

Comparing Evaluation Methods for Encumbrance and Walking on Interaction with Touchscreen Mobile Devices

Alexander Ng, John Williamson and Stephen Brewster

Glasgow Interactive Systems Group

School of Computing Science, University of Glasgow

a.ng.1@research.gla.ac.uk, jhw@dc.s.gla.ac.uk, stephen.brewster@glasgow.ac.uk

ABSTRACT

In this paper, two walking evaluation methods were compared to evaluate the effects of encumbrance while the preferred walking speed (PWS) is controlled. Users frequently carry cumbersome objects (e.g. shopping bags) and use mobile devices at the same time which can cause interaction difficulties and erroneous input. The two methods used to control the PWS were: walking on a treadmill and walking around a predefined route on the ground while following a pacesetter. The results from our target acquisition experiment showed that for ground walking at 100% of PWS, accuracy dropped to 36% when carrying a bag in the dominant hand while accuracy reduced to 34% for holding a box under the dominant arm. We also discuss the advantages and limitations of each evaluation method when examining encumbrance and suggest treadmill walking is not the most suitable approach to use if walking speed is an important factor in future mobile studies.

Author Keywords

Encumbrance; Preferred walking speed; Target acquisition; Mobile interactions.

ACM Classification Keywords

H.5.2. User Interfaces: *Input devices and strategies*

INTRODUCTION

The continuous development of touchscreen mobile devices and apps provide users with a variety of functionality while on the move. As a result, handheld devices are being used in a diverse range of mobile contexts. One important area that deserves more focus from researchers is the effect of encumbrance: for example, interacting while carrying objects such as bags and boxes or using an umbrella when it is raining. People often carry these types of objects in their day-to-day activities and are therefore likely to experience usability problems when their hands are full and attention from the mobile device is required. Users may struggle to input accurately or in an effective manner when their hands are physically hampered by holding objects at the same time which can cause interaction to become frustrating and

prone to errors. As a result, it is important to assess the impact of encumbrance, especially when the user is also on the move, a situation already known to cause interaction problems [21,22]. A greater understanding of the issues caused by encumbrance while walking would mean better interaction techniques and more efficient interfaces can be developed to assist the user in these physically awkward and mentally demanding situations.

To examine the effects of encumbrance and mobility in a controlled study, several experimental design issues need to be addressed. Firstly, what objects should be used to replicate realistic encumbrance scenarios? There are many possible objects that a user may carry during interaction and many ways in which the objects may be carried. Ng *et al.* [21] conducted an observational study to assess the typical objects that users carried when using mobile devices in the public and reported that different types of bags and boxes were the most common objects that users held. In the user study presented in this paper, we used a similar approach and selected a typical shopping bag and a packaging box as the encumbrance objects. Both objects are frequently handled during interaction and present users with physical challenges due to each object's size, weight and cumbersome nature.

Secondly, which mobile evaluation method should be used to assess the effects of walking *and* encumbrance on interaction? In previous studies, two broad methods have been used to examine the effects of mobility on input performance with handheld devices: 1) walking on a treadmill [3,4] and 2) walking around a predefined route [3,13,25]. Barnard *et al.* [3] compared both methods and reported the advantages and limitations of each to examine the impact of walking on interactions with mobile devices. However, it is unclear if the two approaches are still appropriate to examine the effects of encumbrance while the user is on the move. Therefore, an experiment was conducted to compare the two walking evaluation techniques to assess their effectiveness.

Thirdly, should the user's preferred walking speed (PWS) be controlled during an encumbered experiment? A potential confounding issue with not controlling the PWS in walking studies is that input performance can be traded with walking speed. For example, the user could walk significantly slower to select onscreen targets more accurately. It is then difficult to measure the true cost of mobility on

interaction in a controlled experiment. This problem is likely to be further compounded when examining encumbrance while walking as the PWS may be further reduced when input becomes too physically and mentally difficult to interact in an efficient manner. As a result, it is problematic to examine the effects of encumbrance as it is mixed up with walking speed. Therefore, to separate the two physical factors (encumbrance and walking), the users in the experiment presented in this paper maintained PWS during input.

The question then arises as to how to control the PWS for each walking evaluation technique. The treadmill method is straightforward once the user adjusts the speed of the treadmill to the pace at which he/she would normally walk. No extra effort is required to control the PWS. The treadmill guides the user to walk at a constant pace without having to divide attentional resources between interaction, navigation and avoiding nearby obstacles. For walking around a predefined route, we used the method of Ng *et al.* [22] to control the user's PWS by using a human pacesetter. An experimenter who acted as the pacesetter walked at each user's PWS (which was measured at the beginning of the study). The methods we chose to control the PWS allowed us to directly compare walking speed and input performance while encumbered between the two mobile evaluation methods.

BACKGROUND

The Effects of Encumbrance

In one of the few studies that has examined the effects of encumbrance, Ng *et al.* [20] extended the work of Crossan *et al.* [7] by investigating wrist rotation gestures as an alternative form of input with mobile devices when the user was encumbered. The users were walking and either carrying a bag in the dominant hand or clutching a box under the preferred arm while performing the gestures to select targets on a mobile phone. The results from their mobile Fitts' Law pointing experiment showed that input performance when carrying the box under the arm was very similar to unencumbered as users were able to stabilize their forearms against the box which helped to successfully perform the wrist-based gestures. In contrast, accuracy significantly decreased when holding the bag as the swinging motion of the user's arm when walking made it physically difficult to create a steady arm posture to input. This shows that encumbrance can have unpredictable effects on mobile interactions.

Ng *et al.* [21] conducted an observational study across three different types of locations in a major city to identify the objects that users frequently held during interaction. The observations showed that users carried different types of handheld bags the most often (46% of 878 objects). Mainwaring *et al.* [17] and Jain [11] both reported similar findings. Based on this, they then examined the impact of encumbrance while walking by measuring targeting performance on a touchscreen mobile phone. The results showed that the users were significantly less accurate when encumbered compared to holding no objects. Motion capture cameras were used to track body movements during input

and showed a significant increase in hand instability when the dominant hand or arm (which was also doing the target selection) was encumbered.

