Computers in Human Behavior 26 (2010) 427–435

Contents lists available at ScienceDirect

Computers in Human Behavior

journal homepage: www.elsevier.com/locate/comphumbeh

Small-device users situationally impaired by input

Yeliz Yesilada [a], Simon Harper [a,*], Tianyi Chen [a], Shari Trewin [b]

[a] Human Centred Web Laboratory, University of Manchester, Manchester, UK
[b] Accessibility Research, IBM T J Watson Research Center, New York, USA

ARTICLE INFO

Article history:
Available online 4 January 2010

Keywords:
Motor-impaired
Dexterity-impaired
Small-device
Typing
Pointing
Input
Errors

ABSTRACT

Users of small computational devices, such as Mobile telephones or Personal Digital Assistants, are situationally impaired by both the device and the context of the device's use. This paper describes empirical work which makes the link between the behaviour of motor impaired desktop users and non-impaired users of small-devices. This is important because it may, therefore, be possible to leverage existing solutions for motor-impaired users into the small-device domain. We find that there is significant overlap in the extent of the problems encountered, but not the magnitude. Eight of the 11 existing errors made by motor-impaired users were also present in our small-device study in which two additional error types, key ambiguity and landing errors, were also observed. In addition, small-device rates for common error types were higher than those of desktop users with no impairment, but lower than those of desktop users with motor impairments. We suggest that this difference is because all users were seated to maintain constancy between studies and assert that this magnitude difference will equalise once the small-device is used in a mobile context.

1. Introduction

Very often, technology developed in support of people with disabilities proves beneficial to a broader population. 'Small-device'[1] users experience difficulties in typing and pointing because of the limited input facilities of these devices. Manufacturers of these devices have therefore adopted text entry techniques such as T9 and word prediction previously developed within the accessibility community (Pavlovych et al., 2004).

Sears and Young (2003) point out that able-bodied individuals can be affected by both the environment in which one is working and the activities in which that person is engaged, resulting in situationally-induced impairments. For example, an individual's typing performance may decrease in a cold environment in which one's finger does not bend easily due to extended exposure at low temperature. Wobbrock (2006) raises the issue that able-bodied people will have to deal with situationally-induced impairments as small-devices permeate their lives; and it may be useful to investigate whether dexterity impairments and situational impairments affect users in a similar way. Anecdotally, Trewin et al. (2006) suggests that there are strong similarities between physical usability issues in both small-device and 'accessible'[2] desktop Web browsing scenarios. Both small-devices and accessible Web Browsers share the need to support various input techniques; they both benefit from flexibly authored and accessible Web pages; and text entry and navigation in both scenarios are slow and error-prone. However, there is no empirical evidence to show whether these similarities extend to input errors.

Looking at these commonalities and potential solution migrations, we are concerned with gaps or duplicated work, which exist within the mobile and accessibility areas. The interdisciplinary nature of these domains along with the speed at which they are developing suggest to us that there is high possibility of research scientists missing related or overlapping work and therefore 'reinventing the wheel'. The work presented in this paper aims to empirically show that small-device users and 'motor-impaired' users[3] share common input errors and thus that it is worth migrating solutions between these two domains.

Originally, Trewin and Pain (1999) conducted an empirical user study in order to investigate the input problems experienced by desktop users with motor impairments. This study shows that there are categories of errors common to these users, such as pressing a key for too long which generates unwanted keystrokes (long key press), or pressing two adjacent keys in reverse order

* Corresponding author. Tel.: +44 1612750599.
E-mail address: simon.harper@manchester.ac.uk (S. Harper).

[1] By 'Small-Device', we mean a mobile device such as a Mobile Telephone, Personal Digital Assistant, or any other hand-held device with a small keyboard and display. Further, in the context of this paper the user of the small-device will always be in a static seated position; and not in a mobile context or moving environment unless otherwise stated.

[2] By 'Accessible', we mean having the ability to support access by disabled users; in this case motor impaired users with a physical disability which affects dexterity.

[3] By 'motor impairment', we mean a user with a physical disability which affects dexterity.

0747-5632/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.chb.2009.12.001

(transposition error). This study also demonstrates that desktop users with motor impairments have problems in manipulating the mouse, which causes errors in clicking, multi-clicking and dragging. Here we replicate this study with small-device users to investigate whether these common typing and pointing errors are also applicable to small-device users.

When we replicated Trewin and Pain's study we found that there were common typing errors between small-device and desktop users with motor impairments such as: missing certain key presses; pressing keys that are adjacent to the target key; or pressing keys unintentionally. In addition, there are common pointing errors such as: clicking in the wrong places; sliding the stylus during multiple clicks; and making errors when dragging. Some typical errors of motor-impaired users, such as long key press errors, are not very common in small-device users. Finally, the small-device study also confirms that there are some errors specific to small-device users, which were not observed in the original study. On a typical small-device keypad, different characters typically share a key and the user has to use a function key to find the right character. Therefore, the user might press the right key but type the wrong character; key ambiguity error. We also observed that small-device users experienced landing errors, which occur when the pen is placed at one position for a certain period of time without being moved or lifted up. Indeed, the small-device study found that the key ambiguity error is the most frequent typing error and the landing error is the most frequent pointing error.

The following sections describe existing work on input errors in the desktop and small-device context (see Section 2.1), describe the original desktop study (see Section 2.3) and the small-device replication of this study (see Section 3), showing that there are common input errors shared by both small-device and desktop users with motor impairments (see Section 5), Further, we generalise these findings and show that by replicating an empirical investigation across domains, following a similar methodology, we can uncover commonalities between the two (see Section 6). Finally we assert that these commonalities provide a route for solution transfer between domains (see Section 7).

2. Background

Existing literature suggests that there are common experiences between small-device and dexterity impaired users. However, other than reports of 'off-by-one' errors (see Section 2.1), there is little empirical evidence to support such common errors even when a high number of errors both in extent and magnitude are recorded in singular user domains (Trewin, Keates, & Moffatt, 2006; Keates, Trewin, & Paradise, 2005) (see Section 2.3).

2.1. Common user experiences

Existing literature suggest that there are strong similarities between the usability challenges faced by people with motor impairments and small-device users. Dexterity impairment causes a number of challenges that affect users' interaction such as pointing to a target or clicking on a target (Trewin et al., 2006). However, for small-device users, the challenge is not because of the severe physical disability, but because of the environment in which users are working in and the current context. We can say that compared to dexterity impaired users, small-device users experience situationally-induced impairments which occur temporarily (Sears & Young, 2003).

