Ng, Alexander Wing Ho (2016) *The effects of encumbrance and mobility on interactions with touchscreen mobile devices.* PhD thesis.

THE EFFECTS OF ENCUMBRANCE AND MOBILITY ON INTERACTIONS WITH TOUCHSCREEN MOBILE DEVICES

ALEXANDER WING HO NG

SUBMITTED IN FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF

Doctor of Philosophy

SCHOOL OF COMPUTING SCIENCE

COLLEGE OF SCIENCE AND ENGINEERING

UNIVERSITY OF GLASGOW

MARCH 2016

Abstract

Mobile handheld devices such as smartphones are now convenient as they allow users to make calls, reply to emails, find nearby services and many more. The increase in functionality and availability of mobile applications also allow mobile devices to be used in many different everyday situations (for example, while on the move and carrying items). While previous work has investigated the interaction difficulties in walking situations, there is a lack of empirical work in the literature on mobile input when users are physically constrained by other activities. As a result, how users input on touchscreen handheld devices in *encumbered* and *mobile* contexts is less well known and deserves more attention to examine the usability issues that are often ignored.

This thesis investigates targeting performance on touchscreen mobile phones in one common encumbered situation - *when users are carrying everyday objects while on the move*. To identify the typical objects held during mobile interactions and define a set of common encumbrance scenarios to evaluate in subsequent user studies, Chapter 3 describes an observational study that examined users in different public locations. The results showed that people carried different types of bags and boxes the most frequently.

To measure how much tapping performance on touchscreen mobile phones is affected, Chapter 4 examines a range of encumbrance scenarios, which includes holding a bag in-hand or a box underarm, either on the dominant or non-dominant side, during target selections on a mobile phone. Users are likely to switch to a more effective input posture when encumbered and on the move, so Chapter 5 investigates one- and two- handed encumbered interactions and evaluates situations where both hands are occupied with multiple objects. Touchscreen devices afford various multi-touch input types, so Chapter 6 compares the performance of four main one- and two- finger gesture inputs: tapping, dragging, spreading & pinching and rotating, while walking and encumbered.

Several main evaluation approaches have been used in previous walking studies, but more attention is required when the effects of encumbrance is also being examined. Chapter 7 examines the appropriateness of two methods (ground and treadmill walking) for encumbered and walking studies, justifies the need to control walking speed and examines the effects of varying walking speed (i.e. walking slower or faster than normal) on encumbered targeting performance.

The studies all showed a reduction in targeting performance when users were walking and encumbered, so Chapter 8 explores two ways to improve target selections. The first approach defines a target size, based on the results collected from earlier studies, to increase tapping accuracy and subsequently, a novel interface arrangement was designed which optimises screen space more effectively. The second approach evaluates a benchmark pointing technique, which has shown to improve the selection of small targets, to see if it is useful in walking and encumbered contexts.

Contents

List of Tables

List of Figures

Acknowledgements

I am extremely grateful to have Professor Stephen Brewster as my supervisor and would like to thank him for giving me this opportunity. Without his guidance and patience, this Phd would not have been possible. I am also very fortunate to have Dr John Williamson as a co-supervisor and would like to thank him for his expertise and input during our weekly discussions.

Many thanks to Dr Andrew Crossan for getting me excited about HCI and subsequently introducing me to the Multimodal Interaction Group. I would like to thank my colleagues (in no particular order): Euan, Graham, Ross, Mark, Matt, Ioannis, Marilyn, Julie, Dong Bach and previous members for their friendship and support and making MIG a truly awesome research group. Thank you all for making me a better researcher.

I am grateful to all the participants that took part in my demanding user studies. This thesis would not have been possible without their participation. Many thanks to my good friend Vito for keeping me sane during my PhD and being a good actor when required. I am hugely thankful to have been brought up by amazing parents. Thank you for your never-ending love and support and for making me a better person.

Lastly, I would like to thank my examiners for their helpful feedback on my thesis.

Declaration

I declare that this thesis was completed by myself and that the work contained therein is my own, except where explicitly stated otherwise in the text.

Alexander Wing Ho Ng

List of Contributing Publications

Ng, A., Williamson, J.H., and Brewster, S. The Effects of Encumbrance and Mobility on Touch-Based Gesture Interactions for Mobile Phones. In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices & Services - MobileHCI '15* (2015), ACM Press, pp 536-546.

Ng, A., Williamson, J.H., and Brewster, S.A. Comparing Evaluation Methods for Encumbrance and Walking on Interaction with Touchscreen Mobile Devices. In *Proceedings of the 16th International Conference on Human-Computer Interaction with Mobile Devices & Services - MobileHCI '14* (2014), ACM Press, pp 23-32.

Ng, A., Brewster, S.A., and Williamson, J.H. Investigating the Effects of Encumbrance on One- and Two- Handed Interactions with Mobile Devices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '14* (2014), ACM Press, pp 1981-1990.

Ng, A., Brewster, S.A., and Williamson, J.H. The Impact of Encumbrance on Mobile Interactions. In *Proceedings of the 14th IFIP TC 13 International Conference on Human-Computer Interaction - INTERACT 2013* (2013), Springer Berlin Heidelberg, pp 92-109.

Chapter 1

Introduction

Touchscreen mobile devices such as smartphones and tablets are now very popular. Users are no longer solely restricted to traditional desktop computers to complete daily activities since mobile devices provide an array of functionality in the form of mobile applications (apps) available at all times and while on the move. For example, apps allow users to communicate with friends by text messaging, broadcast a particular event on social media, find nearby services and respond to work activities via email. As a result, mobile devices are used in a wide variety of everyday situations and when developing mobile apps and new interactions techniques, researchers and user interface designers can no longer assume or take for granted that their end-users are always stationary (i.e. seated or standing still) or not physically constrained by other activities (e.g. carrying objects, holding children, opening doors). While a considerable amount of research has examined the effects of mobility, for example interacting with mobile devices while walking (e.g. [5,7,8,40,53,65]), there is a lack of empirical work in the literature that has examined the effects of *encumbrance* on the way users input with touchscreen handheld devices.

The word '*encumbrance*' is defined as "*an impediment or burden*"[1] and in the context of using mobile devices, users could be impeded in many different situations such as operating machinery, holding onto things in a bus and carrying bags. The research presented in this thesis investigates mobile interactions in one specific encumbrance situation: when users are carrying everyday objects while on the move. With the continuous development of mobile devices, users are no longer limited to interacting with computing technology in a fixed location, so when mobile apps and input techniques are used out of this context, the interaction difficulties and potential usability issues are unknown.

People frequently use or hold physical objects while on the move (for example, carrying shopping bags after a visit from supermarkets or using an umbrella when it is raining). In these contexts, users are likely to experience input difficulties as mental and visual

[1] Oxford Dictionaries: http://www.oxforddictionaries.com/

resources are required to attend to situational distractions (e.g. navigating and keeping personal space) and the physical effects of carrying objects can further limit interaction with mobile devices. There is a shortage of studies investigating what input performance is like and how users adapt under these demanding circumstances. As a result, it is difficult for researchers to design more appropriate and effective interaction techniques for encumbered situations.

In comparison, the effects of mobility on interactions with mobile devices have been well studied. Walking is a common daily physical activity for most people and frequently performed to travel between places (e.g. to go to work, shops and return home), but it is also one context that can cause usability issues with handheld devices. For instance, visual attention can no longer be fully used for interaction, as users have to divide its resources for situational distractions, such as avoiding nearby obstacles or ensuring personal and other peoples' safety when crossing a road, for example. As stated by Oulasvirta *et al.* [58], interaction with handheld devices in public spaces is often fragmented into short 4 - 8 second periods, so it is often difficult to maintain continuous attention for mobile usage. Furthermore, extraneous body movements caused when walking makes input with mobile devices physically challenging and therefore selecting on-screen elements is likely to be more error prone. Sears *et al.* [67] conceptualised this problem as "*situational impairment*", meaning a user's inability to interact or input with his/her personal devices efficiently due to situationally-induced constraints. As a result, a body of empirical research now exists that shows the negative impact walking has on mobile interactions (e.g. [5,7,45,66]) and some researchers have also designed ways to improve input in walking situations (e.g. [22,35]).

However, little research has examined the effects of walking and encumbrance together on input performance with touchscreen devices. As the popularity of touchscreen mobile devices continues to grow and along with the number and variety of mobile apps, it is important to understand how basic interactions are affected under different encumbered contexts so that researchers can explore more suitable ways to improve usability when input becomes problematic. A well designed mobile application or interaction technique should be effective in a range of different conditions and should attempt to correct and assist the user if performance declines. Therefore, the research presented in this thesis examines the effects of encumbrance and mobility on abstract targeting performance with touchscreen mobile phones and provides researchers with the basis to design better interaction techniques to improve input in these highly demanding situations.

Mobile phones were evaluated in the thesis because they are most likely to be used while in motion and their physical characteristics (i.e. limited screen space) make designing better user interfaces and interaction techniques a greater challenge. The main goal of the research described in this thesis is to measure abstract input performance on touchscreen mobile phones in order to gain a better understanding of the usability issues in walking and encumbered contexts, and give suggestions as to how researchers can improve mobile input in these situations. While basic tapping to select on-screen icons, buttons, text, etc. is still the most fundamental method of interaction with mobile devices, the introduction of touchscreens has given users other touch-based inputs to manipulate on-screen objects and user interfaces. A variety of single- and multi- finger on-screen gestures are now common to interact with mobile apps. Therefore, the performance of these standard on-screen gesture inputs was also examined while users were walking and encumbered to provide baseline measurements for future work to build upon.

As the research discussed in this thesis was the first to extensively examine the effects of encumbrance and mobility, little was known as to which evaluation methods were the most appropriate to use. This is a non-trivial issue since there are two physical factors (encumbrance and mobility) that could have an impact on input performance and to ensure any changes observed can be correctly quantified and reasoned, additional care is required when designing user studies. The evaluation methods used in the studies presented in this thesis build on ones from the literature and take them further. Furthermore, later on in this thesis, a comparison is made between two ground and treadmill evaluation methods that are commonly used for walking studies to standardise an approach for researchers to conduct future mobile and encumbered user studies.

The remainder of Chapter 1 is as follows. The next section defines the term "*encumbrance*" in the context of the research presented in this thesis. Then, a set of research questions is defined, which are answered by the findings of the user studies described in the experimental chapters. Chapter 1 concludes by discussing the main contributions of the research presented in this thesis.

1.1 The Definition of Encumbrance

While there are many different ways in which users might be physically constrained during mobile usage, this thesis focuses on one type of encumbered situation: the effects of carrying everyday objects while interacting with mobile phones. More precisely, the term *encumbrance* in the context of the research described in this thesis is defined as:

"Carrying objects which physically impede the user's hands during interaction with touchscreen mobile devices."

There are many different types of objects that users might carry while using handheld devices and the way in which these objects are held can vary greatly. Therefore, this thesis focuses on a set of common everyday encumbrance scenarios, where each situation directly limits the user's hands to input on a touchscreen mobile phone. Encumbrance scenarios where users carry objects but do not have an immediate physical effect on their hands or arms (for example, carrying a backpack) are not evaluated in this thesis. The objects and how they were held, which formed the set of encumbrance scenarios, were defined by observing the way users interacted with their mobile devices in public when encumbered.

Other forms of encumbered situations where the user is performing different manual tasks at the same time are not examined in the research presented in this thesis. For example, using mobile devices in driving contexts, while operating machinery in workplaces and holding onto hand-rails and poles while in a train or subway. The effects on interaction and input performance with mobile devices in these environments are out of this thesis' scope, which extends the literature on interactions while walking by evaluating the usability issues when users are carry objects in hand or underarm while on the move.

1.2 Research Questions

This section defines a set of the research questions. Each in turn is answered by the empirical findings from the user studies discussed in the experimental chapters.

Q1 *What are the most common types of encumbrance scenarios?*

Q1.1 *What are the typical objects held during interaction with mobile devices?*

Q1.2 *How are these typical objects held during interaction with mobile devices?*

The number of different types of objects that could be held during interaction is vast and the number of possibilities is further increased by the various ways in which the objects could be held, for example a shopping bag might be held by its handles in the hand, around the wrist or between the forearm and upper-arm. Answering *Q1.1* and *Q1.2* means that a specific set of common encumbrance scenarios is defined and evaluated in subsequent user studies to measure the changes in touch performance on mobile phones when users are walking and encumbered.

Q2 *How do encumbrance and mobility affect input performance on touchscreen mobile phones?*

Q2.1 *How do encumbrance and mobility affect tapping performance?*

Q2.1.1 *How does the change of input posture affect tapping performance when walking and encumbered?*

Q2.2 *How do encumbrance and mobility affect the performance of other standard touch-based gestures?*

Answering *Q2* means making measurements of input performance with a touchscreen mobile phone in different walking and encumbered situations to understand how users interact when their hands are busy carrying objects while on the move. To begin, tapping performance is examined since a majority of daily tasks completed on mobile phones require a simple tap e.g. selecting icons, buttons and characters on a virtual keyboard. Next, the position in which the mobile device is held is varied to examine the effects of changing input posture, for example portrait/landscape, one-handed/two-handed, on encumbered targeting performance. Then, the focus is switched to investigating the performance of other forms of on-screen gesture interactions. More specifically, standard types of touch-based gestures that have become a common way to interact with mobile applications: one-finger dragging and two-finger pinching, spreading and rotating are examined to see how well users perform more complex actions on touchscreens in walking and encumbered situations.

Q3 *How to evaluate the effects of encumbrance and mobility in controlled user studies?*

There are several evaluation approaches described in the literature, which have been used to examine the effects of walking on interactions with handheld devices. However, no related work has examined the impact of walking and encumbrance at the same time, so little is known if these approaches are still effective, or if other experimental designs are required. Answering *Q3* means that more suitable evaluation methods can be recommended for future walking and encumbered studies.

Q4 *Can interactions with touchscreen mobile phones be improved for encumbered and walking contexts?*

Q4.1 *What are the appropriate target sizes and target placements for encumbered and walking interactions?*

Q4.2 *Can pointing techniques improve targeting performance while walking and encumbered?*

After examining abstract targeting performance, one key aim of the research presented in this thesis is to use the results and to improve touch input with mobile phones in walking and encumbered contexts. To begin, the results obtained from earlier studies are used to define an appropriate target size, which is likely to increase the probability of selecting a target accurately in physically demanding encumbered situations. Next, one established finger-based pointing technique, which has shown to improve target selections on handheld devices under normal static conditions, is evaluated to see if it is still effective when users are walking and encumbered. Answering *Q4* will show if these two approaches are useful at improving input performance in mobile and encumbered conditions, and suggest research directions for future work. The findings will motivate researchers and designers to examine alternative interactions and develop more effective interfaces for mobile phones to be used in a wider range of everyday contexts.

1.3 Thesis Statement

This thesis examines how input performance with touchscreen mobile phones is affected in walking and encumbered contexts. Because mobile handheld devices are frequently used in physically demanding multitasking situations and there is a lack of empirical findings in the literature to show how users interact under these conditions, on-screen target selections were measured in a variety of typical encumbered situations where users carried objects while walking. This thesis showed a reduction in targeting performance, demonstrated the effectiveness of alternative input techniques to improve usability and suggests more effective evaluation approaches for walking and encumbered studies.

1.4 Contributions

The research described in this thesis makes several contributions. The user studies presented fill a gap in the literature and extend current research that has examined the effects of walking by also investigating the impact of encumbrance at the same time. The quantitative results from the studies demonstrate the extent to which touch targeting on mobile phones is affected and gives researchers the foundations to design more effective input techniques to improve input and usability in these contexts. Comparisons could also be made with other types of encumbrances that were not evaluated in this thesis.

With respect to increasing input performance in walking and encumbered situations, the results from the studies were used to define an appropriate target size of 22.4 x 22.4mm to improve target accuracy and selection times, and subsequently design an icon arrangement for mobile phones, which optimises screen space using a systematic approach. Using this target size increased selection accuracy to almost 100% when users were encumbered and walking. A similar level of accuracy for selecting smaller targets could be achieved in these physically demanding contexts *if* a well-developed input technique is used. Vogel and Baudisch's input technique *Shift* [73], which is arguably the benchmark pointing technique, was evaluated and showed its effectiveness even when used in walking and encumbered situations as it reduced inaccurate target selections to almost 0%. This demonstrates the advantage of *Shift* but other established input techniques might not perform as well under similar encumbered and mobile conditions, thus, the need to re-evaluate them in the future.

This thesis makes an important contribution regarding the use of more appropriate experimental designs for both walking only and walking and encumbered studies. This thesis advocates (i) the use of ground walking over treadmill walking when evaluating the effects of mobility, (ii) the appropriate speed participants should walk when performing experimental tasks and (iii) the need to control the participant's walking speed with a pacesetter to eliminate the trade-off with input performance and to isolate the effects of encumbrance, which will yield more valid results that are a better representation of the user's actual performance in these mobile and encumbered contexts.

1.5 Outline of Thesis

The remainder of this thesis is organised so that each subsequent chapter builds on the previous and answers at least one of the main research questions defined in Section 1.2. Chapter 2 provides an in-depth review of previous work that is related to the research discussed in this thesis. The chapter will be split into four main sections. The first section discusses observational studies that have surveyed the types of objects that are frequently carried during mobile usage. In addition, an extensive analysis is presented of one key user study that investigated the effects of different multitasking scenarios on interactions with mobile devices while users were not walking. The second section discusses the wealth of research in the literature that has examined the effects of mobility (but unencumbered) on targeting performance and interactions with mobile devices. The third section reviews related work that has measured targeting performance with mobile devices in different hand postures and device orientations to discuss the implications encumbrance and mobility might have on these common input postures. The fourth section discusses non-standard pointing techniques that have shown to improve target accuracy on mobile devices and how these alternative input techniques might perform when users are encumbered and on the move.

Chapter 3 describes an observational study that was conducted to identify the typical objects that users frequently carry in their everyday lives and how they differ or are similar to those reported in previous surveys. The chapter also defines a set of common encumbrance scenarios, based on the results from the observational study, that are evaluated in subsequent experiments to investigate the effects of encumbrance on mobile interactions.

Chapter 4 describes a user study which measured tapping performance on a typical touchscreen mobile phone with the main emphasis placed on evaluating a range of the encumbrance scenarios defined in Chapter 3. Chapter 5 presents a user study which examined abstract tapping input on a touchscreen mobile phone but the focus was on how changing input posture affected targeting performance while users are walking and encumbered. Chapter 6 shifts the attention of this thesis to other common forms of touch-based inputs and discusses a user study which examined the performance of standard on-screen gesture interactions when users were carrying objects and on the move.

Chapter 7 presents an evaluation approach that is recommended for future research into the effects of encumbrance and mobility. The proposed approach was designed based on the findings of a user study which compared two main mobile methods to see which was more suitable for walking and encumbered experiments.

Chapter 8 describes several methods to improve targeting performance in encumbered and walking contexts. First, the chapter defines a target size based on the results from the previous user studies described in this thesis to improve target selection accuracy when users are encumbered and on the move. Then, a systematic approach is described to utilise screen space efficiently when the defined target size is used and as a result, an alternative icon arrangement was created. The chapter also explains why *Shift* [73] is an appropriate input technique to use to improve the selection of small targets when users are walking and encumbered. Chapter 8 ends by presenting two user studies that evaluated the effectiveness of these methods. Chapter 9 concludes this thesis by summarising the main findings from previous chapters, discusses the main contributions and limitations of the research presented in thesis, and suggests research directions for future work.

Chapter 2

Background

This chapter discusses previous work related to the research presented in this thesis and is divided into four main sections. The first section starts by reviewing one key study that examined the effects of multitasking on a range of pointing tasks with different computing devices. Then, the first section reviews previous work that has surveyed the types of objects that are frequently used or held during mobile usage and discusses the limitation of these studies and why an observational survey was required to define a set of common encumbrances to evaluate in subsequent user studies.

The second section reviews previous research that has investigated the effects of walking on interactions with mobile devices and discusses one main limitation in that none of these studies have examined how users perform when objects are being carried while on the move. The research reviewed in the second section covers targeting performance on handheld devices while walking but also includes a discussion on other types of interactions, such as reading performance, in mobile contexts. Furthermore, the experimental designs of the walking studies reviewed are discussed to consider suitable evaluation methods for the walking and encumbered studies presented in the experimental chapters of this thesis.

The third section of this chapter reviews previous work that has examined the effects of changing input posture on mobile interactions and discusses how users might switch to a different input position when encumbered and how targeting performance might be affected. Lastly, the fourth section reviews non-standard pointing techniques that were developed to improve targeting performance on mobile devices and discusses how these alternative interaction techniques might perform when used in walking and encumbered situations.

2.1 The Effects of Encumbrance

This section begins with an in-depth review of one main user study that examined how well users performed different pointing tasks with various computing devices in a range of multitasking scenarios. Then, the section describes related work that observed the objects people use in their everyday activities.

2.1.1 Input Performance when Multitasking

In encumbered situations where the user's hands are occupied with other activities, it is likely to be more difficult to maintain normal levels of performance with mobile devices. Oulasvirta and Bergstrom-Lehtovirta [57] examined a set of multitasking situations where users held a range of objects while inputting on mobile devices at the same time. Their study explored how successfully users readjusted in order to handle the object and operate the mobile device simultaneously. Oulasvirta and Bergstrom-Lehtovirta defined this type of encumbered interaction as "*the user's ability to maintain high performance upon the introduction of secondary manual demands*". A set of non-walking manual tasks, named as the Manual Multitasking Test (MMT), was developed to replicate common encumbrance scenarios where the user's hands and arms are hindered by holding the during input on the device. The differently sized objects were selected to reproduce physical constraints to the shoulder and upper-arm (holding a basketball underarm), hand (grasping a cigarette packet), wrist (mug), and fingers (pinching a pair of tongs/pegs). The test also assessed handedness (dominant vs. non-dominant) and support (device either held in-hand or rested on a table). An overall performance ratio was created to measure how well users multitasked, named as the Manual Multitasking Index (MMI) and was calculated as:

$$\frac{average\ performance\ of\ the\ constraint\ conditions}{average\ performance\ of\ the\ baseline\ condition}$$

MMI ranged from 0 - 1, where a value of '0' suggested that "*performance floors in every condition*" while '1' represented "*top performance reachable with each constraint*". MMT was deployed on two pointing tasks: (1) target selections on a laptop via three different input devices (mouse, trackpad and trackpoint[2]) and (2) text entry on a mobile phone by using three different input modalities (physical keyboard, finger or stylus input on a soft/virtual keyboard).

The results from Oulasvirta and Bergstrom-Lehtovirta's target acquisition study showed that using the mouse was faster at selecting targets than both the trackpad and trackpoint when users were unconstrained. However, mouse input was affected the most when multitasking since it had the lowest MMI (0.6) compared to the trackpad (~0.85) and trackpoint (~0.76). Interestingly, the MMI for each input device was similarly matched when the large object (basketball) was held underarm. However, as the size of the object was reduced (holding the pen and cigarette packet for example), MMI was much lower for the mouse than both the trackpad and trackpoint.

The results from Oulasvirta and Bergstrom-Lehtovirta's text entry study on a mobile phone showed that typing with a physical keyboard performed better than both touch-based and stylus typing on an on-screen keyboard while unencumbered. Interestingly, touch typing on the virtual keyboard performed better than using the stylus and the physical keyboard when objects were held. Oulasvirta and Bergstrom-Lehtovirta suggested that less force is required to select keys on the virtual keyboard while more energy is needed to press the keys on the physical keyboard, which was more difficult to perform when the user was constrained by the object. Text entry using the stylus had the worst performance in both unconstrained and encumbered situations because fine motor-control is required to grasp the stylus and input simultaneously. Users also found it difficult to re-position the stylus when encumbered with an object, which caused awkward input postures and therefore pointing to be less effective and prone to errors.

The work of Oulasvirta and Bergstrom-Lehtovirta has extensively examined the effects of multitasking and showed that users were able to maintain almost normal levels of performance in certain encumbered situations while input suffered greatly in others. The main limitation of their study is that it only focused on encumbered interactions when users were either sitting down or standing still. Users frequently operate mobile devices while walking and when objects are held at the same time, input is likely to become even more difficult and error prone. The user studies presented in the experimental chapters of this thesis fill this gap in the literature by investigating a different set of encumbrances that are commonly performed when users are on the move. In addition, MMI was not used as a performance metric when analysing the results from the user studies described in this

[2] http://shop.lenovo.com/gb/en/laptops/thinkpad/

thesis because the tasks measured abstract target selections using one input method only (touch), unlike Oulasvirta and Bergstrom-Lehtovirta who compared three different input modalities.

2.1.2 The Objects Held During Mobile Interactions

One important issue needs to be addressed when examining the effects of encumbrance - what are the typical objects that users frequently carry or interact with? In Oulasvirta and Bergstrom-Lehtovirta's study [57], holding objects such as scissors, pens are mugs were evaluated but these object types are unlikely to be held in public settings when users are on the move. Several ethnographic studies have observed the way users interact with mobile devices in their everyday lives and reported different forms of objects.

Mainwaring *et al.* [49] conducted an ethnographic study across three major cities (London, Los Angeles and Tokyo) to examine the types of personal belongings that young professionals (aged between 22 - 32 years) carried in everyday life. One main theme from their observations was that participants who took part in the survey carried various type of bags the most often to hold mobile devices (such as cell phones, cameras and music players) and other personal items (like keys, wallets and bank cards). Some users also carried multiple bags to group similar kinds of items together for ease of access or personal preference. Jain [34] also carried out an ethnographic study to explore the use of mobile phones in day-to-day activities. In one particular scenario, a participant was seen struggling to carry heavy boxes when using a mobile phone. The bulky characteristics of boxes are likely to consume the user's upper body which makes input on mobile devices physically challenging and restricted. Similar to the findings of Mainwaring *et al.* [49], participants placed mobile devices in personal bags for safety and security while providing ease of access when required.

Tamminen *et al.* [70] conducted an ethnographic study in Helsinki to observe users in day-to-day mobile contexts to provide design guidelines for future context-aware mobile technologies. Participants who took part in their survey were seen performing a variety of different activities. In one instance, a user performed a series of secondary activities while making a call on a mobile phone: from getting money from a bag to pay for goods to selecting the correct doorbell to enter a building. The latter task was tested by Oulasvirta and Bergstrom-Lehtovirta [57] who found that users were able to press buttons placed on a wall while typing on a mobile phone using an on-screen keyboard (MMI = 0.8). In another example, a user was pushing a pram while making a call with a mobile phone. In Oulasvirta and Bergstrom-Lehtovirta's multitasking test [57], a similar pushing scenario was simulated while performing a text entry task on a mobile phone and results showed that users were able to perform these activities with relative success (MMI was between 0.6 - 0.7).

The bags and boxes described in the mentioned studies are just some of the objects that users could hold when interacting with mobile devices while on the move. With the exception of the work by Oulasvirta and Bergstrom-Lehtovirta [57], one limitation of the mentioned ethnographic studies is that they do not describe *how* the objects are held which could have a different impact on the user's ability to input. Therefore, an observational study was conducted to identify the common types of objects and the way they are held by users while walking. The results from the study and subsequently the definition of encumbrance scenarios that were evaluated in following user studies are discussed in Chapter 3. The next section discusses related work which have examined interactions with mobile devices when users are situationally impaired while on the move.

2.2 The Effects of Mobility

Mobile devices are frequently used while walking and researchers have acknowledged the potential interaction difficulties with handheld devices in mobile contexts. Some research have also examined the risks of using mobile devices while on the move and have reported that a large number of users were unaware of surrounding events during interaction as their attention was focused on the device. For example, in Hyman *et al.*'s observational study [33], people who were engaged with their mobile phones made more abrupt directional changes and weaved to take sudden evasive action than those who were not using handheld devices. Furthermore, Hyman *et al.* deployed an unusual visual distractor in a public area and found that 75% of mobile phone users did not perceive the stimulus. Hatfield and Murray [26] conducted an in-depth observational study to examine pedestrian behaviour at road crossings and found that mobile phone users were less likely to check on-coming traffic than people who were not using mobile devices. Nasar *et al.* [55] reported similar findings to Hatfield and Murray and reported that 43.2% of pedestrians observed were using some form of mobile device when crossing over roads. Perhaps the extraneous movements caused when walking, which is likely to make input on mobile devices more physically demanding, is playing a part in utilising too much of the user's visual attention and mental resources. Therefore, many studies in the literature have examined the effects of walking on interactions with mobile devices. Some researchers have also designed more effective interaction techniques for mobile contexts to improve targeting performance and perhaps this will allow users to make better judgements and divide their attention between activities in a more efficient manner while on the move.

This section begins by reviewing related research that investigated targeting performance in mobile contexts. While this thesis examines finger-based input on touchscreen mobile phones, the walking studies discussed in this section also discusses stylus-based target selections. Before the popularity of touchscreen mobile devices, personal digital assistants (PDA), which require a stylus for input, were commonly used since they provide similar applications to those found on modern handheld devices. At the time when PDAs were gaining popularity, researchers have already started examining the effects of walking and the important findings from these studies are worth discussing. The section also reviews studies that investigated text entry performance on handheld devices while walking. Text entry is one form of continuous target selection and is perhaps one of the most important and common pointing tasks performed on mobile devices. This section will end with a short review on previous work that assessed reading comprehension and visual search performance with mobile devices while users were on the move. Visual capacity could have an impact on targeting performance in mobile situations and therefore deserves a brief discussion.

2.2.1 Stylus-based Input in Mobile Contexts

In one of the earlier experiments that assessed the effects of mobility, Brewster [8] conducted several user studies to see if sonically-enhanced feedback could improve button selections on PDAs. In Brewster's mobile study, users walked along a quiet outdoor pathway and performed a data entry targeting task. The results showed that auditory feedback increased the number of correct data codes entered regardless of the size of the buttons. When comparing targeting performance between standing and walking, the negative impact of mobility was shown. For example, the number of data codes entered using 8 *x* 8px buttons without the aid of audio feedback decreased by 40% when walking compared to sitting down. Data entry performance also declined with larger 16 *x* 16px buttons as the mean number of codes entered while seated was 45 compared to 32 when walking, a decrease of 28.9%. The findings of Brewster's study have highlighted the poor input performance in walking situations but target selections can be improved with the support of auditory feedback.