More recently, Ng *et al.* [22] assessed the impact of encumbrance on one- and two-handed interactions. Users performed a target acquisition task in three common input postures (two-handed index finger, one-handed preferred thumb and two-handed both thumbs) while walking and carrying multiple objects (a bag in each hand). The results showed that encumbrance caused target accuracy to decrease while selection time took significantly longer than interaction alone. Interestingly, target accuracy using both thumbs was not significantly better than single thumb or single index finger input when both hands were encumbered. The index finger was significantly quicker at selecting the targets than both thumb-based input postures. The advantage of having an extra finger to input is likely to diminish when both hands are busy carrying cumbersome objects during interaction.

Oulasvirta and Bergstrom-Lehtovirta [23] evaluated the user's physical ability to multitask when interacting with different types of computing devices. The physical tasks covered a range of hand postures and holding differently sized objects, ranging from small items such as a ballpoint pen to larger things like a basketball, while either performing a pointing task on a laptop or text entry on a mobile device. Their findings showed that mouse input on a laptop was affected more than using a trackpad and trackpoint when holding small and medium sized objects at the same time. Furthermore, switching from two-handed to one-handed typing on a mobile device while multitasking caused a decrease in text entry performance. Users performed the tasks while stationary therefore mobility was not examined.

The Effects of Mobility

Many previous studies have examined the impact of walking when using mobile devices. One common method to evaluate the impact of mobility is to use a treadmill [3,4,13]. The user sets the preferred pace on the treadmill by estimating PWS. The recorded walking speed is then used for each trial during the experiment. No additional effort is required to control the PWS as the user walks at a constant pace. However, walking on a treadmill might not be adequate to assess mobility since studies such as [1,19] have found that people's physical movements are different to normal walking on the ground.

In addition, walking on a treadmill is cognitively less demanding than walking in realistic outdoor settings due to the loss of environmental distractions. Consequently, the recorded PWS on the treadmill may not be an accurate representation of the user's real world walking pace as interferences such as avoiding nearby obstacles and keeping a personal distance from others increase mental workload. Oulasvirta *et al.* [24] found that in everyday situations, interaction with mobile devices was fragmented into periods between four to eight seconds as interruptions constantly

competed for the user's attention. Users are also exposed to constantly changing surfaces and other outdoor features that are difficult to replicate on the treadmill which could have an adverse impact on PWS and the user's ability to input accurately. Hyman *et al.* [10] reported that users walked slower and were more unaware to nearby events when using mobile phones. Users could be further affected if they also had to carry bulky objects during interaction, as shown by [21,22].

Mobile studies in the real world present their own challenges. For example, it is difficult to control outdoor elements (weather, lighting, number of surrounding people etc.) to keep the experimental conditions consistent and not introduce noise in the results. As an alternative, researchers have replicated real world settings by creating walking routes indoors [3,14,25] and outdoors [5,12,13] for users to navigate. Some studies [3,14,15] have also added obstacles to the walking path to increase the user's cognitive workload. Barnard *et al.* [3] varied room illumination to simulate realistic lighting changes to reproduce "an actual user experience" when comparing treadmill and ground walking. In the mobile study by Mackay *et al.* [16], users were instructed to stop walking and stand still at random intervals to replicate scenarios where the user is suddenly interrupted.

The majority of studies that use the predefined route evaluation method have not controlled the user's walking speed since the purpose of this experimental setup is to replicate walking in the real world. As mentioned earlier, the user can then trade walking speed to improve input accuracy by walking slower. Bergstrom-Lehtovirta *et al.* [4] showed that reducing PWS by 20% (on the treadmill at least) caused targeting performance on a touchscreen device to begin to level off. If walking speed is controlled during input, then the results will give a more accurate reflection of the user's targeting performance while walking without introducing confounding factors. An 'ideal' device would also not slow users down when using it on the move. A human pacesetter is typically used to regulate the user's walking speed when navigating around a predefined route [8,12]. In the outdoor mobile study by Kane *et al.* [12], all users walked at the same pre-set pace by following an experimenter and kept within a certain distance. We used a similar method to Ng *et al.* [22] to control walking speed in our experiment as users walked at their own measured PWS.

Previous research has shown the impact of using mobile devices on the user's walking speed. In the treadmill study by Bergstrom-Lehtovirta *et al.* [4], the mean PWS decreased by 24% to 2.97km/h when targeting on a touchscreen mobile phone compared to walking alone. Similarly, Lin *et al.* [15] stated that users were able to maintain 90% accuracy at the expense of dropping 36% of their walking speed when users walked around a predefined route. Lim and Fera [14] reported a decrease in walking speed of 22% when conducting a visual searching experiment on a handheld device. Mizobuchi *et al.* [18] recorded a low average walking speed of 1.77 km/h when examining

different button dimensions to identify the optimal size for text entry via stylus input. The encumbrance and mobile targeting experiment by Ng *et al.* [21] reported a mean PWS of 2.3 km/h during input (unencumbered) which resulted in a mean target accuracy of 42%. When users were carrying a broad box under the non-dominant arm, PWS decreased to 1.6 km/h (a difference of 41% compared to normal walking speed) which caused a mean target accuracy of 37%. The cited mobile studies have all reported the cost of interaction on mobility and found users to reduce walking speed as a compromise to improve input performance.

EXPERIMENT

A mobile target acquisition experiment was carried out to examine the effects of encumbrance while PWS was controlled. Two established methods used in the current literature to evaluate the effects of mobility were compared: walking on a treadmill and walking around a predefined route on the ground. In previous encumbrance studies [20,21,22], the predefined route method was used therefore no studies have yet to use a treadmill to evaluate the impact of walking while encumbered. The participants were given a target selection task on a touchscreen mobile phone while walking and carrying typical everyday objects. We also measured targeting performance at various levels of PWS to simulate situations where the user is walking slower (the user is being held up) and faster (the user is in a rush) than the preferred pace.

Measuring and Controlling PWS

A calibrated Woodway Bari-Mill treadmill with handrail support (see bottom images of Figure 1) was used for the treadmill walking conditions. Each participant's PWS on the treadmill was recorded before the experiment began and was measured by increasing the speed of the treadmill at 0.1 km/h increments up to the speed the user would normally walk. Like Barnard *et al.* [3] and Bergstrom *et al.* [4], participants were asked to think about the speed he/she would usually walk while not in a hurry when estimating the PWS. Once the PWS was recorded, the experimenter adjusted the pace accordingly for all the treadmill conditions for each participant.