A dexterity impairment can have high impact on accuracy, completion times and error rates of clicking and pointing to a target. For example, arthritis, which is an inflammation of the joints, might prevent a user from being able to make necessary movements. Similarly, spasms, which are a sudden, involuntary contraction of a muscle, can cause unwanted diversions or mouse clicks during pointing movements. People with Parkinson's disease click more slowly, giving even greater opportunity for slippage (Trewin et al., 2006). Further, a situationally-induced impairment can also have an high impact on accuracy, completion times and error rates of clicking and pointing to a target. For example, Barnard, Yi, Jacko, and Sears (2005) show that lighting changes have highly noticeable effects on small-device users' behaviour. In a follow-on study, Lin, Goldman, Price, Sears, and Jacko (2007) show that the overall task completion time, error rate, and several measures of workload differ significantly between various mobility conditions, which include seated, walking on a treadmill, and walking through an obstacle course. They show that error rates increased dramatically when the participants walked through the obstacle course.

In general, dexterity impairments may cause users to press unwanted keys, or miss the keys they intended to press. Bradykinesia, describes slowness in the execution of movement, and causes typing rates in a physical keyboard to be greatly reduced (Sears & Young, 2003). Similarly, Brewster (2002) shows that when a small-device is used in a more realistic situation (whilst walking outside), usability is significantly reduced (with increased workload and less data entered) than when used in a laboratory.

Finally, Trewin and Pain (1999) conducted a study of keyboard and mouse input errors made by 20 people with motor impairments affecting dexterity.The focus of the study was physical input errors, dubbed performance errors, and not cognitive and task errors such as misspellings. This study serves as the baseline for our work, and hence, will be analysed in detail in the following section; Section 2.3.

2.2. Key ambiguity error

It is useful to single out Key ambiguity as it occurs most frequently in Small-Devices (see Section 4 — *key ambiguity error*). Key ambiguity errors occur on devices where each key may represent multiple characters and users fail to distinguish the target letter from the other letters. Assigning more than one character to a key is a common approach for designing small keyboards, especially when space is a concern or when language uses thousands of characters (MacKenzie & Tanaka-Ishii, 2007). Therefore, key ambiguity is a typical error of small-device users (Butts & Cockburn, 2002; James et al., 2001; Wigdor et al., 2003; Damerau, 1964). Typing errors on both desktop and small-devices are often classified as insertion errors (an extra character was added), deletion errors (a character is missing), substitution errors (deletion and insertion), or transposition errors (two characters are swapped) (MacKenzie & Soukoreff, 2002; Clawson, Lyons, Rudnick, Iannucci, & Starner, 2008); this classification includes both physical and cognitive errors. A number of techniques have been introduced to reduce key ambiguity errors, which include:

Multi-tap, a user has to press a key repeatedly to pick up the target letter from the other ones on the same key (Pavlovych et al., 2004);
Dictionary-based text entry, such as T9, iTap and eZitext, disambiguate input based on a dictionary (Pavlovych et al., 2004);
LetterWise, works like dictionary based methods and allows entry of non-dictionary words without switching to another input mode (MacKenzie, Kober, Smith, Jones, & Skepner, 2001);
TiltText, requires the user to tilt the phone in one of four directions while pressing a key in order to select one character from a group (Wigdor et al., 2003); and
Pen-based text entry, allows user to write letter shapes with a stylus. One such technique uses the edges of the writing area for stabilization (Wobbrock, Myers, & Kembel, 2003).

'Off-by-one' errors, in which a key adjacent to the intended key is pressed, have been reported as a major source of typing error on mini-qwerty keyboards (Clawson et al., 2008). This error definition includes insertions, substitutions, and also repeats of the same letter. A solution based on keystroke timing has been shown to correct up to 32% of mini-qwerty off-by-one typing errors (Clawson et al., 2008).

2.3. Original study

The Trewin and Pain (1999) study was conducted on the desktop and compared 20 people with motor impairments to six with no impairment.[4] The authors used a Macintosh 475 8/150 desktop machine with either a domestic or ISO extended Apple keyboard and a single button Macintosh mouse. It involved a typing task, a mouse task, and an editing task. The typing task required participants to copy a text passage that was presented on paper next to the monitor; requiring a minimum of 547 key presses. Copy typing was used to reduce the opportunity for spelling and language errors, and to make error identification easier. The passage was typed twice, once with and once without error correction. The pointing task consisted of 17 pointing, clicking, double clicking and dragging activities and was also presented twice. The activities were performed on a text passage in which all but the target words were obscured. They involved targets of varying sizes in different locations on the screen. The third task was a more realistic text editing task involving both typing and mouse use. It consisted of 27 typing and mouse operations, including the use of scroll bars, selection from hierarchical menus, use of a dialog box and use of the 'Command' key. The activities were based on a text passage containing errors to be corrected.

The study identified seven performance errors related to typing and also observed that Desktop-Impaired users had problems in manipulating the mouse. This caused errors in aiming, clicking, multi-clicking and dragging and are further described in Table 1, in order of frequency of occurrence, with error rates for the typing and pointing tasks are reported in Table 5.

In the Desktop-Impaired group, 14 participants[5] completed both typing tasks both with[6] and without[7] error correction. The mean keystrokes for *T*1 was 585 and for *T*2 was 649. The mean time to type the given passage for *T*1 was 840 s, and for *T*2 was 1008 s. The average time spent on corrections of performance errors in *T*2 was 7.3% of the total time spent.

In the Desktop-Not-Impaired group, five participants completed both typing tasks.[8] The mean keystrokes for *T*1 was 566 and for *T*2 was 569. The mean time to type the given passage for *T*1 was 294 s, and for *T*2 was 360 s. The average time spent on corrections of performance errors in *T*2 was 1.7% of the total time spent.

The study did not report average aiming error rates (missing data in Table 5), but found that 14 of 20 participants in the Desktop-Impaired group had aiming error rates greater than 10%, and eight had error rates over 20%. In contrast, the highest aiming error rate observed in the Desktop-Not-Impaired group was 6.3%. While the Desktop-Not-Impaired group performed the pointing task significantly faster the second time, the Desktop-Impaired group did not. Clicking, multi-clicking and dragging error rates were also high in the Desktop-Impaired group, with dragging being especially difficult. However, we note that the clicking rates reported here are not errors but the proportion of clicks in which the mouse was moved during the click. Usually this did not produce an error, but it provides an indication of the accuracy of clicking actions. In addition, the dragging errors here excluded aiming errors at the start of the drag.

3. Small-device users

When we replicated Trewin and Pain's original study, we followed the same methodology, which is summarised in Section 2.3, with some minor updates to account for the change from desktop computer to small-device (described throughout this section, particularly in Section 3 — *Tasks*).

Hypotheses: To examine whether input errors for small-device users and desktop users with dexterity impairment are similar, the following four overarching hypotheses were investigated.