Pirhonen *et al.* [62] also explored the use of auditory feedback on PDAs in conjunction with gestural input as an alternative eye-frees interaction technique for mobile situations. One advantage of using non-visual interfaces over standard touch-based input while walking is that visual attention can then be fully used for navigation and to attend environmental obstacles. A music application was developed to compare gesture-based selections with audio feedback against standard visual input using a stylus. Unlike the outdoor experimental setup of Brewster's walking study [8], a figure of eight path was arranged in a corridor for users to navigate while performing the task. The results from Pirhonen *et al.*'s experiment showed that overall subjective workload (using the NASA TLX method[3]) was reduced when their gestural interface with auditory support was used while walking. Furthermore, overall task completion times were faster when the gestural interface was used compared to basic tapping, which illustrates the potential usefulness of eyes-free interaction in mobile contexts.

The Percentage Preferred Walking Speed (PPWS) metric was also used in Pirhonen *et al*'s study to evaluate and compare the different interfaces. PPWS was previously introduced by Clark-Carter *et al.* [10] and later used by Petrie *et al.* [61] to evaluate the efficiency of mobile technologies for visually-impaired users. An increase or decrease from the ideal walking speed can indicate that there are usability issues with the interface or technology being examined. An increase in PPWS of 9% and 24% was reported for the gesture- and tapping- based interfaces respectively. This showed that users were more able to maintain their preferred walking speed with the gestural interface.

While non-touch gesture-based input modalities and non-visual interfaces can be beneficial for mobile contexts, tapping on handheld devices is still the *de facto* input method because eyes-free interactions have shown to be slow [14,16] and it can be difficult to design a set of intuitive gestures for applications that have many functions. Therefore, research has

[3] NASA TLX: Task Load Index: http://humansystems.arc.nasa.gov/groups/tlx/

generally examined standard tapping performance in mobile contexts to make better design choices and develop more suitable interaction techniques without changing input modality.

Mackay *et al.* [45] investigated the effects of mobility on stylus-based target selections and on-screen scrolling behaviour. Three different scrolling techniques were compared: standard scroll bars at the side of the screen, *tap-and-drag* (tap anywhere on the screen and dragged the stylus in the required direction) and *touch-n-go* (scroll speed and direction were determined by the distance and the touch position relative to the centre of the screen respectively). The scrolling techniques were evaluated while users were either sitting down without resting their arms on a surface, "sudden stop" walking where users were instructed to stand for a predetermined time period when a beep was played by the PDA, and "continuous" walking where users walked at their own pace without stopping.

The results from Mackay *et al.*'s experiment showed that target selections took longer when walking than sitting down. The overall mean selection times were 3445ms and 3777ms for sitting and continuous walking respectively, an increase of ~9.5%. Standard scrollbars performed poorly regardless of the user's mobility while both *tap-n-drag* and *touch-n-go* had better and very similar target selection times when users were walking. In terms of subjective preference and ease-of-use, *tap-n-drag* and *touch-n-go* were both ranked higher than the standard scrollbars. Scrolling is a fundamental and common on-screen activity on mobile devices since the amount of the information that can be presented at any given time is limited by the screen space. Therefore, designing better scrolling techniques could have a major impact on targeting performance. In cognitively demanding mobile contexts where it is difficult to maintain continuous visual interaction due to situational interruptions [58], choosing the most appropriate input technique can substantial improve usability.

Lin *et al.* [43] assessed stylus-based target selections on a PDA and used various evaluation methods to measure input performance while walking. Three walking contexts were examined: walking slower (80%) or faster (120%) than the Preferred Walking Speed (PWS) on a treadmill and natural walking around a pre-defined indoor path with obstacles. Users completed a Fitts' Law [20] style targeting task where target size ranged from 1.9mm and 6.4mm. The results from Lin *et al.*'s experiment showed no statistical significances between the different levels of mobility for target selection time. The mean selection times were very similar between the three walking contexts (1.07 - 1.09 seconds) while input speed when sitting down was only marginally faster (1.05 seconds). In terms of error rate, a statistical significant effect was reported between the different levels of mobility. On-ground walking around the path resulted an overall mean error rate of 29.6% which was higher than both slow and fast walking on the treadmill, 17.8% and 20.1% error rates respectively with the latter being statistical significant. The mean error rate for sitting down was 17.1% which is comparable to the error rate for slow walking on the treadmill. As reported by the authors, participants were poor at selecting the smallest 1.9mm targets while seated which shows that very small target sizes should be avoid regardless of the user's position. While increasing target size reduced error rate across all walking conditions, the number of inaccurate selections for the largest 6.4mm targets while walking the obstacle course was still relatively high at 8.9%. Target size would need to increase even further to improve tapping accuracy when users are walking and encumbered.

Performing target selections also had an effect on walking speed in Lin *et al.*'s experiment. The walking speed for navigating the obstacle course decreased from 3.53km/h (non-interacting) to 2.25km/h (during target selections), a statistical significant difference of 36.3%. Target accuracy while walking slower than the preferred pace on the ground still could not match the input performance of when seated, which demonstrates the level of inaccurate input that mobility can cause. Interestingly, slow walking on the treadmill, which had the same mean PWS as walking the obstacle course, resulted in a near identical error rate as sitting down perhaps due to visual attention not being interrupted by situational distractions. In contrast, Musić and Murray-Smith [53] found that tapping performance was poorer when users walked on the ground slower than their ideal walking speed. Furthermore, research into the biomechanical movements of the human body (such as Murray *et al*. [52] and Alton *et al.* [1]) has found physical differences in gait behaviour between treadmill and ground walking. For example, Murray *et al.* reported that participants had a faster cadence and shorter stride length when walking on a treadmill than on the ground. Therefore, using treadmills for mobile studies might not give an accurate reflection of input performance of real-world walking behaviour. The user study presented in Chapter 7 compares the treadmill walking and ground walking evaluation approaches to investigate the effectiveness and limitations of both methods for walking and encumbered experiments.

To sum up, this sub-section has discussed research that has examined stylus-based target selections and interactions in walking situations. While the brief discussion on auditory feedback and alternative gesture-based input methods have shown that they could improve targeting performance in mobile contexts, standard tapping on screens is still favoured but results from the mentioned studies showed a decline in stylus-based targeting performance.

The next section reviews related work which has investigated touch-based targeting performance on handheld devices while walking.

2.2.2 Touch-based Input in Mobile Contexts

Schildbach and Rukzio [66] examined thumb-based targeting on a touchscreen mobile phone while users walked around an outdoor path. Three different target sizes were evaluated: 6.74 x 6.74mm, 8.18 x 8.18mm and 9.50 x 9.50mm. The results showed that selection time for the smallest target width increased by 31.25% when walking (603ms) compared to the baseline targeting speed when standing still (459ms). The 8.18mm and 9.50mm target widths were not tested when standing. The 9.50 x 9.50mm targets took longer to select when users were walking than selecting the smaller 6.74 x 6.74mm targets when standing, which demonstrates the negative impact of walking on input speed. For error rate, users made more inaccurate selections when walking than when standing still. The mean error rates for selecting the 6.74 x 6.74mm targets when walking and standing were 30% and 23% respectively, an increase of 7%. The number of incorrect selections when users were walking reduced as target diameter increased. An error rate of 21% was measured for selecting the 8.18 x 8.18mm targets when walking which was lower than tapping on the 6.74 x 6.74mm targets. The largest target size reduced inaccurate selections further as an error rate of 16% was reported, which shows the performance gains of using bigger on-screen objects. The implications of increasing target size on small touchscreen mobile devices are discussed in Chapter 8.

Similar to Lin *et al.* [43], Schildbach and Rukzio's mobile study also revealed the cost of interaction on the user's walking speed and walking behaviour. Users walked slower during target selections than when walking alone without interaction. The mean PWS without interaction was 4.26km/h and decreased to 3.11km/h when selecting the 6.74mm target widths. As target size increased, walking speed was marginally faster, as speeds of 3.27km/h and 3.38km/h were reported for the 8.18mm and 9.50mm target widths respectively. Video recordings also revealed abrupt walking behaviour during input as users stopped walking frequently. The smallest target size evaluated in Schildbach and Rukzio's study was based on design recommendations from Apple to define button sizes for iPhones[4]. These guidelines are perhaps suitable for ideal conditions but when users are walking, Schildbach and Rukzio showed that selecting 6.74mm target widths was highly inaccurate. More appropriate recommendations are required to allow researchers and designers to develop better user interfaces that are effective in a range of contexts.

The studies of Lin *et al.* [43] and Schildbach and Rukzio [66] have shown the benefits of using larger target sizes (or wider target diameter) to improve stylus- and touch- based tapping accuracy while walking. However, increasing the size of on-screen components can cause design challenges when developing user interfaces due to limited screen space on mobile devices, which restricts the amount of information that can be presented. One common problem as a result of making targets bigger is the increase in on-screen navigation (i.e. scrolling). While tapping accuracy is expected to improve, target search and selection times are likely to increase considerably which might not be ideal in walking situations since continuous interaction has shown to be difficult in mobile contexts [58].

Kane *et al.* [35] tried to address the speed vs. accuracy trade-off with selecting large targets on touchscreen mobile devices while walking by exploring adaptive user interfaces that dynamically changed depending on the user's mobility. A music application was developed on an ultra-mobile personal computer (UMPC) where targets enlarged in walking situations while the design of the interface stayed the same when users were standing. In their pilot experiment, users were given a music selection task to complete by selecting songs from a list. The size of each item in the list ranged from 3.43mm and incrementally increased to 13.73mm. Users walked along a corridor at a pace of 112 steps per minute and kept within this speed by using an audio pacesetter. Kane *et al.* found no statistical significant difference for average task time between standing and walking, and reported that the longest task times occurred when target size was either very small or very big. In terms of selection errors, no statistical significant effect between standing and walking was found. Errors per trial for walking dropped steeply between 3.43mm and 5.15mm but then quickly levelled off for each target size up to 13.73mm. Small targets can cause selection difficulties, whether the user is standing or walking, as shown by the studies discussed so far. However, non-standard pointing techniques can improve the selection of small targets, even in walking and encumbered situations, which will be discussed in greater detail in Chapter 8.

After the pilot study, Kane *et al.* conducted a main experiment to evaluate their adaptive user interface where the size of targets adjusted according to the user's context. The

[4] Apple Inc. (2010). iOS Human Interface Guidelines: https://developer.apple.com/library/ios/navigation/

dynamic user interface was compared to a static layout with either 3.81 x 3.81mm or 11.43 x 11.43mm targets. A music application was developed and users were given a series of music selection tasks to perform which varied in difficulty. Similar to Brewster [8] and Schildbach and Rukzio [66], participants walked around a pre-defined outdoor route. However, walking speed was controlled as each participant walked at the same predetermined walking pace by following a human pacesetter. The results from the main experiment showed no statistical significant difference for task time between standing and walking. However, as noted by Kane *et al.*, removing their adaptive interface from the analysis showed that walking increased task time when compared to standing. The mean task time when using the static interface with the larger 11.43 x 11.43mm buttons was 2864ms when standing, compared to 3396ms for walking, an increase of ~18.6%. The longest task time of 3988ms occurred when the small buttons were used on the static interface while walking, compared to 3412ms when standing. Input speed between standing and walking for the adaptive user interface was near identical, 3453ms and 3402ms respectively. The results for selection error showed that very few incorrect selections were made across all conditions. Kane *et al.* suggested that their implementation of adaptive user interfaces did not performed as well as expected and that using a simpler interface with larger buttons was more suitable for walking contexts. Logically, expanding target size is likely to increase selection accuracy but at the expense of longer selection time for scrolling tasks. However, choosing accuracy over selection time might be more preferable in walking (and encumbered) contexts since an inaccurate selection is likely to take even longer to recover the error.

The user studies mentioned this far have illustrated the negative impact walking has on targeting performance. One strategy that users might adopt when input becomes too difficult is to compromise walking speed (i.e. walking slower) to aim at the touchscreen more accurately. Bergstrom-Lehtovirta *et al.* [7] found an optimal trade-off between the user's PWS and targeting performance on a touchscreen mobile phone. In order to control and vary the user's PWS, a calibrated treadmill was used. Participants performed an abstract targeting task where on-screen crosshair targets were selected as quickly and as accurately as possible. The results from their experiment showed the impact of mobility on input, as walking at 100% of PWS caused target accuracy to drop to approximately 80% while reducing PWS by 20% resulted in a slightly higher target accuracy of approximately 83%. Between 30% - 90% of PWS, a near linear relationship was observed between walking speed and the reduction in targeting accuracy. However, the drop in performance was gradual and an optimal performance trade-off for walking between 40 - 80% of PWS was reported. Interestingly, data from the mobile phone's accelerometer showed that despite extraneous movements in the user's non-dominant hand when walking, the inputting hand was able to negate the instability and therefore maintain a reasonable level of targeting performance. A mean PWS of 3.90km/h was reported for walking alone (without interaction) while PWS during target selections dropped to 2.97km/h, a difference of 24%. Walking speed may decrease further when users are also carrying objects to trade with targeting performance.

To sum up, previous work reviewed in this sub-section has showed a decrease in finger-based targeting performance on touchscreen mobile devices when users are walking. While increasing the size of targets has shown to improve selection accuracy for walking situations, limited screen space on mobile phones present other issues and design challenges such as more on-screen navigation and longer task times. The studies discussed this far have also showed a reduction in walking speed during input and is likely to decline further when users are also encumbered.

2.2.3 Text Entry Performance in Mobile Contexts

Text entry is one form of continuous target selection and is a task commonly performed on mobile devices. Researchers have realised the difficulties of typing with on-screen soft keyboards while walking since extraneous body movements make it physically challenging to accurately select small and densely packed keys. Therefore, user studies have investigated the effects of mobility on text entry performance with mobile devices.

Mizobuchi *et al.* [50] examined the effects of walking on text entry performance on a PDA by varying key size gradually from 2.0 *x* 2.5mm to 5.0 *x* 6.3mm. Users were asked to type short phrases on the PDA while either standing still or walking along a corridor. The results showed no statistical significant difference for typing speed between standing and walking but text entry using the smallest key size of 2.0 *x* 2.5mm was slower than the larger key sizes. In terms of error rate (number of uncorrected errors divided by the number of characters in each phrase), walking caused a higher number of errors than standing. The smallest keys caused the highest error rate while walking and increasing target size reduced typing errors. The results from Mizobuchi *et al.*'s experiment also reported a mean walking speed of 1.77km/h during input, which is slower than those figures reported in the user studies discussed so far in this chapter. This finding suggests that walking speed might have been traded with typing accuracy since no error rate, for each key size, while walking was greater than 3.5%. An effective text entry system should allow users to input in an efficient manner without compromising walking speed greatly.

Yatani and Truong [80] also investigated text entry on a PDA while walking, but explored the use of the non-dominant thumb as a supplement to improve stylus-based typing performance. Their hybrid chord keyboard was compared to three other types of keyboards: mini-qwerty, hand-writing and quikwriting (a gesture-based typing method). Each keyboard type was evaluated in three poses: sitting down, walking along an indoor path and walking a stairway. The results showed that mobility had a negative impact on input as typing speed using the baseline mini-qwerty keyboard decreased from 20 words per minute (wpm) while seated to ~15wpm when walking around the path and further declined to ~12wpm when walking the stairway. Typing speed for the hybrid chord keyboard was near identical across all three contexts at ~10wpm. The results for error rate (number of backspace key presses and uncorrected errors divided by the length of the phrase) showed that walking caused a higher number of typing errors than standing across all keyboards. Hand-writing performed particularly poor, especially when walking as error rate increased to ~50%, which suggests it is less effective in walking situations. Users found it difficult to "write" on the screen when walking and therefore caused ambiguous letter detection by the device. Interestingly, Mizobuchi *et al.* [50] reported that slow walkers preferred hand-writing input since it required less visual attention than the mini-qwerty and chorded keyboards. The chorded keyboard resulted in fewest errors when walking and therefore shows potential in using the non-dominant thumb to supplement typing performance. A mean PWS without interaction of ~3.9km/h was reported, while walking speed dropped to ~2.0km/h when typing on the PDA, a noticeable decrease of 48.7%. Similar to Mizobuchi *et al.* [50], users walked much slower during input but typing performance was poor. If users are also encumbered, typing performance is likely to decline even further.

Nicolau and Jorge [56] examined the effects of walking and input posture on text entry performance on a touchscreen mobile phone. Participants in their experiment were asked to walk around an indoor path and maintained walking speed by following a human pacesetter, like Kane *et al.* [35]. Two predetermined walking speeds were chosen: 2 steps per second based on the findings of Barnard *et al.* [5] to represent normal walking speed and 1.3 steps per second to replicate slower walking at 65% of normal. Typing performance while walking was examined in three input postures: one-handed single thumb or both thumbs with the device held either in portrait or landscape mode. Overall typing speed between slow walking, normal walking and when seated were near identical at ~25wpm. One-handed typing using the preferred thumb was slower than two-handed text entry which shows the advantage of having an extra finger for input to increase typing speed when walking and perhaps also when encumbered. The participants in their study also subjectively preferred and were faster at typing when the device was held in landscape mode during two-handed input perhaps due to the wider keys. The error rates for walking at the normal predetermined pace and sitting down were comparable, 7.5% and 5.1% respectively. Nicolau and Jorge [56] also found that users made more walking errors during text entry and at times could not keep up with the experimenter when walking at the normal predetermined pace.

Clawson *et al.* [11] conducted an extensive user study to examine thumb-based typing on a physical QWERTY keyboard in three settings: sitting, standing and walking. Unlike the walking studies discussed this far, more sophisticated hardware was used for the walking conditions, as a figure of eight path was created indoors with motion sensors used to track the users' movement for accurate walking speed and distance measurements. An average of 991 meters walked per 20-minute session was reported, which resulted in an average walking speed of 2.73km/h. As expected, the results showed that typing speed was near identical between sitting and standing (56.79wpm and 56.61wpm respectively), while text entry was slower when walking at 52.51wpm, a decrease of ~7.5% when compared to the static positions. Typing accuracy between sitting down and standing was again very similar (95.36% and 95.25% respectively). Surprisingly, typing accuracy decreased by less than 1% while walking when compared to the sitting and standing. Clawson *et al.* reasoned that improvements in keyboard technology on mobile devices and increasing typing expertise on physical keyboards could negate the decline in performance when walking. However, walking speed was not controlled in their experiment so it is difficult to know if walking speed was compromised for typing accuracy. It is also uncertain whether the high level of typing performance can be maintained when users are also encumbered since Oulasvirta and Bergstrom-Lehtovirta [57] reported difficulties with pressing physical buttons on mobile phones when the user's hands were busy handling objects.

To sum up, this sub-section reviewed previous work that has investigated text entry performance in mobile contexts. Text entry is one form of continuous target selections and is a very common task that users perform on mobile phones. Although this thesis does not examine text entry in walking and encumbered contexts, the user studies presented in the experimental chapters inform the level of targeting performance under these conditions and could help researchers and designers to develop more effective text entry systems that are useful in a range of everyday situations.

2.2.4 Visual Search and Reading Performance in Mobile Contexts

The related work discussed so far in this section has examined targeting performance on mobile devices when users are walking. While tapping on touchscreens or pressing physical buttons on handheld devices is challenging when walking, part of being able to select targets and other user interface elements successfully is also due to the amount of visual attention that can be used for interaction. As reported by Oulasvirta *et al.* [58] and Tamminen *et al.* [70], continuous interaction is difficult to maintain since situational distractions compete for the user's visual attention and mental resources. As a result, interaction is often fragmented in mobile settings as visual awareness is divided between different activities. This sub-section reviews previous work that has examined reading and visual search performance with handheld devices in mobile settings.

Lim and Feria [42] conducted an experiment to measure visual searching performance on a mobile phone while users walked around an indoor course. The task required users to locate the orientation of the target letter among a number of distractors. The experiment evaluated two letter sizes (6.74mm and 9.5mm), two contrasts (black and grey) and also the placement of letters. The results from their experiment showed that reaction time took longer when walking than when standing. The mean reaction times for standing and walking were 1641ms and 1957ms respectively, an increase of 19.3%. Users also took longer to locate the larger targets than the smaller ones. The mean reaction times reported for the 6.74mm and 9.5mm targets were 1770.37ms to 1828.ms respectively, a statistical significant increase of 3.2%. Lim and Feria suggested that because there was less spacing between the large letters and its distractors, this made it more difficult to quickly distinguish the on-screen targets. This finding is supported by non-mobile studies such as

Everett and Byrne [17] and Tseng and Howes [72] who examined visual searching performance of graphical user interfaces on desktop computers and found that users were much slower when icons and on-screen information were densely packed together.

Changing the contrast ratio of the letters did not have any statistical significant effect on reaction time, although the experiment took place in a controlled lab setting whereas constant variations in lighting conditions in the real world is likely to have an impact on visual ability. The placement of the targets had a statistical significant effect on reaction time. In general, reaction time was quicker when targets were located near the centre of the screen than those placed close to the edges. However, once users were walking, the difference in reaction time between targets placed at the centre and the edge of the touchscreen was much smaller, which suggests that it is more challenging to optimise target placement to improve visual searching performance in mobile contexts.

Mustonen *et al.* [54] examined the effects of walking on reading performance and pseudo-text searching on a mobile phone that had a small screen (display resolution of 176 x 208px) when compared to the size of modern touchscreen mobile devices. Users performed the reading tasks in four mobile settings: walking along a corridor and on a treadmill at the measured PWS (a mean walking speed of 3.7km/h was reported) and predetermined speeds of 1.5km/h and 3.0km/h. Walking faster not only degraded targeting performance as shown by Bergstrom-Lehtovirta *et al.* [7], but also had a negative impact on visual interaction as Mustonen *et al.* reported that "*Visual performance deteriorates with increased walking speed*". Mean error rates of approximately 11% and 12% for walking at 1.5 and 3.0km/h respectively on the treadmill were reported, while error increased to 14% when walking on the treadmill at 3.7km/h. On-ground walking caused the highest error rate of approximately 17% since users also had to navigate the route. Walking faster also increased subjective workload, which illustrates high mental demands on users during visual interaction with small handheld devices in mobile contexts.

Barnard *et al.* [5] investigated the impact of walking on reading comprehension and word searching on a PDA. The reading comprehension task required users to read a short passage of text and followed by answering a multiple-choice question. The word searching task asked users to locate a specific word in a short section of text by tapping anywhere on the line that the word appeared on. Barnard *et al.* also compared treadmill and ground walking, and were one of the first to introduce obstacles within the path to replicate realistic walking environments. In additional, lighting condition was varied to simulate contextual changes that users are likely to experience in outdoor settings.

The results from Barnard *et al.*'s reading task showed no statistical significant difference for reading time (the duration taken to read the passage), response time (the duration taken to answer the question), the number of correctly answered questions and the number of scrolls made to read the passage between treadmill and ground walking. For each dependent variable, the results were similar between the two walking methods. The mean reading times ranged from 27.93 - 29.79 seconds and the average score (out of 10) ranged from 7.68 - 8.02 between the different walking and lighting conditions. Further analysis from the reading task found a statistical significant difference for both response time and the number of scrolls made between the two lighting conditions. Users made more scrolls and took longer to answer the questions in low lighting conditions.

The results from Barnard *et al.*'s word searching task showed that users took less time for treadmill walking than ground walking. The same mean word searching time of 3.9s was reported for both lighting conditions while walking on the treadmill. Searching time increase to 4.03s and 4.67s for the high and low light conditions respectively while walking around the obstacle course. There was no statistical significant difference for the number of correct selections as users scored surprisingly higher across all walking and lighting conditions (the score was at least 9.55). In terms of walking speed, users dropped their PWS (mean = 2.19km/h) by 30.1% and 37.9% during the reading comprehension and word searching tasks respectively when walking around the obstacle course, which further illustrates the compromise in walking speed required for interaction.

The study by Schildbach and Rukzio [66] also examined reading comprehension while users walked around an outdoor track. A similar reading task to Barnard *et al.* [5] was used and three text sizes were evaluated: 2.20mm (height of an uppercase letter "H"), 2.64mm (20% increase) and 3.08mm (40% increase). The results showed that reading speed was slower by 18.6% when walking (190 wpm) than when standing (155 wpm). No statistical significant differences were found for reading speed between the different text sizes when walking since the advantage of having larger text for better readability was negated by the amount of scrolling required to read the entire text passage.

Performing the reading task in Schildbach and Rukzio's study also caused PWS to decrease by 26.3%. The walking speed reported for no interaction was 3.88km/h and dropped to 2.86km/h when reading text passages in the smallest text size. Increasing text size did not have a statistical significant effect on walking speed as the measured walking speeds were near identical across all three sizes. Video analysis of the users' walking behaviour showed varied results. For example, the largest text size caused users to slow down their walking speed less often than the smaller text sizes. However, users stopped to a standstill and looked up at the environment more often when reading the largest text size. Schildbach and Rukzio's findings suggest there is no clear advantage of increasing text size for better visual performance in mobile contexts but when on-screen elements are too small, it is also difficult to read and to select accurately.

To sum up, this sub-section reviewed previous work that has examined reading and visual search performance on mobile handheld devices while users were walking. The mentioned studies have reported longer task times when walking compared to when standing still. Furthermore, increasing target or text size for better visual representation in mobile contexts does not give the expected performance gains if the spacing between on-screen elements is not carefully designed and increases the amount of on-screen scrolling activity.

2.3 The Effects of Input Posture

One advantage of dynamic user interfaces on touchscreen mobile devices is that they can be used in either portrait or landscape orientations and allow users to interact in different input posture: one- vs. two- handed input and one vs. multiple finger interactions. When '*input posture*' is discussed in this thesis, no emphasis is placed on the user's mobility (i.e. sitting, standing or walking), only the way users hold and interact with mobile devices. This section reviews previous work that has observed the common input postures that users adopt with interacting with mobile devices and how these psotures affect targeting performance.

The observational study of Karlson *et al.* [36] found that users, when seated, were more inclined to use both hands for interaction while one-handed input was preferred when standing and especially when walking. Furthermore, Karlson *et al.* reported that 60% of users were encumbered (their hands were occupied in some way) while walking, which could explain why a majority of users interacted with mobile devices with one hand. In a follow-up survey, participants stated that they were most likely to use both hands to input if the handheld device had a touchscreen. In addition, device form factor determined which input posture was the most efficient and comfortable to use. Two-thirds of participants favoured one-handed interaction to perform most tasks, while only 9% of participants preferred using both hands to input regardless of the type of device.

Whether one-handed mode is preferred because it frees the non-input hand for other activities or simply down to personal preference, designing interfaces for single handed interactions can be advantageous for encumbered situations. However, issues can arise with one-handed thumb-based interactions due to on-screen targets that are difficult for the thumb to reach without adjusting hand grip or switching to a two-handed input posture. Karlson *et al.* [36] conducted a following study to examine the optimal screen areas for thumb input and reported that for right-handed users, movement between the north-west and south-east directions were the most difficult and slowest to perform. The centre area of the device was found to be the best for the thumb to reach in terms of ease of access.

Hoober [32] carried out an observational study in different public settings to investigate how mobile devices are held. In the 780 observations where a user was seen interacting with a mobile device, 49% of them were one-handed interactions and therefore the thumb was the primary input modality. In the remaining 51%, a two-handed input posture was used and 36% of them were in the “cradle” position where only preferred thumb or index finger was used for input. The final 14% of users who held their mobile device with both hands used both thumbs for input. The observations of Hobber [32] and Karlson *et al.* [36] have identified the different types of one- and two- handed input postures that are commonly used with mobile devices in different contexts. Other research has evaluated these input modes to examine how targeting performance varies on handheld devices.

Parhi *et al.* [59] also examined one-handed thumb input by conducting several Fitts’ Law targeting tasks on a PDA. The results from their discrete targeting task found that selection time decreased as target size increased. Error rate dropped as target size increased while the smallest targets of 3.8 x 3.8mm and 5.8 x 5.8mm caused the most inaccurate target selections. Users also subjectively found targets near the centre of the screen easier to select during discrete pointing. Targets located in the top and bottom far side of the input hand were the most uncomfortable regions to tap since those areas where more difficult for the thumb to reach. The results from their continuous targeting task showed that in general, input speed and error rate both improved as target size increased. Parhi *et al.* recommended targets sizes of 9.2mm and 9.6mm for discrete and continuous targeting tasks respectively. However, it is unknown if these target sizes are still effective when users are walking and encumbered since their study did not examine these contexts.

Wobbrock *et al.* [77] measured target selections in eight different input postures on a PDA which covered handedness (one vs. two hands), input modality (thumb vs. index finger) and input surface (front vs. back of device). A Fitts’ Law pointing task was used to compare the performance of each input posture while users were non-mobile and unencumbered. The results showed that movement time was faster when both hands were used for targeting than single-handed input. Movement time was faster when the index finger was used to select on-screen targets than the thumb. The results also revealed that two-handed input caused fewer inaccurate target selections than one-handed input. Error rate was lower when the index finger was used for targeting in two-handed mode than when the dominant thumb was used in one-handed mode. Perhaps using the index finger meant all areas of the screen can be easily accessed whereas for the thumb to select targets in difficult to reach areas, users would need to adjust their hand grip or even change to a two-handed input posture. Shifting the device to select targets with one hand only is likely to be difficult when the user is encumbered.

Azenkot and Zhai [3] investigated the effects different input postures had on continuous targeting performance by examining text entry on a touchscreen mobile phone. Three input postures were examined: one-handed mode using the preferred thumb and two-handed input using either both thumbs or the index finger. An initial online survey was conducted to investigate the frequency of use of each typing posture and the results showed that no particular input mode prevailed, which suggests the importance of understanding different input postures for tapping tasks. The results from Azenkot and Zhai's main text entry experiment showed that using both thumbs to type was faster than using one thumb or the index finger. A mean typing speed of 50.03wpm was reported for two thumb typing while speed of input reduced to 33.78wpm and 36.34wpm for the preferred thumb and index finger postures respectively. Single thumb typing was only marginally slower than using the index finger. If users can maintain relatively stable input performance with one hand, then it can be advantageous in mobile and encumbered situations since it releases the non-input hand for secondary tasks such as holding objects. It remains to be seen if typing speed will be similar between the preferred thumb and index finger typing postures when users are encumbered and walking as these contexts were not evaluated. In terms of typing accuracy, using both thumbs were better than both single finger input postures. However, despite statistical significant results, the error rates between the three input postures were comparable; 10.80% for two thumb typing, 8.17% when using the index finger and 7% for single thumb input.