For the ground walking conditions, an oval-shaped path was marked out using small plastic cones in a spacious and quiet room. The total length of the route was 20 meters long by 1.5 meters wide, as shown in the top image of Figure 1. The distance between each outer cone (red) was two meters to assist the pacesetter to walk at each participant's PWS. The PWS on the ground was measured by asking each participant to walk the route for six laps. The total time from lap two to lap six was recorded and since the distance was known, the average walking speed was calculated to determine the PWS. The duration of the first lap was not included in the calculation to allow the participants to build up to their normal walking speeds.

To control the PWS for the ground walking conditions, the participants walked side-by-side with an experimenter who acted as a pacesetter. A metronome application was devel-

oped in Android 4.1 and ran on a HTC One X mobile phone. The pacesetter used this to tune the metronome speed for each participant once the PWS was calculated. Audio feedback from the application kept the pacesetter at the appropriate walking speed for each participant. Noise levels from the application were kept to a minimum to avoid distracting the participants. Vibrotactile feedback was considered but during initial testing, the experimenter had difficulties walking at the desired pace thus auditory feedback was used. For each experimental condition, the pacesetter and the participant started walking and once the participant was satisfied with the pace and was comfortable with carrying the objects, he/she began a targeting task on a touchscreen mobile phone. Participants were instructed to avoid drifting out of the boundaries of the path during the experiment. Participants were also told to keep up or slow down if they failed to keep pace with the experimenter.

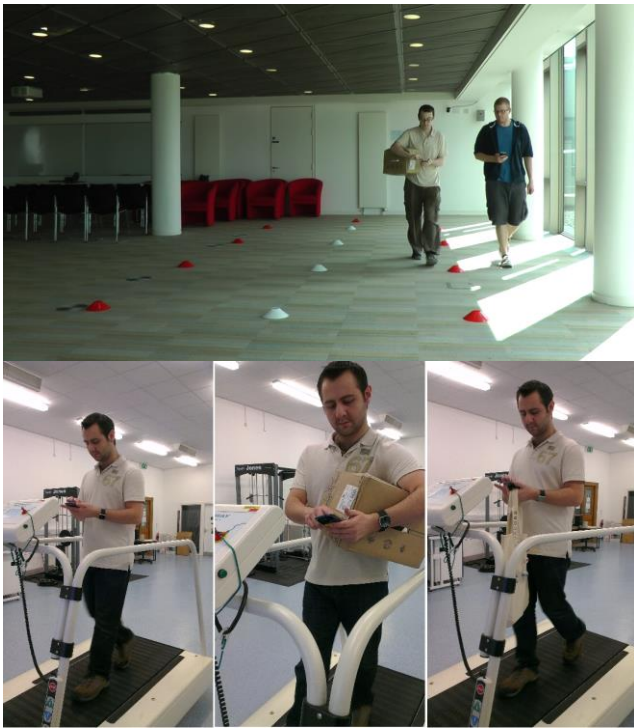


Figure 1. Top image illustrates part of the predefined oval route used in the ground walking conditions. The route was marked with plastic cones. The participant (inside) maintained PWS by walking alongside a pacesetter (outside). The bottom images show a participant performing the task while encumbered and walking on the treadmill.

We also measured targeting performance at various levels of PWS to simulate situations where the user walks slower (for example, when using a mobile device and keeping personal distance from other people) and faster (for instance, the user is in a hurry to get to a meeting). This allowed us to see if the two evaluation methods were practical at controlling different levels of walking speed and to see what effects these different speeds would have on input performance. Based on the findings from Bergstrom-Lehtovirta *et al.* [4], who reported that targeting performance began to level-off when the normal walking speed was decreased by

20%, the PWS was reduced to 80% in our experiment for the slow walking conditions. The walking speed was increased to 120% of the PWS for the fast walking conditions.

Encumbrance Scenarios

Based on the observational study by Ng *et al.* [21], a standard supermarket carrier bag and a packaging box were selected as the encumbrance objects. The dimensions (width x height x depth) of the bag were 45 x 55 x 25 cm, while the measurement of the box were 39 x 30 x 29 cm. Similar to [21,22], the bag and box both weighed 3kg to simulate the physical effects of carrying realistic objects yet reduce the amount of fatigue and tiredness on the participants. The bag was held in the hand while the box was held underarm as these are common ways to hold the objects. Each object was either held in the non-dominant (ND) hand/arm or the dominant (D) hand/arm during input. Figure 2 shows the bag and box that we used and the ways the objects were held.

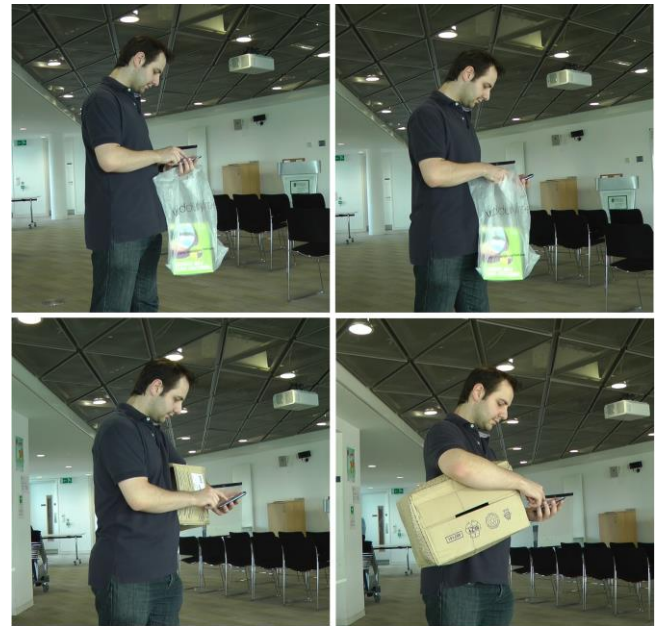


Figure 2. The encumbrance scenarios that were evaluated in the experiment: holding the bag in ND hand (top left), holding the bag in the D hand (top right), carrying the box under the ND arm (bottom left), and carrying the box under the D arm (bottom right). ND = non-dominant and D = dominant.