1. The same forms of typing and pointing performance error will be observed on the small-device as were found for the standard keyboard and mouse in the desktop study.
2. Error rates for people with no motor impairment will be higher on the small-device than the desktop, but desktop users with motor impairments will have the highest error rates.
3. The relative frequencies of these errors will be the same.
4. Practice will reduce error rates on the small-device, though it does not on the desktop for people with motor impairments.

Participants: A total of 15 participants (5 female and 10 male) aged between 19 and 44 took part. None had any dexterity impairment. Participants were asked to rate their previous experiences on text entry and the mobile Web (1 = none; 5 = expert). On average, they rated their text entry experience 3.47 (sd = 0.99) and mobile Web experience 2.07 (sd = 1.22). Participants were also asked to specify the pointing techniques that they are familiar with and all our participants except one indicated that they are familiar with the stylus. Only one of the participants was lefthanded (Chen, Yesilada, & Harper, 2008).[9]

Tasks: The sequence of tasks and tasks themselves were the same as Trewin and Pains study. However, the differences between the desktop and small-devices meant that the desktop study could not be precisely replicated. In adapting the desktop study to the small-device environment, the aim was not to assess the usability of an application, but to identify users' errors and experience when performing typical tasks. Therefore, where a desktop task referred to mouse-specific or keyboard-specific items or actions, this was either adapted to the most reasonable equivalent on the small-device, or dropped from the study. The following three tasks were used in the original study:

1. *Typing tasks*: Participants were given a text passage to type. The small-device study used the original passage, which deliberately included characters on all parts of the keyboard (see Appendix C in Chen et al. (2008)). It required a minimum of 553 key presses on the small-device.
2. *Pointing tasks*: Participants were asked to conduct 15 sub-tasks, which included tapping, multi-tapping and dragging with a stylus on the PDA. Three changes were made to the original pointing sub-tasks: a click on a ClarisWorks application button in the top left of the screen was replaced by a tap on a Wiki button of

[4] Now referred to as the 'desktop study', with the desktop users referred to as the 'Desktop-Impaired' group, and the comparison group as the 'Desktop-Not-Impaired' group.

[5] An additional participant completed only the first Task due to fatigue and lack of time.

[6] Task 1—*T*1.

[7] Task 2—*T*2.

[8] Again, an additional participant completed only the first Task due to slow typing speeds and a lack of time.

[9] Detailed information about the study, can be found at: http://hcw-eprints.cs.man.ac.uk/81/.

Table 1
Errors affecting desktop users with motor impairments.

Typing errors
Long key press error: an alphanumeric key was pressed for longer than the active key repeat delay, producing unwanted characters
Additional key error: a key near to the intended key was activated by the same finger, either instead of, or in addition to the intended key
Missing key error: an attempt to press a key that failed, either because the participant's aim is off target, or because the key is not pressed with sufficient force
Dropping error: failure to press two keys simultaneously, e.g. when using the Shift key
Bounce error: a key is unintentionally pressed more than once, producing unwanted characters
Remote error: accidental key presses not caused by poor aiming. Includes leaning on part of the keyboard
Transposition error: two characters are typed in each others place
Pointing errors
Aiming error: a click at the wrong position
Clicking error: a slip while clicking, producing mouse movement that may cause an error
Multi-clicking error: a multi-clicking error occurs when the cursor or pointer slides between clicks, too long a gap is left between clicks, or the wrong number of clicks are made
Dragging error: the mouse button is released before reaching the target

similar size and position; one mouse-specific sub-task (repositioning the mouse on the mousepad) was omitted; and one multi-clicking sub-task that had no specific target was omitted.

3. *Editing tasks*: Participants were asked to make a set of edits to a given text passage. The original tasks and passage used in the desktop study were used, with some modifications. The text passage to be edited was reflowed to fit the width of the PDA screen. One of the original tasks required the use of the Apple modifier key on the Macintosh keyboard. This was replaced with a use of the Control modifier key on the PDA keyboard (see Appendix E in Chen et al. (2008)).

Apparatus: An HP iPAQ hw 6900 was used as a *typical*[10] PDA. Fig. 1 shows the overall layout of this iPAQ keypad. Keys on the keypad repeat in the same way as those on a standard keyboard, with an initial delay of 500 ms before the repeat starts. Modifier keys operate in latch mode. The keyboard itself has three modes—one for letters, one for numbers, and one for punctuation. The mode determines what character is produced for a given key press. Thus, switching between typing letters, numbers and punctuation provides an opportunity for a new form of error—the key ambiguity error, discussed later. Selection was done by tapping a stylus on the touch screen and is referred to as 'tapping' to distinguish it from clicking, which is selection using a mouse button. A double tap is performed by tapping the stylus twice on the screen with a period of less than 500 ms between taps.

In the original study a Macintosh Desktop computer was used with the ClarisWorks word processor along with recording software called InputLogger. The small-device study used an instance of a Wiki (i.e., MediaWiki) which is open source and a proxy logger which is called UsaProxy (Atterer, Wnuk, & Schmidt, 2006). UsaProxy sat between the Web server and the browser, and logged every input event a participant made on the webpages. The Wiki page was designed to behave like a word processor and included all the functionality tested in this study.

Procedure: This study was conducted in a usability lab so that the use of the small-device could be examined in a stable and static position. The overall study was divided into three sessions: a background session, the main task session, and a feedback session. In the background session, an experimenter asked a participant background questions, such as age and previous experience. After the background session, the participant had 5 min to practice with the PDA. This was to ensure that the participant was familiar with the device before conducting the main tasks. In the main task session, the participant was asked to conduct a typing task, a pointing task and an editing task in sequence, and then repeat the pointing task and the typing task. Therefore, in total, there were five tasks to conduct. The repetition was to minimize other errors such as a misunderstanding of the task. Participants were not allowed to correct their errors in the first trial of typing and pointing tasks. When an error occurs, the participant would ignore it and carry on. However, in the repetitions, error correction was allowed. Finally, the participant was asked to rate the typing and pointing performance.

[10] See 'Default Delivery Context', http://www.w3.org/TR/small-device-bp/.

4. Results

This section presents the errors observed in the small-device study with the typing and pointing tasks.

4.1. Typing errors

All participants completed both typing tasks, once with ($T1$) and once without error correction ($T2$). The mean keystrokes for $T1$ was 569.4 (sd = 41.01) and for $T2$ was 590.87 (sd = 45.20). The mean time to type the given passage for $T1$ was 667.47 (sd = 183.77) and for $T2$ was 549.87 (sd = 213.98). The average time spent on corrections in $T2$ was 17.15 sec, which was about the 3% of the total time spent. Although participants were asked not to correct their errors in $T1$, three participants still did so. Overall, six types of typing error were observed in this study: key ambiguity, missing key, additional key, bounce, transposition and long key press error (see Table 2 and Fig. 2).

