Getting Off the Treadmill: Evaluating Walking User Interfaces for Mobile Devices in Public Spaces

Shaun K. Kane,[1,2] Jacob O. Wobbrock[1] and Ian E. Smith[2,3]

[1]The Information School
DUB Group
University of Washington
Seattle, WA 98195 USA
{skane, wobbrock}@u.washington.edu

[2]Intel Research Seattle
1100 NE 45th Street, 6th Floor
Seattle, WA 98105 USA

[3]Transmutable Networks LLC
4742 42nd Ave. SW #326
Seattle, WA 98126 USA
iansmith@acm.org

ABSTRACT

Using a mobile device while moving limits attention and motor ability and can result in reduced performance. Mobile devices that can sense and adapt to contextual factors such as movement may reduce this performance deficit. We performed two studies evaluating the feasibility of *walking user interfaces* (WUIs) that adapt their layout when the user is moving. In a pilot study with 6 users, we evaluated the effects of different button sizes on performance when walking while using a portable music player. Results showed significant interactions between size and movement. In the second study, 29 users evaluated the performance of a WUI that dynamically changed button sizes as the user moved. Results show that our dynamic user interface performs at the level of its component static interfaces without any additional penalty due to adaptation. This work adds to our design knowledge about walking user interfaces and provides lessons learned in evaluating mobile devices while walking in public spaces.

Categories and Subject Descriptors: H.5.2. [**Information interfaces and presentation**]: User interfaces — *Input devices and strategies*.

General Terms: Design, Experimentation, Human Factors

Keywords: Mobile device, media player, walking user interface, adaptive user interface, situational impairments.

1. INTRODUCTION

Mobile devices are used across an increasing range of places and contexts. Many users carry mobile devices at all times and use them at home, at work, on the street, in the car, and in other places [2,11]. As mobile devices become more powerful and portable, they will be used in an increasing variety of situations.

Most mobile user interfaces are designed for a person who is standing still and paying full attention. When these devices are used outside of home or office contexts, however, users must adapt their use of the device to an external environment which places high demands on their ability to interact [8]. For example, a user who is typing a text message while walking down the street must maintain awareness of his or her surroundings, avoid obstacles, and use a device that is itself in motion. The user's ability to read text on screen or to accurately press buttons may be compromised. We call these effects *situational impairments* [21] because they are contextual factors that reduce a user's ability to interact, which may be comparable in nature to users with physical or sensory impairments [18]. Just as technology may be modified to work for users with limited physical capabilities, mobile technology may provide *situational accommodations* to better support users' capabilities as they move about in the world.

Figure 1. A participant interacting with our adaptive walking user interface (WUI) on an ultra-mobile PC.

Situational impairments may be triggered by a range of contextual factors, and may exert a range of effects on performance. Those situations that are most pervasive present the most compelling opportunities for research. For example, it is known that walking can reduce a user's ability to read text [1] and enter data into a device [15]. Therefore, it is worthwhile to quantify the negative effects on use due to walking, and to explore potential design changes that may ameliorate these effects.

In this paper, we describe two investigations that explore the effects of walking on interaction with a mobile device. In the first study, we examine the effects of walking on performance with soft buttons when using a mobile device. In the second study, we develop a prototype *walking user interface* (WUI) music player that changes the screen layout based on the user's movement (Figure 1). Performance of the dynamic WUI prototype was comparable to its component static interfaces, indicating there was no additive penalty for adaptation. We conducted these studies in open public spaces in order to fully observe the effects of using mobile devices while walking. Our contributions include providing (1) an increased understanding of the real-world effects of mobile use while walking, (2) reflections on the difficulties posed in performing mobile device experiments in public places, and (3) design implications of the prototype WUI application.

2. RELATED WORK

2.1 Studies of Use While Walking

Walking while using a mobile device is a common activity that has measurable effects on the user's ability to read text and interact with user interface elements. Several studies have systematically examined these effects. Barnard et al. [1] evaluated reading comprehension and word search tasks while walking with a PDA, and found that walking significantly increased task times over a sitting position. Mustonen et al. [17] evaluated reading while walking, and found that both objective performance and subjective ease-of-use measures decreased while walking. Mizobuchi et al. [15] examined text entry performance using a PDA and stylus while walking, and found that text entry performance decreased as users walked and as the interface size decreased, but found no interaction effects between size and movement. Lin et al. [9] conducted a Fitts' law study of stylus-tapping while walking, and found that tapping performance decreased while walking, but only for smaller-sized targets. A subsequent study by Lin et al. [10] found that time to complete single target tapping tasks did not increase while walking, but that subjective workload and overall task completion times did increase while walking. Mackay et al. [11] compared multiple scrolling techniques on a PDA with stylus while walking in a public area on campus. They found that the task was slower while walking, but that the effect of walking did not differ across techniques.

Although Lin et al. [10] suggest that interfaces designed to be used while walking should have substantially larger buttons than their standing counterparts, they do not develop a prototype to test this hypothesis. Therefore, our work carries out the prototyping and evaluation of a walking user interface (WUI). Furthermore, we performed outdoor studies in pubic places to measure real-world effects of mobile device use. This is in contrast to all studies described above except Mackay's [11]; the others simulated walking conditions using a treadmill or on a closed course.

2.2 User Interfaces for Walking

Prior research projects have addressed some of the situational impairments that may be caused by walking. A number of projects have developed interaction techniques that are usable with one hand [4,6,19]. Other efforts have focused on supporting field workers in mobile environments: Kristofffersen [8] provided a set of design guidelines for mobile technologies to support workers in the field, and Pascoe et al. [20] introduced *minimal attention user interfaces* to support ecologists in the field. Zhao et al. [24] and Brewster et al. [3] developed eyes-free mobile interaction techniques that use audio feedback and gestures to allow use without looking at a screen. While these systems provide an alternative or reduced-functionality interface that is always on, our work explores interfaces that provide different granularities of interaction for standing and walking.

3. SITUATIONAL ACCOMMODATIONS AND WALKING USER INTERFACES

3.1 Situational Accommodations

We situate this work within our larger investigation of the effects of situational impairments on mobile use [21]. These impairing effects may be caused by a range of contextual factors, including:

- *Environmental factors:* low light, glare, ambient noise, vibration tremor, extreme temperatures, rainwater, uneven terrain;
- *Attentional factors:* Physical obstacles, social interactions, divided attention, abrupt distraction, device out-of-sight;
- *Physical factors:* Impeding clothing, baggage, occupied hands, user or device movement, posture or grip, user fatigue.