Experimental Task

The participants selected a series of targets on a touchscreen mobile phone as quickly and as accurately as possible. This type of tapping task has been used in previous studies [4,6,21,22] to measure general targeting performance on a mobile device. There were nine target positions located in a 3 x 3 grid: one centre and eight outer targets. Each outer target was selected ten times in a random order and every second selection was the centre target. As a result, there were 160 target selections per condition for each participant. Like Crossan *et al.* [6], a random interval ranging from 0.5 and 1.5 seconds was placed in between each target selection to reduce any rhythm created between

the participant's walking and input. A Samsung Galaxy S3 smartphone with a touchscreen resolution of 720 x 1280 pixels (~12 pixels/mm) was used. Each target was 60 pixels (5mm) wide and 96 pixels (8mm) long with a central crosshair measuring 30 pixels (2.5mm) in both directions (the same size as a key on the standard onscreen Android keyboard for the S3). The device was held in the non-dominant hand while the index finger of the dominant hand was used for input. We chose this type of two-handed input because it is a familiar interaction posture [2,9,22] and to remove the problem of selecting targets that are difficult to reach (in the one-handed preferred thumb posture, for example). The target selection task and input posture are shown in Figure 3.

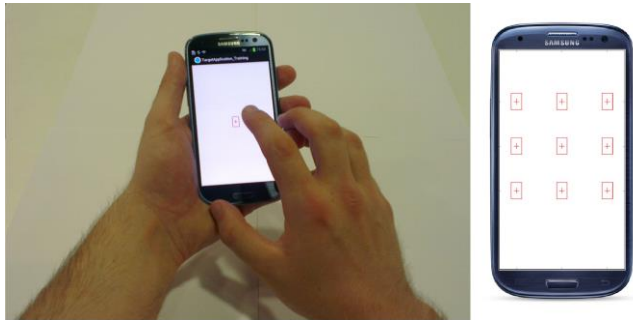


Figure 3. The two-handed index finger input posture used to select onscreen targets (left). The nine target positions on a Samsung Galaxy S3 mobile phone (right).

Experimental Design

A within-subjects design was used for the experiment. Twenty right-handed students (10 male, 10 female) aged between 18 – 41 years (mean = 22.4, SD = 5.3) were recruited from the university. The experiment was split into two sessions and took place on different days as we were concerned with the participants' fatigue and wellbeing. The participants were also given a sufficient resting period between each condition. All the treadmill conditions were completed in one session while all the ground walking conditions were done in the other session. Half of the participants (randomly chosen) completed the treadmill walking conditions first while the other half of the participants began with the ground walking conditions. Each session lasted an approximately one hour (*introduction + training + performing the task + debriefing*) and participants were paid £12 upon completing both sessions.

The Independent Variables were *type of encumbrance* (5 levels - unencumbered, holding the bag either in the *ND* or *D* hand, and holding the box either under the *ND* or *D* arm), *walking method* (2 levels – walking on the treadmill and walking around the predefined route on the ground) and *walking speed* (3 levels - walking at 80%, 100% and 120% of PWS). As a result, there were 30 conditions in total (15 for each session). The conditions in each session were randomised for each participant to reduce learning effects and bias in the results.

The Dependent Variables were target accuracy, target error and selection time. A target was accurately selected if the

recorded touch up position was within the target borders. Target error (in millimetres) was the absolute distance from the centre of the target crosshair to the recorded finger touch up position on the screen. Selection time (in milliseconds) was the duration from the display of the current target to the instant that a press up event was logged.

The hypotheses were:

- H1: Both error and selection time will be significantly increased while accuracy will be significantly decreased when encumbered compared to unencumbered;
- H2: Holding the bag/box in the dominant hand/arm will cause accuracy to decrease significantly and selection time to increase significantly more than holding the objects in the non-dominant side.
- H3: Target accuracy will be significantly higher and selection time will be significantly faster when walking at 80% of PWS compared to walking at 100% of PWS;
- H4: Target accuracy will be significantly reduced and selection time will be significantly slower when walking at 120% of PWS compared to walking at 100% of PWS;
- H5: Target accuracy will be significantly decreased while both error and selection time will be significantly increased for walking on the ground when compared to walking on the treadmill;
- H6: The PWS will be significantly faster for walking on the treadmill than walking around the predefined route on the ground.

RESULTS

There were 160 target selections per condition therefore 4800 targets per participant (30 conditions). Twenty participants resulted in a total of 96000 targets for the whole experiment. To filter out unintentional screen taps, targets that took less than 100ms to select were removed from the data set. As a result, 21 targets were eliminated from the analysis. Three factor (type of encumbrance, walking method and walking speed) repeated-measures ANOVAs were conducted to analyse target accuracy, target error and selection time.

Target Accuracy

The mean target accuracy for each walking condition is shown in Figure 4. The ANOVA for accuracy showed a significant main effect for walking method, $F(1, 19) = 10.906$, $p < 0.01$. Target selection was significantly more accurate for walking on the treadmill than walking on the ground (mean difference = 4.85%). A significant main effect was observed for walking at the different levels of PWS, $F(2, 38) = 49.106$, $p < 0.001$. *Post hoc* pairwise comparisons with Bonferroni corrections showed that target accuracy was significantly higher when walking at 80% of PWS than 100% of PWS (mean difference = 3.55%). The participants were significantly less accurate at targeting when walking at 120% of PWS than walking at 100% of PWS (mean difference = 2.63%). A significant main effect was also found for type of encumbrance, $F(4, 76) = 60.933$

, $p < 0.001$. *Post hoc* pairwise comparisons with Bonferroni corrections showed that accuracy was significantly higher when unencumbered compared to holding the objects during input. Target accuracy while carrying the bag in the *D* hand was significantly lower than the *ND* hand. There was no significant difference for accuracy between the *D* and *ND* sides when holding the box underarm. All type of encumbrance pairwise comparisons for accuracy are shown in Table 1. No significant interaction was observed between the factors ($p > 0.05$).