Key ambiguity error: This was the most frequent typing error in the small-device study, previously reported in other studies (Silververg, MacKenzie, & Korhonen, 2000) (see Section 2.2 for a full explanation). A key ambiguity error occurs when the participant fails to switch to the appropriate keyboard mode before pressing a key. The key pressed is correct but the mode error means that the wrong character is generated. All participants experienced key ambiguity errors in both $T1$ and $T2$, except one participant did not make any key ambiguity error in $T1$. The mean key ambiguity error per person per passage for $T1$ was 9.33 (sd = 1.423) and for $T2$ was 6.33 (sd = 1.355). These error rates were even higher than the sum of the error rates of all the other types of typing errors (see Table 1). There was also a significant inverse correlation between participants' key ambiguity errors in $T1$ and their subjective ratings on previous experiences in text entry (Spearman $\rho = -0.519$, $p = 0.047$, 2-tailed). This suggests that the more experienced small-device users rated themselves, the less key ambiguity errors they made.

Missing key error: All participants experienced missing key error both in $T1$ and $T2$. The mean missing key error per person per passage was 3.20 for $T1$ (sd = 0.745) and 2.53 for $T2$ (sd = 0.593). These results also include the characters that the participants simply forgot to type.

Fig. 1. HP iPAQ hw 6900 keypad.

Table 2
Typing errors observed in the small-device study; average error per person.

Typing error	Task 1	Task 2
Key ambiguity	9.33 (sd = 1.423)	6.33 (sd = 1.355)
Missing key	3.20 (sd = 0.745)	2.53 (sd = 0.593)
Additional key	1.40 (sd = 2.257)	1.40 (sd = 1.595)
Bounce	0.73 (sd = 0.284)	0.33 (sd = 0.159)
Transposition	0.33 (sd = 0.126)	0.20 (sd = 0.107)
Long key press	0.07 (sd = 0.260)	0.60 (sd = 1.300)

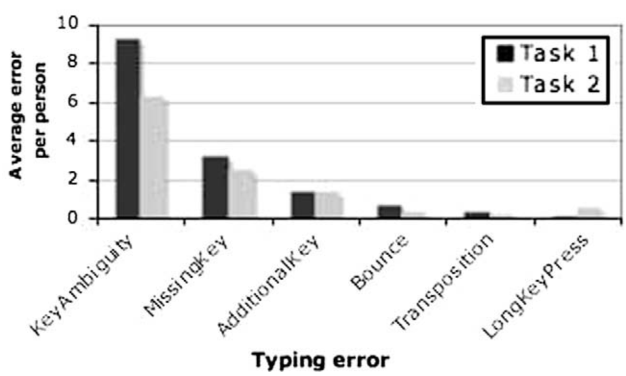

Fig. 2. Typing errors in the small-device study.

Additional key error: Eleven participants in $T1$ and nine participants in $T2$ experienced additional error. The mean error per person per passage was the same for $T1$ and $T2$, which was 1.4 (sd = 2.257 for $T1$ and sd = 1.595 for $T2$). There was a significant positive correlation between participants' subjective ratings on avoiding additional key errors and the actual errors they made in $T1$ (Spearman $\rho = 0.517$, $p = 0.048$, 2-tailed). This indicates that the more confident participants were, the more errors they made.

Bounce error: Six participants in $T1$ and four participants in $T2$, in overall nine participants experienced bounce error. The mean error rate per person per passage was 0.73 (sd = 0.284) for $T1$ and 0.33 (sd = 0.159) for $T2$.

Transposition error: Five participants in $T1$ and three participants in $T2$, overall, seven participants experienced transposition error. The mean error rate per person per passage was 0.33 for $T1$ (sd = 0.126) and 0.20 (sd = 0.107).

Long key press error: Overall, only three participants experienced long key press error, which was only 20% of the participants. The mean error rate per person per passage was 0.07 for $T1$ (sd = 0.26) and 0.6 for $T2$ (sd = 1.30).

4.2. Pointing errors

The average time to complete the pointing tasks without corrections ($P1$) was 359.13 s (sd = 128.12); and with corrections ($P2$) was 238.73 s (sd = 82.89). Participants spent significantly less time in $P2$ than in $P1$ (Wilcoxon Signed Ranks Test, $Z = -3.351$, sig. = 0.001). In addition, participants spent 757.40 s (sd = 258.33) to complete the editing tasks that involved pointing. Although error correction was not recommended in $P2$, participants still spent 9.38% of their time to correct pointing errors. However, this dropped to 8.53% when error correction was introduced in $P2$. According to the feedback gathered after pointing tasks, participants thought dragging was the most difficult and error-prone, followed by tapping and multi-tapping. Five types of pointing errors were observed: aiming, landing, drag breaking, multi-tap moving, tap count (see Table 3 and Fig. 3).

Aiming: This error happened when the pen was off target. It affected tapping and dragging. Eight participants experienced a total of 27 aiming errors in three tasks. Participant N11 had eight aiming errors in-all. He particularly struggled with the task that requested a single tap between the letter 'i' and the letter 'l' in the word 'April'.

Landing error: Occurred when the pen was placed at one position for a certain period of time without being moved or lifted up. On the device used in this study, a pop-up menu would be displayed upon landing. This would interrupt a tapping or dragging action. Landing error was the most frequent pointing error. A total number of 126 landing errors were observed. Participant N9 had 29 landing errors in $P2$. He tried more than 10 times to drag the pen to select a word 'a', and failed all of them because of the landing error.

Drag breaking: This error happened during a dragging action when a participant lifted up the pen before reaching the end. We observed 12 drag breaking errors. Eight of them occurred in the first trial of the pointing tasks ($P1$).

Multi-tap moving: This error happened when the pen slid between taps. This interrupted the multi-tap action. 29 multi-tap moving errors were observed in $P1$ and $P2$, none was observed in the editing tasks. We observed that participants were more likely

Table 3
Mean pointing error rates observed in the small-device study; $P1$: without corrections & $P2$: with corrections.

