City, NY (2002)
13. Guiard, Y.: The Kinematic Chain as a Model for Human Asymmetrical Bimanual Cooperation. In: Colley, A.M., Beech, J.R.B.T.A.i.P. (eds.) *Cognition and Action in Skilled Behaviour*, vol. 55, pp. 205–228. North-Holland (1988). [https://doi.org/https://doi.org/10.1016/S0166-4115\(08\)60623-8](https://doi.org/https://doi.org/10.1016/S0166-4115(08)60623-8)
14. Harrison, C., Benko, H., Wilson, A.D.: OmniTouch: Wearable Multitouch Interaction Everywhere. In: Proceedings of the 24th Annual ACM Sympo-

- sium on User Interface Software and Technology. pp. 441–450. ACM (2011). <https://doi.org/10.1145/2047196.2047255>
15. Harrison, C., Tan, D., Morris, D.: Skinput: Appropriating the Body As an Input Surface. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. pp. 453–462. ACM (2010). <https://doi.org/10.1145/1753326.1753394>
 16. Holleis, P., Schmidt, A., Paasovaara, S., Puikkonen, A., Häkkinä, J.: Evaluating Capacitive Touch Input on Clothes. In: Proceedings of the 10th International Conference on Human Computer Interaction with Mobile Devices and Services. pp. 81–90. ACM (2008). <https://doi.org/10.1145/1409240.1409250>
 17. Hsieh, Y.T., Jylhä, A., Orso, V., Gamberini, L., Jacucci, G.: 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. pp. 4203–4215. ACM (2016). <https://doi.org/10.1145/2858036.2858436>
 18. Hühn, A.E., Khan, V.J., Lucero, A., Ketelaar, P.: On the Use of Virtual Environments for the Evaluation of Location-based Applications. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. pp. 2569–2578. ACM (2012). <https://doi.org/10.1145/2207676.2208646>
 19. IPSOS: Distracted Walking Study: Topline Summary Findings. american academy of orthopedic surgeons pp. 1–6 (2015)
 20. Jain, M., Balakrishnan, R.: User Learning and Performance with Bezel Menus. In: CHI '12. pp. 2221–2230. ACM (2012). <https://doi.org/10.1145/2207676.2208376>
 21. Jalaliniya, S., Mardanbeigi, D., Pederson, T., Hansen, D.W.: Head and Eye Movement as Pointing Modalities for Eyewear Computers. In: 2014 11th International Conference on Wearable and Implantable Body Sensor Networks Workshops. pp. 50–53 (jun 2014). <https://doi.org/10.1109/BSN.Workshops.2014.14>
 22. Katsuragawa, K., Pietroszek, K., Wallace, J.R., Lank, E.: Watchpoint: Freehand Pointing with a Smartwatch in a Ubiquitous Display Environment. In: Proceedings of the International Working Conference on Advanced Visual Interfaces. pp. 128–135. ACM (2016). <https://doi.org/10.1145/2909132.2909263>
 23. Kim, W., Choo, K.T.W., Lee, Y., Misra, A., Balan, R.K.: Empath-d: Vr-based empathetic app design for accessibility. In: Proceedings of the 16th Annual International Conference on Mobile Systems, Applications, and Services. pp. 123–135. MobiSys '18, ACM (2018). <https://doi.org/10.1145/3210240.3210331>
 24. Kjeldskov, J., Skov, M.B.: Was It Worth the Hassle?: Ten Years of Mobile HCI Research Discussions on Lab and Field Evaluations. In: Proceedings of the 16th International Conference on Human-computer Interaction with Mobile Devices & Services. pp. 43–52. ACM (2014). <https://doi.org/10.1145/2628363.2628398>
 25. Koelle, M., El Ali, A., Cobus, V., Heuten, W., Boll, S.C.J.: All About Acceptability?: Identifying Factors for the Adoption of Data Glasses. In: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems. pp. 295–300. ACM (2017). <https://doi.org/10.1145/3025453.3025749>
 26. Koh, F.: Singapore trials LED lights on pavements: What other places are doing to keep smartphone zombies safe. The Straits Times (4), 789–792 (mar 2017)