The work of Azenkot and Zhai [3] has shown the variation in text entry performance with mobile devices between different input postures. Other researchers have explored ways to detect the user's input posture to improve input performance. Goel *et al.* [23] examined the possibility of detecting input posture based on the user's touch behaviour when typing on touchscreens in order to improve text entry accuracy and speed. They found that each individual input posture had its own unique touch patterns and developed a model to predict and classify four typing modes: left thumb, right thumb, both thumbs and index finger. All participants were right-handed so the left thumb was examined to measure text entry performance of the non-dominant hand for situations where the preferred hand is not available for input, like when encumbered. The results from Goel *et al.*'s non-mobile and unencumbered experiment showed that text entry was performed quickest using both thumbs. Expectedly, text entry was slowest using the left non-preferred thumb. The input posture detection system managed to reduce typing errors than unassisted text entry. As expected, using the non-preferred thumb for typing caused the highest error rate while text entry using both thumbs was less accurate than using the index finger. The results from both Goel *et al.* and Azenkot and Zhai [3] studies have illustrated a trade-off between speed and accuracy when both thumbs are used for typing on touchscreen mobile devices. One advantage of using predictive models to infer the user's context is that no or minimal visual changes are required. Potential performance improvements are done in the background without the user having any knowledge. This can be beneficial for mobile and encumbered contexts where situational distractions constantly compete for the user's visual attention and unfamiliar changes to normal user interfaces might not be effective, as shown by Kane *et al.* [35].

Touchscreen mobile devices now feature a range of built-in inertial sensors and researchers have used them to detecting the user's current input posture. Goel *et al.* [24] combined tapping behaviour with data from the built-in gyroscope to measure changes in the orientation of the device and the amount of force applied to the touchscreen. This allowed their system to predict the current hand posture out of three possible modes: index finger, left thumb or right thumb. Two tasks, one selecting items from a list and the other entering text, were used to test their prediction model. Results showed that the correct mode was detected 81.11% and 87.4% of the time for the text entry and scroll-to-select tasks respectively. The work of Goel *et al.* [23] and Goel *et al.*[24] have shown promise in detecting the user's input posture and other research (e.g. [22,60,82]) has reported the possibility of predicting the user's type of mobility (i.e. standing or walking). If these methods could accurately detect when the user is encumbered then mobile devices can make better decisions to improve usability in physically and mentally demanding situations.

Musić and Murray-Smith [53] conducted a study to examine the effects of the user's gait, walking speed and type of input posture on tapping performance with handheld devices. Five input postures and five walking speeds were examined. Similar to Bergstrom-Lehtorvirta *et al.* [7], tapping performance in terms of error rate (the percentage of correct selections) decreased when users started walking. Overall error rate was lower when the device was held in landscape mode than portrait mode due to wider targets and the index finger outperformed the thumb-based pointing methods. An interesting observation from their results is that the median error rate is higher when walking slower than the normal walking pace. Musić and Murray-Smith reported that the participants subjectively required more concentration when walking at slower speeds. Examining the distribution of the index finger tapping behaviour also supported this finding as users tapped closer to the centre of the target and had less variation when walking at 100% of PWS than both 50% and 75% of PWS.

Musić and Murray-Smith [53] also examined the effects of walking speed and input posture when building offset models to correctly predict the user's on-screen touch position. They showed that with the use of offset models, a reduction in error rate of up to 17% and 30% are achievable when users are standing and walking respectively. However, one limitation stated by the authors is that a specific model is required for each input posture – walking speed combination in order to attain the improvements. Prior work such as [9,74,75] have also used offset models to infer touch behaviour and improve tapping accuracy while users are seated or stationary. No work has yet to explore the possibility of using such techniques to improve tapping performance when users are walking and encumbered.

To sum up, this section reviewed related work that has examined the common types of input postures and its effects on targeting performance. Three broad input postures were identified: one-handed mode using the preferred thumb and two-handed mode using either the index finger or both thumbs. The mentioned studies suggested that most users preferred one-handed input but can cause problems when selecting targets that are out of the thumb's reach without adjusting handgrip. Two-handed mode using both thumbs resulted in faster input than single finger tapping but at the expense of lower accuracy.

2.4 Non-standard Touch-based Input Techniques

One of the main reasons for tapping difficulties on mobile handheld devices is the size of the input finger relative to the size of on-screen widgets (*the fat finger problem*) which causes uncertain selections especially when many targets are closely placed together. The loss of haptic feedback when tapping on smooth touchscreen surfaces also contributes to ambiguous target selections and although research has shown the effectiveness of using vibro-tactile feedback to aid touch input (e.g. [6,27,81]), selecting small on-screen elements accurately is still a challenge for users.

However, Holz and Baudisch [30] suggested that the fat finger problem is not the primary cause of inaccurate target selections on touchscreen mobile devices. They suggested that the input finger's angle (the roll, pitch and yaw of the fingertip) plays a big part in how users point and that user-specific mental models vary the way onscreen targets are selected. A user study was conducted which examined a range of pitch and roll angles of the input finger tapping on crosshair targets and results showed that different users had different touch patterns. Holz and Baudisch [31] examined the problem further by identify both vertical and horizontal visual features of the pointing finger that users use to align with crosshair targets. They found that users point using visual cues along the fingertip while most capacitive touchscreens detect touch events much lower down at the bottom of the finger. However, the user studies of [30,31] were conducted while participants were seated so it is unknown if the same visual cues are used to point while walking and encumbered. The authors also suggested that the use of overhead camera-based sensing to track the pointing finger could substantially reduce error rate. However, such hardware setups are likely to be a challenge when users are walking and encumbered.

Despite these suggestions, novel input techniques have been designed to reduce the number of incorrect target selections when tapping on touchscreen handheld devices. The evaluation of these input techniques has shown promising results as accuracy is increased across a range of target sizes. However, one main limitation of these input techniques is that none have been examined in walking and/or encumbered situations and as a result, it is unknown if these alternative pointing methods are still effective when used in more physically demanding contexts. In section reviews relevant touch-based input techniques and discusses whether these methods could be beneficial to improve touch accuracy when users are encumbered and on the move.

One main cause of uncertain input when targeting on touchscreens is that it is difficult to know if a selection has been made successfully or not when the input finger occludes the target area. Vogel and Baudisch [73] described this occlusion problem as one of the main reasons why finger performs worse than a stylus for input because firstly, users tend to incorrectly estimate the actual selection point and secondly, the input finger blocks any visual feedback. Therefore, Vogel and Baudisch tried to address this "*fat finger*" problem by developing *Shift*, a pointing technique that uses a circular callout to show the occluded area underneath the input finger and allows the user to make precise on-screen adjustments to select the required target accurately, as illustrated in Figure 2.1. The adjustment gives users a second chance to correct their initial selection at the expense of longer completion time.

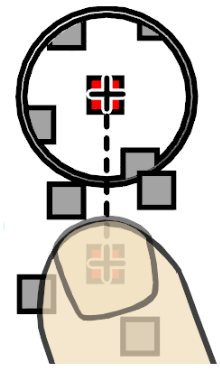

Figure 2.1: Vogel and Baudisch's input technique Shift uses a callout to show the area occluded by the input finger. This illustration was copied from [73]**.**

In Vogel and Baudisch's study, *Shift* was used to select six different target sizes, ranging from 2.6mm and incrementally increased to 41.6mm, on a PDA by using either the fingertip or fingernail. The results showed that error rate was higher when using simple tap for target selections than when *Shift* was used for both fingertip and fingernail. The highest mean error rate of 81% occurred when selecting the smallest 2.6mm targets using simple tap with the fingernail. For the same target size, *Shift* reduced error rate to ~5% which shows its effectiveness for selecting very small targets. The error rates for both simple tap and *Shift* were evenly matched (< 5%) for targets above 10.5mm and is perhaps the threshold in which *Shift* does not gain any performance benefits, for non-mobile interaction at least. In terms of target selection time, *Shift* was slower than unassisted target selections for targets between 5.25mm - 10.5mm.

Prior to *Shift*, Potter *et al.* [63] implicitly explored the target occlusion problem on touchscreens by comparing three different touch-based selection strategies. The first technique *land-on*, either selected a target or not based on the initial touch down position.

The second technique *first-contact* was similar to *land-on* but selected the first selectable target. The third technique *take-off* used an offset cursor placed 12.7mm above the current touch point which allowed for continuous movement and correction. *Take-off* reduced any occlusion problems of blocking targets by the input finger since selection was made indirectly by positioning the cursor over the required target. The results from their target selection experiment showed that *first-contact* was faster than *take-off* while no statistical significant results were found between *land-on* and the other two techniques. In terms of error rate (selecting the wrong target or a white space), *take-off* caused fewer incorrect selections than both *land-on* and *first-contact*. This finding was later supported by Vogel and Baudisch [73] and illustrates the effectiveness of allowing users to correct their initial selection if it is unsuccessful.

Sears and Shneiderman [68] extended the work of Potter *et al.* [63] by comparing finger and mouse input modalities for selecting a range of target sizes. For touch targeting, the *take-off* technique was used and enhanced by using cursor stabilisation to improve precise control. The results from their experiment showed that for selection time, touch input was faster than using the mouse for 6.9 x 9.0mm targets. However, mouse input was quicker than finger input for very small, single pixel 0.4 *x* 0.6mm targets. For error rate, touch input without cursor stabilisation caused more inaccurate target selections than with stabilisation and mouse input. Target selection using the mouse resulted in fewer errors than finger input for single pixel targets. One main drawback of using an offset cursor is that is not possible to select targets placed at the bottom of the screen without shifting the interface upwards or switching to an alternative input method. Vogel and Baudisch [73] reported that users found it too confusing when the entire interface was shifted during the development of *Shift*. Also, in walking and encumbered contexts where the user's attention and resources are overworked, adding complexity to the interface could have a negative impact on usability.

Yatani *et al.* [79] tried to improve targeting time of *Shift* by developing their own input technique *Escape*, which also eliminated the finger occlusion problem but without the need of an inset to show the enclosed area beneath the finger. Targets within the touch position were colour coded and had visual "beaks" (see Figure 2.2) attached to them to illustrate the gesture direction to select the required target if the initial selection was incorrect. One advantage of *Escape* is that it requires less precision to select the appropriate target than *Shift*, which could make it easier to use when users are walking and encumbered. The results from Yatani *et al.*'s experiment showed that *Escape* was faster than *Shift* for all target sizes tested, which ranged from 6 to 24 pixels. In terms of error rate, no statistical significant differences were found between *Escape* and *Shift*. One limitation of *Escape* is that when targets get very small, the advantage of the visual indicators diminishes since users had difficulties seeing the targets clearly. Also, error rate was generally higher when gesturing towards areas that were difficult for the thumb to reach, which can be even more challenging when users are also carrying objects while on the move.

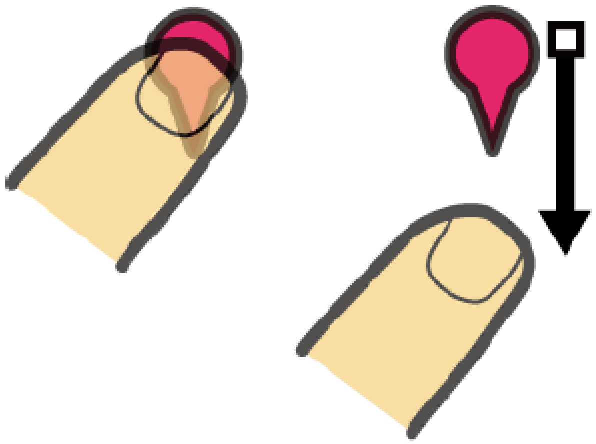

Figure 2.2: Yatani et al.'s input technique Escape. Visual 'beaks' were used to determine the gesture direction to select a target. This image was copied from [79]**.**

Au *et al.* [2] developed *LinearDragger*, a pointing technique similar to *Escape* but mapped a two-dimensional target selection problem to a simplified one-dimensional task. Similar to *Shift*, a circular callout was used to show the region occluded by the input finger. *LinearDragger* calculates the targets within the initial touch region which are then ordered depending on the direction of movement. Each target is then placed into equal regions and requires the finger to move in one direction to select a particular target area. *LinearDragger* requires less precise motor control than *Shift*, so perhaps it could be more effective for selecting targets when users are walking and carrying cumbersome objects. *LinearDragger* was evaluated on a touchscreen tablet and was compared to standard touch, *Shift*, *Escape* and *Bubble* [25], which is an area-cursor input technique but will not be discussed since it was designed for mouse input.

The target sizes examined in Au *et al.*'s experiment were very small, ranging from 0.8 to 3.2mm. The results showed that target selections were faster when *LinearDragger* was used than all other input techniques. An overall mean selection time of 2557ms was reported for *LinearDragger*, compared to 4384ms for *Escape* and 3574ms for *Shift*, which were increases of 71.5% and 39.7% respectively. Unassisted touch targeting performed the worse in terms of input speed as an overall mean selection time of 8336ms was reported. *LinearDragger* also had the lowest mean error rate of 7.2%, followed by *Shift* (9.6%) and then *Escape* (29.1%). Unassisted touch input caused the highest mean error rate of 35.3%, which demonstrates the usefulness of pointing techniques to improve targeting performance.

The input techniques discussed so far in this section have mainly addressed the target occlusion problem to improve selection accuracy and were mostly designed for two-handed interaction. As discussed earlier, one-handed input is also commonly used but it is difficult to select targets that are out of the thumb's reach. Therefore, input techniques have been developed to improve one-handed thumb-based target selections.

Karlson and Bederson [37] developed an input technique called *ThumbSpace*, which made difficult to reach targets more accessible for the thumb to select. A screenshot of the current interface was replicated into a smaller overlay area that was positioned to where the user felt was most comfortable to interact with their thumb. The proxy area was then used as a control to select any targets on the screen, which meant movement distance for the thumb is greatly reduced (see Figure 2.3). A target acquisition experiment on a PDA was conducted to evaluate the effectiveness of *ThumbSpace* on two target sizes (4.8 x 4.8mm and 9.6 x 9.6mm).

The results showed that *ThumbSpace* caused an overall mean selection time of 2068ms, which was slower than direct touch (811ms). No statistical significant differences were found between direct touch and *ThumbSpace* for the number of correct target selections as both had accuracy rates over 90%. There was small 2% difference in accuracy between direct touch (92%) and *ThumbSpace* (94%) for targets placed at difficult to reach regions. One reason why direct touch performed well could have been due to the relatively small 3.5" touchscreen, which meant users could access most areas without too much difficulty. The size of touchscreens found on modern mobile phones has greatly increased (for example, the Google Nexus 6 mobile phone has a 5.96" touchscreen). *ThumbSpace* is likely to be more effective for such devices but its effectiveness remains to be seen when used in walking and encumbered contexts.

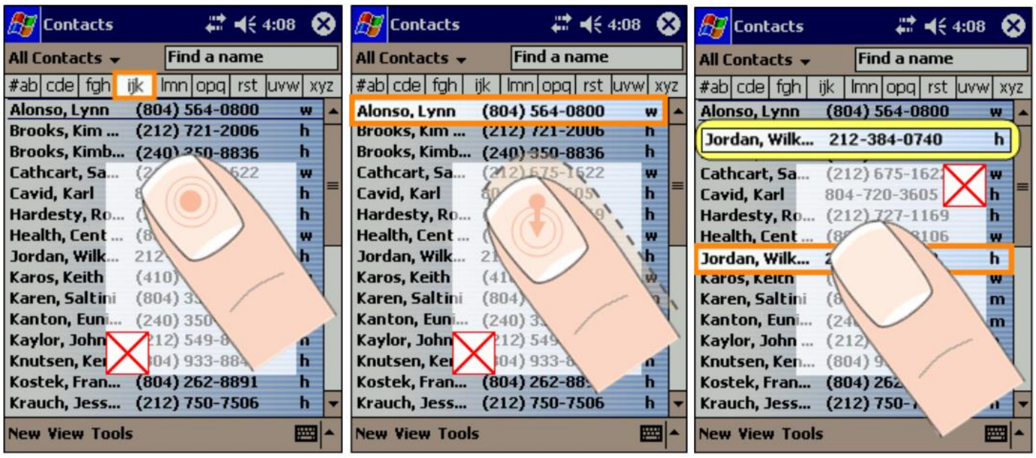

Figure 2.3: Karlson and Bederson's thumb-based input technique ThumbSpace. This illustration was copied from [37]**.**

Later, Karlson and Bederson [39] examined different input techniques for one-handed interactions with handheld devices. They compared *ThumbSpace* to *Shift* and target selections using physical buttons. The input techniques were used to select 3.6 x 3.6mm targets on a PDA with a 2.8" screen. The tasks were completed either standing still or walking around a pre-defined route with walking speed uncontrolled. The results showed that target selection time was fastest when direct touch was used (865ms) while *Shift* (1581ms) and *ThumbSpace* (2,027ms) were both much slower. However, unassisted touch targeting caused the highest error rate of 32% while users made fewer errors when *Shift* (8%) and *ThumbSpace* (10%) were used. Furthermore, both input techniques had low error rates regardless of where the targets were located whereas direct touch caused more inaccurate selections when targets were placed in difficult to reach areas. Interestingly, no statistical differences were found in targeting performance between standing still and walking. Both thumb-based pointing techniques were also tested on a larger 3.5" touchscreen device and users preferred *Shift* for nearby target selections since it allowed faster input while *ThumbSpace* was more suited for targets that were too far or uncomfortable for the thumb to select. The results from Karlson and Bederson's study have illustrated the efficacy of input techniques to improve targeting performance for one-handed thumb-based interactions. However, it is difficult to predict if these input techniques are still effective and allow users to maintain a standard level of targeting performance in encumbered situations.

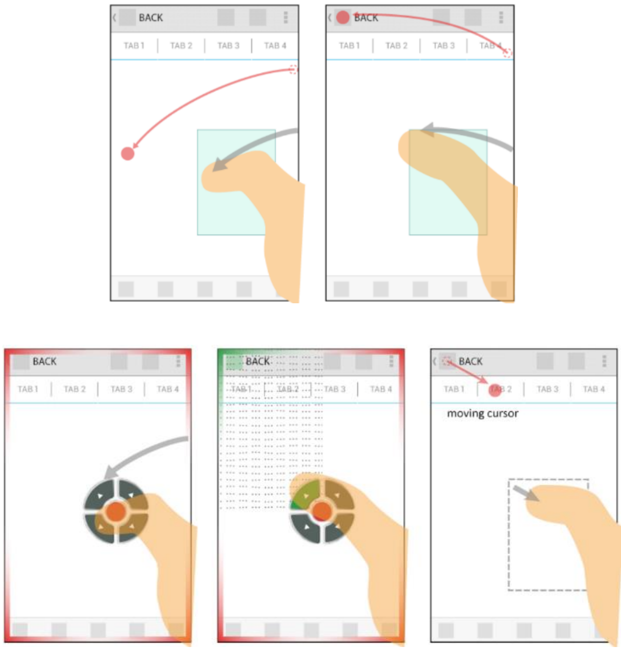

Figure 2.4: Yu et al.'s thumb-based input techniques: BezelSpace (top) and CornerSpace (bottom). These illustrations were copied from [83]**.**

Similar to *Shift*, one main limitation of *ThumbSpace* is the slow target selection time when compared to standard thumb input. Therefore, Yu *et al.* [83] developed two thumb-based pointing techniques: *CornerSpace* and *BezelSpace* (see Figure 2.4) to improve input speed. Both *CornerSpace* and *BezelSpace* determined the thumb's optimal input region and required only small movements within a proxy area like *ThumbSpace* to select distant and difficult to reach targets. A target selection task was used to compare *CornerSpace* and *BezelSpace* with *ThumbSpace*, where 7mm targets were placed at difficult to reach areas on a 5.3 inch touchscreen mobile device. The results showed that selection times were very similar between *ThumbSpace* (2222.4ms) and *CornerSpace* (2213.7ms) while *BezelSpace* (1456.6ms) was faster than both techniques. In terms of error rate, both *CornerSpace* and *BezelSpace* performed similarly (<10%) while *ThumbSpace* caused the higher error rate of 16%. Like *ThumbSpace*, Yu *et al.*'s input techniques minimises the amount of thumb movement necessary to select distant targets by using a proxy area, but it is difficult to predict how these alternative input methods will perform when users are encumbered since they all require precise thumb control, which could be physically challenging to maintain when users are carrying objects while on the move.

2.5 Conclusions

To conclude, this chapter began by reviewing the most related user study to the thesis that investigated the effects of multitasking on targeting performance with mobile devices. Oulasvirta and Bergstrom-Lehtovirta [57] tested a set of multitasking scenarios and found that certain encumbrances caused poorer targeting than others. However, users in their study were not walking while multitasking so the effects of encumbrance and mobility remains to be examined. The chapter then discussed ethnographic studies that investigated how users interacted with mobile devices in everyday situations and one common theme was that different types of bags were commonly used. One limitation of the surveys is that they do not mention how the objects were held so it is difficult to replicate exact encumbrance scenarios for evaluation. The objects examined in Oulasvirta and Bergstrom-Lehtovirta's multitasking study [57] are more unlikely to be held when users are on the move. Therefore, Chapter 3 presents an observational study to answer research questions *Q1.1* and *Q1.2* to identify the typical objects that users carry while walking and *how* these objects are held. This allowed a set of common encumbrance scenarios to be defined and evaluated in the user studies presented in the experimental chapters of this thesis.

This chapter then reviewed related work that examined the effects of mobility on targeting performance with handheld devices. The mentioned studies have shown that walking caused users to make more inaccurate target selections, particular when tapping on small targets. However, none of the walking studies reviewed also examined the effects of encumbrance on input. Therefore, Chapter 4 presents a user study that's main focus was to investigate tapping performance on a touchscreen mobile phone in a range of encumbrance scenarios (defined in Chapter 3) to answer research question *Q2.1* and see how targeting performance is affected. Chapter 6 also examined tapping performance and other standard forms of touch-based gesture inputs to answer research question *Q2.2* and investigate the performance of one- and multi- finger interaction while walking and encumbered.

The experimental design of the walking studies reviewed was also discussed to select appropriate evaluation approaches for subsequent user studies. There were two main methods: ground walking where participants navigated around a pre-defined route or walking on treadmills. Several studies that used ground walking also deployed a human pacesetter to avoid walking speed causing an effect on input performance. It is unclear if these evaluation approaches are appropriate for experiments that examine the effects of walking and encumbrance, so in Chapter 7, a user study is presented which compared ground and treadmill walking to answer research question *Q3*.

This chapter then discussed previous work that examined the main input postures that are commonly used with touchscreen mobile phones and how changing input mode affected targeting performance. Three main input postures were identified: one-handed input using the preferred thumb and two-handed interactions using either the index finger or both thumbs. The studies showed that during one-handed input, targets that were out of the thumbs reach were selected poorly while accuracy was traded with faster input speed when both thumbs were used during two-handed interactions. However, none of the mentioned studies examined the different input postures when users are walking and/or encumbered. Therefore, Chapter 5 presents a user study to answer research question *Q2.1.1* to see if changing input posture had any impact on targeting performance in walking and encumbered situations.

This chapter ended with a discussion on non-standard input techniques that were designed to improve target selections on touchscreen handheld devices. The input techniques reviewed addressed two main issues with tapping on touchscreens: uncertain selection due to the size of the input finger relative to the size of on-screen targets and the selection of targets that are difficult to reach during one-handed thumb input. While the mentioned input techniques have shown to improve target accuracy at the expense of longer completion times when compared to unassisted touch targeting, again, none of them were tested in encumbered contexts. Therefore, Chapter 8 presents a user study to answer research question *Q4.2* by examining a benchmark input technique to test its effectiveness when used while walking and encumbered. Chapter 8 also answer research question *Q4.1* by defining and evaluating an appropriate target size to allow users to maintain a high level target accuracy when walking and encumbered.

Chapter 3

User Study 1: Defining Encumbrance Scenarios

3.1 Introduction

This chapter answers research questions *Q1.1* (*What are the typical objects held during interaction with mobile devices?*) and *Q1.2* (*How are these typical objects held during interaction with mobile devices?*) by presenting an observational study which examined the types of objects that users frequently carried while on the move and *how* these objects were held. The results from the observational study were used to define a set of common encumbrance scenarios that were examined in the studies presented in the experimental chapters. This chapter begins by describing the method used to conduct the observational study in terms of the types of locations surveyed and the encumbrance characteristics that were recorded. Then, the results from the study are presented and the chapter ends by stating a range of encumbrance scenarios. For ease of reference, the observational study will be denoted as *User Study 1* and all subsequent studies discussed in the following chapters will have a designated user study identifier.

3.2 Method

In *User Study 1*, the public was observed in three different types of locations in the city centre of Glasgow: a popular main street, a key train station and a major supermarket as shown in Figure 3.1. The locations were chosen because they are well-populated areas and surveying them would increase the variety of objects that are frequently held while on move and how users adapted in these situations could be observed. *User Study 1* was conducted in December 2011 and each type of location was examined at peak times due to the expected high influx of people. The train station was observed between the hours of 08:00 – 10:00 due to the number of commuters travelling to work in the morning. The main street location was examined between the hours of 16:00 – 18:00 due to people completing errands after work. And lastly, the supermarket location was surveyed from 12:00 – 14:00 due to people going for their lunch.

Figure 3.1: The three types of location examined in *User Study 1*. Glasgow Central station - a major railway terminal (left), Sauchiehall Street - one of the main streets in the city centre of Glasgow (middle) and 24-hour major supermarket (right).

The observations were noted down from a quiet and unobtrusive area at each location to avoid distracting nearby people. No video recording equipment was used. For each instance where a person was seen to hold or use an object (or objects), the following characteristics were documented:

- Type of object (i.e. bags, boxes, disposable cups etc.)
- The way the object was held or used (i.e. directly in-hand, underarm etc.)
- Number of objects held at each instance (one or multiple objects being held)
- Number of hands encumbered (one or both hands encumbered)
- The input posture used if interacting with mobile devices (one- or two- handed interactions)
- Type of mobility (seated, standing, walking or intermediately switching between positions)

Before presenting the results from *User Study 1*, several limitations of the observations need to be discussed. The exact size and dimensions of each object observed could not be precisely measured. Therefore, the main focus was to identify the main object types. The weight of the objects observed also could not be measured, so later in this chapter when the encumbrance scenarios have been defined, a pilot study is presented which measured the weight of objects that user can carry comfortably while interacting with mobile phones. And lastly, the type of interaction (i.e. replying to an email, reading a webpage etc.) performed on the mobile devices could not be seen, so the observations could only document the input posture used.

3.3 Results

Each location was examined for two hours giving a total of six hours of observational data. A total of 878 objects was recorded for the entire survey (an average of 2.4 objects observed per minute). As shown in Figure 3.2, 554 (63.1%) of the total number of objects recorded were held when interacting with mobile devices while the remaining 324 (36.9%) objects were carried alone without interaction. The following sub-sections present the results for each object characteristic stated in Section 3.2.

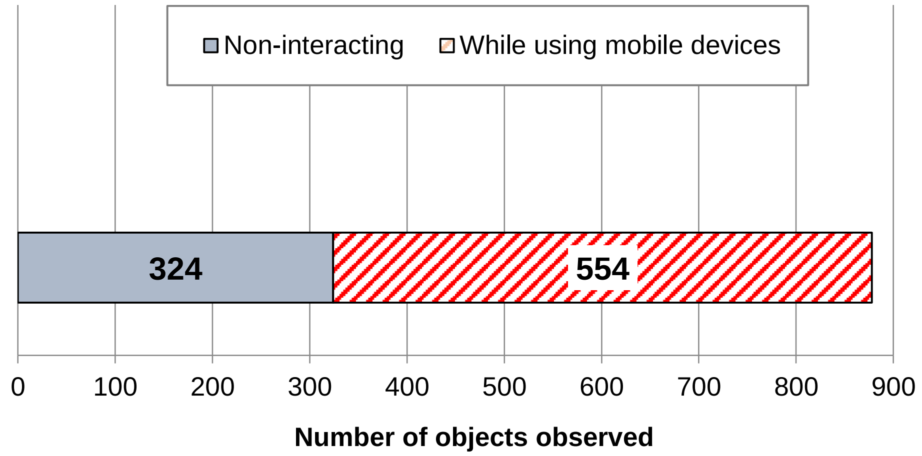

Figure 3.2: The number of objects held while not using mobile devices (blue solid bar) and during interaction (red striped bar).

3.3.1 Type of Object

The entire data set was then grouped by object type and seven main object categories were identified, as shown in Table 3.1. The most common object type observed was different forms of bags, which made up of 49% of the total number of objects recorded. 409 out of 430 log entries (95.1%) when bags were held, users were interacting with mobile devices. The second most frequent type of object surveyed was different types of boxes, which made up 35% of the total number of objects recorded. 41.7% of the boxes documented were held during mobile usage and were mainly grasped underarm.

Child-related objects such as prams, pushchairs, baby baskets and holding children took up 7.4% of the total number of situations recorded. Only one out of 65 entries did a user interact with a mobile device while standing to holding onto a pushchair to prevent it from moving. Beverages bought from food outlets made up 4.2% of the total number of entries recorded and 13.5% of those instances were mobile devices used at the same time. At the train station, people were seen pulling luggage along the ground, which took up 2.4% of the total number of recordings. In one third of those entries, users were looking at or inputting to mobile devices. During one session observing the main street location, the change in weather meant umbrellas were used, which made up of 1.3% of the total number of observations documented and mobile devices were used in 3 out of 11 instances. The last main type of object observed was pet-related, which took up 0.8% of the total number of entries logged and only in one instance was a mobile device used at the same time.

	Non-interacting	**While using mobile devices**	**Total**
Bags	21	409	430
Boxes	179	128	307
Child-related *	64	1	65
Disposable cups	32	5	37
Luggage **	14	7	21
Umbrellas	8	3	11
Pets-related ***	6	1	7
Total	**324**	**554**	**878**

Table 3.1: The count of each main object type when held alone and when using mobile devices. *Child-related objects included prams, pushchairs, baby baskets and holding children. **Luggage that was pulled along the ground. ***Animals guided by leads.

3.3.2 Type of Mobility

Figure 3.3 illustrates the entries sorted by type of mobility. There were three main forms: stationary (either sitting down or standing still), continuous walking, and occasionally switching between standing and walking. Objects were held while walking in 70.4% of the total number of entries documented. Of these recordings, 62.4% of users were walking, carrying objects and interacting with mobile devices all at the same time. Objects held in a stationary position took up 17.8% of the total number of observations recorded. In 62.2% of these situations, mobile devices were used while holding objects. In the remaining 11.8% of the total number of entries recorded, people switched between walking and almost to a standstill to avoid nearby people and obstacles. Of these trials, 68.2% of them were interacting with mobile devices.