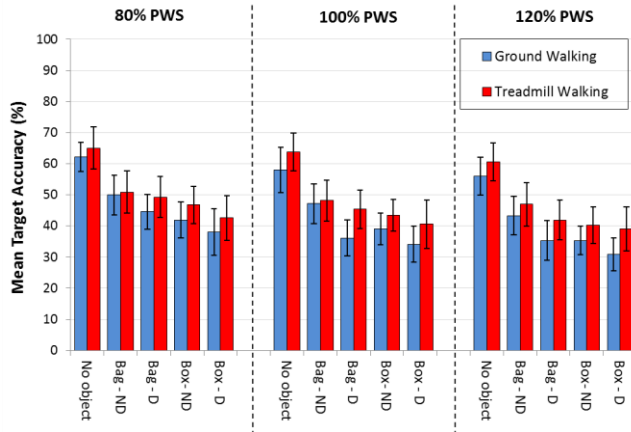


Figure 4. The mean target accuracy (%) for each walking condition, grouped by the % of PWS. The blue and red bars represent the ground and treadmill walking conditions respectively. Error bars denote 95% CI.

Comparison		Mean Difference (%)	Sig*
No object	Bag ND	13.19	0.000
No object	Bag D	18.82	0.000
No object	Box ND	19.79	0.000
No object	Box D	23.39	0.000
Bag ND	Bag D	5.64	0.007
Bag ND	Box ND	6.61	0.002
Bag ND	Box D	10.20	0.000
<i>Bag D</i>	<i>Box ND</i>	<i>0.97</i>	<i>1.000</i>
<i>Bag D</i>	<i>Box D</i>	<i>4.56</i>	<i>0.423</i>
<i>Box ND</i>	<i>Box D</i>	<i>3.60</i>	<i>0.253</i>

Table 1. The pairwise comparisons of type of encumbrance for target accuracy (%). The mean difference that are non-significant are italicised and highlighted in bold. *Adjustment for multiple comparisons: Bonferroni.

Target Error

The mean target error for each walking condition is shown in Figure 5. The ANOVA for target error showed a significant main effect for walking method, $F(1, 19) = 5.198$, $p < 0.05$. Target error was significantly higher when walking on the ground than walking on the treadmill (mean difference = 0.4mm). A significant main effect was also observed for walking speed, $F(2, 38) = 24.286$, $p < 0.001$. *Post hoc* pairwise comparisons with Bonferroni corrections

showed that error was significantly higher for walking at 100% of PWS compared to 80% (mean difference = 0.3mm). On the other hand, error at 100% of PWS was significantly lower than walking at 120% of PWS (mean difference = 0.3mm). A main effect was found for type of encumbrance, $F(4, 76) = 32.595$, $p < 0.001$. *Post hoc* pairwise comparisons with Bonferroni corrections (Table 2) showed that error was significantly higher when holding the objects than unencumbered interaction. Target error was significantly higher when the bag was held in the *D* hand than the *ND* hand. Likewise, holding the box under the *D* arm resulted in a significant increase in error compared to the *ND* arm. No significant interaction was found between the factors ($p > 0.05$).

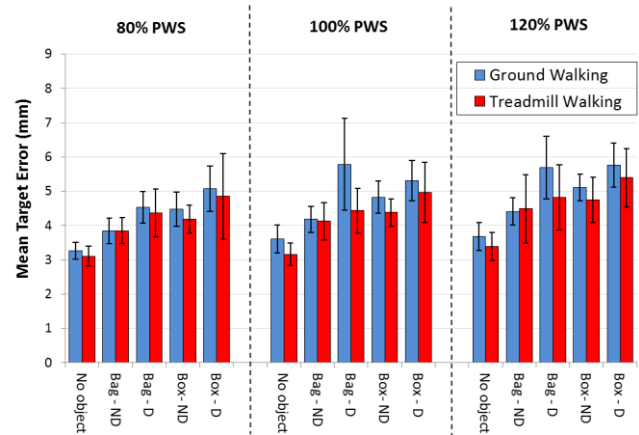


Figure 5. The mean target error (mm) for each walking condition. Error bars denote 95% CI.

Comparison		Mean Difference (mm)	Sig*
No object	Bag ND	-0.8	0.000
No object	Bag D	-1.6	0.000
No object	Box ND	-1.3	0.000
No object	Box D	-1.9	0.000
Bag ND	Bag D	-0.8	0.004
Bag ND	Box ND	-0.5	0.003
Bag ND	Box D	-1.1	0.000
<i>Bag D</i>	<i>Box ND</i>	<i>0.3</i>	<i>1.000</i>
<i>Bag D</i>	<i>Box D</i>	<i>-0.3</i>	<i>0.605</i>
Box ND	Box D	-0.6	0.05

Table 2. The pairwise comparisons of type of encumbrance for target error (mm).

Selection time

Figure 6 shows the mean selection time for each walking condition. The ANOVA for selection time showed a significant main effect for walking method, $F(1, 19) = 23.320$, $p < 0.001$. Selection time was significantly quicker for walking on the treadmill than walking on the ground (mean difference = 26.1ms). A significant main effect was observed for walking at the three different levels of PWS, $F(2, 38) = 7.543$, $p < 0.05$. *Post hoc* pairwise comparisons with Bon-

ferroni corrections showed that selection time was not significantly quicker for walking at 80% of PWS than 100% of PWS. There was also no significant difference for selection time between walking at 100% and 120% of PWS. However, selection time was significantly quicker when walking at 120% of PWS than 80% of PWS, a mean difference of 14.3ms. A significant main effect was found for type of encumbrance, $F(4, 76) = 27.198$, $p < 0.001$. *Post hoc* pairwise comparisons (Table 3) showed that target selections were significantly quicker when unencumbered compared to carrying the objects, except for holding the bag in the *ND* hand. The participants were significantly quicker at selecting the targets when the bag/box was held in the *ND* hand/arm than holding the objects in the *D* side. There was no significant interaction between the factors.