 27. Lauber, F., Butz, A.: In-your-face, Yet Unseen?: Improving Head-stabilized Warnings to Reduce Reaction Time. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. pp. 3201–3204. ACM (2014). <https://doi.org/10.1145/2556288.2557063>
 28. Liu, M., Nancel, M., Vogel, D.: Gunslinger: Subtle Arms-down Mid-air Interaction. In: Proceedings of the 28th Annual ACM Symposium on User Interface Software and Technology. pp. 63–71. ACM (2015). <https://doi.org/10.1145/2807442.2807489>

29. Lopes, P., Baudisch, P.: Muscle-propelled Force Feedback: Bringing Force Feedback to Mobile Devices. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. pp. 2577–2580. ACM (2013). <https://doi.org/10.1145/2470654.2481355>
30. Lucero, A., Vetek, A.: NotifEye: Using Interactive Glasses to Deal with Notifications While Walking in Public. In: Proceedings of the 11th Conference on Advances in Computer Entertainment Technology. pp. 17:1—17:10. ACM (2014). <https://doi.org/10.1145/2663806.2663824>
31. Lumsden, J., Brewster, S.: A Paradigm Shift: Alternative Interaction Techniques for Use with Mobile & Wearable Devices. In: Proc. CASCAN. pp. 197–210. IBM Press (2003)
32. Lumsden, J., Kondratova, I., Durling, S.: Investigating Microphone Efficacy for Facilitation of Mobile Speech-based Data Entry. In: Proc. of British HCI. pp. 89–97. British Computer Society (2007)
33. McCallum, D.C., Irani, P.: ARC-Pad: Absolute+Relative Cursor Positioning for Large Displays with a Mobile Touchscreen. In: Proceedings of the 22Nd Annual ACM Symposium on User Interface Software and Technology. pp. 153–156. ACM (2009). <https://doi.org/10.1145/1622176.1622205>
34. McCullough, M.: On Attention to Surroundings. Interactions (2012)
35. Melzer, J.E.: Head-mounted displays: designing for the user. McGraw-Hill Professional (1997)
36. Mistry, P., Maes, P., Chang, L.: WUW - Wear Ur World: A Wearable Gestural Interface. In: CHI '09 Extended Abstracts on Human Factors in Computing Systems. pp. 4111–4116. ACM (2009). <https://doi.org/10.1145/1520340.1520626>
37. Montero, C.S., Alexander, J., Marshall, M.T., Subramanian, S.: 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. pp. 275–278. ACM (2010). <https://doi.org/10.1145/1851600.1851647>
38. Mustonen, T., Berg, M., Kaistinen, J., Kawai, T., Häkkinen, J.: Visual task performance using a monocular see-through head-mounted display (HMD) while walking. (2013). <https://doi.org/10.1037/a0034635>
39. Nancel, M., Pietriga, E., Chapuis, O., Beaudouin-Lafon, M.: Mid-Air Pointing on Ultra-Walls. *ACM Trans. Comput.-Hum. Interact.* **22**(5), 21:1—21:62 (aug 2015). <https://doi.org/10.1145/2766448>
40. Nasar, J.L., Troyer, D.: Pedestrian injuries due to mobile phone use in public places. *Accident Analysis & Prevention* **57**, 91–95 (aug 2013). <https://doi.org/10.1016/j.aap.2013.03.021>
41. Ng, A., Williamson, J.H., Brewster, S.A.: Comparing Evaluation Methods for Encumbrance and Walking on Interaction with Touchscreen Mobile Devices. In: Proc. of ICMI. pp. 23–32. ACM (2014). <https://doi.org/10.1145/2628363.2628382>
42. Oulasvirta, A., Tamminen, S., Roto, V., Kuorelahti, J.: Interaction in 4-second Bursts: The Fragmented Nature of Attentional Resources in Mobile HCI. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. pp. 919–928. ACM (2005). <https://doi.org/10.1145/1054972.1055101>
43. Pirhonen, A., Brewster, S., Holguin, C.: Gestural and Audio Metaphors As a Means of Control for Mobile Devices. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. pp. 291–298. ACM (2002). <https://doi.org/10.1145/503376.503428>