Pointing error	$P1$	Editing	$P2$
Aiming	0.53 (sd = 0.92)	0.07 (sd = 0.26)	1.20 (sd = 1.78)
Landing	4.40 (sd = 7.36)	2.80 (sd = 3.00)	3.33 (sd = 3.27)
Drag breaking	0.53 (sd = 0.64)	0.20 (sd = 0.41)	0.07 (sd = 0.26)
Multi-tap moving	1.00 (sd = 0.93)	0.00 (sd = 0.00)	0.93 (sd = 1.22)
Tap count	0.47 (sd = 0.83)	0.13 (sd = 0.35)	0.27 (sd = 0.59)

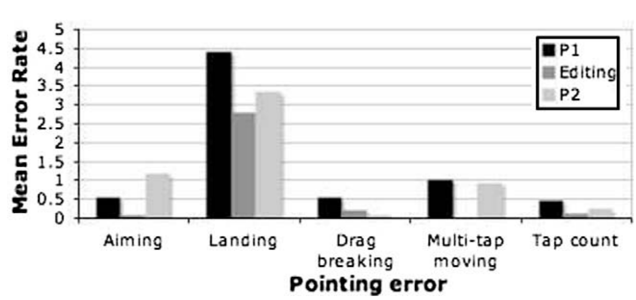

Fig. 3. Pointing errors of small-device users ($P1$: without corrections and $P2$: with corrections).

to slide the pen when conducting triple-tapping than double-tapping.

Tap count: In some cases, participants tapped the wrong number of times. For example, participant N15 performed double-taps on the word 'April' when he was asked to tap between the letter 'i' and letter 'l' of the word. Thirteen tap count errors were observed, seven in *P*1, four in *P*2 and the other two in the editing tasks.

We found that our participants had significantly less error rate when they repeated the tasks that required multi-taps (Wilcoxon Signed Rank Test, $Z = -2.170$, sig. = 0.030). This indicates that practice helps multi-taps performance. However, this does not hold for tasks that required a single tap or dragging.

5. Comparison

The desktop and small-device studies involved very different users and devices but a similar set of tasks. This allows some comparison of the results for the three participant groups:

1. Desktop users with dexterity impairment (Desktop-Impaired);
2. desktop users with no impairment (Desktop-Not-Impaired); and
3. small-device users with no impairment (Small-Device).

5.1. Typing errors

Participants typed an identical text passage in both studies. On the small-device, the passage required an additional six key presses to complete (total 553) because of the need to switch keyboard modes. When we look at the time taken to type the given passage, the Desktop-Impaired group spent the longest time and the Desktop-Not-Impaired grouped spent the shortest time both with and without corrections (see Fig. 4). Therefore time-wise the Small-Device group is in between. The Small-Device group typed the passage without correction (*T*1) in 79% of the time taken by the Desktop-Impaired group, and 227% of the Desktop-Not-Impaired groups time. When error correction was introduced, Small-Device participants took 54% of the time taken by the Desktop-Impaired group, and 153% of the Desktop-Not-Impaired group's time (see Table 4). The total number of keystrokes typed in each task were similar, with the exception that the Desktop-Impaired group typed more characters when correcting errors (see Fig. 5).

Hypothesis 1 Section 3 (Error types): Hypothesis one is supported by the data. Of the seven types of performance error observed in the desktop study, five were also observed in the small-device study: long key press errors, additional key errors, missing key errors, bounce errors and transposition errors. Dropping errors were not possible in the small-device study because there is no requirement to hold down one key while pressing another. Remote errors were only observed in the Desktop-Impaired group. One new error type specific to the small-device was observed: the key ambiguity error. Thus, hypothesis one is largely supported, but some error types are platform/population specific.

Table 4
Comparison of typing tasks without (*T*1) and with corrections allowed (*T*2).

Metric	Desktop-Impaired		Desktop-Not-Impaired		Small-Device	
	*T*1	*T*2	*T*1	*T*2	*T*1	*T*2
Total time (s)	840	1008	294	360	667	549
Total keystrokes	585	649	566	569	569	590
Correction time (s)	5.88	73.58	0.00	6.12	6.43	17.15

Hypothesis 2 Section 3 (Error rates): Hypothesis two is supported by the data. Table 4 compares error rates for the observed typing error types, showing the mean number of errors per person per text passage typed for the three participant groups. Error rates of the Small-Device group are typically higher than for the Desktop-Not-Impaired group, but lower than the Desktop-Impaired group. This pattern is also reflected in the time spent correcting performance errors in the second typing task, which was 3% for the Small-Device group, compared with 7.3% for the Desktop-Impaired and 1.7% for the Desktop-Not-Impaired group.

Hypothesis 3 Section 3 (Relative frequencies): Hypothesis three is not supported; however see Section 7. The relative frequencies of the observed errors showed important differences. Key presses longer than the default key repeat delay were the overriding source of performance error for the Desktop-Impaired group, but in the Small-Device group a very low error rate was observed. The key ambiguity error was by far the most common small-device error. Additional key errors and missing key errors were the second and third most frequent forms of error in the Desktop-Impaired and Small-Device groups, and the two most frequent in the Desktop-Not-Impaired group.

Hypothesis 4 Section 3 (Practice effect): Hypothesis four is not supported. Typing performance did not significantly increase for the Small-Device group when it was repeated. Although the error rates were in general lower when the small-device users repeated the typing tasks, the results were not significant. Therefore, we cannot support this hypothesis.

5.2. Pointing errors

The overall pointing task times are not directly comparable, as the desktop study had two additional pointing tasks. However,

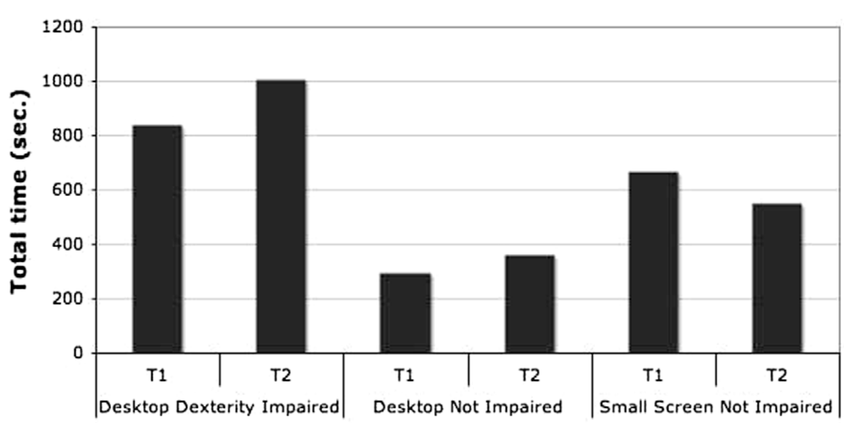

Fig. 4. Total time to complete typing tasks with (*T*1) and without (*T*2) corrections.