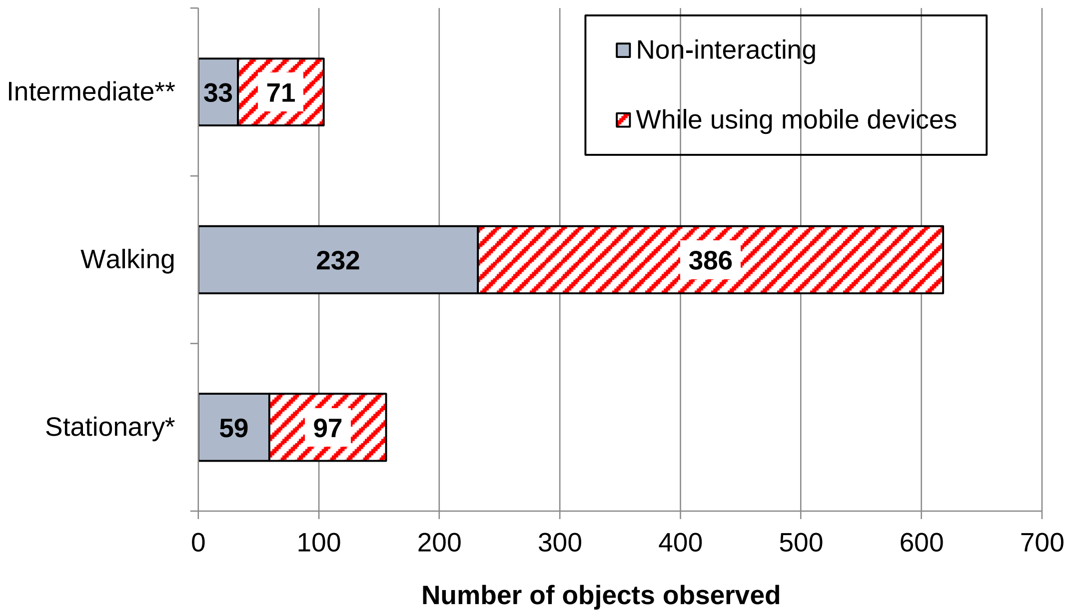

Figure 3.3: The number of objects recorded when held alone and during interaction for each type of mobility. * Either sitting down or standing still. ** Alternating between walking and almost to a standstill.

3.3.3 Number of Objects Held Per Instance

Figure 3.4 shows the observations ordered in terms of the number of objects held in each instance. For simplicity, two categories were defined: holding a single object or multiple objects. When multiple objects were held, the category was split further into instances where all the objects were held in one hand or shared between both hands. After arranging the data in this way, there were a total of 634 entries. In 432 cases (68.1%), a single object was held and 278 of those instances were during interaction with mobile devices. In the remaining 202 cases (31.9%), multiple objects were held in one hand, which took up 90 instances. The unencumbered hand was used to interact with mobile devices in 56.6% of these situations. In the remaining 112 cases, objects were held in both hands and Mobile devices were also used in 60 of these instances.

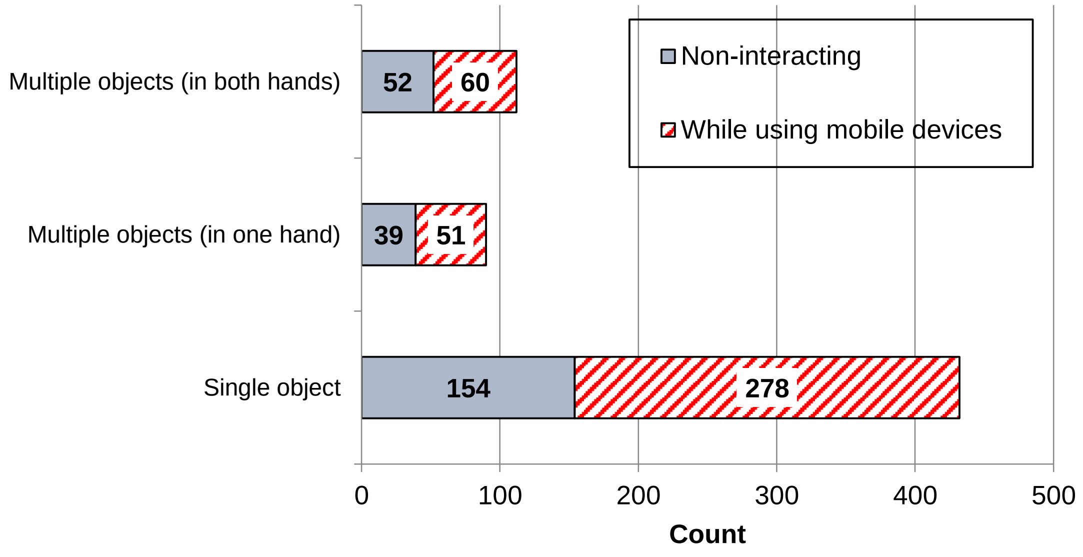

Figure 3.4: The count of either a single or multiple objects being held. The occurrences for holding multiple objects were grouped into situations where one hand held all the objects and cases where the objects were carried in both hands.

3.3.4 Type of Interaction

Of the 554 occurrences out of the total number of entries recorded where mobile devices were used and objects were held at the same time, 378 (68.2%) of these instances were one-handed interactions. Two-handed interactions accounted for 176 (31.8%) recordings but micro-level observations such as the orientation of the mobile device and the exact two-handed input posture (if either the preferred index finger or both thumbs were used for input) could not be clearly recorded due viewing difficulties.

3.3.5 How Object Types Were Held

In the instances where bags were carried, two main approaches were observed as shown in Figure 3.5. The bags were either held directly in-hand or placed along the forearm. In both of these situations, the object had a direct physical impact on the user's hands or arms. However, bags were also seen carried over the shoulders or across the body by using the attached straps. In these circumstances, no hindrance is caused to the user's hands or arms and therefore, these types of encumbered situations were not recorded during *User Study 1*. Of the 430 observations when bags were being carried, 317 (73.7%) instances were held in-hand while in the remaining 113 cases (26.3%), the bags were carried along the forearm.

Figure 3.5: The two main ways different types of bags were held during *User Study 1*. Directly in-hand (left) or placed along the forearm (right).

In terms of the way different types of boxes were held, two main methods were noted which mainly depended on the size of the object. Boxes were either carried underarm or grasped using both arms if the object was too large to be held single handed, as shown in Figure 3.6. Out of the 307 trials where boxes were being carried, 149 (49%) instances were held underarm while the remaining 158 (51%) cases were carried using both hands. The remaining five object types were mainly one-hand encumbered situations. Beverages, umbrellas, pet-related objects and luggage were all operated single-handed. Children were typically held hand-in-hand while prams and pushchairs were normally operated using both hands.

Figure 3.6: The two main ways boxes were held during *User Study 1*. Underarm (left) or grasped using both arms (right).

3.4 Discussion

To answer research questions *Q1.1* and *Q1.2*, the results from *User Study 1* identified seven main object types. Different types of bags were held most frequently and accounted for 49% of the total number of objects recorded. Mainwaring *et al.* [49] also reported that bags were often used to hold personal items for day-to-day activities. The bags observed in *User Study 1* were typically held in-hand or carried along the forearm by using the attached handles of the bags.

The second most surveyed object type during *User Study 1* were different types of boxes, which took up 35% of the total number of trials recorded. Carrying boxes gave a different form of physical impedance than holding bags due to the lack of dedicated handles, which makes boxes potentially more challenging to grasp securely and causes a more awkward body posture. The boxes were held underarm (especially using mobile devices at the same time) or carried using both arms (if the object was too large to hold under one arm).

Based on these frequent occurrences, bags and boxes were chosen to define a set of encumbrance scenarios to evaluate in the studies presented the following experimental chapter. The other five object types: beverages, umbrellas, luggage, child-related and pet-related, accounted for the remaining 16% of the total number of trials observed in *User Study 1*. These object types will not be examined in this thesis since they are less frequently held when users are on the move and the difficulty of replicating them in a controlled experiment.

3.5 Encumbrance Scenarios

In this section, a set of encumbrance scenarios is defined based on the observations from *User Study 1*. The encumbrance scenarios were replicated and evaluated in the user studies described in the following chapter to measure targeting performance on touchscreen mobile phones in these physically demanding encumbered contexts. Bags and boxes were the two main object types commonly carried during *User Study 1*. The bags observed were typically held in-hand while boxes were carried underarm, so the following six encumbrance scenarios were created:

Encumbrance Scenario 1A: A bag held in the non-dominant hand during two-handed input

Figure 3.7: Encumbrance Scenario 1A - a bag held in the non-dominant hand during two-handed interaction.

Encumbrance Scenario 1B: A bag held in the dominant hand during two-handed input

Figure 3.8: Encumbrance Scenario 1B - a bag held in the dominant hand during two-handed interaction.

Encumbrance Scenario 1C: A bag held in each hand during two-handed input

Figure 3.9: Encumbrance Scenario 1C - a bag held in each hand during two-handed interaction.

Encumbrance Scenario 1D: A bag held in each hand during one-handed input

Figure 3.10: Encumbrance Scenario 1D - a bag held in each hand during one-handed interaction.

Encumbrance Scenario 2A: A box held in under the non-dominant arm during two-handed input

Figure 3.11: Encumbrance Scenario 2A - a box held under the non-input arm during two-handed interaction.

Encumbrance Scenario 2B: A box held in under the dominant arm during two-handed input

Figure 3.12: Encumbrance Scenario 2B - a box held under the input arm during two-handed interaction.

Encumbrance Scenario 1A (Figure 3.7) evaluates situations where a bag is held in the non-dominant hand when both hands are required for input. *Encumbrance Scenario 1B* (Figure 3.8) is similar to *Encumbrance Scenario 1A*, but the bag is now held in the dominant hand. In both scenarios, the mobile phone is held in the non-dominant hand while input is completed using the preferred index finger of the dominant hand. *Encumbrance Scenario 1C* (Figure 3.9) is a continuation to scenarios *1A* and *1B* but a bag is held in each hand during two-handed interactions to simulate situations when the user is carrying multiple objects and both hands are encumbered. *Encumbrance Scenario 1D* (Figure 3.10) is similar to scenario *1C* but tests one-handed interaction instead of using both hands for input, so only the dominant hand is required. *Encumbrance Scenario 2A* (Figure 3.11) and *Encumbrance Scenario 2B* (Figure 3.12) simulate situations where a box is either carried under the non-dominant or the dominant arm respectively during two-hand input. Both scenarios test input performance when the user's arms are awkwardly constrained. All the encumbrance scenarios were performed either walking or standing still in the studies discussed in the following chapters.

Now that the encumbrance scenarios have been defined, the size and the weight of the bags and boxes require further discussion. Many different types and sizes of bags and boxes were observed in *User Study 1*. It is not feasible or possible to evaluate all the different kinds of bags and boxes, nor is it the intention of the research presented in this thesis to solely focus on evaluating different forms of each object type. Therefore, to keep the number of experimental conditions to a reasonable level in the studies discussed in the following chapters, two different sizes for each object type were chosen.

The selected bags were based on the size of a typical shopping bag in the UK[5]. The dimensions (w *x* d) of a typical medium plastic carrier bag is 279 *x* 533mm as shown in Figure 3.13. Therefore, the medium-sized bag used for *Encumbrance Scenarios 1A - 1D* was approximately the same dimensions as these guidelines and measured 450 *x* 550mm. The size of the smaller bag was 400 *x* 250mm, which is approximately half the height of the medium bag. The depth or thickness of the bags depended on the items placed inside them to add weight so the exact figures are not stated but was no more than 250mm. Another method to replicate the effects of carrying bags is to attach a weight to a handle (a pivot) to simulate the swinging motions of a pendulum. One limitation of this approach is the loss of the physical size of a bag which may have an effect on interaction, especially when the user is walking.

The dimension (l *x* w *x* d) of the boxes for *Encumbrance Scenario 2A* and *Encumbrance Scenario 2B* was based on the guidelines for sending a small parcel with the Royal Mail delivery service[6], as shown in Figure 3.14. Therefore, the standard and wider boxes measured 370 *x* 300 *x* 150mm and 390 *x* 300 *x* 290mm respectively.

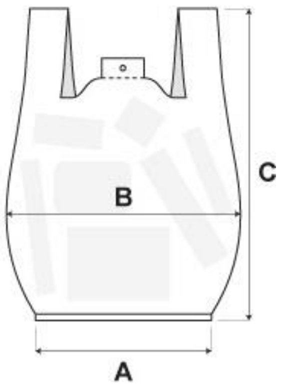

Figure 3.13: The dimensions of a typical plastic carrier bag in the UK[5]. The bottom width (A) is 279mm, the mid-width (B) is 432mm when the gusset expands and the height from the bottom of the bag to the top of the handle (C) is 533mm.

[5] Trade Supplies UK: http://www.tradersupplies.co.uk/medium-blue-recycled-vest-carrier-bags-100-per-pack.html
[6] Royal Mail. Size and weight formats for UK mail: http://www.royalmail.com.

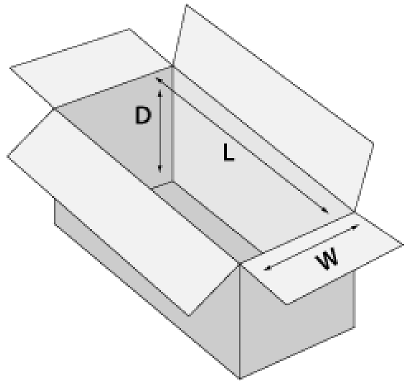

Figure 3.14: The dimensions of a Royal Mail small parcel box[6]: length (L) = 450mm, width (W) = 350mm and depth (D) = 160mm. Please note, the diagram is not drawn to scale and for illustration purposes only.

To choose an appropriate weight for the bags and boxes, a small pilot test was conducted after *User Study 1*. Eight students (two females) aged between 18 - 45 years (mean = 27.5, SD = 5.2) were recruited from the University. Each participant, who was standing still, was given a shopping bag to hold in the preferred hand which varied in weight from 1kg and incrementally increased by 0.5kg to 7kg. The participants were asked to think about the maximum weight that they would happily carry while using mobile devices.

The mean weight recorded was 3.4kg (SD = 0.8) so the bags and boxes used to replicate the encumbrance scenarios did not exceed this limit when carrying a single object. For *Encumbrance Scenario 1C* and *Encumbrance Scenario 1D* when a bag was held in each hand, the maximum weight of each object was halved to 1.7kg. The participants' well-being and physical condition during the user studies was a main concern, so the weight chosen for the encumbrance scenarios is a suitable compromise between simulating the effects of carrying realistic objects and keeping fatigue to a minimum.

3.6 Conclusions

To conclude, this chapter has presented an observational study to answer research questions *Q1.1* (*What are the typical objects held during interaction with mobile devices?*) and *Q1.2* (*How are these typical objects held during interaction with mobile devices?*). The observations were made at three different locations in the city centre of Glasgow and users in the public were surveyed. The results from the *User Study 1* suggested that different forms of bags and boxes were the two most frequently held object types. The bags were commonly held in-hand or along the forearm while boxes were grasped underarm or carried using both arms. Based on these findings, six different encumbrance scenarios were defined which included carrying bags in-hand, boxes underarm, situations where one or multiple objects are being held and when either one or both hands are encumbered.

A discussion was made in choosing the size of each object type and based on guidelines, two sizes of bags and boxes were selected. To select an appropriate weight for the bags and boxes to add realism of carrying objects but keep fatigue to a minimum during user studies, a small pilot test was conducted. The results suggested that the maximum weight of holding a single object should not exceed 3.4kg. The limit was halved to 1.7kg for each object for the encumbrance scenarios that evaluate the effects of carrying a bag in each hand.

The user studies presented in Chapter 4 - 8 evaluates at least one encumbrance scenario defined in Section 3.5 to examine the effects of carrying objects while walking on targeting performance on touchscreen mobile phones. The encumbrance scenarios can also be clearly replicated by researchers who want to investigate the physical effects of carrying objects in future work.

Chapter 4

User Study 2: Examining Tapping Performance While Walking and Encumbered

4.1 Introduction

Touchscreen mobile phones are often used while on the move and previous user studies (e.g. [7,8,43,45,53,66]) have reported a decline in input performance when users are walking. However, users also frequently carrying objects such as shopping bags while on the move as found in *User Study 1*. There is a lack of research in the literature that has examined targeting performance in these physically demanding encumbered and mobile situations. This chapter answers research question *Q2.1* (*How do encumbrance and mobility affect tapping performance*) by presenting a user study that was conducted to assess tapping performance on a touchscreen mobile phone while users were walking and encumbered. Users performed an abstract target selection task in several encumbrance scenarios that were defined in Section 3.5 to measure how tapping performance is affected under these conditions while on the move.

4.2 Method

4.2.1 Encumbrance Scenarios

To examine the effects of encumbrance, participants in *User Study 2* performed a target selection task on a touchscreen mobile phone while either carrying a bag in-hand or a box underarm. Two different sizes of each object type were held as discussed in Section 3.5. The bags were either held in the dominant or non-dominant hand (*Encumbrance Scenarios 1A* and *1B*). The boxes were either held under the dominant or non-dominant arm.

(*Encumbrance Scenarios 2A* and *2B*). Each object weighed 3kg to minimise fatigue and tiredness but simulate the realism of carrying objects. Figure 4.1 illustrates the way the bags and boxes were held during *User Study 2*.

Figure 4.1: The encumbrance scenarios evaluated in *User Study 2*. Holding a bag in-hand (left and left-middle) and carrying a box underarm (right-middle and right).

4.2.2 Tracking Body Movements

The participants' upper body and hand movements were tracked with motion capture hardware to measure the physical effects caused by encumbrance and mobility. Twelve Vicon[7] infrared cameras sampling at 120Hz were used to record body movements (to the thousandth of a millimetre) in three-dimensions. The cameras were located in a quiet room and a pre-defined 3.0 *x* 2.8m rectangular route was setup for the participants to navigate for each walking condition, as shown in Figure 4.2.

[7] Vicon Motion Capture: www.vicon.com.

Figure 4.2: The Vicon motion capture hardware setup is shown in the top image. The walking route created within the capturing volume is shown in the bottom image. The outer part of the route measured 3 x 2.8m.

For each participant, a total of 15 reflective markers were securely placed on the front and back of the upper torso, the hands and the arms as shown in Figure 4.3. The marker attached to the back of the neck was used to calculate the total distance walked and the average walking speed for each participant. The markers placed on the left shoulder, right shoulder, left index finger and left thumb (all participants were right-handed so the mobile phone was held in the left, non-dominant hand) were used to determine the relative position between the mobile phone and the user, which meant the amount of hand movement along each axis could be calculated.

The x-, y- and z- axis of motion represented left – right, forward – backward and upwards – downwards movements respectively. The marker attached to the intermediate phalanx section of the right index finger was used to track the motion of input. It would have been more appropriate to place the marker on the tip of the index finger but due to the size of the markers, it would have obscured part of the touchscreen and therefore made targeting more difficult. The remaining markers were used to define sections of the upper body and to track additional arm movements. The participants avoided wearing loose clothing to prevent excessive marker movements.

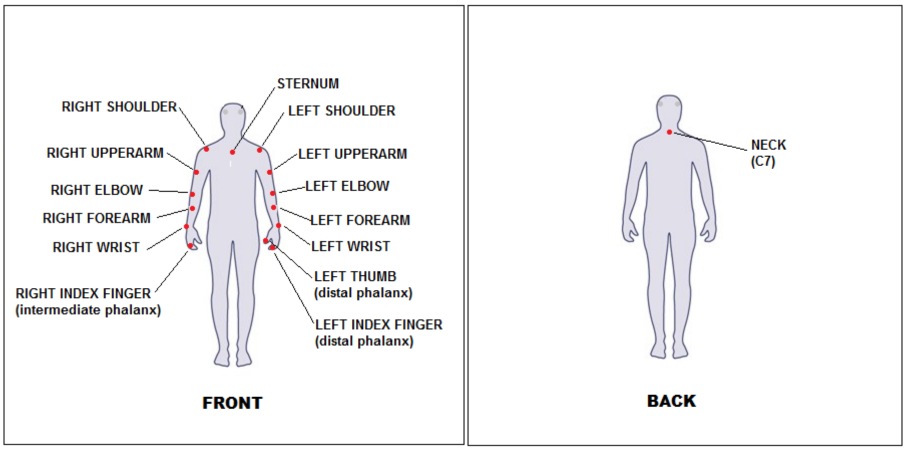

Figure 4.3: The location of the 15 reflective markers (red dots) attached on each participant.

Figure 4.4: The markers placed on the right index finger used for input and on the index finger and thumb of the left hand holding the mobile phone.

4.2.3 Task

To measure tapping performance while walking and encumbered, participants completed an abstract targeting task on a touchscreen mobile phone. For each trial, two targets were displayed on-screen as shown in Figure 4.5. The green target was the current target to select while the red target illustrated where to tap next until the last target selection was reached. This design was selected so that participants always knew where to tap next and the positions of the two targets for each trial were randomly chosen. The targets measured 40 *x* 60px (4 *x* 6mm) with a crosshair placed at the centre of the target. The size of each target was approximately the same as a key on a standard Android keyboard on the mobile phone used. There were 100 target positions aligned in a 10 *x* 10 grid and each was selected once per condition.

A gap was created between the last row of the targets and bottom of the touchscreen to prevent the participants from accidentally tapping the soft keys on the phone. The participants were instructed to select each target as quickly and as accurately as possible. The task ran on a Google Nexus One (Android 3.1) mobile phone which had a 3.7″ touchscreen and a resolution of 480 *x* 800px (9.94px/mm). All participants performed the task using the two-handed index finger input posture therefore the mobile phone was held in portrait mode in the non-dominant hand while the preferred index finger of the dominant hand was used for input.

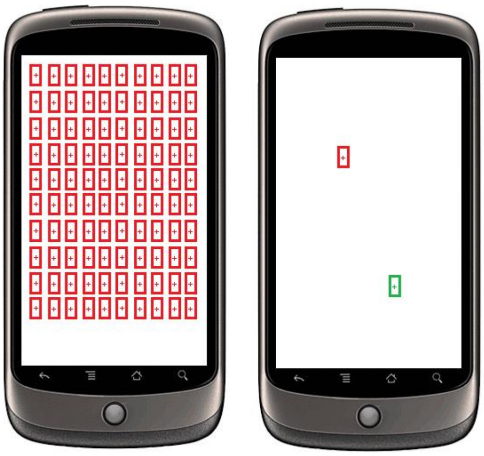

Figure 4.5: The target selection task illustrating the target positions (left) and the interface presented to the user (right).

4.2.4 Experimental Design

Eighteen participants (4 males, 14 females) aged between 19 - 38 years (mean = 25.3, SD = 3.1) and all preferred using their right hand to input with mobile devices were recruited from the University to take part. Each participant was paid £6 and the study lasted approximately 60 minutes. Sufficient resting periods were given between conditions and whenever necessary to reduce fatigue.

There were two Independent Variables:

Type of Mobility (2 levels) -

- Standing still
- Walking around the pre-defined route

Type of Encumbrance (9 levels) -

- Unencumbered
- Holding the small bag in either the non-dominant or dominant hand
- Holding the medium bag in either the non-dominant or dominant hand
- Holding the standard box under either the non-dominant or dominant arm
- Holding the wider box under either the non-dominant or dominant arm

As a result, there were 18 conditions and a within-subject design was used. The conditions were counter-balanced by type of mobility and further pseudo-randomised by type of encumbrance to reduce learning and order effects. There were three Dependent Variables: *Target accuracy (%)* - if the on-screen touch up position was within the target border, *Target error (millimetres)* - the absolute distance from the touch up position to the centre of target and *Selection time (milliseconds)* - the time taken to select each target. The hypotheses were:

H1A: Target accuracy will be significantly lower when walking than standing;

H1B: Target error will be significantly higher when walking than standing;

H1C: Selection time will be significantly longer when walking than standing;

H2A: Target accuracy will be significantly lower when encumbered than holding no objects;

H2B: Target error will be significantly higher when encumbered than holding no objects;

H2C: Selection time will be significantly longer when encumbered than holding no objects;

H3A: Carrying the bag/box in the dominant hand/arm will cause target accuracy to significantly decrease when compared to holding the objects on the non-dominant side;

H3B: Carrying the bag/box in the dominant hand/arm will cause target error to significantly increase when compared to holding the objects on the non-dominant side;

H3C: Carrying the bag/box in the dominant hand/arm will cause slower selection time than holding the objects on the non-dominant side.

4.3 Results

A total of 32,400 target selections (*100 targets x 18 conditions x 18 participants*) was recorded for *User Study 2*. To filter out likely unintentional taps, target selections that took less than 100ms were discarded. As a result, 21 targets were not included in the final data analysis.

4.3.1 Target Accuracy

Shapiro-Wilk tests were conducted to examine the distribution of target accuracy. The Shapiro-Wilk normality test is recommended when the sample size (N) is small [18] and in this case, N = 20 for each condition. The results are shown in Table 4.1 and one condition, holding the standard box under the dominant arm while standing, was significant[8] therefore the data deviates from a normal distribution. The uneven distribution of this condition is also illustrated in Figure 4.6 which shows the box-and-whisker plots for all conditions. Furthermore, the box-and-whisker plots suggests that the data for other conditions might not be normally distributed (e.g. holding the medium bag in the dominant hand while standing), which resulted in non-significant results from conducting a Shapiro-Wilks test.

[8] When the word “significant” (along with “significance” and “significantly”) is use in this thesis, it refers to *statistical* significance. Results are significant when the *p*-value is less than 0.05 for all statistical tests reported in this thesis.

Because not all data for target accuracy conforms to a normal distribution, the Aligned Rank Transform (ART) method described by Wobbrock *et al.* [76] was used prior to conducting factorial data analysis. ART aligns the data (target accuracy in this case) and then averaged ranks are applied. Once this step is completed, factorial data analysis can be performed (repeated-measure ANOVA in this case). The mean target accuracy for each condition is shown in Figure 4.7.

Type of Mobility	Type of Encumbrance	W Statistic	Sig.
Standing	No object	0.952222	0.460937
Standing	Small bag (ND)	0.945155	0.354127
Standing	Small bag (D)	0.917261	0.115612
Standing	Medium bag (ND)	0.940595	0.296653
Standing	Medium bag (D)	0.950105	0.426672
Standing	Standard box (ND)	0.9423	0.317114
Standing	Standard box (D)	0.816456	0.002619*
Standing	Wider box (ND)	0.95203	0.457752
Standing	Wider box (D)	0.922171	0.141238
Walking	No object	0.956074	0.527977
Walking	Small bag (ND)	0.975987	0.899541
Walking	Small bag (D)	0.925394	0.161058
Walking	Medium bag (ND)	0.945536	0.35933
Walking	Medium bag (D)	0.919695	0.127675
Walking	Standard box (ND)	0.949644	0.419465
Walking	Standard box (D)	0.959126	0.584851
Walking	Wider box (ND)	0.958707	0.576878
Walking	Wider box (D)	0.953015	0.474259

Table 4.1: Shapiro-Wilk normality tests performed on target accuracy for each condition in *User Study 2*. Significant results (*p*-value < 0.05 i.e. the data deviates from a normal distribution) are shaded in grey and highlighted and with '*'. ND = non-dominant and D = dominant.

The results from conducting a two-factor (**Type of Mobility** and **Type of Encumbrance**) repeated-measures ANOVA for target accuracy showed a significant main effect for **Type of Mobility**, $F(1, 289) = 275.1364$, $p < 0.001$. Target accuracy was significantly lower when the participants were walking than standing. The overall mean accuracies for standing and walking were 54.0% and 36.7% respectively, as shown in Figure 4.8.

A significant main effect was observed for **Type of Encumbrance**, $F(8, 289) = 14.4396$ $p < 0.001$. *Post hoc* pairwise comparisons that are relevant to the experiment hypotheses are shown in Table 4.2. Target accuracy was significantly higher when holding no objects than when encumbered with the bags and boxes. There was no significant difference between the dominant and non-dominant hands when holding either the small bag or the medium bag. Target accuracy was significantly lower when carrying the standard box under the dominant arm than the non-dominant arm. However, holding the wider box under the dominant arm did not caused target accuracy to significantly decrease when compared to holding the same object in the non-dominant arm. The overall mean target accuracy for each type of encumbrance are shown in Figure 4.9.

The interaction between the two factors was significant, $F(8, 289) = 2.7148$, $p < 0.01$. Despite significant results for the interaction, it is not required to support or reject hypotheses *H1A*, *H2A* and *H3A* therefore *post hoc* comparisons are not conducted.

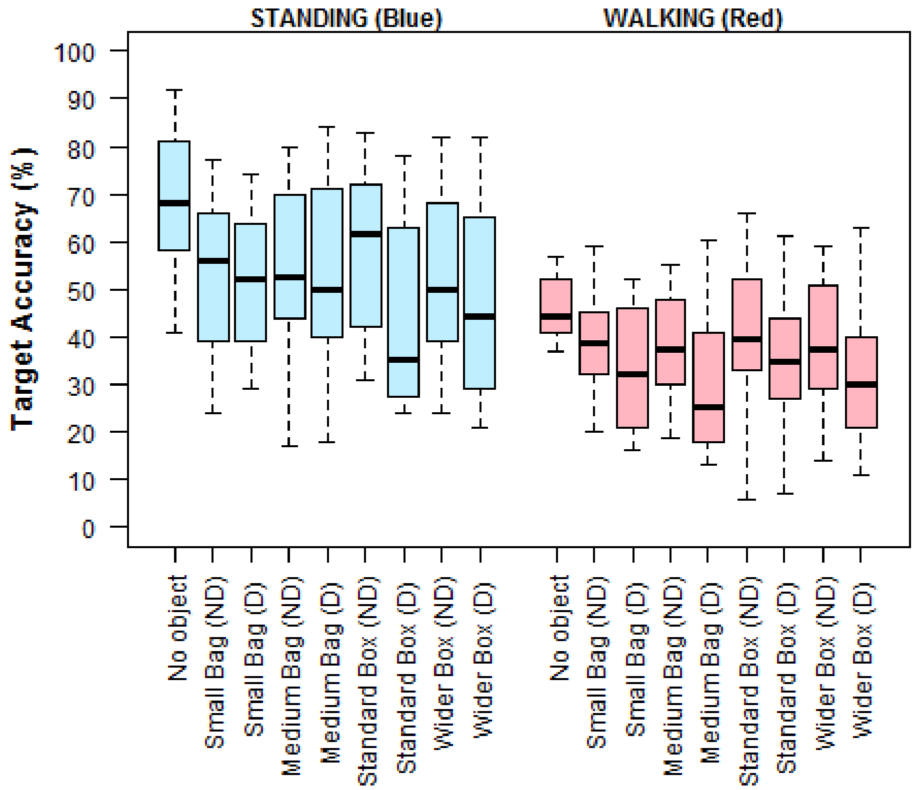

Figure 4.6: Box-and-whisker plots for target accuracy (%) for each condition in *User Study 2*. ND = non-dominant and D = dominant.