These situational factors may appear at various times while a user is operating a mobile device in the world and around other people, and may interact with each other. For example, a rainstorm may cause a user to walk faster in order to find shelter, further impairing his or her ability to interact accurately with a mobile device. In some cases, users may be able to temporarily reduce the impairing effects of the environment by taking action. For example, a user who is interacting with a PDA while walking may stop walking for a moment in order to concentrate on the device. However, it may not be possible to stop walking in the middle of a busy train station or while crossing the street, and it may be uncomfortable to stop walking in a snowstorm.

Current devices are largely blind to a user's context. It is therefore important for designers to consider incorporating *situational accommodations* to provide some compensation for these contextual influences. Accommodations may include changing the user interface in some way, such as making on-screen text larger, or providing an alternative interface optimized for a specific context. Accommodations may be invoked automatically by device sensors that infer a user's context, or may be explicitly activated by the user. As mobile interactions become more deeply intertwined with and affected by the outside world, we envision a wide range of possible situational accommodations that address the three different categories of situational impairments and that occur at different system levels (e.g., operating system, application, widget, or perhaps even the physical device).

3.2 Walking User Interfaces (WUIs)

In this paper, we focus on the specific subset of situational impairments caused by using a mobile device while walking, and on situational accommodations that address these impairments. We chose walking because it is a common activity, and because there exists evidence that walking affects performance when using mobile devices.

We introduce the term *walking user interface* (WUI) to denote user interfaces that are designed specifically to compensate for the effects of walking on mobile device usability. WUIs may take a range of forms, and may use a variety of methods to accommodate walking. Here we introduce a simple WUI that alters text size and widget size in order to be easier to use while walking. Admittedly, this is just one of many potential accommodation strategies, and a simple one at that. By starting with simple strategies, we can determine whether more elaborate adaptations are necessary.

Although we define WUIs as interfaces that accommodate the effects of walking, we note that walking in the real world does not consist simply of locomotion, but also introduces other concerns, such as the need to maintain situational awareness and avoid walking into obstacles.

3.3 Evaluating Walking User Interfaces

As Kjeldskov and Stage [7] have shown, mobile device field studies uncover different issues than lab experiments. For this reason, evaluations of WUIs should use experimental conditions that accurately reflect the constraints and risks of using a device while walking in public whenever possible. For this reason, we chose to perform our evaluations in public spaces with pedestrian traffic.

While other studies have asked participants to walk freely, measuring both task performance and walking speed, we chose to control for walking speed. We made this choice based on the assumption that users will often be unable to slow down or stop walking in order to use a mobile device. This choice was also intended to maximize impairing walking effects, as users were not able to slow down if the task became difficult. Finally, this choice allowed us to use deviations from the walking course as a dependent variable, and as an additional measure of accuracy.

Walking pace is easy to control using a treadmill, but is more difficult to control on an outdoor course. We accomplished this in the first study using a click track. In the second study, participants needed to both maintain pace and follow a predefined path, so we used a human pacesetter.

We made the following assumptions about typical mobile device use, and used these to define the parameters of our studies: (1) that users will attempt to maintain a fixed pace when walking, and will not stop to use a mobile device, (2) that users will be required to avoid both stationary and moving obstacles, and (3) that users will often be unable to retrieve a stylus while walking, and thus may interact with their fingers only if possible.

We expect that other evaluations of WUIs will make different decisions regarding experimental conditions. This is reasonable given the range of mobile device usage scenarios; however, we recommend that researchers consider situational factors carefully when evaluating technology that provides situational accommodations. Key situational factors that we considered when designing the experiments are summarized in Table 1.

Situational factor	*Possible choices*
Walking path	Straight, curved, variable, participant-chosen
Walking speed	Set by participant, suggested by experimenters, fixed pace
Walking task	Walk freely, walk at fixed pace, follow pacesetter
Distractions	Sound, light level, conversation
Interruptions	Between tasks, within tasks
Location	Indoors, outdoors
Obstacles	None, stationary, moving
Hands	One hand, two hands

Table 1. Key situational factors for WUI evaluation studies.

4. PROTOTYPE WUI APPLICATION

In this section, we introduce a prototype WUI used to explore the feasibility and effects of a WUI in a realistic setting. This prototype uses the situational accommodation of enlarging user interface elements in order to reduce the effects of walking on performance. As stated above, WUIs may take a variety of forms in order to address user activities and situational effects. This is just one of many possible forms that WUIs may take.

One prototype, a *WUI music player*, is an application that uses interface scaling to minimize the effects of walking. The player mimics the user interface layout of common portable media players, containing a scrollable list of songs (Figure 2). Tapping a song title causes that song to be played. As the user begins walking, interface elements including text and buttons become larger, so fewer items are shown on the screen. For experimental purposes, interface size was controlled by the researchers using a Wizard of Oz configuration, and was not changeable by the user.

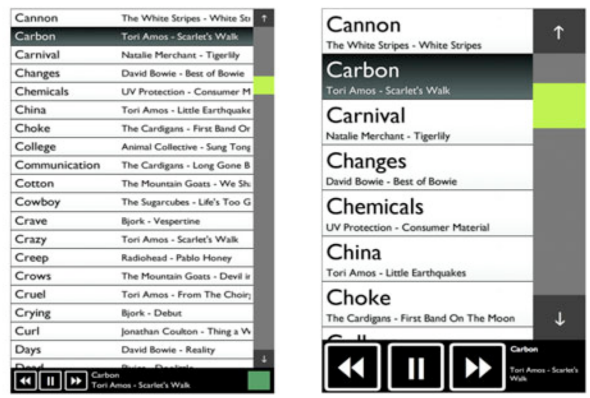

Figure 2. Our music player user interface in two sizes. (*left*) The player while standing; (*right*) the player while walking.

The music player was developed using Adobe Flash and was installed on a Sony UX2 ultra-mobile PC (UMPC) running Microsoft Windows XP. This device weighed 1.2 pounds and had a 4.5 inch (114.3 mm) touch screen with 296 dpi. The device was held in portrait orientation at a screen resolution of 600×1024 pixels. Users interacted with the device with their fingers and thumbs on the touch screen; neither styli nor hard buttons were used. While the UX2 is larger than some mobile devices, mid-sized devices are good candidates for WUIs, as they are usable for both mobile tasks as well as more complex dedicated tasks.

In the following sections, we describe two experiments that use the music player prototype to measure the effects of use while walking, and to determine the feasibility of a WUI that adapts automatically based on walking.

5. EXPERIMENT 1: EFFECTS OF WALKING AND SOFT-BUTTON SIZE

We performed a pilot evaluation of the WUI music player at different sizes in order to determine the effects of walking on performance, and how these effects varied with soft-button size.