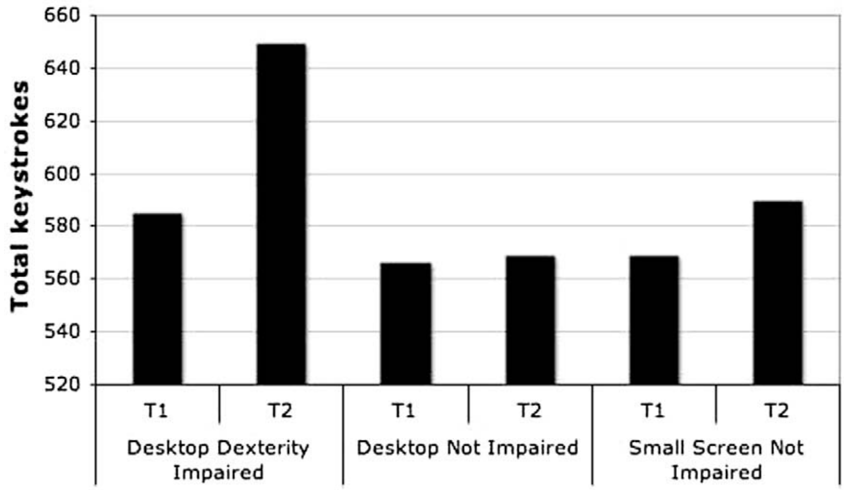

Fig. 5. Total keystrokes to complete typing tasks with ($T1$) and without ($T2$) corrections.

performance on individual tasks and observed errors can be compared. Because the same tasks were used, relative target sizes are comparable between the studies. Target locations are not, as the pointing and editing passages were reformatted to fit the width of the small-device screen.

Hypothesis 1 Section 3 (Error types): Hypothesis 1 is partially supported, with three of the four classes of desktop pointing performance error also observed in the small-device study, and one small-device-specific error type. These were aiming errors, multi-click errors, and drag breaking errors. Multi-click errors can be further broken down into errors caused by movement while clicking or tapping, pausing between clicks or taps, and clicking or tapping the wrong number of times. All three of these multi-clicking error types were observed in both studies. Stylus movement while performing a single tap was not observed in the small-device study, whereas mouse movement while clicking was a source of error in the desktop and other studies of mouse users with motor impairment (Trewin et al., 2006). Landing errors were specific to the small-device, as a pause on mouse down has no effect within the word processing application used in the desktop study.

Hypothesis 2 Section 3 (Error rates): Hypothesis two is supported by the data for aiming, drag breaking and multi-tap moving errors. Mean error rates for $P1$ and $P2$ reported in Table 3 are higher than the overall error rates for the Desktop-Not-Impaired group, but lower than those of the Desktop-Impaired group, as reported in Table 5. Mean data for aiming errors for the desktop study was not published, but the reported number of Desktop-Impaired participants with very high aiming error rates, and low maximum error rate for the Desktop-Not-Impaired group supports the hypothesis. However, hypothesis two did not hold true for clicking errors, which were only observed in the desktop study, or for tap count errors, which were higher in both desktop study groups than the small-device study.

Hypothesis 3 Section 3 (Relative frequencies): The data does not support the hypothesis that error frequencies are comparable across the groups; again see Section 7. The most difficult operations for the Desktop-Impaired group were dragging, multiple clicking, clicking and aiming, all of which had high error rates in the Desktop-Impaired group. For the Desktop-Not-Impaired group, multiple clicking had the highest error frequency, followed by dragging, then aiming. Note that the click movement reported here as a clicking error did not necessarily lead to an actual error in the click, and should be understood as an indication of the accuracy with which the device is being used rather than an error category.

Table 5
Comparison of typing and pointing error rates in keyboard and mouse tasks.

Type	Desk-Impaired	Desktop-Not-Impaired	Small-Device
Typing errors; mean errors per person per typing task			
Long key press	Projected 79.1	0.0	0.3
Additional key	8.0	1.0	1.4
Missing key	5.4	0.9	2.9
Dropping	0.6	0.1	–
Bounce	1.3	0.0	0.5
Remote	1.1	0.0	0.0
Transposition	0.1	0.2	0.3
Key ambiguity	–	–	7.8
Pointing errors; percentage (%) pointing error rate			
Aiming	–	–	5.8
Click/tap movement	28.1	6.3	0.0
Multi-click/tap moving	12.4	0.0	6.4
Tap count	9.9	6.0	2.5
Landing	–	–	26
Dragging	32.8	0.0	2.0

Actual click errors are dependent on the target size, location of the click and direction of slip. Landing errors were by far the most common problem in the Small-Device group but did not occur on the desktop. Among those errors that were shared, dragging on the small-device was less problematic than aiming or multi-clicking (see Table 6).

Hypothesis 4 Section 3 (Practice effects): Hypothesis 4 is not strongly supported for the pointing tasks. Pointing performance did not improve significantly when the Desktop-Impaired group repeated the pointing tasks, but it did for the Desktop-Not-Impaired group. For the Small-Device group, only multi-clicking error rates were significantly reduced when the task was repeated.

6. Discussion

Many error types were common across both studies, specifically the additional key and missing key errors, aiming, multi-clicking and dragging errors. Such errors could be expected in any device that requires users to press buttons or make selections on a screen. Given the small key size of the device (see Fig. 1), error rates were surprisingly low. However, this study examined the use of the small-device in a stable, static position. This does not produce the same level of situational impairment that is typically described in the literature when comparing small-device and accessibility

Table 6
Frequency of errors per person per task.

Order	Desktop–Impaired	Desktop-Not-Impaired	Small-device
Typing errors			
1	Long key press	Additional key	Key ambiguity
2	Additional key	Missing key	Missing key
3	Missing key	Transposition	Additional key
4	Bounce	Long key press, bounce	Bounce
5	Transposition	–	Transposition, long key press
Pointing errors			
1	Aiming	Clicking	Landing
2	Clicking	Tap count	Multi-click/tap
3	Dragging	Aiming	Aiming
4	Multi-click	–	Tap count
5	Tap count	–	Dragging

scenarios (Barnard et al., 2005; Brewster, 2002). It is likely that in this context the error rates found here would be higher, especially, if the task was repeated in typical mobile contexts such as on a train where both environment, device, and the users' motion are all factors. If error rates were significantly higher in this environment then existing accessibility solutions may be beneficial. For instance, a dynamic keyboard which self-adjusts its input mechanism such as key repeat and debounce time, may help small-device users to reduce long key press errors and bounce errors (Trewin, 2004). Further, other desktop systems for automatic spelling correction that take into account common typing error patterns and keyboard layout may also be applicable (Kane, Wobbrock, Harniss, & Johnson, 2008). In addition, an ability-based interface (Gajos, Wobbrock, & Weld, 2008), which automatically changes the size and layout of on-screen items to improve the mouse performance of desktop users with motor impairments, may also reduce the aiming, multi-tapping and landing errors of small-device users.