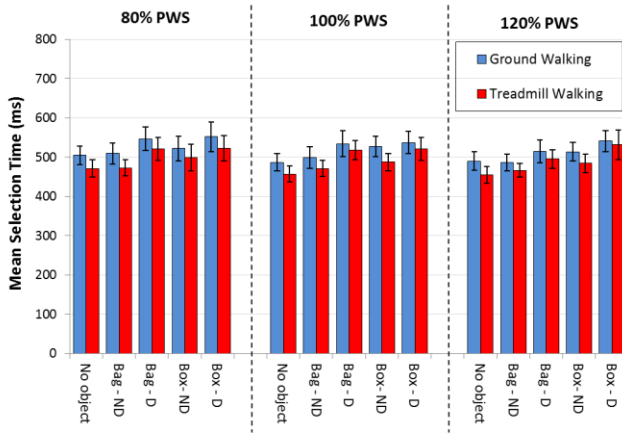


Figure 6. The mean selection time (ms) for each walking condition. Error bars denote 95% CI.

Comparison		Mean Difference (ms)	Sig*
No object	Bag ND	-6.7	0.865
No object	Bag D	-44.0	0.000
No object	Box ND	-28.1	0.000
No object	Box D	-56.7	0.000
Bag ND	Bag D	-37.3	0.000
Bag ND	Box ND	-21.4	0.012
Bag ND	Box D	-50.0	0.000
Bag D	Box ND	15.9	0.252
Bag D	Box D	-12.7	1.000
Box ND	Box D	-28.6	0.009

Table 3. The pairwise comparisons of type of encumbrance for selection time (ms).

PWS, Distance Walked and Experiment Completion Time for Treadmill and Ground Walking

A paired t-test was conducted to compare the measured PWS between the two walking methods. There was a significant difference in walking speed (km/h) for ground walking (mean = 4.88, SD = 0.70) and treadmill walking (mean = 3.57, SD = 1.03); $t(19) = 6.556$, $p < 0.05$. The participants walked significantly faster on the ground than

on the treadmill (a drop in walking speed of 26.8%). Furthermore, each participant's PWS on the treadmill was slower than walking on the ground. In the encumbrance and ground walking study by Ng *et al.* [22], who used the same method to control the user's walking speed, a similar mean PWS of 4.9km/h was reported.

Table 4 shows the approximate mean distance walked and total interaction time to complete all 15 conditions for each walking method. Please note, the interaction times do not include any resting periods or the time required to switch between conditions in each session. For comparison, Ng *et al.* [22] reported a mean distance of 1.480km and a total interaction time of 18.3mins to complete their ground walking encumbered study.

	Ground Walking	Treadmill Walking
Mean distance walked (km)	1.722 (SD = 0.166)	1.193 (SD = 0.108)
Mean interaction time (mins)	21.17 (SD = 2.04)	20.05 (SD = 1.82)

Table 4. The approximate mean walking distance and total interaction time for each walking method.

DISCUSSION

The results for encumbrance showed that target selections while carrying the bag or box caused accuracy to decrease while error increased when compared to unencumbered. However, there was no significant difference in terms of selection time between unencumbered and carrying the bag in the *ND* hand, despite a significant difference observed for all other pair comparisons. Therefore, hypothesis H1 can only be partially supported.

The results also revealed that holding the bag or the box in the *D* hand or arm caused selection times to significantly increase when compared to holding the objects in the *ND* side. However, it is worth noting that the difference is minimal (28.6ms for the box and 37.3ms for the bag). There was no significant difference for target accuracy between the *D* and *ND* arms when carrying the box thus, hypothesis H2 can only be partially supported. We anticipated significant results since Ng *et al.* [21] found that users experienced more hand movements when the *D* hand or arm was physically hampered which caused accuracy to significantly decrease when compared to the *ND* side.

As previous walking studies have shown, there is always a cost to performance when the user is on the move as target accuracy dropped to 58.0% and 63.8% for ground and treadmill walking respectively (when walking at 100% of PWS). Once the users were also encumbered while maintaining 100% of PWS on the ground, target accuracy decreased to 34.1% for carrying the box under the *D* arm. For comparison, target accuracy for the same walking condition was 36.1% when the bag was held in the *D* hand. For treadmill walking at 100% of PWS, target accuracy was 40.6% and 43.4% for carrying the box under the *D* and *ND*

arm respectively. While walking on the treadmill at 100% of PWS, the bag caused target accuracy to drop to 45.4% and 48.1% for the *D* and *ND* hand respectively. From our observations and comments from the participants, the size of the box made it awkward and physically challenging to input accurately, while the participants felt that the weight of the bag caused them more problems in aiming at the touchscreen to input precisely despite both objects weighing the same.

In terms of the impact encumbrance had on selection time, the box (held under either arm) caused significantly longer target selections than unencumbered input. The bag also resulted in significantly slower selection times when it was held in the *D* hand compared to unencumbered interaction. Again, the significant selection times were marginal. The longest mean selection times while walking on the ground at 100% of PWS was 534ms and 537ms for carrying the bag (*D* hand) and the box (*D* arm) respectively. For comparison, treadmill walking at 100% of PWS resulted in mean selection times of 518ms and 521ms for the same encumbrances. While selection times were similar, a greater difference was shown for target accuracy between unencumbered and holding the objects, so the participants might have traded input speed for accuracy.

We measured targeting performance at different levels of the PWS to simulate situations where the user was walking slower and faster than normal. No previous studies have examined the effects of carrying cumbersome objects during interaction while varying walking speed. The results showed that reducing PWS by 20% significantly improved target accuracy and decreased targeting error. However, walking slower at 80% of PWS did not cause significantly quicker selection time than walking at 100% of PWS therefore hypothesis H3 is partial supported. Despite a significant effect being found for target accuracy, the mean accuracy rate only improved by 3.55% when walking at 80% of PWS. Similarly, the difference in mean targeting error between 80% and 100% of PWS was also small at 0.3mm. This suggests that the user would need to reduce walking speed further in order to gain better input accuracy. In the unencumbered treadmill study by Bergstrom *et al.*[4], an ideal trade-off was reported between targeting performance and walking at 40-80% PWS.