5.1 Method

5.1.1 Participants

We recruited 6 participants (2 male, 4 female) through university mailing lists. Participants were students at the university. The experiment took approximately 30 minutes for each participant.

5.1.2 Apparatus

Participants used a version of the music player running on the Sony UMPC as described in Section 4. The playlist consisted of 100 songs ordered alphabetically by title. We selected songs with simple, single-word titles to minimize the possibility that participants would forget a song title during a task. Songs were referred to by title only, so participants did not gain any advantage by being familiar with the songs or artists.

Log files were recorded locally on the UMPC in XML format. Logs were later parsed using Python scripts, and the parsed files were analyzed with a commercial statistics package.

5.1.3 Location

The experiment took place in a corridor in an academic building on the university campus. This was a public area in which other people were frequently standing and walking. All participants were familiar with the location before participating in the study.

5.1.4 Procedure

Participants performed a set of music selection tasks using the music player. For each trial, users were required to scroll through the music list to find the given song and to tap the song to play it. Before each trial, users viewed an information screen containing the task to be performed and indicating whether the trial was to be performed standing or walking. If the trial involved walking, the prototype played a recorded track of walking footsteps. Users were instructed to keep the pace set by the audio footsteps track. The speed of the audio track was initially set at 112 steps per minute based on prior human factors research [16], but was adjusted until the participant was comfortable with the pace.

Participants were instructed to complete the trials quickly and accurately. Each trial began when the user dismissed the information screen, and ended when the user pressed the correct song. During walking trials, participants were instructed to walk from one end of the corridor to the other, and to turn around without interrupting the task when they reached the opposite end.

Conditions were presented in random order to counteract order effects. Within each condition, songs were presented in random order, but the same songs were used for each participant. For each song, the playlist was scrolled to either the top or bottom of the list. Scroll positions were randomly selected, but were the same for all participants.

5.1.5 Design and Analysis

The experiment was a 7×2 within-subjects factorial design with the following factors and levels:

- *Size (mm)* {3.43, 5.15, 6.86, 8.58, 10.30, 12.01, 13.73}
- *Movement* {standing, walking}

Participants completed 9 music-finding trials in each of 14 conditions for 126 trials total. With 6 participants, this meant that our study comprised 756 trials all together. Our dependent measures were task time and number of task errors. We defined *task errors* to be when the user played an incorrect song.

As is common with temporal dependent measures, task time did not fit a normal distribution (Shapiro-Wilk W=8.27, $p<.0001$). Similarly, task errors were rare and non-normal (W=0.63, $p<.0001$). This made ANOVAs inadequate. Therefore, task time was analyzed using an exponential regression model, and trial error data was analyzed using a Poisson regression model. Both analyses are suited to these types of data [22].

5.2 Results

5.2.1 Task Time

Figure 3 shows the average task time for levels of *Size* and *Movement*. There was a significant effect of *Size* on task time ($\chi^2_{(6,N=756)}$=88.60, p<.01), and a *Size*Movement* interaction ($\chi^2_{(6,N=756)}$=19.56, p<.01), but no main effect of *Movement* on task time ($\chi^2_{(1,N=756)}$=1.03, n.s.). This shows that interface size affects task time, and that the effect of size on task time varies with movement. The lack of a main effect of *Movement* contradicts prior work in this area that suggests walking increases task time (e.g. [10,15]), but this may be due to having a small sample size and high variance: walking did raise the total task time for 5 of 7 interface sizes.

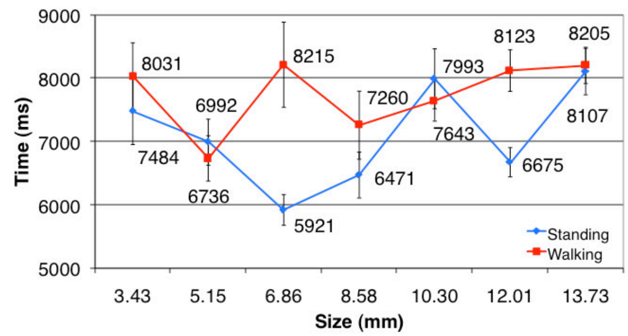

Figure 3. Standing and walking task times for each value of *Size*. Lower is better. Error bars show ±1 SE.

Examination of the mean time values suggests that task time is longest when the interface is very small or very large. This is intuitive because targets that are very small take more time to acquire, while lists of larger targets require more scrolling. Overall, the 6.86 mm interface was fastest while standing, and the 5.15 mm was fastest while walking. However, the fastest interface varied across users. While standing, 2 participants were fastest with 8.58 mm targets, 3 with 6.86 mm and 1 with 12.01 mm. While walking, 3 participants were fastest with 5.15 mm, 1 with 13.73 mm, 1 with 8.58 mm and 1 with 3.43 mm. This suggests that the ideal interface size may differ between users. However, this is a small study with high variance, and further research is needed to better measure these effects.

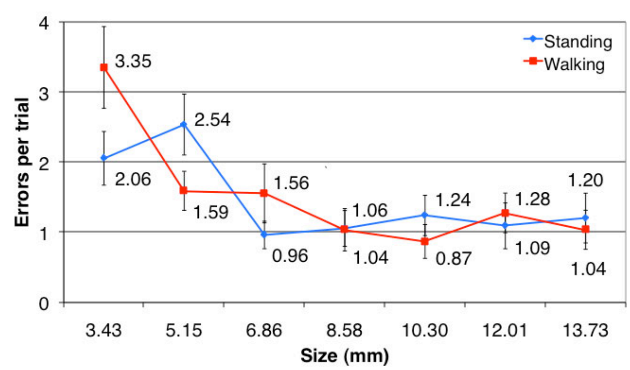

Figure 4. Mean errors per trial for each value of *Size* and *Movement*. Lower is better. Error bars show ±1 SE.

5.2.2 Task Errors

As with task time, there were significant effects of *Size* ($\chi^2_{(6,N=756)}$=21.34, p<.01) and *Size*Movement* ($\chi^2_{(6,N=756)}$=19.56, p<.01), but no main effect of *Movement* ($\chi^2_{(1,N=756)}$=1.03, n.s.) on task errors. This suggests that the number of errors varies by size, and that movement unevenly affects how size affects errors. Figure 4 shows the number of errors for each size.

6. EXPERIMENT 2: ADAPTIVE WALKING USER INTERFACE PERFORMANCE

The first experiment revealed an interaction between target size and movement, which suggests that changes to target size may have a positive effect on performance if a device can provide the best-sized interface for any walking speed. The second experiment therefore evaluated a WUI prototype in the field.