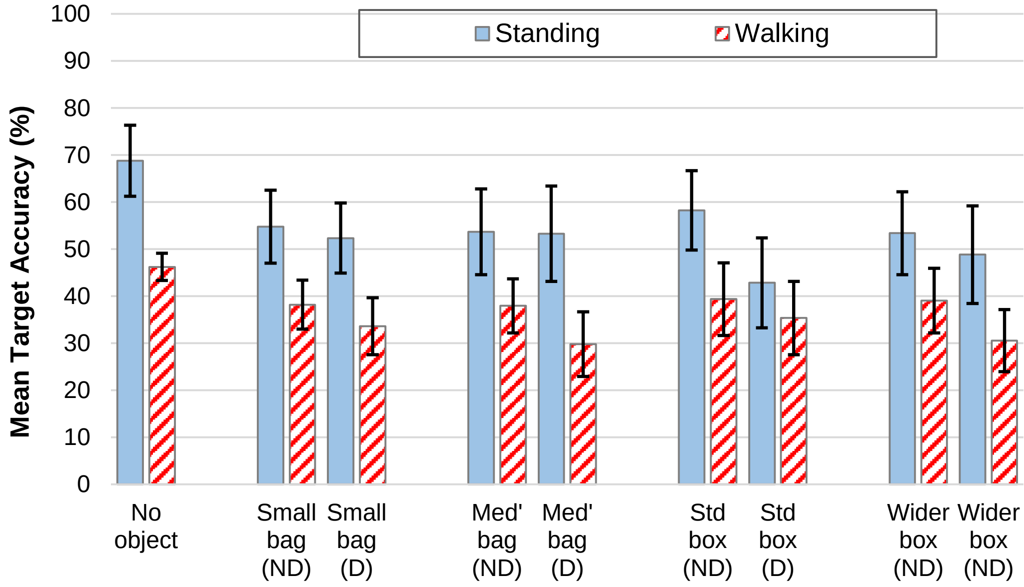

Figure 4.7: The mean target accuracy (%) for each condition in *User Study 2*. The standing and walking conditions are represented by the solid blue and striped red bars respectively. ND = non-dominant and D = dominant. Error bars denote Confidence Interval (95%).

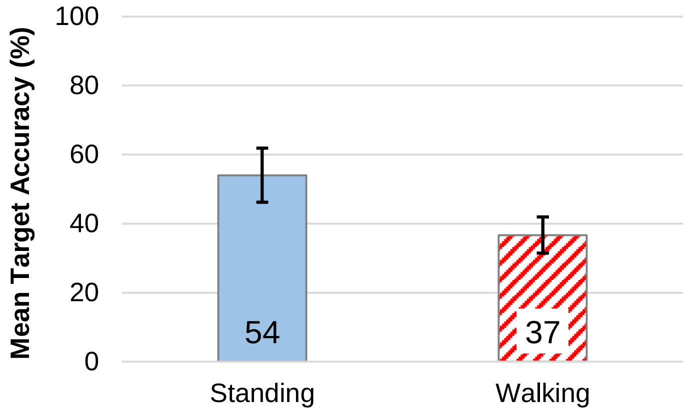

Figure 4.8: The overall mean target accuracy for standing (solid blue) and walking (striped red). Error bars denote Confidence Interval (95%).

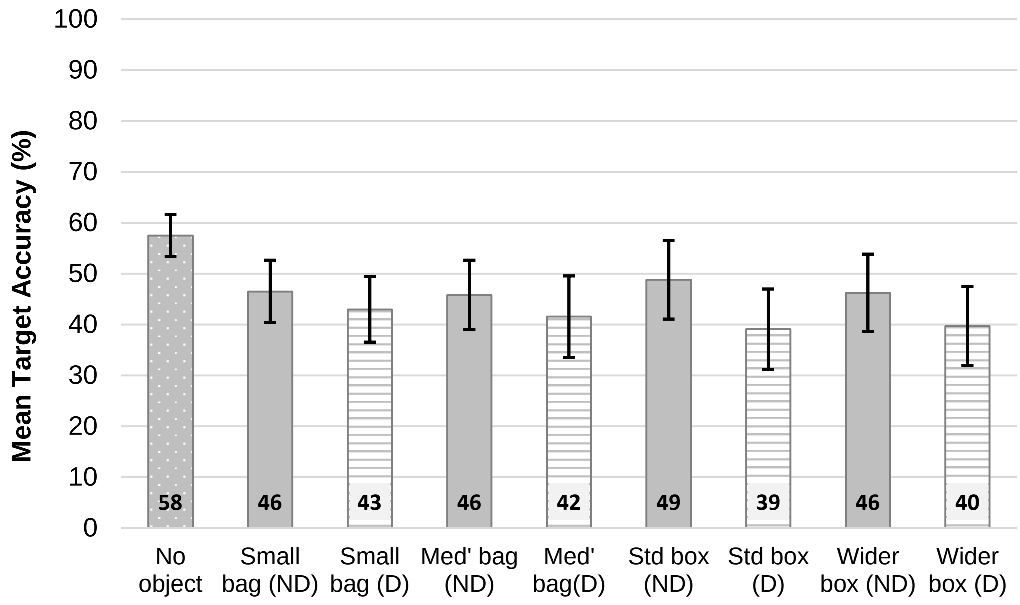

Figure 4.9: The overall mean target accuracy for each type of encumbrance. ND = non-dominant (solid) and D = dominant (horizontal stripes). Error bars denote Confidence Interval (95%).

Comparison		t ratio	*p*-value
No object	Small bag (ND)	7.028	<.0001*
No object	Small bag (D)	5.168	<.0001*
No object	Medium bag (ND)	5.304	<.0001*
No object	Medium bag (D)	7.589	<.0001*
No object	Standard box (ND)	4.008	0.0025*
No object	Standard box (D)	8.746	<.0001*
No object	Wider box (ND)	5.283	<.0001*
No object	Wider box (D)	8.381	<.0001*
Small bag (ND)	Small bag (D)	1.860	0.6418
Medium bag(ND)	Medium bag (D)	2.285	0.3548
Standard box (ND)	Standard box (D)	4.739	0.0001*
Wider box (ND)	Wider box (D)	-3.097	0.0543

Table 4.2: The relevant Type of Encumbrance pairwise comparisons for target accuracy. *P*-value adjusted using Tukey method. Significant results are shaded in grey and highlighted with '*'. ND = non-dominant and D = dominant.

4.3.2 Target Error

Shapiro-Wilk tests were conducted to assess the normality of the data recorded for target error and the results are shown in Table 4.3. Three conditions were significant: (1) carrying the small bag in the dominant hand, (2) holding the medium bag in the non-dominant hand while standing and (3) carrying the standard box under the non-dominant arm while walking. The median and distribution of target error for all conditions are shown from the box-plots in Figure 4.10. Because target error violated normality, the data was transformed using ART [76] prior to conducting an ANOVA. A two-factor (**Type of Mobility** and **Type of Encumbrance**) repeated-measures ANOVA was carried out for target error. The mean target error for each condition is shown in Figure 4.11.

Type of Mobility	Type of Encumbrance	W Statistic	Sig.
Standing	No object	0.916705	0.113021
Standing	Small bag (ND)	0.903554	0.066285
Standing	Small bag (D)	0.896298	0.049548*
Standing	Medium bag (ND)	0.78994	0.001099*
Standing	Medium bag (D)	0.964563	0.69147
Standing	Standard box (ND)	0.923458	0.148847
Standing	Standard box (D)	0.931046	0.202585
Standing	Wider box (ND)	0.961371	0.628307
Standing	Wider box (D)	0.968517	0.769389
Walking	No object	0.940961	0.300941
Walking	Small bag (ND)	0.941482	0.307154
Walking	Small bag (D)	0.946321	0.370246
Walking	Medium bag (ND)	0.940011	0.289915
Walking	Medium bag (D)	0.969748	0.792934
Walking	Standard box (ND)	0.774796	0.000684*
Walking	Standard box (D)	0.968556	0.77014
Walking	Wider box (ND)	0.933307	0.221934
Walking	Wider box (D)	0.928837	0.185247

Table 4.3: Shapiro-Wilk normality tests performed on target error for each condition in *User Study 2*. Significant results (i.e. the data deviates from a normal distribution) are shaded in grey and highlighted and with '*'. ND = non-dominant and D = dominant.

The ANOVA results showed a significant main effect for **Type of Mobility**, $F(1, 289) = 221.5676$, $p < 0.001$. Walking caused tapping error to significantly increase when compared to standing. The overall mean target errors for standing and walking were 3.3mm and 4.3mm respectively, as illustrated in Figure 4.12.

A significant main effect was found for **Type of Encumbrance**, $F(8, 289) = 17.2441$, $p < 0.001$. The appropriate *post hoc* pairwise comparisons to the experiment hypotheses are shown in Table 4.4. The results showed that targeting error was significantly lower when unencumbered than holding the objects, with the exception of carrying the standard box under the non-dominant arm. There was no significant difference in targeting error between the dominant and non-dominant hand when the small bag was held. In contrast, targeting error was significantly lower when the medium bag was held in the non-dominant hand than the same object being held in the dominant hand. Both the standard and wider boxes, when held under the dominant arm, caused a significant increase in targeting error than carrying either object under the non-dominant arm. The overall mean target error for each encumbrance type is illustrated in Figure 4.13. The interaction between the two factors was not significant, $F(8, 289) = 1.8799$, $p > 0.05$. The interaction effect is not required for the experiment hypotheses.

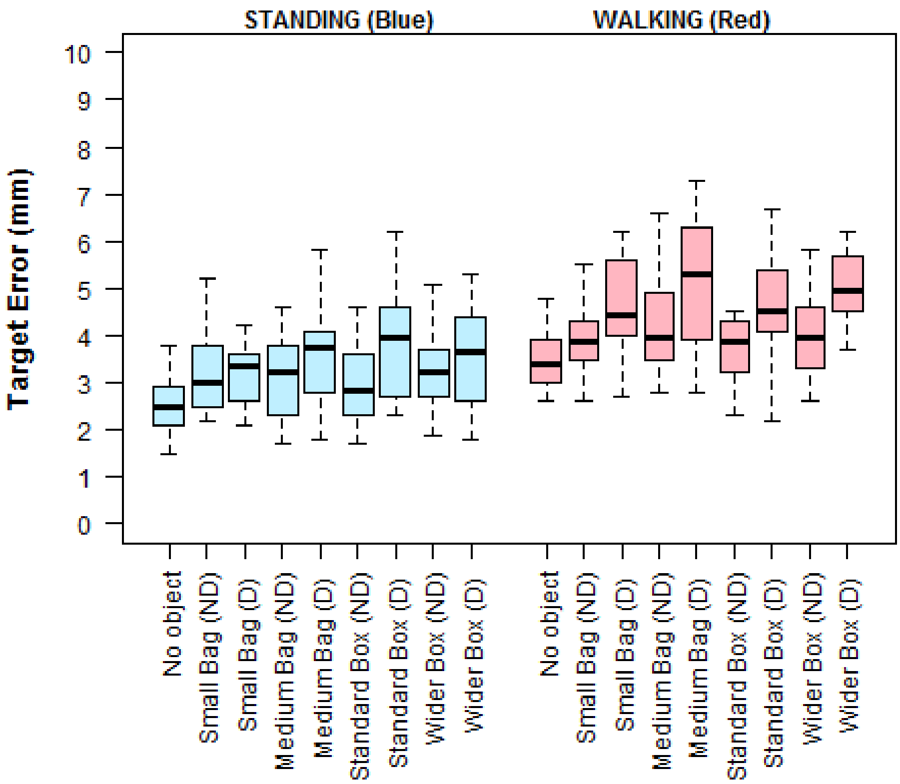

Figure 4.10: Box-and-whisker plots for target error (mm) for each condition in *User Study 2*. ND = non-dominant and D = dominant.

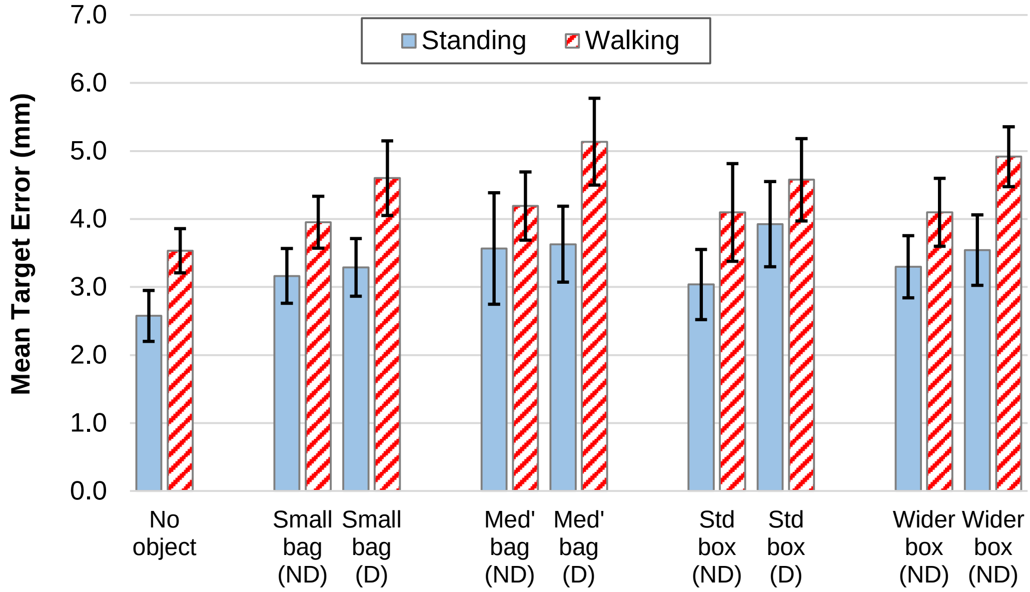

Figure 4.11: The mean target error (mm) for each condition in *User Study 2*. The standing and walking conditions are represented by the solid blue and striped red bars respectively. ND = non-dominant and D = dominant. Error bars denote Confidence Interval (95%).

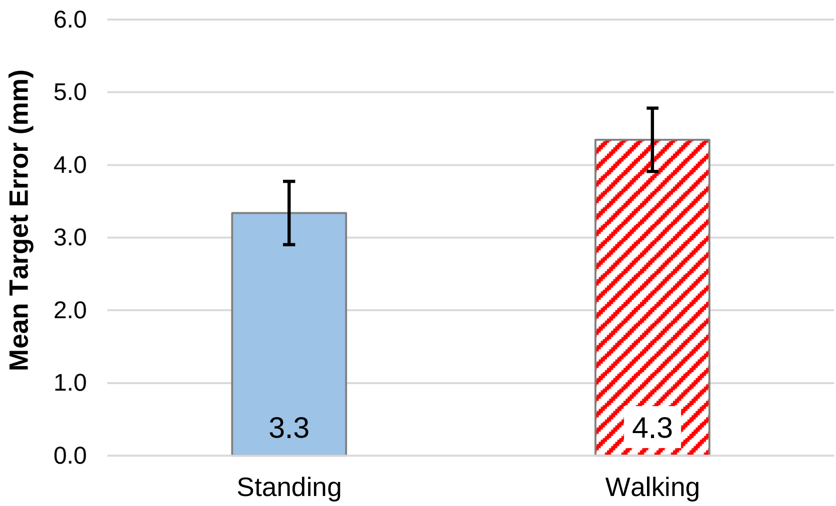

Figure 4.12: The overall mean target error for standing (solid blue) and walking (striped red). Error bars denote Confidence Interval (95%).

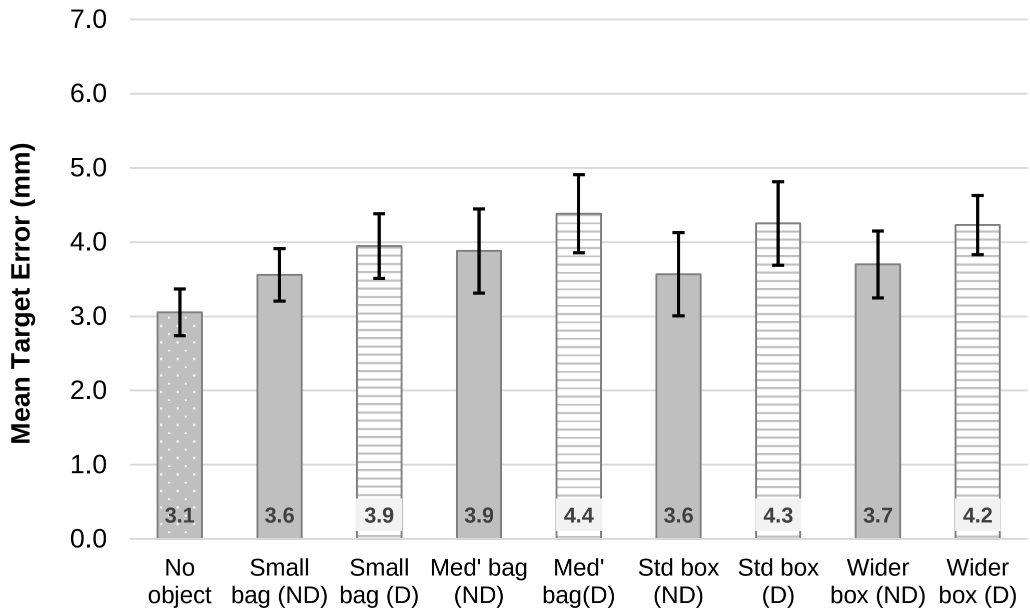

Figure 4.13: The overall mean target error for each Type of Encumbrance. ND = non-dominant (solid) and D = dominant (vertical stripes). Error bars denote Confidence Interval (95%).

Comparison		**t ratio**	***p*-value**
No object	Small bag (ND)	-3.807	0.0053*
No object	Small bag (D)	-6.490	<.0001*
No object	Medium bag (ND)	-5.095	<.0001*
No object	Medium bag (D)	-8.867	<.0001*
No object	Standard box (ND)	-3.014	0.0686
No object	Standard box (D)	-8.051	<.0001*
No object	Wider box (ND)	-4.759	0.0001*
No object	Wider box (D)	-8.694	<.0001*
Small bag (ND)	Small bag (D)	-2.682	0.1587
Medium bag(ND)	Medium bag (D)	-3.772	0.0060*
Standard box (ND)	Standard box (D)	-5.037	<.0001*
Wider box (ND)	Wider box (D)	-3.934	0.0033*

Table 4.4: The relevant Type of Encumbrance pairwise comparisons for target error. *P*- value adjusted using Tukey method. Significant results are shaded in grey and highlighted with '*'. ND = non-dominant and D = dominant.

4.3.3 Selection Time

Shapiro-Wilk tests were performed to examine the distribution of the data recorded for selection time and the results are shown in Table 4.5. One condition was significant: carrying the standard box under the non-dominant arm while standing. The median and distribution of each condition for selection time are shown in Figure 4.14. Because selection time was not normally distributed, ART [76] was performed on the data before conducting a two-factor (**Type of Mobility** and **Type of Encumbrance**) repeated-measures ANOVA. The mean selection time for each condition is shown in Figure 4.15.

Type of Mobility	Type of Encumbrance	W Statistic	Sig.
Standing	No object	0.985441	0.988713
Standing	Small bag (ND)	0.96025	0.606462
Standing	Small bag (D)	0.926973	0.171744
Standing	Medium bag (ND)	0.905085	0.070507
Standing	Medium bag (D)	0.899494	0.056304
Standing	Standard box (ND)	0.874218	0.020895*
Standing	Standard box (D)	0.963487	0.670081
Standing	Wider box (ND)	0.948265	0.398448
Standing	Wider box (D)	0.940041	0.290258
Walking	No object	0.920795	0.133533
Walking	Small bag (ND)	0.975217	0.887968
Walking	Small bag (D)	0.921282	0.13621
Walking	Medium bag (ND)	0.960466	0.610662
Walking	Medium bag (D)	0.903196	0.065335
Walking	Standard box (ND)	0.928633	0.183722
Walking	Standard box (D)	0.947732	0.390562
Walking	Wider box (ND)	0.977963	0.926583
Walking	Wider box (D)	0.96865	0.77195

Table 4.5: Shapiro-Wilk normality tests performed on selection time for each condition in *User Study 2*. Significant results (i.e. the data deviates from a normal distribution) are shaded in grey and highlighted and with '*'. ND = non-dominant and D = dominant.

The results from conducting an ANOVA showed a significant main effect for **Type of Mobility**, $F(1, 289) = 78.5897$, $p < 0.001$. Target selections took significantly longer when walking than standing, a difference of 12.13%. The overall mean selection times for both standing and walking were 449.8ms and 504.4ms respectively, as shown in Figure 4.16.

A significant main effect was also observed for **Type of Encumbrance**, $F(8, 289) = 12.5610$, $p < 0.001$. The relevant *post hoc* pairwise comparisons to the experiment hypotheses are shown in Table 4.6. The results showed that target selections took significantly longer when unencumbered than holding the medium bag in the non-dominant hand and both types of boxes under the non-dominant arm. Holding the small and medium bags in the dominant hand caused slower selection times than carrying the objects in the non-dominant hand. Likewise, carrying both the standard and wider boxes under the dominant arm caused significantly longer selection time when compared to carrying the boxes under the non-dominant arm. The overall mean selection time for each type of encumbrance is illustrated in Figure 4.17. No significant difference was observed for the interaction between the two factors, $F(8, 289) = 1.5377$, $p > 0.05$. The interaction effect between the factors is not required to support or reject the experiment hypotheses on selection time.

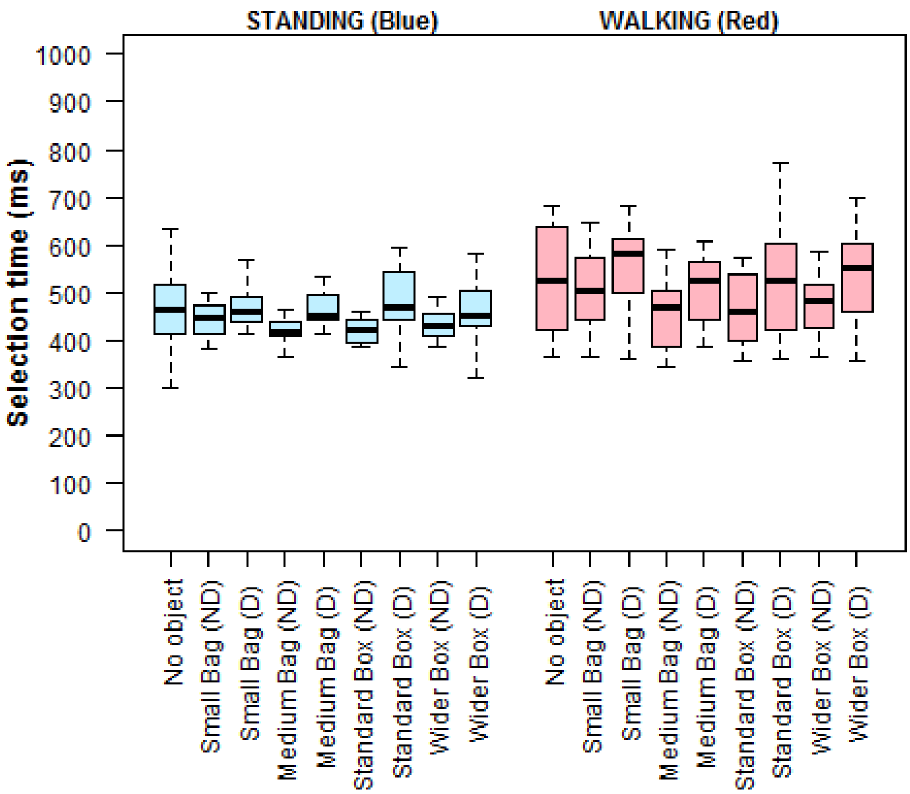

Figure 4.14: Box-and-whisker plots for selection time (ms) for each condition in *User Study 2*. ND = non-dominant and D = dominant.

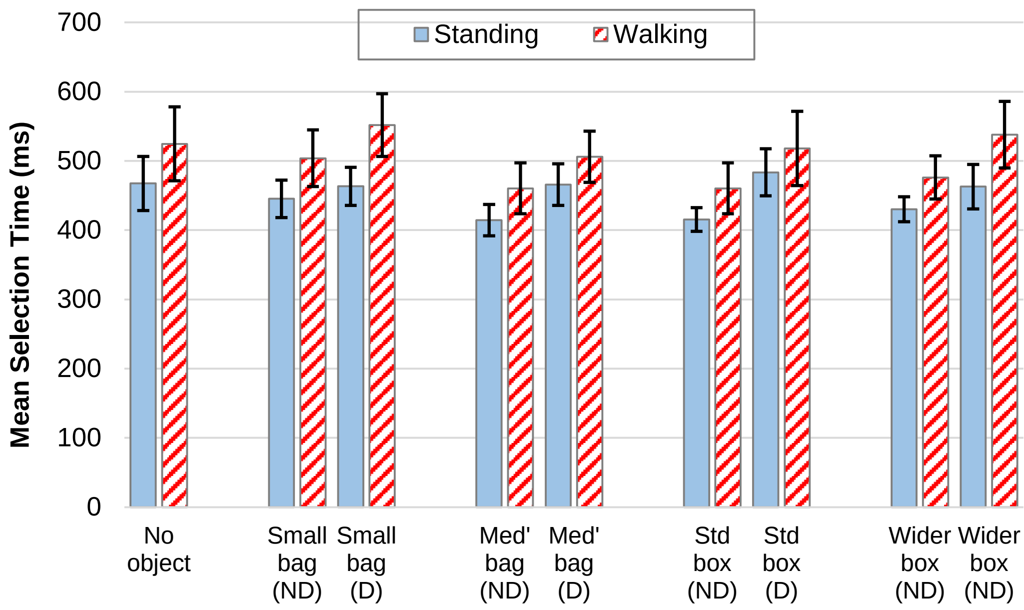

Figure 4.15: The mean selection time (ms) for each condition in *User Study 2*. The standing and walking conditions are represented by the solid blue and striped red bars respectively. ND = non-dominant and D = dominant. Error bars denote Confidence Interval (95%).

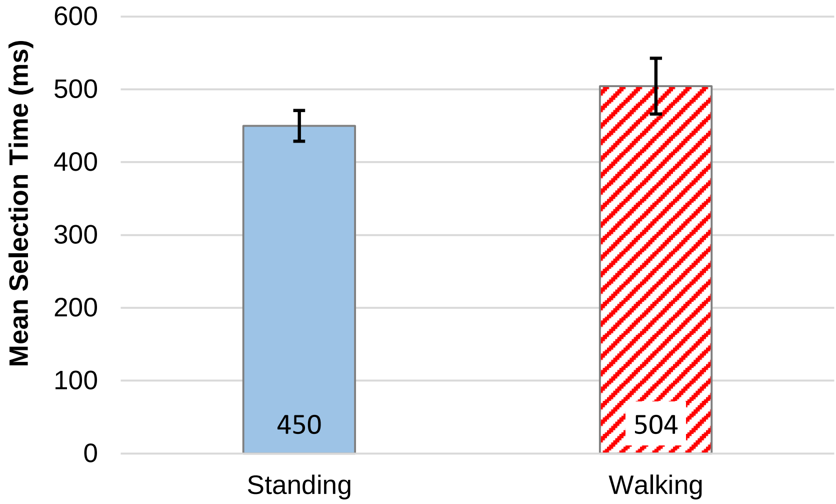

Figure 4.16: The overall mean target selection time for standing (solid blue) and walking (striped red). Error bars denote Confidence Interval (95%).

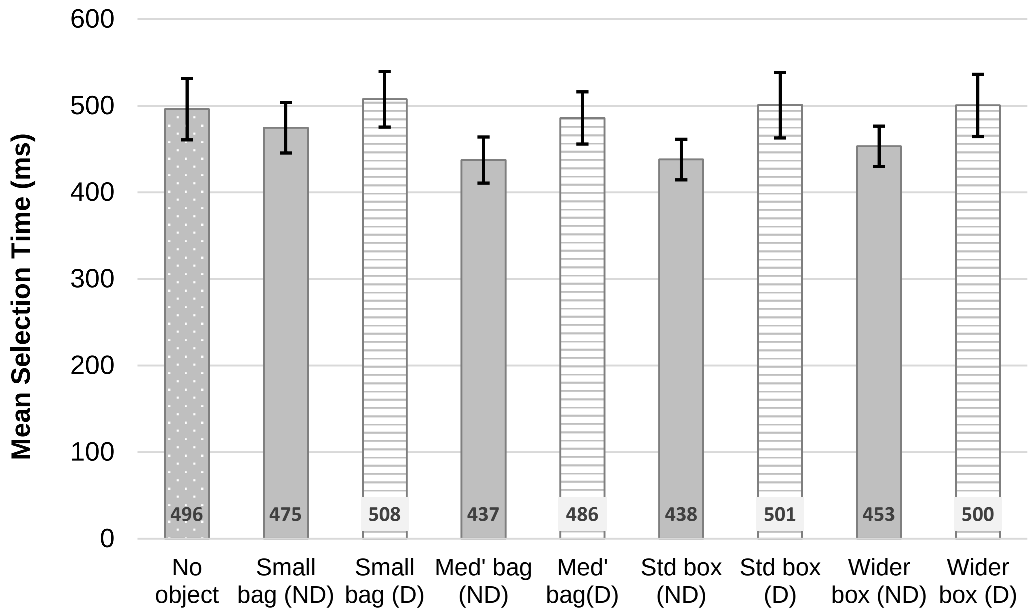

Figure 4.17: The overall mean selection time (ms) grouped by Type of Encumbrance. ND = non-dominant (solid) and D = dominant (vertical striped). Error bars denote Confidence Interval (95%).