The results for walking faster at 120% of PWS showed that target accuracy significantly decreased when compared to 100% of PWS. However, selection time did not take significantly longer than 100% PWS, hence hypothesis H4 is partially supported. Similar to walking slower, the difference in mean target accuracy between 100% and 120% of PWS was small at 2.63%. The highest mean targeting error of 5.8mm was caused by holding the bag in the *D* hand while walking at 100% of PWS on the ground, which was near identical to walking at 120% of PWS (5.7mm). One surprising result from the experiment was that the mean selection time significantly decreased as the PWS increased. In other words, the faster the user walked, the faster the input. We expected the user would need more

time to aim accurately at the screen when carrying the cumbersome objects and walking faster than normal. A random delay was placed between each target selection to negate any rhythm between targeting and the user's gait. However, we can only assume that when walking faster, users felt more rushed to input, perhaps to keep up with the body movements when walking.

The comparison between the two evaluation methods showed that target accuracy significantly decreased for ground walking when compared to walking on the treadmill. The mean target accuracy for each ground walking condition was lower than the equivalent treadmill one. Target error was significantly higher for walking on the ground than on the treadmill. As predicted, selection time was shorter for walking on the treadmill than walking on the ground. Therefore, hypothesis H5 is supported. Previous studies [3,13] have compared the two walking methods and reported that walking on the ground and around a pre-defined route will increase the participant's mental workload more than walking on a treadmill. Furthermore, participants in our study maintained their PWS which reduced the possibility of trading walking speed with targeting performance. Therefore, input is likely to differ between the two walking methods, as shown in our study.

The mean PWS for both walking methods showed that participants walked significantly faster on the ground than on the treadmill. Thus, hypothesis H6 is rejected. We expected the PWS for walking on the ground to be slower than on the treadmill because the participants had to navigate and keep within the path. However, further data analysis showed that all participants walked slower on the treadmill than on the ground. Seven out of twenty participants reduced their PWS on the treadmill by more than 25% when compared to walking on the ground. One particular participant dropped PWS by 71%. The difference in PWS between the two walking methods could have been one reason why targeting performance was significantly better for walking on the treadmill than on the ground.

At the end of the experiment, we asked each participant to walk at the PWS (measured for ground walking) on the treadmill. A majority of the participants were surprised by the difference in walking speed and commented that it was difficult to judge the pace that he/she would normally walk on the treadmill because there was no point of reference. Comments also suggested that the participants walked at a more conservative pace to prevent them from getting close to the edge of the treadmill. This implies that there are possible confounding psychological factors as well as physical factors [1,19] associated with treadmill-based evaluations that cause participants to walk differently from their normal walking speed. Our results and observations showed that the treadmill method should be used cautiously for assessing interactions on the move. The ground walking method is better at measuring the user's PWS more accurately than using a treadmill and should be used if natural walking speed is an important factor in future mobile studies.

Despite the disparity in PWS between the two evaluation techniques, both methods are appropriate to examine the effects of walking while encumbered if extra care is taken when planning a mobile experiment. From our observations, we make the following recommendations and discuss the limitations of each walking method.

The treadmill method is suitable if limited space is available to setup a walking route indoors. Also, it is simpler to arrange than the ground walking method and no additional effort is required from an experimenter to control the PWS once the user estimates his/her walking speed on the treadmill. The user walks at a consistent pace during the experiment without variation in walking speed. On the other hand, the ground walking method requires more initial work to map out and measure the walking route. A human pacesetter is needed to control each participant's PWS. It is a challenging task for the pacesetter to walk at the required walking speed consistently for each participant across all the conditions. Training is required for the pacesetter to walk consistently across a wide range of walking speeds and we advise taking regular resting periods during the experiment to reduce tiredness. Another issue with the ground walking method is that the participants might struggle to keep pace with the pacesetter especially when encumbered. As a result, a difference in walking speed between the user and the pacesetter may occur during input. In our study, the participants were asked to speed up or slow down if they failed to walk at the correct speed with the pacesetter. From our observations, all the participants were able to maintain walking speed with the pacesetter with the exception of a few minor instances where the participant struggled to walk at the correct pace.

In terms of using the evaluation methods to examine the impact of encumbrance while walking, a potential issue with the treadmill is the limited space due to the safety sidebars. In our study, we ensured that carrying the bag and the box on the treadmill would not cause the user any input problems. One constraint of the treadmill method is that it restricts the types of encumbrance scenarios that can be assessed. For example, it is likely to be difficult to evaluate the effects of carrying multiple cumbersome objects [22] and new encumbrance scenarios that require more complex movements such as pushing objects (e.g. a pram) [21,23]. There is no such problem with the ground walking method as the user is not restricted in upper body movements and has more physical space to carrying the objects and input at the same time.

One of the main goals of our experiment was to compare two evaluation methods to examine the effects of walking while encumbered. The other was to widen our understanding of the effects of encumbrance on mobile interactions and make researchers more aware of the usability problems in these contexts. Touchscreen mobile devices such as smartphones present developers with design challenges due to the limited screen space. Interface components such as buttons and icons are small in size which makes them difficult to select accurately when the user is walking and en-

cumbered. A possible solution is to make touchable elements larger to give users bigger target areas to improve accuracy. Interface guidelines from Apple [27] state the size for keys should be 6.7 mm for touchscreen interfaces. Schildbach and Rukzio [26] increased target size to 9.5 mm and showed this improved tapping performance when users were walking. However, these recommended target sizes are less effective if the user is also encumbered. Our results for targeting error suggest a larger target size is required to assist interaction while carrying objects and on the move. Also, increasing target size on screen limited mobile devices can cause other interaction issues such as increasing the amount of scrolling needed to reach all items, as reported by Kane *et al.* [12]. Designers will also need to carefully consider the trade-off between input accuracy and selection time when developing new interaction techniques for encumbrance. The cost of an incorrect selection and subsequent reattempts at selecting a target can result in a frustrating user experience when the user has to multitask while on the move.

CONCLUSIONS

The main contribution of the study presented in this paper is the comparison of two mobile evaluation methods to see if these approaches were appropriate to examine the impact of walking while encumbered. The aim was to measure targeting performance on a touchscreen mobile phone while encumbered and, since users can trade walking speed with input performance, the PWS was controlled. The results and observations from our experiment showed the ground walking method yielded a better representation of the user's PWS than using a treadmill. Despite the variation in walking speed, the differences in terms of targeting performance between the two evaluation methods were small. Both walking techniques are suitable to examine the effects of encumbrance while on the move but each approach has its limitations and care is required when designing encumbered walking experiments.