For this experiment, we developed an adaptive version of the music player described in Section 4. This prototype scales its target size based on the user's movement. When the user stands still, targets shrink, allowing more to be displayed on the screen. When the user is walking, the interface expands so that text is easier to read and targets are easier to hit. Some prior work shows that adaptive user interfaces have provided minimal performance benefits and have been disliked by users [13,14]. However, since a user's own abilities may change while using a mobile device in different contexts, a user interface that adapts to match a user's capabilities may be desirable. Prior work by McGrenere et al. [12] suggests that users may be willing to switch between simplified and complex versions of a single interface in some situations.

We evaluated the performance of the adaptive media player compared with two versions of a static target-size media player. In addition, we extended the design of the previous study to include different levels of task difficulty, added a pace-setting experimenter to reduce the difficulty of paced walking, and moved the experiment to an outdoor course in a public square.

6.1 Method

6.1.1 Participants

We recruited 30 participants through university mailing lists and a web-based community bulletin board. One participant was excluded from analysis due to a previously undisclosed health condition, leaving 29 participants in our analysis. Participants' average age was 27.1 (*SD*=8.9). Nineteen (65.5%) were female. Twenty-five (86.2%) were right-handed. The experiment took approximately 90 minutes for each participant.

6.1.2 Apparatus

This experiment used a modified version of the music player described previously. The prototype was improved to include a more robust scrolling control and song navigation buttons. The playlist was extended to 190 songs.

Our "adaptive condition" was designed to switch between two interface sizes: a complex (small-button) size and a simple (large-button) size as shown in Figure 2. Our "static conditions" were fixed at either small- or large-button sizes. During transitions, the interface played a soft chime to indicate its transition to a new interface size. Different chimes were used to identify each size.

Button sizes for the static-simple and static-complex interfaces were informed by pilot studies and based on the sizes of the device *hardware* buttons. A value of 150% was chosen for the static-simple interface, and 50% for the static-complex interface. In the static-simple condition, song titles were 11.43×46.99 mm, interface buttons were 11.43×11.43 mm, song title text was 6.1 mm tall, and there were 7 songs on screen at any one time. In the static-complex condition, song titles were 3.81×52.83 mm, other buttons were 3.81×3.81 mm, song title text was 3.05 mm tall, and there were 19 songs on screen at any one time.

The experimenter controlled changes to the user interface in the adaptive condition, using another Sony UX2 device with custom software. Both devices were connected wirelessly using an ad-hoc 802.11b network. The experimenter's UMPC ran an application that allowed him to change button size, start and end tasks, and record participants' walking speed and events. Log files were transmitted wirelessly from the prototype to the experimenter's console and recorded in a single XML file. XML logs were later parsed using Python scripts, and the parsed comma-separated files were analyzed with a commercial statistics package.

6.1.3 Location

The experiment took place in an open plaza on a university campus (Figure 5). This was a public area in which other people were frequently standing, walking, and interacting with one another. This area was familiar to some participants, although none had walked the specific experimental course. Sessions were held in the afternoon for day-to-day consistency.

Figure 6. Participant following the pacesetter along the outdoor walking course.

6.1.4 Procedure

This experiment followed the procedure of the previous experiment, but with some new additions. First, the audio track used to set the walking pace in the first experiment was replaced with a human pacesetter. Second, we replaced the indoor corridor with an outdoor path in a public square with trees and pedestrians as obstacles. Third, the walking course was changed from a straight line to a curved path to increase walking difficulty.

Participants performed a set of music navigation and playing tasks while standing and walking with the UMPC. The course consisted of two points, *A* and *B*, marked by trees at either end of the square (Figure 6). Navigating the course involved weaving through a series of trees. In order to maintain a consistent walking pace among all participants, another experimenter served as the pacesetter. We chose to keep a fixed pace rather than measure pace as a dependent variable in order to ensure a comparable level of walking load across trials. Participants were instructed to follow and keep pace with the pacesetter as he walked along the path. The experimenter instructed participants to stay within 3 feet of the pacesetter as he walked. If the participant fell behind the pacesetter by more than 6 feet, the experimenter logged a walking deviation for that trial.

Each participant completed 3 sets of 18 trials. Each set corresponded to a specific interface: *static-simple*, *static-complex*, or *adaptive*. A single trial consisted of either playing a target song or pressing a button on the user interface. Each trial began at point *A* or *B*. The experimenter initiated trials using the console. When a trial began, the participant's device displayed a pop-up window that featured the trial instruction and whether to stand still or walk.

Trials had one of three difficulty levels. *Easy* tasks required the participant to tap one of the fixed buttons on the screen, such as the "play" or "next track" button. *Medium* tasks required the participant to play a track located in the list that was currently visible on-screen. *Hard* tasks required the participant to scroll through the song list to find and play a song. Note that while participants in this experiment performed an equal number of tasks for each level of difficulty, in real-world use it is likely that *Hard* tasks will be most common with large-button interfaces, as controls become larger and force items off-screen.

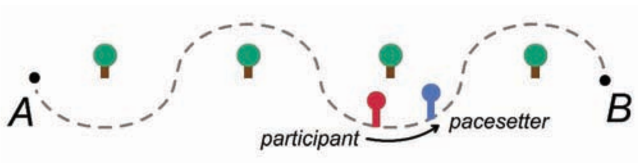

Figure 6. Experimental course. Half of the circuit is shown.

Standing and walking trials were randomly ordered within each set. For standing trials, the participant would stand at the current endpoint and complete the task. For walking trials, the participant would begin to walk with the pacesetter, and then would tap an OK button to begin the task while they were in motion. Participants were instructed to take at least five steps before beginning a walking task. When the user successfully completed the task, the system played a sound and the screen went blank. If the participant tapped an incorrect entry, the system logged the error but did not blank the screen.

At the start of the experiment, participants were given a demonstration of the interface followed by at least 10 practice trials. Participants were encouraged to continue practicing until they felt comfortable with the device. Participants rested for at least one minute between each set of trials, but they were permitted to rest for as long as they wished. Participants completed a brief questionnaire after the experiment.

6.1.5 Design and Analysis

The experiment was a 3×3×2 within-subjects factorial design with the following factors and levels:

- *Interface* {static-simple, static-complex, adaptive}
- *Difficulty* {easy, medium, hard}
- *Movement* {standing, walking}

Participants completed 3 music-finding trials for each of the 18 conditions, or 54 trials total. With 29 participants, this meant that our study comprised 1566 trials all together.