Comparison		t ratio	*p*-value
No object	Small bag (ND)	1.417	0.8908
No object	Small bag (D)	-1.769	0.7024
No object	Medium bag (ND)	4.883	0.0001*
No object	Medium bag (D)	0.276	1.0000
No object	Standard box (ND)	4.986	<.0001*
No object	Standard box (D)	-0.210	1.0000
No object	Wider box (ND)	3.630	0.0100*
No object	Wider box (D)	0.9987	-0.704
Small bag (ND)	Small bag (D)	-3.186	0.0419*
Medium bag(ND)	Medium bag (D)	-4.606	0.0002*
Standard box (ND)	Standard box (D)	-5.196	<.0001*
Wider box (ND)	Wider box (D)	-4.334	0.0007*

Table 4.6: The relevant Type of Encumbrance pairwise comparisons for selection time. *P*-value adjusted using Tukey method. Significant results are shaded in grey and highlighted with '*'. ND = non-dominant and D = dominant.

4.3.4 Analysing Body Movements While Encumbered

The data recorded from tracking the participants' movements were processed firstly, to examine the activity of the input finger and secondly, to analyse the change in movement of the non-dominant hand which held the mobile phone. After examining the motion of the index finger, it was difficult to make clear conclusions as to whether excessive movement was due to carrying the objects or users were adjusting their hand and input position to tap on the targets across the touchscreen. In addition, since the marker was not placed on the tip of the index finger to reduce visual difficulties of selecting the targets, it was difficult to be certain of the precise motion of the index finger when input was made. Therefore, the results to show how walking and carrying the objects physically affected the input hand and caused a change in targeting performance are not stated or discussed further.

The results from analysing the data logged from the non-dominant hand holding the mobile phone revealed clearer findings about the physical effects of encumbrance and mobility. To measure the level of hand activity (mm) during input, the mean change in movement was computed. This was calculated by averaging the absolute change in motion along each axis from 300ms prior to a target being selected to the instance a touch down event was recorded on the mobile phone. The encumbrance types were grouped by handedness because the main focus was to examine the level of movement when the dominant or non-dominant hand or arm was encumbered rather than the type of object being held.

After sorting the data in this way, there were three levels for **Type of Encumbrance**: unencumbered and either the non-dominant or dominant side was encumbered, and two levels for **Type of Mobility**: standing and walking. As a result, there were six conditions and separate analysis was conducted to compare the level of movement caused by mobility and encumbrance on each axis. The mean change in movement (mm) of the non-dominant hand along each axis is shown in Figure 4.18. Shapiro-Wilk normality tests were performed on each condition and on each hand movement axis. The results are shown in Table 4.7.

For the x-axis, the data did not violate normality, therefore a two-factor, repeated-measures ANOVA was conducted. Mauchly's Test of Sphericity was used to see if the assumption of sphericity has been violated. If the results were significant, corrections to the degrees of freedom were made. Greenhouse-Geisser correction was used when the estimated epsilon (ε) is less than 0.75 while Huynd-Feldt correction was used when ε is greater than or equal to 0.75 [18][9]. The results showed a significant mean effect for **Type of Mobility**, $F(1,17) = 936.724$, $p < 0.01$. The mean changes in movement for standing and walking were 2.5 and 5.7mm respectively. Walking caused significantly more movements to the non-dominant hand than standing. A significant main effect was also observed for **Type of Encumbrance**, $F(1.1,18.1) = 409.230$, $p < 0.01$. *Post hoc* pairwise comparisons with Bonferroni corrections showed that unencumbered caused a significantly smaller change in movement than holding the objects in both the non-dominant and dominant sides. However, the difference between holding no objects and when the non-dominant side was encumbered was small (mean difference = 0.2mm). Carrying the bags and boxes in the dominant hand and arm respectively caused a significant increase of 67.6% in movement when compared to holding the objects in the non-dominant side. A significant effect for the interaction between the factors was found, $F(1.0,17.5) = 65.225$, $p < 0.01$.

For the y-axis, the data violated normality, as shown in Table 4.7. Therefore, ART [76] was performed on the data before conducting a two-factor, repeated-measures ANOVA. The results from the ANOVA showed a significant main effect for **Type of Mobility**, $F(1, 85) = 376.8181$, $p < 0.001$. The mean changes of movement for standing and walking were 1.3 and 4.2mm respectively. Walking caused significantly more change in movement to the non-dominant holding the mobile phone than standing still. A significant main effect was also found for **Type of Encumbrance**, $F(2, 85) = 76.179$, $p < 0.001$. *Post hoc* Tukey HSD tests showed that holding no objects caused significantly less change in movement on the y-axis than encumbering the dominant side ($t = 11.519$, $p < 0.001$). There was no significant difference between unencumbered and carrying objects on the non-dominant side ($t = 1.918$, $p < 0.1399$). The difference in movement between dominant and non-dominant sides was significant ($t = 9.601$, $p < 0.001$). Holding the objects in the dominant hand or arm caused a significant increase of 54.2% in movement along the y-axis than carrying the objects in the non-dominant side. The interaction between the factors was also significant, $F(2, 85) = 6.4281$, $p < 0.01$.

For the z-axis, the data was normally distributed, as shown in Table 4.7. Therefore, a two-factor repeated-measures ANOVA was conducting on the change in movement along the z-axis. The results showed a significant main effect for **Type of Mobility,** $F(1, 17) = 940.907$, $p < 0.01$. Walking caused significantly more movements in the z-axis in the non-dominant hand holding the mobile phone than standing still. The mean changes in

[9] This method is used each time Mauchly's Test of Sphericity is significant for the ANOVAs conducted in this thesis.

movement for standing and walking were 1.8 and 5.9mm respectively. A significant main effect was also observed for **Type of Encumbrance**, $F(1.2, 19.6) = 225.470$, $p < 0.01$. *Post hoc* pairwise tests with Bonferroni corrections showed that carrying the objects on both the non-dominant and dominant sides caused significantly more movement in the z-axis than unencumbered input. The difference between unencumbered and carrying the objects in the non-dominant side was small (mean difference = 0.1mm). Holding the bags and boxes in the dominant hand and arm respectively caused significantly more movements than carrying the objects in the non-dominant side. The interaction between the factors was also significant, $F(1, 17.6) = 60.701$, $p < 0.01$.

Type of Mobility	Type of Encumbrance	Axis	W Statistic	Sig.
Standing	No object	X	0.953182	0.477089
Standing	(ND)	X	0.950363	0.430744
Standing	(D)	X	0.890839	0.039892
Walking	(ND)	X	0.975139	0.886761
Walking	(D)	X	0.954152	0.493794
Walking	(ND)	X	0.944667	0.347557
.	.	.	.	.
Standing	No object	Y	0.930841	0.20091
Standing	(ND)	Y	0.920833	0.133737
Standing	(D)	Y	0.974426	0.875522
Walking	(ND)	Y	0.805482	0.001817*
Walking	(D)	Y	0.894739	0.046566*
Walking	(ND)	Y	0.982523	0.972247
.	.	.	.	.
Standing	No object	Z	0.978257	0.930259
Standing	(ND)	Z	0.95501	0.508873
Standing	(D)	Z	0.950345	0.430458
Walking	(ND)	Z	0.922249	0.141687
Walking	(D)	Z	0.94395	0.338079
Walking	(ND)	Z	0.937727	0.264858

Table 4.7: Shapiro-Wilk normality tests performed on hand activity on each axis for Type of Mobility and Type of Encumbrance. Significant results are shaded in grey and highlighted and with '*'. ND = non-dominant and D = dominant.

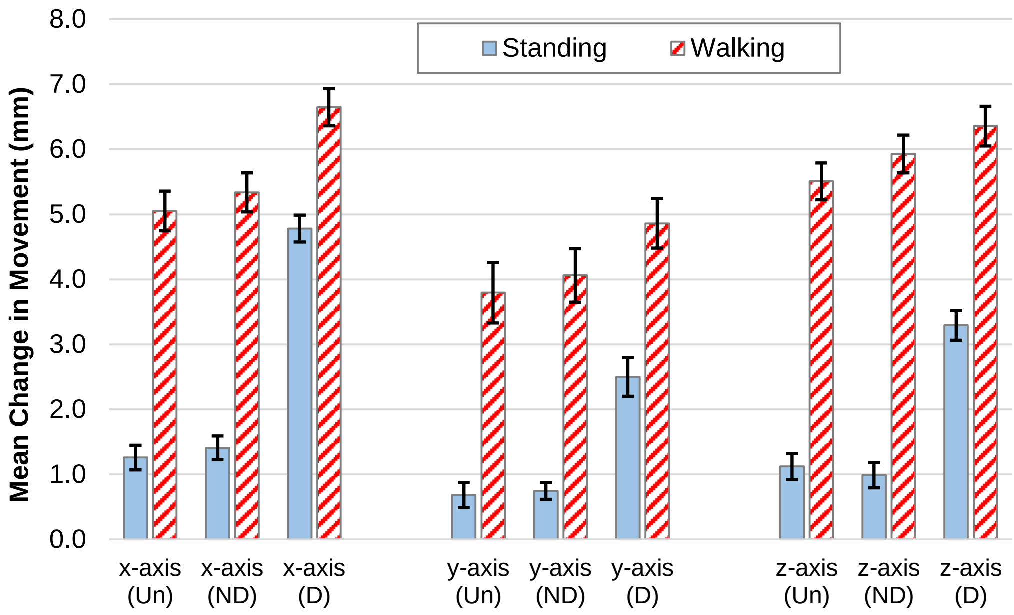

Figure 4.18: The mean change in movement along each axis. The standing and walking conditions are represented by the solid blue and striped red bars respectively. Un = unencumbered, ND = non-dominant and D = dominant. Error bars denote Confidence Interval (95%).

4.3.5 The Effects of Encumbrance and Input on PWS

The mean Preferred Walking Speed (PWS) for each walking condition is shown in Figure 4.19. The baseline PWS represented the normal walking speed (i.e. no interaction or holding objects) and was measured before the study began. Each participant walked around the pre-defined route for five laps and the total time required was recorded and the average walking speed calculated.

Shapiro-Wilk tests were conducted to assess the normality of the PWS data for each walking condition. The results are shown in Table 4.8 and no condition was significant. Therefore, a one-factor, repeated-measures ANOVA was conducted to compare the different levels of PWS. A significant main effect was found, $F(2.8, 47.4) =$, $p < 0.01$. *Post hoc* pairwise comparisons with Bonferroni corrections showed that targeting on the mobile phone whether unencumbered or carrying the objects caused PWS to significantly decrease than walking alone. Carrying the wider box under the non-dominant arm resulted in the slowest mean PWS of 1.7km/h, a drop in walking speed of 41.2% when compared to the baseline PWS. The results were also significant when comparing the walking speeds of unencumbered target selections to input when carrying the bags and boxes. PWS declined further when the bags and boxes were held during target selections on the touchscreen phone.

Walking condition	W Statistic	Sig.
Baseline	0.959977	0.601187
No object	0.978613	0.934582
Small bag (ND)	0.9562	0.530249
Small bag (D)	0.96763	0.752135
Medium bag (ND)	0.927534	0.175702
Medium bag (D)	0.943984	0.338521
Standard box (ND)	0.95549	0.517444
Standard box (D)	0.957682	0.557579
Wider box (ND)	0.971001	0.816294
Wider box (D)	0.961345	0.6278

Table 4.8: Shapiro-Wilk normality tests performed on PWS for the walking conditions in *User Study 2*. ND = non-dominant and D = dominant.

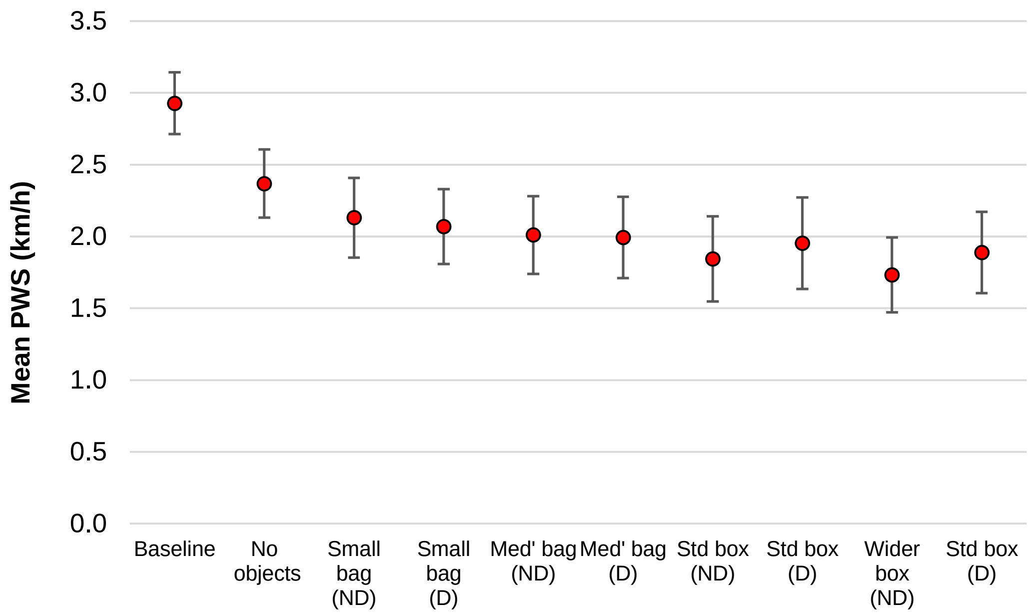

Figure 4.19: The mean PWS for each walking condition in *User Study 2*. Error bars denote Confidence Interval (95%).

4.4 Discussion

The results showed that the participants made significantly more inaccurate tapping selections when walking than standing therefore hypothesis *H1A* is supported. The mean target accuracy of unencumbered tapping while standing was 68.8%. It was somewhat surprising to see that almost one third of the target selections were inaccurately selected despite the participants being non-mobile and were not physically impaired by holding the cumbersome bags and boxes. For comparison, Schildbach and Rukzio [66] reported an accuracy of 77% when users in their study performed a comparable target acquisition task on a mobile phone while standing and unencumbered. However, the smallest target size tested in Schildbach and Rukzio's experiment was 6.74 x 6.74mm, which is bigger than the 4 *x* 6mm targets in *User Study 2*. The cost of mobility on tapping performance was evident as target accuracy dropped to 46.2% when the participants were walking while unencumbered. This illustrates the difficulties of selecting small targets and on-screen elements accurately when users are on the move despite not carrying objects.

Target accuracy was significantly higher when unencumbered than holding the bags and boxes., therefore hypothesis *H2A* is supported. The overall mean target accuracy when unencumbered was 57.5%. Holding the small and medium bags in the dominant hand caused significantly fewer accurate target selections as the overall mean target accuracies were 43% and 41.5% respectively. In addition, holding the medium bag in the dominant hand while walking resulted in the lowest mean accuracy of 29.8%. The standard and wider boxes held under the dominant arm caused the lowest overall mean accuracies were 39.1% and 39.7% respectively. These results show poor touch performance in terms of single finger tapping accuracy when the dominant hand or arm is encumbered.

Comparing target accuracy between holding the objects in the dominant and non-dominant sides revealed mixed results. There was no significant difference for accuracy between holding the small bag in the dominant and non-dominant hands. Likewise, holding the medium bag in the dominant hand did not decrease target accuracy when compared to holding the same bag in the non-dominant hand. Conversely, carrying the standard box under the dominant arm caused target accuracy to decrease when compared to holding the same box under the non-dominant arm. Likewise, target accuracy declined when the wider box was held under the dominant arm than carrying the same object under the dominant arm. Therefore, hypothesis *H3A* can only be partially supported and is rejected. It was observed that carrying the boxes meant that the encumbered arm had limited agility as users tried to securely grasp the bulky object in place against the side of their torso. While the non-dominant hand holding the mobile phone could adjust to make input physically easier for the input finger of the dominant hand, it was observed that tapping on the touchscreen was challenging especially when the boxes were held under the dominant arm. In contrast, the bags were less awkward to hold due to the dedicated handles but comments from the participants suggested that the bags were more tiring to carry than the boxes since most of the weight of the bags was focused directly on the users' hands.

Target error was measured to examine the level of tapping precision so that more appropriate target sizes could be defined to improve input performance in walking and encumbered contexts. The results for targeting error showed that the participants were less precise at tapping on the touchscreen mobile phone when walking than standing. The overall mean error when standing was 3.3mm and significantly increased to 4.3mm when walking. Therefore, hypothesis *H1B* is supported. As expected, walking had a negative impact on target accuracy so tapping precision is also affected. Thus, the targets tested in *User Study 2* are less effective when users are walking and need to be enlarged to improve tapping performance.

Comparing the targeting errors for the different types of encumbrances showed that tapping was more precise when unencumbered than when holding the objects with the exception of carrying the standard box under the non-dominant arm. Therefore, hypothesis *H2B* cannot be fully supported and is therefore rejected. The medium bag held in the dominant hand caused the highest overall mean error of 4.4mm, an increase of 44% when compared to unencumbered tapping precision. Both types of boxes carried under the dominant arm resulted an identical targeting error of 4.2mm.

Holding the small bag in the dominant hand did not significantly increase targeting error when compared to holding the same object in the non-dominant hand. In contrast, targeting error was significantly higher when the medium bag was held in the dominant hand than holding the same bag in the non-dominant hand. Despite the significant result, the difference in error between holding the medium bag in the dominant and non-dominant hands was small (0.5mm). The standard box held under the dominant arm caused a significantly higher targeting error than carrying the same object under the non-dominant arm. The comparison of targeting error between the dominant and non-dominant arms when the wider box was held was not significant. Based on these results, hypothesis *H3B* is only partially supported and is rejected.

Regardless of the user's physical context, there is always a cost to tapping precision when fingers are used to select small touchscreen components. The offset between the perceived and actual selected location on the touchscreen is known as the "fat finger problem" [73] as selection becomes ambiguous when the input finger either completely or partially occludes the target. Tapping precision is further affected when users are walking and encumbered. The mean error measured for the baseline condition of unencumbered target selections while standing was 2.6mm. Error doubled to 5.2mm when participants were walking and carrying the medium bag in the dominant hand. Even when standing, the standard box held under the dominant hand caused a mean targeting error of 3.9mm, an increase of 52.6% when compared to the baseline unencumbered condition.

Based on the highest mean targeting error measured, target width would need to be a minimum of 10.4mm in order to substantially improve target accuracy in walking and encumbered situations. The efficacy of increasing target size for touch input has been well documented. For example, in Schildbach and Rukzio's unencumbered walking study [66], the highest accuracy of 84% was reported when participants selected the largest 9.50 x 9.50mm targets which was greater than selecting the smaller targets tested in their experiment. However, there is a limit as to how much target size could increase on mobile phones due to restricted screen space and applications that have many compact targets (for example, text entry).

***

The results for target selection time showed that tapping speed was significantly slower when walking than standing and therefore hypothesis *H1C* is supported. As predicted, walking caused users to input slower when compared to standing but it is worth noting that the significant difference between the two means was small at 54.6ms. For comparison, Schildbach and Rukzio [66] reported a mean target selection time of 459ms for standing which significantly increased by 31.4% to 603ms when walking.

The comparisons for selection time between the different types of encumbrances showed varied results. Selection time while unencumbered was slower than holding the medium bag in the dominant hand and carrying both boxes under the dominant arm. The remaining pairwise tests comparing the selection times between holding no object and the other encumbrance scenarios were not significant. Thus, hypothesis *H2C* is rejected. It was anticipated that more time would be required for targeting when encumbered than when holding no objects. Perhaps the participants felt more rushed to complete the target selections when objects were held while when unencumbered, there was less hurry and more time was spent to input more accurately. But overall, the difference in mean selection time between the different types of encumbrance was small which suggests that encumbrance did not have a strong effect on input speed.

As predicted, selection times showed that the participants took longer to input when the dominant hand or arm was encumbered compared to holding the objects on the non-dominant side. The small bag held in the dominant hand caused slower targeting speed than holding the same bag in the non-dominant hand. Similarly, target selections took longer for the medium bag when held in the dominant hand than the non-dominant. Both types of boxes when held under the dominant arm resulted in slower targeting speed than carrying them under the non-dominant arm. Therefore, hypothesis *H3C* is supported. Again, it is worth noting that the significant differences between the means were small. No significant increase in selection time between the dominant and non-dominant sides was greater than 15%. It was anticipated that encumbrance would have had a larger effect on targeting speed when the dominant hand was encumbered.

Using motion capture cameras to track the participants during input showed that walking caused a significant increase in movement along all three axes to the non-dominant hand which held the mobile when compared to standing still. The extraneous movements caused while walking made it difficult for users to maintain a stable input posture and consequently, input becomes prone to errors as shown by the target accuracy results. Despite the significant results revealing that there were more movements in all three axes in the non-dominant hand when it was encumbered compared to unencumbered, the differences were very small (< 0.3mm). This finding suggests that in terms of unsteadiness, there were less physical effects of carrying objects in the non-dominant hand that also held the mobile phone despite of the decrease in targeting accuracy. When the objects were held in the dominant hand or arm, a significant increase in movement to the non-dominant hand was recorded. Even though the non-dominant hand was not encumbered, further adjustments to the input posture might have been made to make input easier to select the targets when the dominant hand also had to carry the bags or hampered by the boxes.

The results for walking speed showed that the mean baseline PWS was 2.9km/h, which is slower than those reported in related walking studies [5,7,66]. Due to the location and setup of the motion capture cameras, a shorter walking route was used in *User Study 2*, which might have caused the participants to walk slower than normal. However, interaction and encumbrance had an impact on walking speed. The mean walking speed was significantly reduced by 19.1% to 2.4km/h when the participants were targeting on the mobile phone while unencumbered. This finding supports previous research as Bergstrom-Lehtovirta *et al.* [7] found that walking speed dropped by 24% when users in their study performed a similar targeting task on the mobile phone. Schildbach and Rukzio [66] also reported a significant drop in walking speed of 27% when users selected the smallest 6.74mm target widths in their target acquisition task.

Targeting on the mobile phone while encumbered caused walking speed to decrease further. Carrying the bags and boxes on both the dominant and non-dominant sides during input resulted in a significant drop in walking speed when compared to walking alone. The slowest recorded mean PWS of 1.7km/h occurred when the wider box was held under the non-dominant arm, a decline of 41% when compared to the baseline PWS. Even though walking speed was reduced further when the participants were carrying bags and boxes, the number of inaccurate target selections still increased, which illustrates the interaction difficulties when encumbered. The participants might have tried to trade walking speed (i.e. walking slower) to input more accurately when encumbered but the results showed that the reduction in momentum was not enough the improve target selections.

To answer research question *Q2.1* (*How do encumbrance and mobility affect tapping performance on touchscreen mobile phones?*), the results from *User Study 2* have shown that walking decreased tapping accuracy while all encumbrance scenarios tested caused a reduction in the number of accurate target selections when compared to unencumbered input. There were no statistical significant differences for accuracy when holding the bags between the dominant and non-dominant hands. Conversely, both types of boxes caused a statistical significant decrease in accuracy when the objects were carried under the dominant arm. Target accuracy decreased by as such as 70% when the larger of the two types of bags was held in the dominant hand. As a consequence of poor target accuracy, walking and encumbrance also had a negative effect on tapping precision. The highest mean error recorded was 5.2mm when users were walking and carrying the medium bag in the dominant hand. In terms of tapping speed, walking had a stronger effect than carrying the objects, as the differences in selection time across all encumbrance types were comparable. Tapping speed was significantly slower when walking by 12% compared to standing still.

4.5 Conclusions

To sum up, this chapter has answered research question *Q2.1* by presenting a user study which was conducted to examine the impact encumbrance and mobility have on tapping performance on a typical touchscreen mobile phone. Four encumbrance scenarios defined in Section 3.5, which evaluated two different sizes of bags (held in-hand) and two different widths of boxes (grasped underarm) were used to investigate the impact of physically constraining the dominant and non-dominant sides during selections of 4 *x* 6mm targets. The results showed that walking and carrying the objects caused both target accuracy and tapping precision to decrease when compared to standing still and unencumbered. The results from *User Study 2* also support previous walking studies that reported a decrease in PWS during interaction and when users were also encumbered, walking speed dropped even further.

The work described in this chapter extends current research on mobile interactions by discussing the effects of encumbrance and walking on target selections and show the extent targeting performance declines in these physically demanding situations. *User Study 2* primarily focused on examining tapping performance in a range of encumbrance scenarios while walking and the task was completing only in the two-handed index finger posture. However, other input postures are commonly used with touchscreen mobile phones and when encumbered, users might switch to a more suitable position to input more effectively. The next chapter presents a user study that investigated tapping performance in three different input postures while walking and encumbered.

Chapter 5

User Study 3: The Effects of Encumbrance, Walking and Input Posture on Targeting Performance

5.1 Introduction

In the previous chapter, the effects of encumbrance and walking on tapping performance were examined and results showed that carrying the selected bags and boxes while walking caused a decrease in target accuracy. The main focus of *User Study 2* was to evaluate a variety of different encumbrance scenarios, so to keep the number of experimental conditions feasible without added further complexity to the experimental design, target selections were completed only in the two-handed index finger posture. In encumbered situations where interaction becomes too physically demanding, users are likely to change to another input posture to try to improve performance. For example, switching to one-handed input so the non-interacting hand could be used to hold objects or for other activities. Therefore, this chapter answers research question *Q2.1.1* (*How does the change of input posture affect tapping interactions when walking and encumbered?*) by presenting a user study that was conducted to investigate the effects of encumbrance and walking on one- and two- handed tapping performance.

Three main input postures, as shown Figure 5.1, were evaluated: *two-handed index finger* (mobile phone held in the non-dominant hand while the index finger of the dominant hand was used for input), *one-handed preferred thumb* (the device and input was both completed using the dominant hand only) and two-handed both thumbs (both hands were used to hold the device and both thumbs were used for input). The input postures were selected due to the survey by Hoober [32], found that 49% of people used mobile devices in the one-handed preferred thumb posture, while 36% held the device in the two-handed index finger position to input. The remaining 15% of users held the device in the two-handed both thumbs posture. Prior to Hoober’s survey, Karlson *et al.* [36] conducted a field study in an airport to examine user behaviour with mobile devices and their results suggested that 60% of users engaged with their mobile phone in a one-handed posture when walking. A follow-up survey suggested 45% of participants would use one hand only for all device interactions compared to 19% for two-handed interactions.

Previous research has also measured tapping performance in the three input postures mentioned. For example, Azenkot and Zhai [3] compared text entry performance using the index finger, preferred thumb and both thumbs on a touchscreen mobile phone. Their results showed that typing with both thumbs was faster than using one thumb or the index finger while error rate was higher when using both thumbs for input. Musić and Murray-Smith [53] examined the effects of walking on the three input postures and a further two more position: using both thumbs to tap while the mobile device is held in portrait mode and pointing with the index finger while the device was held in landscape position. They reported that overall the index finger outperformed the thumb-based pointing methods. However, little research has compared targeting performance in the three input postures while walking and carrying objects. Therefore, *User Study 3* investigated how tapping on touchscreen mobile phones is affected while encumbered and on the move when the input posture is varied.

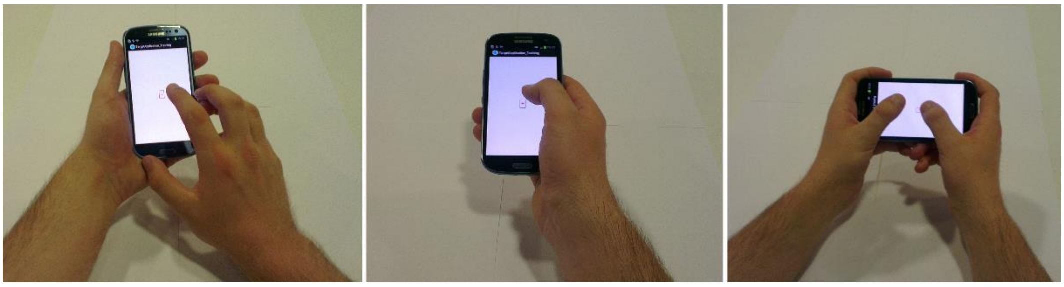

Figure 5.1: Three common input postures: two-handed index finger (left), one-handed preferred thumb (middle) and two-handed both thumbs (right).

5.2 Method

5.2.1 Encumbrance Scenarios

Encumbrance Scenarios 1C and *1D* from Section 3.5 were evaluated in *User Study 3*, so participants held a shopping bag in each hand during input on a mobile phone. These particular encumbrance scenarios were chosen for two reasons. Firstly, because *User Study 3* tested one- and two- handed interactions, the encumbrance scenarios had to have a consistent physical effect across all three input postures. For example, during one-handed input, if all objects were carried in one hand then there would be no direct physical effect on the dominant hand holding and tapping with the mobile device. And secondly, the selected encumbrance scenarios allowed tapping performance to be evaluated in situations where both hands are physically constrained due to carrying multiple objects. Each bag measured approximately 330 *x* 480mm, which was similar to the medium bag used in *User Study 2*. As discussed in Section 3.5, the weight of each bag should not exceed 1.7kg when two objects are held. After pilot testing before *User Study 3* began, this weight was found to be slightly too heavy to carry for prolonged periods, so the weight of each bag was reduced to 1.6kg to minimise fatigue. Figure 5.2 illustrates the bags and the way they were held during one- and two- handed input.

Figure 5.2: The encumbrance scenario evaluated in *User Study 3*. A bag was held in each hand during one- (left) and two- (middle and right) handed input.

5.2.2 Task

A similar targeting task to Crossan *et al.* [15] was used in *User Study 3*. Participants selected a series of targets one at a time on a touchscreen mobile phone as quickly and as accurately as possible. There were nine target positions evenly spaced in a 3 *x* 3 grid. The centre and one of the outer targets were selected in an alternate sequence - every second selection was an outer target - and the order of the outer targets was randomized for each block of trials. Each outer target was selected ten times which resulted in 160 target selections per block and two blocks were completed for each condition per participant. Similar to Crossan *et al.* [15], there was a random interval ranging from 500 to 1500ms between a selection and the next target being shown on-screen to negate any rhythm created between a user's walking and tapping behaviour.

The task ran on a Samsung Galaxy S3 smartphone which had a 4.8″ touchscreen and a resolution of 720 *x* 1280 pixels (~12 pixels/mm). Each target was 60 pixels (5mm) wide and 96 pixels (8mm) long with the central crosshair measuring 30 pixels (2.5mm) in both directions. This was the size of a key on the standard keyboard for this phone. The device was held in portrait orientation for both the two-handed index finger and one-handed preferred thumb input postures. The device was used in landscape mode for the two-handed both thumbs posture with the bottom end of the device always to the right for consistency. Participants were given a short training phase at the start of the study to familiarise them with the targeting task in each input posture. Figure 5.3 shows the target selection task and the orientations of the device used in *User Study 3*.