Existing research suggests that people do not usually use their small-devices to create or edit long documents because of the small screen size and limited input (Cui et al., 2008; Luo, 2004; O'Hara, Perry, Sellen, & Brown, 2002; Waycott & Kukulska-Hulme, 2003). Therefore, the tasks used in this study might not be typical typing and pointing tasks. Indeed, we asked our participants to 'copy-type', this may have introduced different errors compared to a text composition task, because the user must divide their attention between the passage to be typed, and the device itself. However, Oulasvirta, Tamminen, Roto, and Kuorelahti (2005) shows that small-device users usually do multi-tasking, which means users have to divide their attention between a number of tasks. In that respect this task was typical.

While current extended usage is low the flexibility and versatility of PDAs show potential for text editing and creation in education, health care (Luo, 2004), business (O'Hara et al., 2002) and mobile Web interaction (Cui et al., 2008). For example, Waycott and Kukulska-Hulme (2003) indicate that the portability of PDAs meant that students could have access to learning resources at anytime, anywhere, and similarly, Luo (2004) highlights that portability assists convenient document editing in various locations of hospital ward. Although Cui et al. (2008) show that people typically prefer to read emails on PDAs rather than writing emails, they still show that people use their PDAs to access social networking sites which typically require entering text. Furthermore, in many countries such as India and many African countries, small-device penetration is soaring, and many of these users have limited access to a desktop system (Joshi & Avasthi, 2007). In this environment, the desktop-oriented tasks studied here are highly relevant.

Looking at the transfer of solutions from small-device to accessibility settings, the small-device contains a solution to the dropping errors observed on the desktop that could potentially benefit desktop users with motor impairments. Desktop systems employ a software utility, Sticky Keys, that performs a similar function, but this is not well known, and the small-device solution of making all modifier keys latch would likely benefit more users. Given the general tendency for small-device error rates to be lower than those for people with motor impairments, small-device solutions may require modification in order to address the broader range of performance levels of motor-impaired users.

While many error types were common across both studies, the most frequent errors: key ambiguity error; landing error; and long key press error, were either not shared at all, or not frequent in the other study.

It seems certain from these two experiments that Landing Error, the interruption of a taping or dragging task when the 'caret' is placed at one position for a certain period of time without being moved, is specific to the different operating modalities of the devices used; mouse vs. pen. In this case we can see that users without impairment would gain little by transferring desktop solutions to the small-device. However, key ambiguity error, which occurs when the user fails to switch to the appropriate small-device keyboard mode before pressing a key; and the long key press error, which occurs when a key is pressed for longer than the active key repeat delay, both share a possible common explanation. This explanation centres around the use of a 2-dimensional or 3-dimensional scanning, or software, keyboard (Felzer et al., 2005). These devices are common for motor-impaired individuals, and provide a software scanning mechanism for text entry (Baljko et al., 2006); not used in the original study Section 2.3. However, we can see that if this where the case then, for more sever motor impairments, the key ambiguity error rates may be increased based on errors in the scanning selection mechanism (MacKenzie et al., 1999) and long key press error in the original study may have been reduced as a key is selected for a set time.

We can see that the findings of this study highlight the difficulties able-bodied small-device users face when typing with a small-device keypad and pointing with a stylus, even without the challenge of a mobile environment. Our findings, although limited to a specific device, cover the general usage of mobile input devices and thus are important to manufacturers of small-devices and designers of the mobile Web browsers. Further, it is important to note that in this paper our focus was not investigating the usability issues of touch screens. Touch screens have specific usability issues that are out of the scope of this paper (Froehlich, Wobbrock, & Kane, 2007).

Finally, the findings show that common typing and pointing errors exist between desktop users with motor impairments and small-device users. In general, error rates on the small-device are higher than those of desktop users with no impairment, but lower than those of desktop users with motor impairments.

7. Conclusions and future work

There are many commonalities between the needs of small-device users and desktop users with motor impairments. This has lead to the adoption of several accessibility techniques for improving input rate in the mobile context. Accessibility techniques that improve input accuracy have not so far been transferred, and there has been a lack of empirical data to show whether such a transfer would be beneficial.

This paper has presented an empirical study of performance errors on a small-device that replicates an earlier desktop study with people with motor impairments. It shows that most of the observed typing and pointing error types are common in both contexts, but that in a static seated position the magnitude of errors on small-devices is lower. We suggest that this is mainly due to the static nature of the experimentation and that in the mobile context this magnitude difference will equalise. It is however, fair to say that at this point the dominant typing and pointing errors in each context are not present or significant in the other. In this case, while there are a number of accessibility solutions that tackle performance errors observed on the desktop, including long key press errors and bounce errors, these findings suggest that users without impairment would gain little by transferring them to the small-device.

An interesting continuation of this work may be the conjugation of the two defining elements, the impaired user and the constrained operating modality, into an investigation of the requirements of small-device users with motor impairment. Initial work along these lines has already been undertaken (Moffatt & McGrenere, 2007) and observed pen-based tapping errors in older adults that were not present in their younger group. However, these results could be supplemented with additional data that examines different user groups in addition to truly small-device use contexts.

Finally, before suggesting solution migrations, further study is needed to investigate the input errors of small-device users in more challenging environments, a broader range of small-device users, and the extent to which existing solutions are applicable to the mobile context. We suggest that in this context it is likely that additional errors not observed on the small-device in this study we be identified, and the magnitude of those common errors found here will equalise.

Acknowledgements

This work was undertaken as part of the RIAM project funded by the EPSRC (EP/E002218/1) and conducted in the Human Centred Web Lab part of the School of Computer Science at the University of Manchester (UK). About: http://hcw.cs.manchester.ac.uk/research/riam/ Reports: http://hcw-eprints.cs.man.ac.uk/view/subjects/riam.html.

Appendix A. Technical report and experimental data

Further details of the study presented here, including the materials, data and its analysis can be found at the Human Centred Web (HCW) Laboratories' data repository, http://hcw-eprints.cs.man.ac.uk/81/.

References

Atterer, R., Wnuk, M., & Schmidt, A. (2006). Knowing the user's every move—user activity tracking for website usability evaluation and implicit interaction. In *Proceedings of the 15th international conference on world wide web*.

Baljko, M., & Tam, A. (2006). Indirect text entry using one or two keys. In *Assets '06: Proceedings of the eighth international ACM SIGACCESS conference on computers and accessibility* (pp. 18–25). New York, NY, USA: ACM.

Barnard, L., Yi, J. S., Jacko, J. A., & Sears, A. (2005). An empirical comparison of use-in-motion evaluation scenarios for mobile computing devices. *International Journal of Human–Computer Studies, 62*(4), 487–520.

Brewster, S. (2002). Overcoming the lack of screen space on mobile computers. *Personal and Ubiquitous Computing, 6*.