Our dependent measures were task time, number of task errors, number of walking deviations, total number of button presses and number of glances away from the device. We define *task errors* to be when the user played an incorrect song. We define *walking deviations* to be when users stopped along the course, lagged behind the pacesetter by more than 6 feet, or collided with an obstacle. Task time, task errors, and button presses were recorded on the UMPC. Walking deviations and glances were recorded by the experimenter.

The experimenter recorded other variables as well. *Weather* {cloudy, partly cloudy, sunny} was recorded, as were *Handedness* and *Gender*.

For each participant, measures were averaged over the 3 trials in each of the 18 experimental conditions, resulting in 18 measurements per participant. Data for one condition for one participant was lost due to instrument error, resulting in a total of 29×18-1=521 measurements for the experiment.

As in the previous experiment, task time was non-normal (W=0.81, p<.0001), and so was analyzed using an exponential regression model; trial error data was analyzed using a Poisson regression model [22].

A test for condition order on task time was non-significant ($\chi^2_{(1,N=521)}$=2.13, n.s.), indicating adequate counterbalancing of conditions.

6.2 Results

6.2.1 Validation of Experimental Treatments

Although an outdoor study cannot completely control all environmental factors, care was taken to provide consistent conditions for all participants. Prior to the experiment, the pacesetter was trained to walk at a constant speed by walking the path repeatedly until he was able to keep within 2.0 seconds of the mean for two consecutive path lengths. The pacesetter walked the experimental course in an average of 29.96 (*SD*=1.57) seconds. All recorded trials occurred within 30% of this mean trial time, while 98.75% of the trials occurred within 15% of the mean trial time. One participant completed the experiment on an alternate version of the course due to rain. There was no statistical difference in task time for this participant ($\chi^2_{(1,N=521)}$=0.00, n.s.).

6.2.2 Task Time

There was a significant main effect of *Difficulty* on task time ($\chi^2_{(2,N=521)}$=2121.57, p<.0001), indicating that more difficult tasks took longer to complete. This is no surprise, but validates that our tasks were indeed representative of their intended levels of difficulty.

Table 2 shows mean task time for levels of *Interface* and *Movement*. There was a significant main effect of *Interface* on task time ($\chi^2_{(2,N=521)}$=74.49, p<.0001). Overall, the adaptive interface was faster than the static-complex interface ($\chi^2_{(1,N=348)}$=14.77, p<.001), but the static-simple interface was faster than both of the other interfaces ($\chi^2_{(1,N=521)}$=58.47, p<.0001).

There was no main effect of *Movement* on task time ($\chi^2_{(1,N=521)}$=0.001, n.s.); however, this is likely due to the fact that task time does not increase for the adaptive interface while walking, since it switches to the simple interface. Removing the adaptive interface from this analysis does indeed result in a significant main effect for *Movement* on task time ($\chi^2_{(1,N=347)}$=6.28, p<.05), indicating that task time generally did increase while walking.

There was a significant interaction for *Interface*Difficulty* on task time ($\chi^2_{(4,N=521)}$=37.67, p<.0001). Looking closer, we see that while the static-simple interface outperforms the others for easy and medium tasks, there is no significant difference between the interfaces for hard tasks ($\chi^2_{(2,N=173)}$=75.24, n.s.) (Figure 7). This indicates that there may be performance differences among the interfaces that are overwhelmed by the additional time it takes to perform hard tasks, but that remain detectable for easy and medium tasks.

Interface	*Standing (ms)*	*Walking (ms)*
static-complex	3412 (2735)	3988 (3193)
adaptive	3453 (2623)	3402 (3319)
static-simple	2864 (2540)	3396 (3980)

Table 2. Task time for *Interface* and *Movement* conditions. Lower is better. *SD* are in parentheses.

There was a significant interaction for *Interface*Movement* on task time ($\chi^2_{(2,N=521)}$=20.13, p<.0001), showing that movement affected the interfaces differently. This interaction may be caused by the fact that task time does not increase while walking with the adaptive interface; indeed, if we exclude the adaptive interface from this test, we get a non-significant result ($\chi^2_{(1,N=347)}$=1.63, n.s.).

There was also a significant *Difficulty*Movement* interaction ($\chi^2_{(2,N=521)}$=12.74, p<.01), which shows that standing and walking affected task time differently for different levels of task difficulty. Specifically, difficult tasks took proportionally longer when walking than they did when standing as compared to medium and easy tasks. There was also a significant *Interface*Difficulty* Movement* interaction ($\chi^2_{(4,N=521)}$=11.49, p<.05), because the aforementioned interaction was more pronounced for the simple interface than the complex interface.

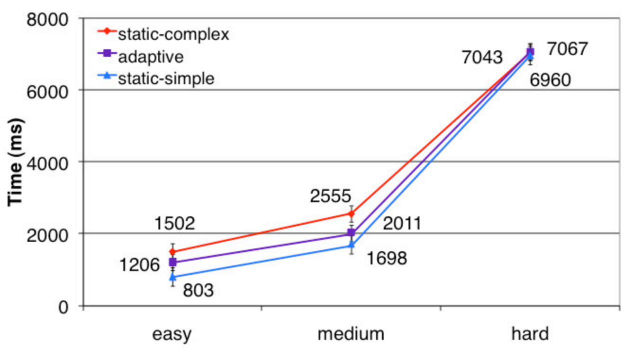

Figure 7. Task time by *Difficulty* and *Interface*. Lower is better. Error bars show ±1 SE. There were no significant differences between interfaces for hard tasks, but there were for medium and easy tasks.

Task time on the adaptive interface was the same as each of the comparable static interfaces; there was no significant time penalty incurred by the adaptive interface ($\chi^2_{(1,N=348)}$=1.24, n.s.) (Figure 8). This is an interesting finding because adaptation itself has been shown to reduce performance [5,13]. Although the adaptive interface performs comparably to its component static interfaces, the static-simple interface is faster overall than the static-complex interface and, for this reason, is also faster than the adaptive interface, which contains both static interfaces as its parts.

Although the adaptive interface was slower overall than the static-simple interface, for some users the adaptive interface was indeed faster. The adaptive interface was fastest for 10 users, while the static-simple interface was fastest for 16 users and the static-complex was fastest for 3 users.

In examining other variables of interest, we found that task time was not significantly affected by *Handedness* ($\chi^2_{(1,N=521)}$=0.01, n.s.) or *Gender* ($\chi^2_{(1,N=521)}$=0.09, n.s.). However, *Weather* nearly exerted a significant effect on task time ($\chi^2_{(2,N=521)}$=5.66, p=.059). Cloudy weather caused about an 18.5% reduction in task time compared to both partly cloudy and sunny conditions. This is likely due to the fact that the device's LCD screen was difficult to see in bright light, indicating another type of situational impairment besides walking *vs*. standing.