Figure 5.3: The target selection task ran on a Samsung Galaxy S3 phone. The nine target positions are shown in portrait (left) and landscape (right) orientations. The dashed circle represents the centre target while the remaining eight positions are denoted as the outer targets.

5.2.3 Measuring and Maintaining Preferred Walking Speed

The study presented in Chapter 4 showed that encumbrance and interaction caused a decrease in walking speed. Despite this naturalistic behaviour, one issue is that it becomes difficult to interpret the targeting results since walking speed is mixed up with the effects of encumbrance. Furthermore, users can trade walking speed with input performance (i.e. walking slower to input more accurately) so it then becomes problematic to measure the true cost of encumbrance. In *User Study 3*, walking speed was controlled which meant the effects of mobility were isolated from the effects of encumbrance. Therefore, any changes observed on tapping performance were likely caused by encumbrance or/and input posture.

All experimental conditions were performed while walking, so a pre-defined oval route (shown in Figure 5.4) was marked out using plastic cones in a spacious and open room. The path measured 20m in total length and was 1.2m wide. Participants were instructed to keep within the path during the experiment. Before the experiment began, three versions of each participant's PWS were recorded:

PWS: Each participant was instructed to walk around the route for five laps at a pace that he/she would normally walk. The total amount of time required was recorded and the average walking speed was calculated, denoted as PWS. This is the standard measure of PWS [62]. No mobile device was used nor bags carried.

PWS&E: The first step was repeated but participants carried one bag in each hand to measure any change in PWS due to encumbrance. The calculated walking speed is denoted as PWS&E and gave a baseline for walking speed when encumbered.

PWS&I: The first step was repeated again but participants also performed the targeting task on the mobile phone to measure the walking speed during interaction (but unencumbered), denoted as PWS&I. This gave a baseline for walking performance when interacting. All participants performed the task described above in Section 5.2.2 in the two-handed index finger posture for consistency. Although the same targeting task was used in the main experiment, targets for each condition were randomly ordered to counterbalance learning effects.

For each condition, the participants maintained a constant walking speed by following a pacesetter around the route. Pacesetters have been used in previous mobile studies (e.g. [22,35,56]) where participants walked at the same pre-defined pace so that walking speed does not become a dependent variable. In *User Study 3* however, each participant walked at their individual PWS&I since it is a more accurate representation of the walking speed that the user naturally walks when using a mobile device. The pacesetter walked at the calculated PWS&I by using a metronome application that ran on a HTC One X phone. The pacesetter used the application to tune the metronome speed for each participant once his/her PWS&I 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. Vibro-tactile feedback was considered but during initial testing, the experimenter had difficulties walking at the desired pace thus auditory feedback was used.

For each 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 the targeting task on the 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. Lastly, at the end of the experiment, both PWS and PWS&E were measured again to assess any fatigue caused by interaction and carrying the bags.

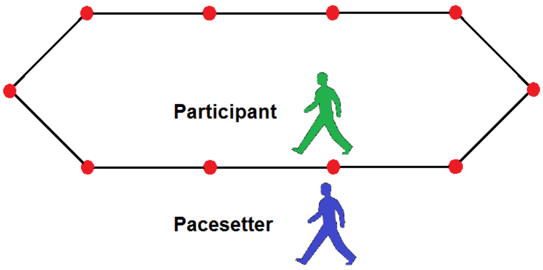

Figure 5.4: The pre-defined oval route that participants walked in *User Study 3* (left). During the experiment, the participant (green figure) maintained their PWS&I by walking side-by-side with the pacesetter (blue figure), as shown in the right image.

5.2.4 Experimental Design

Eighteen participants (11 males, 7 females) recruited from the University took part. The mean age was 25 years (SD = 3.52) and all participants preferred using their right hand for interaction (despite one individual being naturally left-handed). Sixteen participants owned and used a touchscreen mobile phone daily while the remaining two users occasionally interacted with touchscreen mobile devices. Each participant received £6 for completing the study and the experiment took approximately one hour to complete. The participants performed the targeting task either unencumbered or carrying a bag in each hand for each of the three input postures, resulting in a total of six conditions. Each condition was completed while walking at the measured PWS&I by following a pacesetter around the route. A within-subjects design was used and the conditions were counterbalanced to reduce learning and order effects as much as possible.

The Independent Variables were **Type of Encumbrance** (2 levels - unencumbered and carrying the bags) and **Input Posture** (3 levels - two-handed index finger, one-handed preferred thumb and two-handed both thumbs). The Dependent Variables were target accuracy, target error and selection time. Target accuracy was measured as the percentage of successful target selections; the touch up position tapped on the touchscreen was either within the target border or not. Target error (in millimetres) was the absolute distance from the centre of the target to the recorded touch up position. Selection time (in milliseconds) was the duration from the display of the current target to the instance that a press up event was logged. The main hypotheses were:

H1: Participants will be significantly less accurate at target selection when encumbered compared to unencumbered, while walking at their PWS&I;

H2: Participants will be significantly less precise at target selection when encumbered compared to unencumbered, while walking at their PWS&I;

H3: Participants will require significantly more time to target when encumbered compared to unencumbered, while walking at their PWS&I;

H4: Target accuracy using both thumbs will be significantly higher than input using one thumb or the index finger;

H5: Target selections using both thumbs will be significantly more precise than input using one thumb or the index finger;

H6: Target selections using both thumbs will be significantly faster than input using one thumb or the index finger when encumbered;

H7: The PWS will be slower when encumbered or interacting with a mobile device than walking alone.

5.3 Results

Each participant completed 12 blocks of target selections – six conditions and two blocks per condition. As a result, a total of 34,560 targets (*160 targets x 2 blocks x 6 conditions x 18 participants*) was recorded for the whole experiment. To filter out unintentional taps, target selections that took less than 100 milliseconds were not included in the final data analysis. Consequently, 23 targets were eliminated from the data.

5.3.1 Target Accuracy

Shapiro-Wilk tests were performed to examine the normality of target accuracy. The results are shown in Table 5.1 and no condition was significant. Therefore, a two-factor, repeated-measures ANOVA with **Type of Encumbrance** (2 levels) and **Input Posture** (3 levels) as factors was conducted to analyse target accuracy. Mauchly's test for sphericity was significant therefore Huynh-Feldt adjustments were used to correct the degrees of freedom ($\varepsilon > 0.75$). The mean target accuracy for each condition is shown in Figure 5.5.

Type of Encumbrance	Input Posture	W Statistic	Sig.
No objects	Index finger	0.965022	0.700595
No objects	Preferred thumb	0.96046	0.610541
No objects	Both thumbs	0.946322	0.370271
Bags	Index finger	0.939747	0.286913
Bags	Preferred thumb	0.940475	0.295256
Bags	Both thumbs	0.901219	0.060339

Table 5.1: Shapiro-Wilk normality tests performed on target accuracy for each condition in *User Study 3*.

The ANOVA conducted to analyse target accuracy showed a significant main effect for **Type of Encumbrance**, $F(1, 17) = 87.880$, $p < 0.01$. *Post hoc* pairwise comparison with Bonferroni corrections showed that the participants were significantly more accurate when unencumbered compared to carrying the bags. The overall mean accuracies for unencumbered and carrying the bags were 65.4% and 53.7% respectively. No significant main effect was observed for **Input Posture**, $F(2.0, 33.5) = 2.113$, $p > 0.05$. A significant interaction was observed between the factors, $F(2.0, 33.6) = 3.757$, $p < 0.05$. Carrying the bags in both hands caused target accuracy to decrease significantly for each input posture when walking. The significant result for the interaction is not required to support or reject the hypotheses stated in Section 5.2.4.

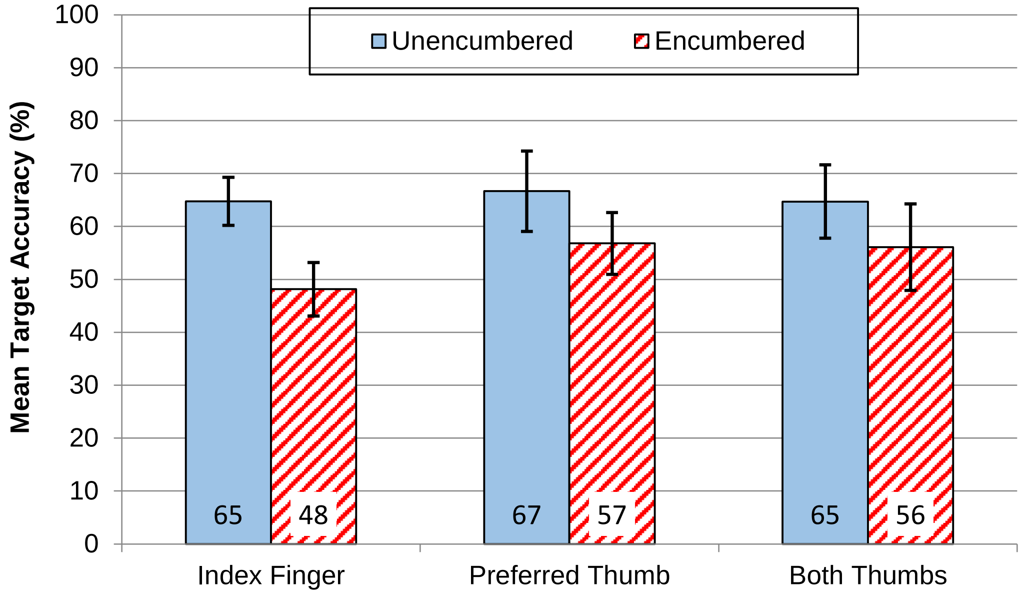

Figure 5.5: The mean target accuracy (%) for each condition in *User Study 3*. The solid blue and striped red bars illustrate the unencumbered and carrying the bags conditions respectively. Error bars denote Confidence Interval (95%).

5.3.2 Target Error

Shapiro-Wilk normality tests were performed to examine the distribution of the data recorded for target error. The results are shown in Table 5.2 and one condition was significant: preferred thumb input while holding the bag. The box-and-whisker plots in Figure 5.6 illustrates the distribution of target error for each condition. Because target error was not normally distributed, the data was transformed using ART [76] before conducting a two-factor (**Type of Encumbrance** and **Input Posture**) repeated-measures ANOVA. The mean targeting error for each condition is shown in Figure 5.7.

The results from the ANOVA showed a significant main effect for **Type of Encumbrance**, $F(1, 85) = 85.2286$, $p < 0.001$. Tapping was significantly less precise when carrying the bags than unencumbered. A significant main effect was observed for **Input Posture**, $F(2, 85) = 3.9199$, $p < 0.05$. *Post hoc* Tukey HSD tests showed a significant difference in error between the index finger and both thumbs ($t = 2.715$, $p < 0.0217$). Target error was higher when using the index finger than using both thumbs. No significant differences were found in the comparisons between one thumb and either the index finger or both thumbs. The interaction between the two factors was not significant, $F(2, 85) = 2.6052$, $p > 0.05$.

Type of Encumbrance	**Input Posture**	**W Statistic**	**Sig.**
No objects	Index finger	0.916422	0.111722
No objects	Preferred thumb	0.937137	0.25871
No objects	Both thumbs	0.960576	0.612793
Bags	Index finger	0.941525	0.307659
Bags	Preferred thumb	0.776066	0.000711*
Bags	Both thumbs	0.954218	0.49495

Table 5.2: Shapiro-Wilk normality tests performed on target error for each condition in *User Study 3*. Significant results are shaded in grey and highlighted with '*'.

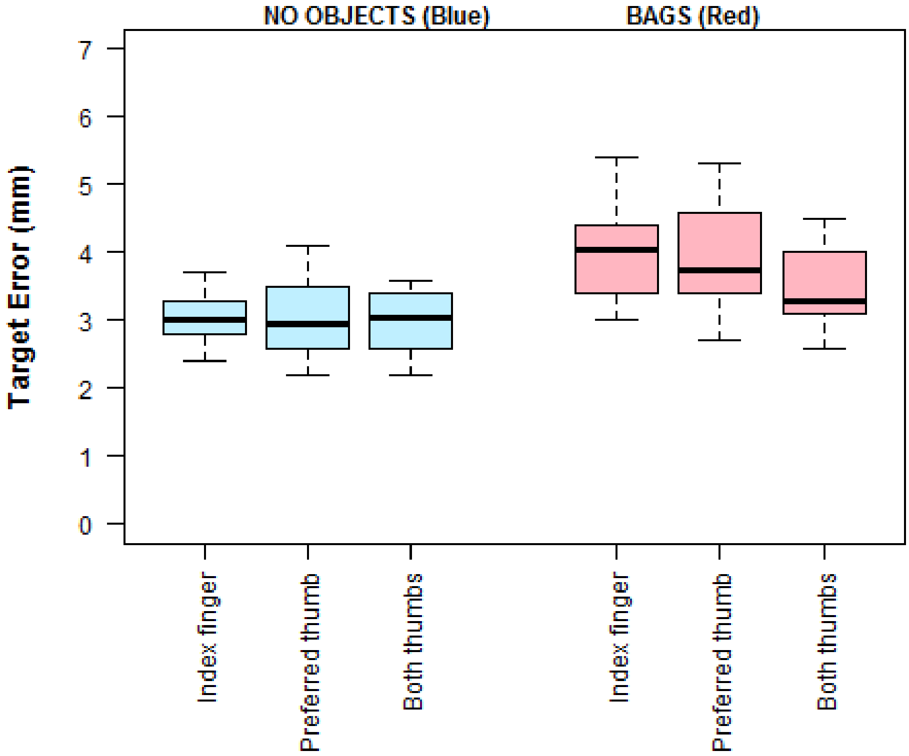

Figure 5.6: Box-and-whisker plots on target error for the conditions in *User Study 3*.

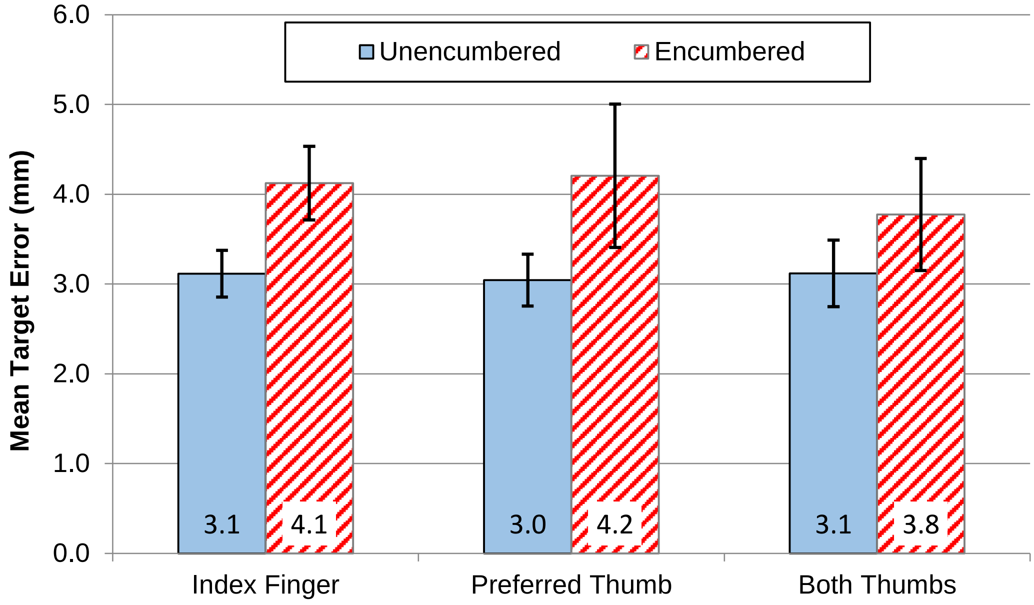

Figure 5.7: The mean target error (mm) for each condition in *User Study 3*. The solid blue and striped red bars illustrate the unencumbered and carrying the bags conditions respectively. Error bars denote Confidence Interval (95%).

5.3.3 Selection Time

Shapiro-Wilk normality tests were conducted to assess the normality of the data collected for selection time. The results are shown in Table 5.3 and no condition was significant, thus selection time did not violate normality. A two-factor (**Type of Encumbrance** and **Input Posture**) repeated-measures ANOVA was conducted to analyse selection time. Mauchly's test for sphericity was significant therefore Huynh-Feldt adjustments were used to correct the degrees of freedom ($\varepsilon > 0.75$). Figure 5.8 illustrates the mean selection times for each condition.

The ANOVA for selection time showed a significant main effect for **Type of Encumbrance**, $F(1, 17) = 11.672$, $p < 0.05$. Target selections took significantly longer when carrying the bags than interaction without carrying the objects. However, it is worth noting that the difference in mean selection time was small at 34.6ms. The overall mean selection times for unencumbered and carrying the bags were 544.7ms and 579.3ms respectively, an increase of 6.3%. A significant main effect was found for **Input Posture**, $F(1.7, 29.6) = 13.646$, $p < 0.05$. *Post hoc* pairwise comparisons with Bonferroni corrections showed a significant difference between all pair combinations, except between the two thumb-based input postures. Target selections using the two-handed index finger posture was significantly quicker than the one-handed preferred thumb and two-handed both thumbs poses. Input using both thumbs was not significantly quicker than using the preferred thumb only. Again, the difference between each comparison was small and no differences were greater than 60ms. A significant effect was also observed for the interaction between the two factors, $F(2.0, 34.0) = 3.924$, $p < 0.05$. Encumbrance caused significantly slower selection time for each input posture than unencumbered. The largest negative effect was on the one-handed preferred thumb posture when encumbered. The interaction effect between the factors is not necessary to support or reject the hypotheses.

Type of Encumbrance	Input Posture	W Statistic	Sig.
No objects	Index finger	0.926133	0.165979
No objects	Preferred thumb	0.931863	0.209382
No objects	Both thumbs	0.979571	0.945472
Bags	Index finger	0.944238	0.341861
Bags	Preferred thumb	0.980336	0.953415
Bags	Both thumbs	0.955844	0.523802

Table 5.3: Shapiro-Wilk normality tests performed on selection time for each condition in *User Study 3*.

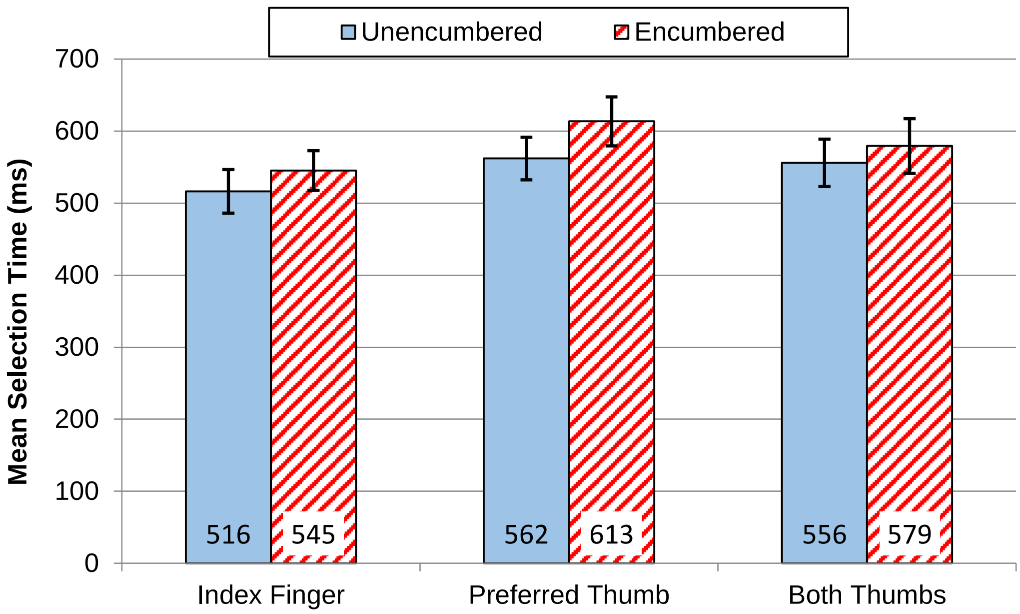

Figure 5.8: The mean selection time (ms) for each condition in *User Study 3*. The solid blue and striped red bars illustrate the unencumbered and carrying the bags conditions respectively. Error bars denote Confidence Interval (95%).

5.3.4 Performance of Individual Target Positions

This section presents the results for each target position for the three input postures to see how users performed across the touchscreen phone. Related work has examined touch locations on mobile phones in similar input postures to build offset models, to predict the user's intended input, to improve selection accuracy while stationary (e.g. [9,74,75]) and walking [53]. This thesis does not examine such predictive models but assess the onscreen performance to see if there are certain areas that users struggle to tap on while walking and encumbered.

The tapping performance of each target position when the two-handed index finger input postures was used is shown in Figure 5.9. Carrying the bags caused accuracy to decrease for all target positions when compared to unencumbered. The left target in the top row had the lowest accuracy of 41% when encumbered. The same target position also had the slowest selection time of 620ms. In general, input speed was slower for all target positions when both hands were carrying the bags than unencumbered input. However, the differences in selection times were small. The spread of the taps at each target position was greater when both hands were carrying the bags than unencumbered as shown by the larger red ellipses. The left targets on the top and middle row had the largest variability of taps when carrying the bags than when unencumbered.

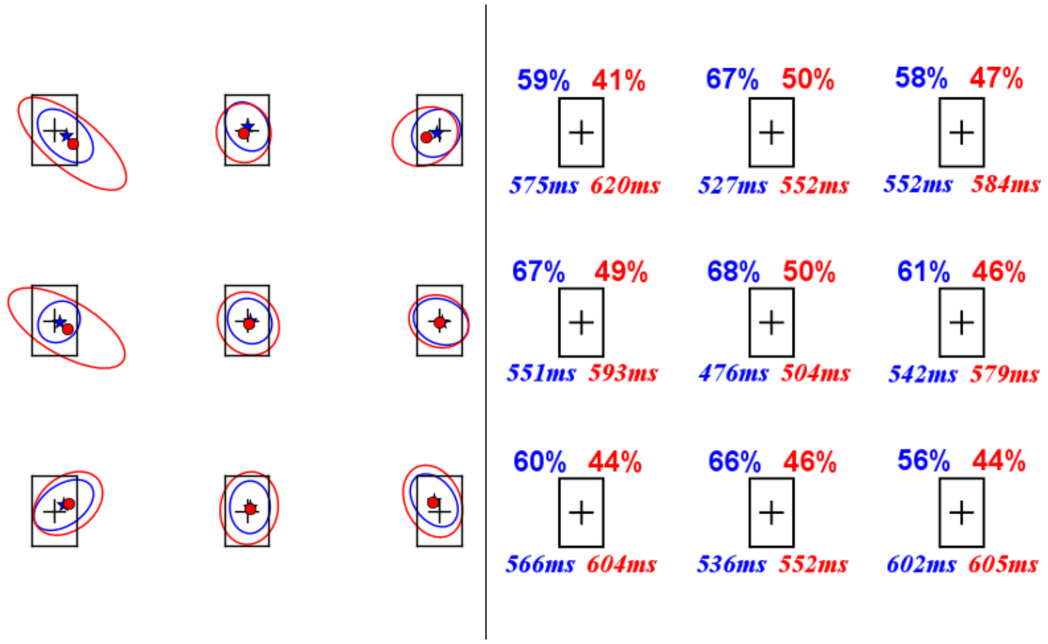

Figure 5.9: The performance of each target position in the two-handed index finger input posture. The left image illustrates the mean and covariance of the x and y targeting error of each target when unencumbered (blue) and encumbered (red). The ellipses represent one standard deviation. The right image shows the mean accuracy (top) and selection time (bottom italic) of each target position when unencumbered (left & blue) and carrying the bags (right & red).

Examining the tapping performance of each target position in the one-handed preferred thumb posture (see Figure 5.10) showed much larger differences between unencumbered and carrying the bags. Accuracy was lower while selection time was slower for all target positions when carrying the bags than unencumbered input. The targets (all targets on the top row and the left targets on the middle and bottom rows) that are furthest away from the thumb (all participants used their right thumb) had the lowest target accuracies regardless of carrying the bags or unencumbered. The left targets on the top and bottom row that where the most difficult to reach had the poorest tapping accuracies when unencumbered and carrying the bags in both hands. Target accuracy was slightly higher for the left target on the bottom row than the left target on the top row but selection time was longer so a trade-off in performance might have occurred. In contrast, the targets (the centre and right targets on the middle and bottom rows) that were closer to the thumb had higher target accuracy for both unencumbered and when carrying the bags. The closer but biomechanically difficult to reach targets (the right targets on the middle and bottom rows) caused higher selection times than the other nearby targets. The left target in the top row caused the greatest spread of taps when encumbered which illustrates the physically difficulties of selecting on-screen components that are out of the thumb's ideal movement range during one-handed input.

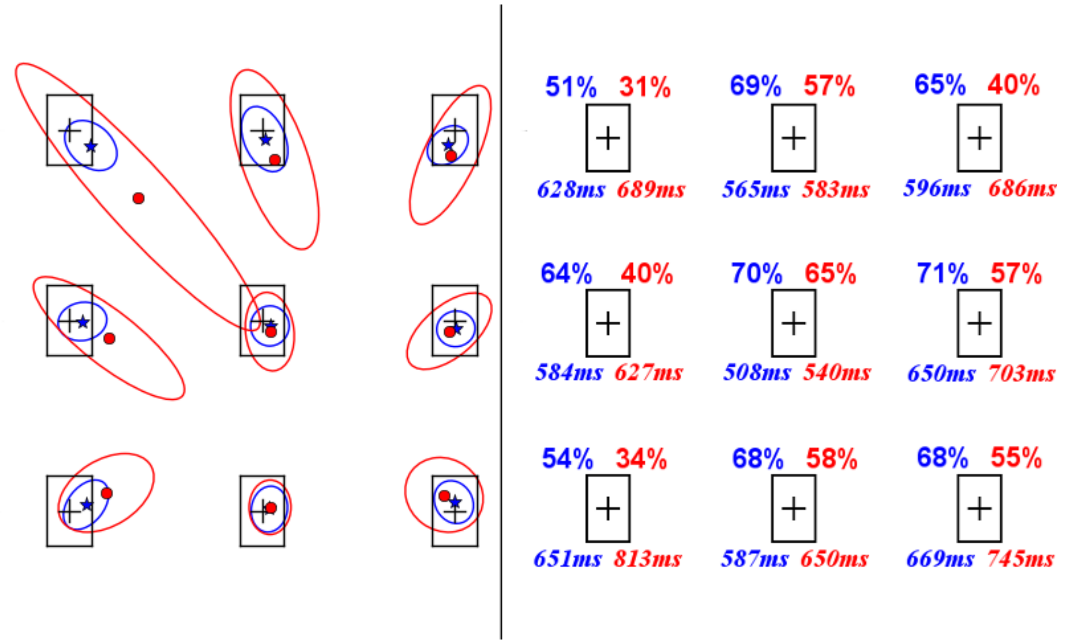

Figure 5.10: The performance of each target position in the one-handed preferred thumb input posture. The left image illustrates the mean and covariance of the x and y targeting error of each target when unencumbered (blue) and encumbered (red). The ellipses represent one standard deviation. The right image shows the mean accuracy (top) and selection time (bottom italic) of each target position when unencumbered (left & blue) and carrying the bags (right & red).

The performance of most target positions was more evenly matched between unencumbered and carrying the bags when both thumbs were used for input (see Figure 5.11). In terms of tapping speed, the selection times for all target positions were similar between holding no objects and the bags. The biggest difference in selection time of 45ms between the two types of encumbrances occurred with the right target in the middle row. The centre target in the middle row had the quickest selection time for both unencumbered and carrying the bags. For accuracy, the number of correct selections for each target position was similarly matched between unencumbered (<7%) and carrying the bags with larger differences occurring with the three targets on the bottom row (> 10%). The right targets on each row had lower accuracy than the left targets on the same row despite all participants preferred using their right hand for input. The distribution of taps at each target position was also comparable between holding no objects and the bags. The spread of taps for the centre targets on each row showed a slight offset to the right which suggests the participants preferred using their dominant right thumb to select those targets.

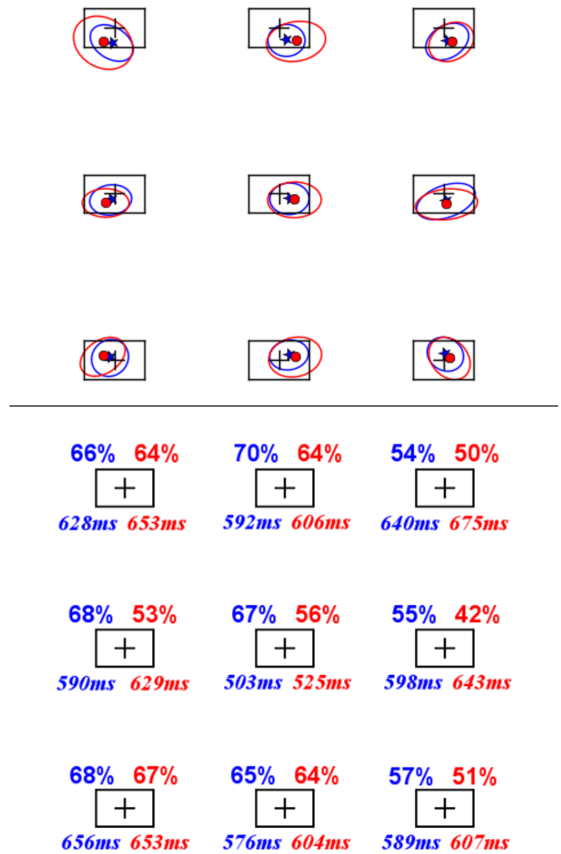

Figure 5.11: The performance of each target position in the two-handed both thumbs input posture. The top image illustrates the mean and covariance of the x and y targeting error of each target when unencumbered (blue) and encumbered (red). The ellipses represent one standard deviation The bottom image shows the mean accuracy (top) and selection time (bottom italic) of each target position when unencumbered (left & blue) and carrying the bags (right & red).

5.3.5 The Effects of Encumbrance and Input on Walking Speed

The different classes of PWS were measured and compared to examine the effects of encumbrance and interaction on walking speed. Table 5.4 shows the results from conducting Shapiro-Wilk normality tests and all three types of PWS were non-significant. The mean walking speeds for baseline PWS, PWS&E and PWS&I recorded at the start of the experiment are shown in Table 5.5.