44. Rahman, M., Gustafson, S., Irani, P., Subramanian, S.: Tilt techniques: Investigating the dexterity of wrist-based input. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. pp. 1943–1952 (2009). <https://doi.org/10.1145/1518701.1518997>
45. Rico, J., Brewster, S.: Usable Gestures for Mobile Interfaces: Evaluating Social Acceptability. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. pp. 887–896. ACM (2010). <https://doi.org/10.1145/1753326.1753458>
46. Ronkainen, S., Häkkinä, J., Kaleva, S., Colley, A., Linjama, J.: Tap Input As an Embedded Interaction Method for Mobile Devices. In: *Proceedings of the 1st International Conference on Tangible and Embedded Interaction*. pp. 263–270. ACM (2007). <https://doi.org/10.1145/1226969.1227023>
47. Saponas, T.S., Harrison, C., Benko, H.: PocketTouch: Through-fabric Capacitive Touch Input. In: *Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology*. pp. 303–308. ACM (2011). <https://doi.org/10.1145/2047196.2047235>
48. Sawhney, N., Schmandt, C.: Nomadic Radio: Speech and Audio Interaction for Contextual Messaging in Nomadic Environments. *ACM Trans. Comput.-Hum. Interact.* **7**(3), 353–383 (sep 2000). <https://doi.org/10.1145/355324.355327>
49. Schwebel, D.C., Gaines, J., Severson, J.: Validation of virtual reality as a tool to understand and prevent child pedestrian injury. *Accident Analysis & Prevention* **40**(4), 1394–1400 (2008). <https://doi.org/http://dx.doi.org/10.1016/j.aap.2008.03.005>
50. Schwebel, D.C., Stavrinou, D., Byington, K.W., Davis, T., O’Neal, E.E., de Jong, D.: Distraction and pedestrian safety: How talking on the phone, texting, and listening to music impact crossing the street. *Accident Analysis & Prevention* **45**, 266–271 (2012). <https://doi.org/http://dx.doi.org/10.1016/j.aap.2011.07.011>
51. Serrano, M., Ens, B.M., Irani, P.P.: 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*. pp. 3181–3190. ACM (2014). <https://doi.org/10.1145/2556288.2556984>
52. Shneiderman, B.: Direct manipulation for comprehensible, predictable and controllable user interfaces. *Proceedings of the 2nd international conference on Intelligent user interfaces - IUI '97* pp. 33–39 (1997). <https://doi.org/10.1145/238218.238281>
53. Soukoreff, R.W., MacKenzie, S.: Towards a Standard for Pointing Device Evaluation, Perspectives on 27 Years of Fitts’ Law Research in HCI. *Int. J. Hum.-Comput. Stud.* **61**(6), 751–789 (dec 2004). <https://doi.org/10.1016/j.ijhcs.2004.09.001>
54. Sutherland, I.E.: A Head-mounted Three Dimensional Display. In: *Proceedings of the December 9-11, 1968, Fall Joint Computer Conference, Part I*. pp. 757–764. ACM (1968). <https://doi.org/10.1145/1476589.1476686>
55. Vogel, D., Balakrishnan, R.: Distant Freehand Pointing and Clicking on Very Large, High Resolution Displays. In: *Proceedings of the 18th Annual ACM Symposium on User Interface Software and Technology*. pp. 33–42. ACM (2005). <https://doi.org/10.1145/1095034.1095041>
56. Yang, X.D., Grossman, T., Wigdor, D., Fitzmaurice, G.: Magic Finger: Always-available Input Through Finger Instrumentation. In: *Proc. of UIST 2012*. pp. 147–156. ACM (2012). <https://doi.org/10.1145/2380116.2380137>
57. Zhao, S., Dragicevic, P., Chignell, M., Balakrishnan, R., Baudisch, P.: Earpod: Eyes-free Menu Selection Using Touch Input and Reactive Audio Feedback. In: *Proc. of CHI*. pp. 1395–1404. ACM (2007). <https://doi.org/10.1145/1240624.1240836>