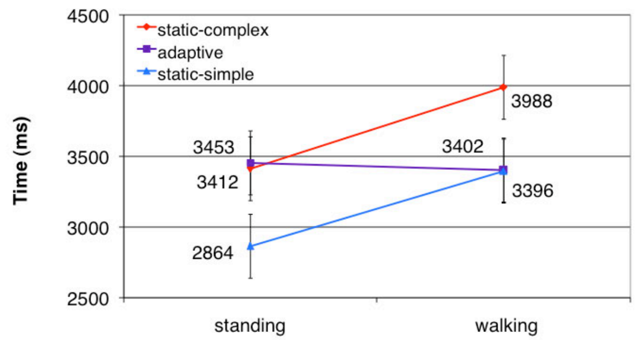

Figure 8. Task time by levels of *Interface* and *Movement*. Lower is better. Error bars show ±1 SE. There was no significant difference between the adaptive interface and comparable static interfaces for their shared levels of movement, indicating no additive penalty for adaptation.

6.2.3 Task Errors

Task errors occurred when a user selected an incorrect target in the user interface and played the wrong song. The average number of task errors per trial was 0.04, indicating that there were very few errors per trial. Table 3 shows mean task errors by *Interface* and *Movement.*

Interface	*Standing errors / trial*	*Walking errors*
static-simple	0.01 (0.08)	0.02 (0.07)
adaptive	0.05 (0.02)	0.02 (0.09)
static-complex	0.07 (0.02)	0.09 (0.02)

Table 3. Task errors for levels of *Interface* and *Movement.* Lower is better. *SD* are in parentheses.

There were significant effects of *Interface* ($\chi^2_{(2,N=521)}$=25.45,p<.0001) and *Difficulty* ($\chi^2_{(2,N=521)}$=45.64, p<.0001) on task errors. Task errors increased for complex interfaces, and increased as task difficulty increased. The number of errors was smallest for the static-simple interface, larger for the adaptive interface, and largest for the static-complex interface. There were no significant interactions. In pairwise comparisons, we found no significant difference in errors between the complex interface and adaptive interfaces while standing ($\chi^2_{(1,N=174)}$=1.74, n.s.), or between the simple interface and adaptive interface while walking ($\chi^2_{(1,N=174)}$=2.71, n.s.), suggesting that the changing adaptive interface did not cause additional errors.

6.2.4 Walking Deviations

The average number of walking deviations per trial was 0.06. There was a significant effect of *Difficulty* on walking deviations

($\chi^2_{(2,N=261)}$=44.74, p<.0001), indicating that more difficult tasks caused more walking errors. There were no significant effects for *Interface* ($\chi^2_{(2,N=261)}$=2.14, n.s.) or for *Difficulty*Interface* ($\chi^2_{(4,N=261)}$=3.37, n.s.).

6.2.5 Button Presses

Easy and medium difficulty tasks could be completed with only one button press. Hard tasks could be completed with two button presses using the scroll bar, but required more button presses for users who did not use the scroll bar. Button presses beyond the minimum required number might indicate task errors such as pressing an incorrect song button, or navigation errors such as scrolling in the wrong direction. The average number of button presses per trial was 2.24. There were no significant differences in the number of total button presses for *Interface*, *Difficulty* or *Movement.*

However, further examination shows a tradeoff between button size and screen resolution as the screen becomes denser. Users of the complex interface hit the *Page Up* and *Page Down* buttons fewer times than users of the other interfaces ($\chi^2_{(2,N=521)}$=17.63, p<.001). However, we see above that task errors increase as the interface becomes more complex and controls become smaller. This suggests that the static-complex interface allows users to more quickly navigate the track list, but with an increased number of errors (Figure 9), resulting in a speed-accuracy tradeoff among interfaces.

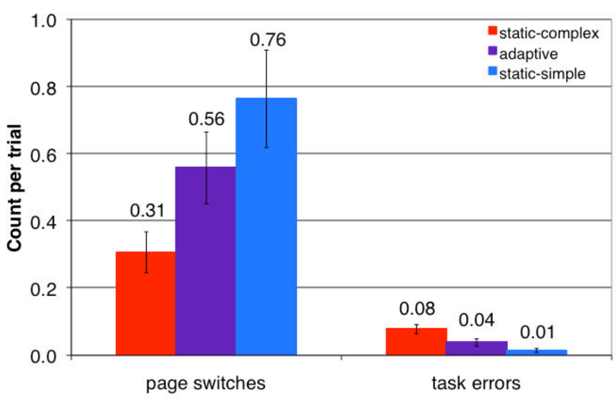

Figure 9. Page switches and task errors by interface. Lower is better. Error bars show ±1 SE. Interfaces with larger buttons had more page switches, but fewer task errors.

6.2.6 Number of Glances

The number of glances was calculated as the number of times that the participant looked up from the device per trial. The average number of glances was 0.51. This indicates that participants were able to complete most trials without glancing up multiple times. There was a significant effect of *Movement* on number of glances ($\chi^2_{(1,N=521)}$=109.72, p<.0001), indicating that participants were more likely to glance up while walking, but did not glance up very often. There were no other significant effects on glances.

6.2.7 Subjective Results

On a Likert scale where 1 was “least difficult” and 7 was “most difficult,” participants rated the walking tasks as being more difficult than standing tasks for all three *Interfaces*, with a mean rating of 3.62 for walking and 2.43 for standing tasks (p<.01).

Participants were also asked to rate each interface using a 7-point Likert scale where 1 was “strongly disagree” and 7 was “strongly agree.” Table 4 contains a summary of participants’ ratings. Using a nonparametric Friedman test, there was a main effect of *Interface* on *easy to use* ($\chi^2_{(2,N=29)}$=20.44, p<.0001), *enjoyable to use* ($\chi^2_{(2,N=29)}$=19.39, p<.0001), *comfortable while walking* ($\chi^2_{(2,N=29)}$=21.81, p<.0001), and *does not require much attention* ($\chi^2_{(2,N=29)}$=18.81, p<.0001). On these questions, the static-simple interface was rated best, the adaptive interface was rated second, and the static-complex interface worst.

Question	*static-simple*	*adaptive*	*static-complex*
Easy to use	5.79 (1.08)	5.14 (1.30)	4.55 (1.64)
Enjoyable to use	5.21 (1.47)	4.83 (1.51)	4.34 (1.56)
Comfortable while walking	5.52 (1.38)	5.14 (1.33)	4.17 (1.61)
Does not require much attention	4.79 (1.57)	4.52 (1.40)	3.59 (1.50)

Table 4. Mean Likert scores for I*nterface*. A 7-point Likert scale was used. Higher is better. *SD* are in parentheses.