Butts, L., & Cockburn, A. (2002). An evaluation of mobile phone text input methods. *Australian Computer Science Communications, 24*(4), 55–59.

Chen, T., Yesilada, Y., & Harper, S. (2008). Investigating the typing and pointing errors of mobile users. Tech. Rep. 51, University of Manchester—School of Computer Science.

Clawson, J., Lyons, K., Rudnick, A., Iannucci, R.A., Jr., & Starner, T. (2008). Automatic whiteout++: Correcting mini-qwerty typing errors using keypress timing. In *CHI '08: Proceeding of the twenty-sixth annual SIGCHI conference on human factors in computing systems* (pp. 573–582). New York, NY, USA: ACM.

Cui, Y., & Roto, V. (2008). How people use the web on mobile devices. In *WWW '08: Proceeding of the 17th international conference on world wide web* (pp. 905–914). New York, NY, USA: ACM.

Damerau, F. J. (1964). A technique for computer detection and correction of spelling errors. *Communications of the ACM, 7*(3), 171–176.

Felzer, T., & Nordmann, R. (2005). How to operate a pc without using the hands. In *Assets '05: Proceedings of the seventh international ACM SIGACCESS conference on computers and accessibility* (pp. 198–199). New York, NY, USA: ACM.

Froehlich, J., Wobbrock, J.O., & Kane, S.K. (2007). Barrier pointing: Using physical edges to assist target acquisition on mobile device touch screens. In *Assets '07: Proceedings of the ninth international ACM SIGACCESS conference on computers and accessibility* (pp. 19–26).

Gajos, K., Wobbrock, J., & Weld, D. (2008). Improving the performance of motor-impaired users with automatically-generated, ability-based interfaces. In *CHI 2008*.

James, C.L., & Reischel, K.M. (2001). Text input for mobile devices: Comparing model prediction to actual performance. In *CHI '01: Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 365–371). New York, NY, USA: ACM Press.

Joshi, D., & Avasthi, V. (2007). Mobile internet ux for developing countries. In *Workshop on mobile internet user experience*.

Kane, S., Wobbrock, J., Harniss, M., & Johnson, K. (2008). Truekeys: Identifying and correcting typing errors for people with motor impairments. In *Proceedings of the 13th international conference on intelligent user interfaces* (pp. 349–352).

Keates, S., Trewin, S., & Paradise, J. (2005). Using pointing devices: Quantifying differences across user groups. In *Proceedings of the third international conference on universal access in human–computer interaction*. Las Vegas.

Lin, M., Goldman, R., Price, K. J., Sears, A., & 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.

Luo, J. (2004). Portable computing in psychiatry. *Canadian Journal of Psychiatry, 49*(1).

MacKenzie, I.S., & Zhang, S.X. (1999). The design and evaluation of a high-performance soft keyboard. In *CHI '99: Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 25–31). New York, NY, USA: ACM.

MacKenzie, I.S., Kober, H., Smith, D., Jones, T., & Skepner, E. (2001). Letterwise: Prefix-based disambiguation for mobile text input. In *UIST '01: Proceedings of the 14th annual ACM symposium on User interface software and technology* (pp. 111–120). New York, NY, USA: ACM Press.

MacKenzie, I. S., & Soukoreff, R. W. (2002). Text entry for mobile computing: Models and methods, theory and practice. *Human–Computer Interaction Volume, 17*, 147–198.

MacKenzie, I. S., & Tanaka-Ishii, K. (2007). Text entry using a small number of buttons. In *Text entry systems: Mobility, accessibility, universality* (pp. 105–121). Morgan Kaufmann.

Moffatt, K.A., & McGrenere, J. (2007). Slipping and drifting: Using older users to uncover pen-based target acquisition difficulties. In *Assets '07: Proceedings of the 9th international ACM SIGACCESS conference on computers and accessibility* (pp. 11–18).

O'Hara, K., Perry, M., Sellen, A., & Brown, B. (2002). Exploring the relationship between mobile phone and document activity during business travel. In *Wireless world: social and interactional aspects of the mobile age* (pp. 180–194). New York, NY, USA: Springer-Verlag New York, Inc.

Oulasvirta, A., Tamminen, S., Roto, V., & Kuorelahti, J. (2005). Interaction in 4-second bursts: The fragmented nature of attentional resources in mobile hci. In *CHI '05: Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 919–928). New York, NY, USA: ACM Press.

Pavlovych, A., & Stuerzlinger, W. (2004). Model for non-expert entry speed on 12-button phone keypads. In *Proceedings of CHI'04. Vienna, Austria.*

Sears, A., & Young, M. (2003). Physical disabilities and computing technologies: an analysis of impairments. In A. Sears & J. A. Jacko (Eds.), *The human-computer interaction handbook: fundamentals, evolving technologies and emerging applications* (pp. 482–503). Mahwah, NJ, USA: Lawrence Erlbaum Associates, Inc.

Silververg, M., MacKenzie, I., & Korhonen, P. (2000). Predicting text entry speed on mobile phones. In *The SIGCHI conference on human factors in computing systems* (pp. 9–16).

Trewin, S. (2004). Automating accessibility: The dynamic keyboard. In *Proceedings of the sixth international ACM SIGACCESS conference on computers and accessibility* (pp. 71–78).

Trewin, S. (2006). Physical usability and the mobile web. In *W4A: Proceedings of the 2006 international cross-disciplinary workshop on Web accessibility (W4A)* (pp. 109–112). New York, NY, USA: ACM Press.

Trewin, S., Keates, S., & Moffatt, K. (2006). Developing steady clicks: A method of cursor assistance for people with motor impairments. In *Assets '06: Proceedings of the eigth international ACM SIGACCESS conference on computers and accessibility* (pp. 26–33).

Trewin, S., & Pain, H. (1999). Keyboard and mouse errors due to motor disabilities. *International Journal of Human Computer Studies Volume, 50*, 109–144.

Waycott, J., & Kukulska-Hulme, A. (2003). Students' experiences with pdas for reading course materials. *Personal and Ubiquitous Computing, 7*(1), 30–43.

Wigdor, D., & Balakrishnan, R. (2003). Tilttext: Using tilt for text input to mobile phones. In *UIST '03: Proceedings of the 16th annual ACM symposium on user interface software and technology* (pp. 81–90). New York, USA: ACM Press.

Wobbrock, J.O. (2006). The future of mobile device research in hci. In *CHI'06 Workshop*. Canada.

Wobbrock, J.O., Myers, B.A., & Kembel, J.A. (2003). Edgewrite: A stylus-based text entry method designed for high accuracy and stability of motion. In *Proceedings of UIST'03* (pp. 61–70).