Our study has also extended the limited literature on the effects of encumbrance. The highest targeting error recorded in our experiment is greater than previous encumbered studies, which shows the extent to which input performance is affected when the user is walking and carrying cumbersome objects at the same time. We hope researchers and user interface designers will be more aware of the usability issues that are caused by the common situation of encumbrance while interacting. At present, there are no input techniques to assist the user to interact in a more effective manner when encumbered yet the popularity of touchscreen mobile devices will continue to grow and present designers with new and unforeseen challenges.

ACKNOWLEDGEMENTS

We thank Dr. Ron Baxendale for his assistance with the treadmill.

REFERENCES

1. Alton, F., Baldey, L., Caplan, S., and Morrissey, M.C. A kinematic comparison of overground and treadmill

- walking. *Clinical biomechanics* (Bristol, Avon) 13, 6 (1998), 434–440.
2. Azenkot, S. and Zhai, S. Touch behavior with different postures on soft smartphone keyboards. In *Proc. MobileHCI 2012*, ACM Press (2012), 251–260.
3. Barnard, L., Yi, J.S., Jacko, J.A., and Sears, A. An empirical comparison of use-in-motion evaluation scenarios for mobile computing devices. *International Journal of Human-Computer Studies* 62, 4 (2005), 487–520.
4. Bergstrom-Lehtovirta, J., Oulasvirta, A., and Brewster, S. The effects of walking speed on target acquisition on a touchscreen interface. In *Proc. MobileHCI 2011*, ACM (2011), 143–146.
5. Brewster, S. Overcoming the Lack of Screen Space on Mobile Computers. *Personal and Ubiquitous Computing* 6, 3 (2002), 188–205.
6. Crossan, A., Murray-Smith, R., Brewster, S., Kelly, J., and Musizza, B. Gait phase effects in mobile interaction. *Ext. Abstracts CHI 2005*, ACM Press (2005), 1312–1315.
7. Crossan, A., Williamson, J., Brewster, S., and Murray-Smith, R. Wrist rotation for interaction in mobile contexts. In *Proc. MobileHCI 2008*, ACM Press (2008), 435–438.
8. Goel, M., Findlater, L., and Wobbrock, J. WalkType: using accelerometer data to accommodate situational impairments in mobile touch screen text entry. In *Proc. CHI 2012*, ACM Press (2012), 2687–2696.
9. Hooper, S. How Do Users Really Hold Mobile Devices? <http://www.uxmatters.com/mt/archives/2013/02/how-do-users-really-hold-mobile-devices.php>.
10. Hyman, I.E., Boss, S.M., Wise, B.M., McKenzie, K.E., and Caggiano, J.M. Did You See the Unicycling Clown? Inattention Blindness while Walking and Talking on a Cell Phone. *Applied Cognitive Psychology* 24, 5 (2010), 597–607.
11. Jain, S.S.L. The means of mobility. *Journal of Consumer Culture* 2, 3 (2002), 419–438.
12. Kane, S.K., Wobbrock, J.O., and Smith, I.E. Getting off the treadmill: evaluating walking user interfaces for mobile devices in public spaces. In *Proc. MobileHCI 2008*, ACM Press (2008), 109–118.
13. Kjeldskov, J. and Stage, J. New techniques for usability evaluation of mobile systems. *International Journal of Human-Computer Studies* 60, 5-6 (2004), 599–620.
14. Lim, J.J. and Faria, C. Visual search on a mobile device while walking. In *Proc. CHI 2012*, ACM Press (2012), 295–304.
15. Lin, M., Goldman, R., Price, K.J., Sears, A., and Jacko, J. How do people tap when walking? An empirical investigation of nomadic data entry. *International Journal of Human-Computer Studies* 65, 9 (2007), 759–769.
16. MacKay, B., Dearman, D., Inkpen, K., and Watters, C. Walk 'n scroll: a comparison of software-based navigation techniques for different levels of mobility. In *Proc. MobileHCI 2005*, ACM Press (2005), 183–190.
17. Mainwaring, S.D., Anderson, K., and Chang, M.F. Living for the global city: mobile kits, urban interfaces, and ubicomp. In *Proc. UbiComp 2005*, Springer-Verlag (2005), 269–286.
18. Mizobuchi, S., Chignell, M., and Newton, D. Mobile text entry: relationship between walking speed and text input task difficulty. In *Proc. MobileHCI 2005*, ACM Press (2005), 122–128.
19. Murray, M.P., Spurr, G.B., Sepic, S.B., Gardner, G.M., and Mollinger, L. a. Treadmill vs. floor walking: kinematics, electromyogram, and heart rate. *Journal of applied physiology* 59, 1 (1985), 87–91.
20. Ng, A., Brewster, S., and Crossan, A. The effects of encumbrance on mobile gesture interactions. In *Body, Movement, Gestures & Tactility in Interaction with Mobile Devices Workshop, MobileHCI 2011*, (2011).
21. Ng, A., Brewster, S., and Williamson, J. The Impact of Encumbrance on Mobile Interactions. In *Proc. INTERACT 2013*, Springer-Verlag (2013), 92–109.
22. Ng, A., Brewster, S., and Williamson, J. Investigating the Effects of Encumbrance on One- and Two- Handed Interactions with Mobile Devices. In *Proc. CHI 2014*, (2014), 1981–1990.
23. Oulasvirta, A. and Bergstrom-Lehtovirta, J. Ease of juggling: studying the effects of manual multitasking. In *Proc. CHI 2011*, ACM Press (2011), 3103–3112.
24. Oulasvirta, A., Tamminen, S., Roto, V., and Kuorelahti, J. Interaction in 4-second bursts: the fragmented nature of attentional resources in mobile HCI. In *Proc. CHI 2005*, ACM Press (2005), 919–928.
25. Pirhonen, A., Brewster, S., and Holguin, C. Gestural and audio metaphors as a means of control for mobile devices. In *Proc. CHI 2002*, ACM Press (2002), 291–298.
26. Schildbach, B. and Rukzio, E. Investigating selection and reading performance on a mobile phone while walking. In *Proc. MobileHCI 2010*, ACM Press (2010), 93–102.
27. Apple Inc. iPhone Human Interface Guidelines. 2009.