Eighteen participants chose the static-simple interface as their favorite interface, while 6 chose the adaptive interface and 5 chose the static-complex interface. Several users stated that the larger buttons were easier to read and press.

7. DISCUSSION

In this section we discuss the results of the second experiment, and provide general discussion about evaluating the effects of WUIs in public spaces.

7.1 Effectiveness of the WUI prototype

Overall, we found that the adaptive version of the WUI music player did not perform as well as the static-simple interface with large buttons. This is probably due to two issues. First, for some users, the small buttons used in the complex interface were simply too small to comfortably press. Second, the design approach that we used–scaling buttons as the user walks–was also subject to a performance tradeoff. While the complex interface allowed more items to be displayed on the screen at one time, and thus should have decreased visual search and scrolling time, the size of the buttons resulted in a higher error rate and thus reduced overall speed. Meanwhile, the simple interface contained screen elements that were easier to hit, but required longer scrolling times to reach them. While it might seem possible to solve the first problem by simply choosing different button sizes, the results of Experiment 1 suggest that *different button sizes may be optimal for different users*, especially when people are using their fingers instead of a stylus. Addressing this problem may require user-specific personalization. The second problem initially seems somewhat less tractable, as it is difficult to make screen elements larger without showing fewer of them. One potential solution is to show only the most important functions in a simple view, allowing users to perform a smaller set of tasks while walking, but with better performance.

Despite the fact that the adaptive interface was not best overall, it did perform best for 10 of 29 users (34.5%). Furthermore, the adaptive interface was no slower or more error-prone than the comparable static interfaces, indicating that there was no additive penalty for switching the interface. In other words, it is possible to have the benefits of a more complex interface while stationary, like needing fewer button presses, and the benefits of a simpler interface while walking, like increased target size, without an apparent switching cost. Although adaptation traditionally has not fared well on the desktop, the reduced attention and interaction capability of mobile users may present an opportunity for user interfaces that can adapt to match users' abilities. While the present study used a Wizard of Oz design to switch between interfaces, prior research has shown that an on-board sensor such as an accelerometer is able to accurately detect whether or not a user is moving [23], so this type of interface could be created relatively easily using current technology. In fact, we found in informal testing that we could create a reliable automatic walking-detector with simple computer vision algorithms using the built-in camera of the ultra-mobile PC, suggesting that automatic adaptation is possible even with built-in device features.

Mackay et al. [11] previously found that the advantages of changing size are small while walking, and suggest that adaptive interfaces may not be worth the added complexity that they bring. Our study is in agreement, showing only small performance differences between the various button sizes. However, we believe that these small differences would be magnified when a user is moving through a demanding public space. For example, a user walking down a city street may only have a few seconds before he or she reaches the next block. If the task is not complete by the end of the block, the user may need to stop the task until he or she has finished crossing the street. Also, the burden of introducing adaptive WUIs seems to be relatively small. Our results indicate that users may be able to handle simple user interface switches without much cognitive, perceptual, or conceptual difficulty.

One surprising result of the second experiment was that participants performed relatively few walking errors and glances during the experiment. We had anticipated that users would need to glance away from the device far more often than they actually did. One explanation for this is that users were able to keep the pacesetter and major obstacles in their peripheral vision throughout the experiment. Using an audio track to keep pace as in Experiment 1 or increasing the number of obstacles might result in more glances, and might further accentuate the differences between interfaces.

7.2 Individual Differences

In contrast to many controlled experiments in the lab, we observed large variations on task time among trials and participants. While this may be due partly to the task that we chose, we suspect that two other factors may be involved. First, running experiments in a public place introduces environmental confounds that cannot be completely controlled. For example, a loud noise in the background might cause a participant to pause for just a moment, and this pause may significantly impact task time. Although we attempted to rerun tasks that were significantly interrupted, it was not possible to monitor and control all distractions. This variability might diminish the precision of our results, but we feel that this tradeoff is acceptable in exchange for the added information gleaned from running an experiment in public. We might overcome this variability by including more participants and more trials per participant.

Furthermore, running an experiment in a natural environment may magnify the effects of individual differences such as eyesight and motor ability. For example, because participants used their fingers to interact with a touch screen, some users with large hands reported difficulty interacting with on-screen buttons. This introduces an additional source of variability that would not be found in a lab study using a stylus. Once again, we felt that it was important that our study reflect natural use conditions.

7.3 Usability Evaluation in Public Spaces

Running usability studies in public spaces reveals issues that cannot be found using lab studies, as studies in public spaces are invariably subjected to environmental effects. Issues included:

Weather. In the second experiment, the amount of glare from sunlight affected task time, and one participant needed to be moved to another location because of rain. We also suspect that user fatigue may increase in hot or cold weather.

Pedestrians. Although we posted signs stating that a study was occurring, we chose not to cordon off the walking path. In several cases, pedestrians walked through the course or attempted to talk with a participant or experimenter. This problem might be reduced through better signage, or by wearing clothing that more clearly indicates that an experiment is taking place.

Public events. In several cases, events occurring in a public space added to the distractibility of the environment, even if they did not directly interfere with the experimental course. For example, one experiment was scheduled at the same time that a concert was occurring on the other side of the public square.

Safety. Participants who are focusing on a task may not pay full attention to the external environment, and may potentially endanger themselves. While no participants were harmed during the experiment, one participant who was wearing a hat lost it in the tree branch while walking past, requiring the trial to be rerun.

Clothing. Some participants asked to wear sunglasses, hats and other items while walking outside. These items may affect factors relevant to the experiment, such as the legibility of text while wearing sunglasses, and should be taken into account when planning an experiment.

In some cases, addressing these issues required repeating trials, pausing the study, or asking pedestrians to move. Our results also show that variable factors such as sunlight increased the variance of our data. This can be a particular problem when the evaluation is dependent on small time measurements. However, these studies provide much richer models of interruption, limited attention, and walking effort that can be achieved in a lab study. Outdoor empirical studies can thus serve as the strictest test of a proposed design, as it provides both the most realistic conditions and the highest degree of variability.

8. FUTURE WORK

Performance was less than optimal using the current prototype. We expect that further revisions to the interface might provide better performance. Two potential improvements are providing more drastic changes to the interface while the user is moving, such as removing uncommon functions, and in personalizing interface element sizes based on the user's hand size.

Extending the current experimental design to cover a wider variety of tasks would provide richer information about the effects of walking on task performance. While prior research has identified the existence of situational impairments caused by walking, an in-depth study of walking effects across several tasks would extend our knowledge about the causes and effects of situational impairments on walking.

Finally, the current study only addresses standing-walking transitions that occur between tasks. A more sophisticated WUI would allow a user to begin a task while standing and complete it while walking, or vice-versa. In the future, we may extend current and upcoming WUI design techniques to handle transitions between standing, walking, and other activities.

9. CONCLUSION

We conducted two experiments that measured the effects of different sized static and adaptive interfaces as a user interacted with a mobile device while standing and walking. Our first study showed an interaction between movement and interface size that suggests that altering size may be an effective strategy for reducing the negative effects of use while walking. Our second study evaluated a prototype that implements this strategy by changing target size as the user walks. The second study showed that walking interfaces are feasible, but they must be parameterized for individual differences, and must target tasks that are affected the greatest by walking. Also, for simple level-of-detail changes, there seemed to be no additive penalty for the adaptation itself.

In this paper, we have focused on walking as an activity that may impair our ability to interact with mobile devices. We introduce the term walking user interfaces (WUIs) to describe design features that explicitly support use while walking, and introduce a prototype media player that supports use while walking through interface scaling. We describe two formative studies that explore the effectiveness of WUIs while walking in public spaces, and illustrate the importance of and issues surrounding mobile device evaluation in the field.

10. ACKNOWLEDGMENTS

The authors thank Sunny Consolvo, Beverly Harrison, and Keith Mosher for assisting with the experiment. This work was supported in part by Intel Research.

11. REFERENCES

[1] Barnard, L., Yi, J.S., Jacko, J.A. and Sears, A. (2007). Capturing the effects of context on human performance in mobile computing systems. *Personal and Ubiquitous Computing, 11 (2*), 81-96.

[2] Brewster, S. (2002). Overcoming the lack of screen space on mobile computers. *Personal and Ubiquitous Computing, 6 (3*), Springer, 188-205.

[3] Brewster, S., Lumsden, J., Bell, M., Hall, M. and Tasker, S. (2003). Multimodal 'eyes-free' interaction techniques for wearable devices. In *Proc. CHI '03*. New York: ACM Press, 473-480.

[4] Dong, L., Watters, C. and Duffy, J. (2005). Comparing two one-handed access methods on a PDA. In *Proc. MobileHCI '05*. New York: ACM Press, 235-238.

[5] Findlater, L. and McGrenere, J. (2004). A comparison of static, adaptive, and adaptable menus. In *Proc. CHI '04*. New York: ACM Press, 89-96.

[6] Karlson, A.K. and Bederson, B.B. (2007). ThumbSpace: Generalized one-handed input for touchscreen-based mobile devices. In *Proc. INTERACT 2007*. New York: Springer, 324-338.

[7] Kjeldskov, J. and Stage, J. (2004). New techniques for usability evaluation of mobile systems. *International Journal of Human-Computer Studies, 60 (5-6)*. Amsterdam: Elsevier, 599-620.

[8] Kristoffersen, S. and Ljungberg, F. (1999). "Making place" to make IT work: Empirical explorations of HCI for mobile CSCW. In *Proc. GROUP '99*. New York: ACM Press, 276-285.

[9] Lin, M., Price, K.J., Goldman, R., Sears, A. and Jacko, J. *(*2005). Tapping on the move – Fitts' Law under mobile conditions. In *Proc. IRMA '05*. Hershey, PA: Idea Group Publishing, 132-135.

[10] Lin, M., Goldman, R., Price, K.J., Sears, A. and Jacko, J. (2007). How do people tap when walking? An empirical investigation of nomadic data entry. *International Journal of Human-Computer Studies, 65 (9)*, 759-769.

[11] Mackay, B., Dearman, D., Inkpen, K. and Watters, C. (2005). Walk 'n scroll: a comparison of software-based navigation techniques for different levels of mobility. In *Proc. MobileHCI '05*. New York: ACM Press, 183-190.

[12] McGrenere, J., Baecker, R.M. and Booth, K.S. (2002). An evaluation of a multiple interface design solution for bloated software. In *Proc. CHI '02*. New York: ACM Press, 164-170.

[13] Mitchell, J. and Shneiderman, B. (1989). Dynamic versus static menus: an exploratory comparison. *SIGCHI Bulletin, 20 (4)*, 33-37.

[14] Mitrovic, N., Royo, J.A. and Mena, E. (2007). Performance analysis of an adaptive user interface system based on mobile agents. In *Proc. DSV-IS '07*. New York: ACM Press, 29-44.

[15] Mizobuchi, S., Chignell, M. and Newton, D. *(*2005). Mobile text entry: relationship between walking speed and text input task difficulty. In *Proc. MobileHCI '05*. New York: ACM Press, 122-128.

[16] Murray, M., Drought, A.B. and Kory, R.C. (1964). Walking patterns of normal men. *The Journal of Bone and Joint Surgery, 46 (2)*, 335-360.

[17] Mustonen, T., Olkkonen, M. and Hakkinen, J. (2004). Examining mobile phone text legibility while walking. In *Proc. CHI '04*. New York: ACM Press, 1243-1246.

[18] Newell, A.F. and Gregor, P. (2000). "User sensitive inclusive design"– in search of a new paradigm. In *Proc. CUU '00*. New York: ACM Press, 39-44.

[19] Parhi, P., Karlson, A.K. and Bederson, B.B. (2006). Target size study for one-handed thumb use on small touchscreen devices. In *Proc. MobileHCI '06*. New York: ACM Press, 203-210.

[20] Pascoe, J., Ryan, N. and Morse, D. (2000). Using while moving: HCI issues in fieldwork environments. *ACM Transactions on Computer-Human Interaction, 7 (3)*, 417-437.

[21] Sears, A., Lin, M., Jacko, J. and Xiao, Y. (2003). When computers fade: Pervasive computing and situationally-induced impairments and disabilities. In *Proc. HCI Int'l '03*. Mahwah, NJ: Lawrence Erlbaum Associates, 1298-1302.

[22] Vermunt, J.K. (1997). *Log-linear Models for Event Histories*. Thousand Oaks, CA: Sage Publications.

[23] Yi, J.S., Choi, Y.S., Jacko, J.A. and Sears, A. (2005). Context awareness via a single device-attached accelerometer during mobile computing. In *Proc. MobileHCI '05*. New York: ACM Press, 303-306.

[24] Zhao, S., Dragicevic, P., Chignell, M., Balakrishnan, R. and Baudisch, P. (2007). Earpod: eyes-free menu selection using touch input and reactive audio feedback. In *Proc. CHI '07*. New York: ACM Press, 1395-1404.