A one-factor ANOVA with walking speed as factor (3 levels) showed a significant main effect, $F(1.1, 19.5) = 52.281$, $p < 0.01$. *Post hoc* pairwise comparisons with Bonferroni corrections showed that the participants walked significantly slower during targeting on the mobile phone (PWS&I) than walking while encumbered (PWS&E) and walking alone (PWS). There was no clear difference between baseline PWS and PWS&E (given the objects chosen for *User Study 3*).

At the end of the study when all the targeting tasks were completed, the baseline PWS and PWS&E were measured again for each participant. The same mean walking speed of 4.9 km/h was once again recorded for both types of walking speeds. The participants took an average of 18.3mins (SD = 1.9) to complete all six conditions (not including time for rests). The mean distance walked to complete all conditions was 1480.2m (~67 laps of the route).

Walking Speed	W Statistic	Sig.
PWS	0.956934	0.543698
PWS&E	0.956061	0.527727
PWS&I	0.947256	0.383612

Table 5.4: Shapiro-Wilk normality tests performed on the three types of PWS in *User Study 3*.

Walking Speed	Mean (km/h)	SD (km/h)
PWS	4.9	0.5
PWS&E	4.9	0.6
PWS&I	4.1	0.4

Table 5.5: The mean baseline PWS (top row), PWS&E (middle row) and PWS&I (bottom row).

5.4 Discussion

The results from the targeting task showed that carrying the bags caused tapping accuracy to significantly decrease by an overall of 11.7% when compared to unencumbered. Therefore hypothesis *H1* is supported. The two-handed index finger posture caused the largest drop in accuracy from 65% when unencumbered to 48% when carrying the bags, compared to difference of 10% and 9% for one-handed preferred thumb and two-handed both thumbs postures respectively. However, no significant differences were observed between the three input postures in terms of target accuracy so hypothesis *H4* is rejected.

Target accuracy between the three input postures when unencumbered and walking was very similar while a larger difference in accuracy occurred when the bags were held in both hands. These results suggest encumbrance had a greater effect on tapping accuracy than the type of input posture.

For tapping precision, target error was significantly higher when encumbered than carrying no objects and therefore hypothesis *H2* is supported. Target error decreased from 4.0mm when unencumbered to 3.1mm when carrying the bags. Only the difference in targeting error between the index finger and both thumbs postures was statistically significant, thus hypothesis *H5* cannot be fully supported and is rejected. Despite significant results, tapping precision between all three input postures was very similar when unencumbered or carrying the bags. The results suggest it is difficult to recommend a suitable input posture if tapping precision is required when walking and encumbered as no one posture substantially outperformed the others.

In terms of tapping speed, selecting the targets took significantly longer when encumbered than holding no objects therefore hypothesis *H3* is supported. Despite the significant result, the difference in selection time between unencumbered and carrying the bags was small (34.6ms). It was anticipated that selecting the targets would have taken more time when both hands were busy carrying the bags than unencumbered. Having an extra thumb for input did not make targeting speed quicker than using the preferred thumb only or the index finger. In fact, target selections took significantly less time when the index finger was used than both thumb-based input postures. Therefore, hypothesis H6 is rejected. Again, despite target selections were significantly quicker when the index finger was used, the differences were marginally when compared to the preferred thumb (57ms) and both thumbs (37ms).

The overall targeting performance showed that having an extra thumb for input did not improve targeting when encumbered and walking as predicted. Overall accuracy was near identical across all three input postures when walking while unencumbered. In addition, using the preferred thumb only had the same accuracy rate as using both thumbs when encumbered although the results were not significant. Likewise, encumbrance had a greater effect on targeting error than the type of input posture as the differences in tapping precision between the three postures were evenly matched. Target selection time was relatively similar across all three input postures as using both thumbs did not significantly or substantially improve tapping speed. Therefore, the performance of each individual target position was analysed to see how the three input postures performed at specific locations of the touchscreen.

When the two-handed index finger posture was used, accuracy for each target position dropped when carrying the bags. The left target on the top row had the lowest accuracy of 41% when both hands were carrying the bags. No individual accuracy when encumbered was greater than 50% when the index finger was used to select the targets. When no objects were held, the targets at the four corners of the grid had lower accuracies than the other five target positions. This suggests more selection difficulties when diagonal movement is required to select on-screen targets. The difference in selection time between unencumbered and carrying the bags for each target position was marginal when the index finger was used for input. No difference in selection time was greater than 50ms. The centre target in the middle row had the quickest selection time when unencumbered (476ms) and carrying the bags (504ms). After the training session and as the experiment progressed, tapping patterns are likely to have occurred as participants learned that every second selection was the centre target whereas the outer targets were randomised. For comparison, Musić and Murray-Smith[53] reported a wider distribution of touch points when users were walking compare to standing for tapping on targets of similar size to the ones examined in *User Study 3* when using the index finger for input.

For the one-handed preferred thumb posture, the target positions that were furthest away or biomechanically difficult to reach had poorer tapping performance than those targets within the optimal input range of the thumb. All the participants used their right hand for input so all three targets on the top row and the left targets on the middle and bottom rows which were all more difficult to reach had lower accuracies than the remaining four target positions that were closer to the thumb. The left target on the top row had the lowest accuracy of 31% when the bags were held. Accuracy for each target position was reduced when encumbered compared to holding no objects.

In terms of tapping speed when the preferred thumb was used, the difference in selection time for each target position between unencumbered and carrying the bags was less than 100ms except for the left target on the bottom row which had a difference of 162ms. Regardless of whether the participants were encumbered or not, the centre target in the middle row was selected quicker than the other eight target positions. As explained earlier, the participants knew every alternative tap return to the centre target and when the preferred thumb was used, the default starting position was always approximately above the centre target so less thumb movement was required. When unencumbered, the right target on the bottom row, which was physically awkward for the thumb to reach, caused the slowest selection time of 669ms. The left target on the bottom row which was further away from the thumb's ideal input range caused the highest selection time 813ms when the bags were carried. Comments from the participants and observations suggested that it was difficult to shift or adjust the position of the phone to make targeting easier during one-handed input when encumbered. The participants said they were more unlikely to adjust hand grip when carrying the objects due to the fear of dropping the phone or the objects, so attempted to input anyway even knowing that the selection was going to be incorrect. This suggests that better input techniques are required for one-handed interactions and the need to see if current thumb-based methods are still effective in walking and encumbered situations.

When both thumbs were used for input, the left targets on each row were selected more accurately than the right targets on the same rows whether the participants were encumbered or not. All participants preferred using their right hand for input, so it was anticipated that targets near the right side of the screen would have been acquired more accurately than those located on the left side. The right target on the top row had the lowest accuracy of 54% when unencumbered while the right target on the middle row caused the least number of accurate selections when the bags were held. Examining the distribution of taps showed a slight offset to the right for the centre targets on each row, which suggests that the right thumb was mainly used to acquire those targets at the centre of the screen. Like the two-handed index finger posture, the variance in selection time between unencumbered and carrying the bags for each target position was small as no difference was greater than 50ms. Again, the centre target was selected the quickest when holding no objects (503ms) and the bags (525ms). The left target on the bottom row caused the slowest selection when unencumbered (656ms) while the right target on the top row took the longest to select when both hands were encumbered (675ms).

The different levels of PWS were measured to examine the impact encumbrance and interaction had on walking speed. The mean baseline PWS measured was 4.9km/h and carrying the bags resulted in the same walking speed so encumbrance (without interaction) did not have a significant effect on typical everyday walking pace. Conversely, interaction (without carrying the objects) caused walking speed to reduce to 4.1km/h, a significant decrease of 16.3% when compared to the baseline PWS. Therefore, hypothesis *H7* can only be partially supported and is rejected. For comparison, PWS dropped by 19.1% when users were performing target selections in *User Study 2*. In Bergstrom-Lehtovirta *et al.*'s treadmill walking study [7], a decrease in PWS of 24% was reported when participants performed target selections on a touchscreen mobile phone. Schildbach and Rukzio [66] found that participants in their mobile study reduced walking speed by approximately 25% when selecting a range of targets on a touchscreen phone. At the end of experiment, the baseline walking speed (PWS) and walking speed while carrying the bags (PWS&E) were measured again to give some indication of tiredness caused by prolonged periods of interaction while encumbered. The same PWS and PWS&E of 4.9km/h and 4.1km/h were measured respectively which suggests the participants were not significantly fatigued during the study and that they could walk with the bags without any major issues.

To answer research question *Q2.1.1* (*How does the change of input posture affect tapping interactions when encumbered and walking?*), the user study presented in this chapter showed that tapping performance between three common input postures were evenly matched when users were walking and both hands were physically constrained. Encumbrance had a greater effect on target accuracy, target error and input speed than the type of input posture used.

The results showed that having an additional thumb for input did not significantly improve targeting performance when both hands were encumbered while users were walking. The advantage of the one-handed input posture when encumbered meant less physical stresses and tiredness is put on the non-interacting arm. However, the disadvantage of using the preferred thumb only for input is the difficulties of selecting targets accurately when they are placed further away from the thumb's optimal input area or require awkward thumb movements to reach whether the user is encumbered or not. Carrying bags only magnified the problem as shown in *User Study 3*. Further work is required to see if input techniques (such as AppLens & LaunchTile [38], ThumbSpace [37] and CornerSpace & BezelSpace [83]) that have shown to improve thumb-based input to select targets that are difficult to reach are still effective in walking and encumbered contexts.

Using the two-handed index finger posture meant there were no problems accessing the full area of the touchscreen but both hands are required for input despite not having the advantage (or at least the choice) of an extra digit for interaction, even though the results showed both thumb tapping gave minimal gains in targeting performance. Based on these findings, it is difficult to recommend a particular input posture that is more appropriate and effective to use in encumbered and walking situations since no one posture significantly outperformed the others in *User Study 3*. What the results did show is the extent targeting performance decreased when carrying cumbersome objects such as bags in both hands while on the move. Changing input posture did not considerably improve tapping accuracy, precision or speed which suggests more effort is required to design better user interfaces and interaction techniques to assist users in physically challenging encumbered and mobile situations.

5.5 Conclusions

To conclude, Chapter 5 answered research question *Q2.1.1* by presenting a user study that examined the effects of changing input posture on tapping performance when walking and encumbered. The study evaluated an encumbrance scenario where both hands were physically hampered by carrying a typical carrier bag in each hand. During the study, PWS was controlled which meant users could not trade walking speed with targeting performance and therefore gave a better representation of the extent input is affected by encumbrance while on the move. The three input postures covered both one- and two-handed input methods and examined situations where an additional thumb is available for selection of on-screen targets. In general, no particular input posture prevailed or noticeably improved tapping performance when users were walking and carrying bags in both hands but the negative effects of encumbrance and mobility are evident. Carrying the bags while walking caused tapping accuracy to decrease while selection time took longer when compared to unencumbered. The work discussed in this chapter makes a contribution towards understanding how tapping on touchscreen handheld devices in common input postures is affected when users are walking and encumbered. Users experienced targeting difficulties across all three commonly used input postures which illustrate the need to design interaction techniques that are effective in different modes and in a range of contexts.

User Study 2 and *User Study 3* have both extensively examined basic tapping performance in a range of encumbrance scenarios. However, touchscreen interfaces provide a range of different on-screen gesture inputs such as two-finger zooming and rotation actions that have become more common and sometimes a necessity to interact with applications on mobile devices. The next chapter examines the effects of encumbrance and mobility on a variety of touch-based gesture interactions.

Chapter 6

User Study 4: The Effects of Encumbrance and Walking on Touch-Based Gesture Interactions

6.1 Introduction

The user studies presented in the previous two chapters have examined abstract tapping performance in a range of different encumbrance scenarios while on the move. However, touchscreen interfaces provide other forms of on-screen gesture interactions from one-finger *dragging* (to pan across a photo) to two-finger *pinching* and *spreading* (zooming in and out of an image) and *rotating* (changing the orientation of a map). These touch-based gestures have become common and from time to time a necessity to interact with applications on mobile phones and other touch-enabled devices. While previous research has studied these on-screen gestures on larger tablet devices (e.g. [12,19,28,29]), more biomechanically complex two-finger actions such as *rotating* gestures are less well studied on smaller touchscreen mobile phones and furthermore, no study has investigated the effects of mobility and encumbrance on their performance.

Therefore, a set of Fitts' Law targeting tasks was designed to evaluate and compare the performance of four main on-screen gestures: *tapping*, *dragging*, *spreading & pinching* and *rotating (clockwise & anticlockwise)* while walking and carrying shopping bags. The results from *User Study 4* answer research questions *Q2.1 (How do encumbrance and mobility affect tapping performance?)* and *Q2.2* (*How do encumbrance and mobility affect the performance of other standard touch-based gestures?*) to fill this gap in the literature.

6.2 Background

A number of studies have examined the performance of *dragging*, *spreading*, *pinching* and *rotating* gesture inputs on touchscreen devices in unencumbered and non-mobile contexts. Tran *et al.* [71] developed a Fitts' Law style targeting task to examine two-finger *spreading* and *pinching* gestures on both a mobile phone (iPhone) and a tablet (iPad) while users were seated. The results from their experiment showed that both gestures took approximately one second to perform on the mobile phone. *Pinching* took longer to perform than *spreading* on both devices. Findlater *et al.* [19] also reported that *pinching* gestures were performed slower than *spreading* on a tablet across different age groups. In general, older users required more time to execute *pinching* and *spreading* actions than younger adults. Hoggan *et al.* [28] also investigated the performance of *spreading* and *pinching* gestures on a touchscreen tablet mimicking multi-touch tabletop interactions. In contrast to [19,71], they found that *pinching* was faster and ergonomically easier to execute than *spreading* actions. Likewise, Kobayashi *et al.* [41] reported that *pinching* gestures were executed quicker than *spreading* when performed by older adult users.

For *rotating* gestures on touchscreen mobile devices, the most related research is by Hoggan *et al.* [29], who examined two-finger rotational motions on a touchscreen tablet, which was placed flat to replicate a tabletop computer, while users were seated. An overall mean execution time of 2.7s for 90° rotations was reported. *Rotating clockwise* took longer to perform than *rotating anticlockwise*, although all users were right-handed so perhaps there was a bias in performance for a particular rotational direction. Hoggan *et al.* also focused on the ergonomics of *rotating* gestures. For example, at a starting position of 0° relative to the horizontal x-axis, rotational motions were faster than starting at 60° and 120°. Furthermore, execution time took longer and failure rate increased as the distance between both fingers increased. Despite these findings, it is difficult to replicate the performance and translate the recommended design guidelines from stationary on-table interactions to smaller handheld devices used in mobile and encumbered contexts, where the user's input posture is uncertain. Furthermore, tabletop interactions are less restricted to those on smaller mobile devices where screen space is much more limited and performing rotational actions is therefore potentially more difficult.

Findlater *et al.* [19] compared four different tasks using either finger input on a touchscreen tablet or mouse input with a desktop computer. The results showed that for young adults, *dragging* movements were performed similarly between the two types of input modalities as movement time took approximately one second. Older adults took around 1.5s to perform *dragging* gestures on the tablet, which was quicker than using the mouse to perform the same action on the computer. Kobayashi *et al.* [41] reported a higher average execution time of 2.17s to complete *dragging* movements for older users. Furthermore, Findlater *et al.* found that *dragging* gestures were subjectively easier to perform than *tapping*. Cockburn *et al.* [12] compared different types of input modalities to perform *tapping* and *dragging* actions on a touchscreen tablet computer. Error rates for touch-based *tapping* and *dragging* were 6.8% and 1% respectively. However, movement time for *tapping* (572ms) was quicker than *dragging* (922ms), which suggests a speed vs. accuracy trade-off between the two gesture inputs.

In summary, the performance of on-screen gestural interactions is less well understood on smaller handheld devices such as smartphones. With the exception of Tran *et al.* [71], the studies discussed in this section examined common types of touch-based gesture inputs on tablet-sized computers where the touchscreen is much larger than those found on mobile phones. With limited screen space, performing more complex multi-fingered gestures could be more difficult, especially in physically challenging situations such as walking and carrying objects. Furthermore, there is a lack of research examining the performance of *rotating* gestures irrespective of the user's context or device-in-use. Therefore, a set of targeting tasks was designed to test four standard touch-based gesture inputs and compared their performance in mobile and encumbered settings.

6.3 Method

6.3.1 Input Posture

The input posture used to perform the gestures in *User Study 4* needs explaining before describing the targeting tasks. A two-handed input posture was used where the mobile device was held in the non-dominant hand in portrait orientation while the dominant hand was used to perform the gestures. Users are likely to switch to one-handed input when encumbered but during pilot testing, it was observed that performing long vertical *dragging* motions or two-finger *spreading*, *pinching* and *rotating* gestures in one continuous motion was difficult with one hand only. In addition, larger mobile phones such as Samsung's Galaxy Note 4[10] and Apple's iPhone 6s Plus[11] (touchscreen sizes of 5.7″ and 5.5″ respectively) are becoming more popular so interaction is likely to require both hands. To avoid this limitation, a two-handed input posture was selected to test the effects of encumbrance and mobility on each gesture (see Figure 6.1).

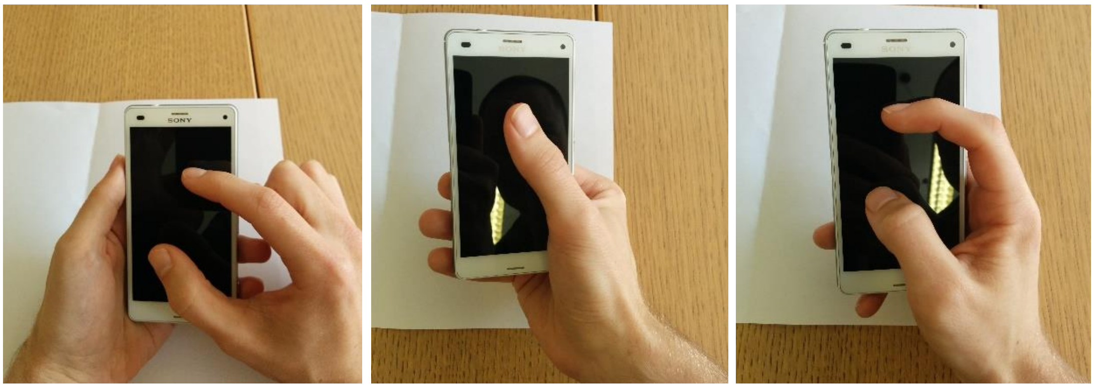

Figure 6.1: The two-handed input posture used to perform all types of gestures in *User Study 4* (left). One-handed input makes one- and two- finger gestures difficult to execute in one continuous movement on a 4.6″ touchscreen phone (middle and right).

6.3.2 Fitts' Law Targeting Tasks

This section begins with an explanation of Fitts' Law, its use in human computer interaction and how previous work has used it to examine and model the performance of touch-based gestures. Then, the section describes a set of Fitts' Law targeting tasks designed to measure the performance of *tapping*, *dragging*, *spreading & pinching* and *rotating clockwise & anticlockwise* on a touchscreen mobile phone while users were 1) walking and 2) walking and encumbered. A two-dimensional (2-D) Fitts' Law task was used to examine *tapping* and *dragging* while one-dimensional (1-D) Fitts' Law tasks were designed to measure *spreading & pinching* and *rotating* gestures. The targeting tasks described below ran on a Samsung Galaxy S3 phone, which has a 4.8″ touchscreen (~65.9% screen-to-body ratio[12]) with a resolution of 720 *x* 1280px (12px/mm). Each trial was completed as quickly and accurately possible. No feedback was given to indicate a correct target selection in any of the tasks to avoid influencing input behaviour [12].

[10] http://www.samsung.com/uk/consumer/mobile-devices/smartphones/galaxy-note/SM-N910FZKEBTU
[11] http://www.apple.com/uk/iphone-6s/
[12] http://www.gsmarena.com/samsung_i9300_galaxy_s_iii-4238.php

6.3.2.1 Using Fitts' Law to Evaluate Touch-based Gestures

Originally, Fitts' Law [20] characterises the performance of a one-dimensional pointing task. Participants move rapidly between two targets where target width and the distance between the targets are controlled. Fitts showed that movement time had a strong correlation with the logarithm of target distance to target width ratio. Fitts' Law has the form:

$$MT = a + b * ID, ID = \log 2 (A/W + 1)$$

Where MT is the Movement Time, a and b are constants determined by linear regression, ID is the Index of Difficulty and, in the Shannon formulation [48] used here, A is the target distance and W is the target width. Since high error rates are predicted when walking and encumbered, and the formula stated above assumes an error rate of 4%, the effective target width (W_e) is used instead. Normally, for 1-D targeting tasks, W_e is calculated as: $SD * 4.133$ [48], where SD is the standard deviation of the endpoint errors in one direction (i.e. the univariate endpoint deviation SD_x). More recently, Wobbrock *et al.* [78] showed that this method of calculating W_e for 2-D targeting tasks is less appropriate and the spread of hits around their centre of mass should be used (i.e. the bivariate endpoint deviation $SD_{x,y}$, see [78] for more details). Therefore, Wobbrock *et al.*'s method was used to calculate W_e for the 2-D tasks in the study. After W_e was adjusted, the effective Index of Difficulty (ID_e) was calculated using $log2 (A/W_e + 1)$.

The Throughput (TP) [46] of each type of gesture was calculated as: ID_{mean}/MT_{mean}, where ID_{mean} is the Index of Difficulty (calculated from the mean ID using W_e) and MT_{mean} is the mean Movement Time. Throughput (in bits/sec) is normally used to compare the performance of different pointing devices but was calculated to show the bandwidth of the communication channel of each touch-based gesture.

The use of Fitts' Law in HCI is not limited to pointing and has been used for other touch-based gesture types. Mackenzie *et al.* [47] used Fitts' Law to evaluate different input devices for *tapping* and *dragging* tasks. Tran *et al.* [71] modelled the performance of *pinching* and *spreading* gestures on mobile devices using a Fitts' Law style targeting task. Zhao *et al.* [84] combined Fitts' Law with the Mahalanobis distance metric [64] to evaluate translocation, rotation and scaling movements on a multi-touch tabletop. Since the gesture types were combined, Zhao *et al.*'s method makes it difficult to discretely measure the performance of *rotating* gestures. The approach used in *User Study 4* allowed two-finger rotational actions to be examined separately from the other gesture forms.

6.3.2.2 Tapping and Dragging

A 2-D targeting task was designed to measure the performance of *tapping* and *dragging*. Two targets (denoted as start and destination) were presented on the screen. The start target was represented by a crosshair and had a diameter of 2.5mm. The size of the start target stayed constant for each trial and the dimensions were chosen after pilot tests. The destination target was shown as a green circle and varied in diameter and distance from the start target depending on the experimental condition. This implementation was chosen to avoid confusion and decrease the chance of participants selecting the wrong initial target. For *tapping*, the index finger selected the crosshair first and then the destination target. As a result, two taps were required to complete each trial. Movement time for each *tapping* trial was the duration from the touch up event of the first tap to the touch up event of the second tap. For *dragging*, the same task was used but instead of two taps, the index finger selected the crosshair first and dragged towards the destination target. Therefore, each *dragging* trial was completed in one action. Movement time for each *dragging* trial was the duration from the touch down to the touch up events.

There were 3 target widths (5.0, 7.5 and 10.0mm), 4 target distances (24, 36, 48 and 96mm), *ID* ranged from 1.8 to 4.3 bits and 8 directions (N-S, NE-SW, E-W, SE-NW, S-N, SW-NE, W-E and NW-SE). There were 90 unique trials instead of 96 because the E-W and W-E directions are not possible for the largest distance of 96mm due to the width of the touchscreen. The distances were selected to cover a wide area of the screen. The target widths represented a range of differently sized icons or buttons that users would typically press on touchscreen mobile phones. Figure 6.2 illustrates the targeting task used for *tapping* and *dragging*.

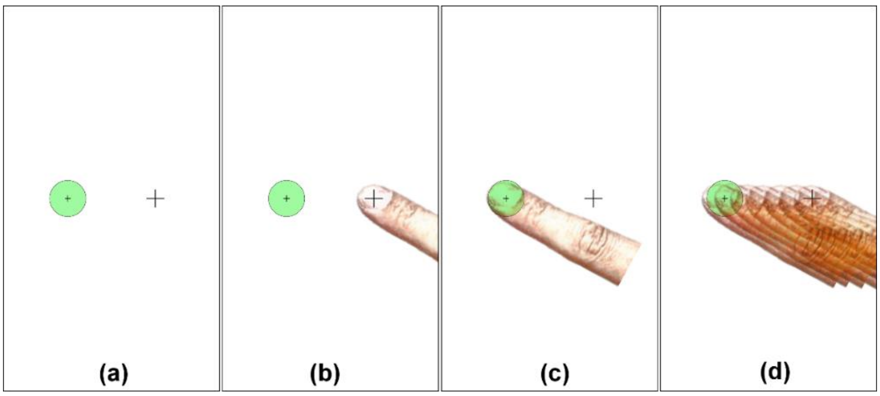

Figure 6.2: The same task used for *tapping* and *dragging* (a). For *tapping*, the first tap selects the crosshair (b) and the second tap selects the destination target (c). For *dragging*, the initial crosshair is selected and the finger dragged towards the destination target (d). The images are not drawn to scale.

6.3.2.3 Spreading and Pinching

To examine two-finger *spreading* and *pinching* gestures, a similar method to Tran *et al.* [71] was developed. Circular targets were used instead of squares for better visual presentation (Findlater *et al.* [19] also used circles). For *spreading*, a circle (denoted as the control) was initially presented at the centre of the screen to show the current trial was ready (yellow). Once the index finger and the thumb (only these digits were used for input) of the dominant hand were placed on the touchscreen, the control circle turned green to show that the trial could begin. Participants were then instructed to move the control circle towards the target (grey ring) by expanding the distance between the digits. Like *Tran et al.* [71], a 1-1 mapping was used to transform the change in distance between the digits to the change in the size of the control circle. A trial ended once either digit was lifted off the screen. There were no fixed start points defined for each digit, but participants were asked to avoid placing their digits too close together at the start to prevent occlusion of the control circle. The task for *pinching* operated in the same way as *spreading*, but the control circle was now initially bigger than the target. Like *spreading*, there were no fixed starting positions but the participants were instructed to touch the outer white area of the target at the start to ensure that there was enough space between the digits to perform the gesture.

Both *spreading* and *pinching* had to be completed in one single action so a gesture was deemed completed when either digit lifted off the touchscreen. In accordance to Fitts' Law tasks, no correction was allowed if the control circle overshot the target area. Movement time for both *spreading* and *pinching* was calculated from the touch down event of the second digit to the first touch up event of either digit. The same target widths and gesture distances as Tran *et al.* [71] were used, thus, there were 3 target widths (1.6, 3.2 and 4.8mm), 3 gesture distances (8, 16 and 24mm) and *ID* ranged from 1.4 to 4.0 bits. Target distance (*A*) in this case was from the edge of the control circle to the centre of the target ring. Figure 6.3 illustrates the tasks used for both *spreading* and *pinching*.

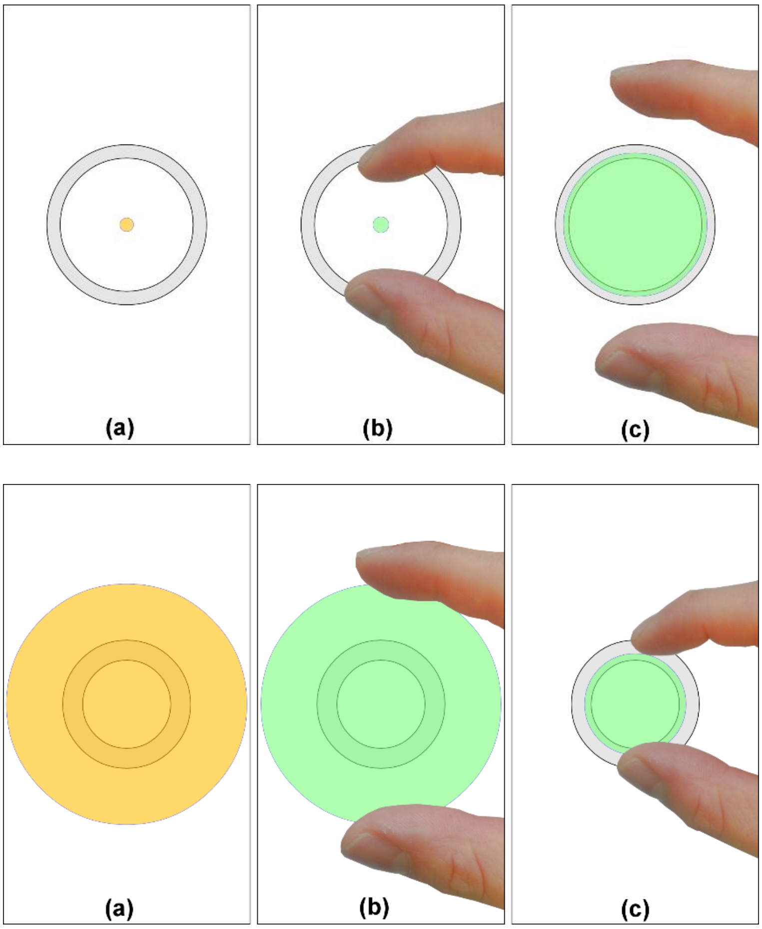

Figure 6.3: The tasks (a) used for *spreading* (top) and *pinching* (bottom). Once both digits were detected, the control circle turned green and the trial began (b). A successful selection when the control circle is within the target area (c). The images are not drawn to scale.

6.3.2.4 Rotating

Figure 6.4 explains the method used to measure two-finger *rotating* gestures. The touchscreen was split into two sections. The upper part was used for visual feedback which consisted of an arc segment (limited to 110°) to show: (1) the amount of angle currently rotated between the fingers and (2) the number of degrees required to reach the green destination target. The lower light blue area was used to perform the rotational gestures (see Figure 6.4a). There were two start points (circle & crosshair) for the index finger (upper left) and the thumb (lower right) for *rotating* clockwise. The index finger and thumb were placed upper right and lower left positions respectively for *rotating* anticlockwise. The distance between the start points was fixed at 42mm for each trial since during pilot testing, this was found to be a comfortable input posture and there was enough room to execute the gesture regardless of the size of the user's hands and fingers. Once both digits had been detected, a red line between the two touch positions appeared to show the task could begin (see Figure 6.4b).

The participants performed the rotations toward the green target area (shown in the feedback arc). The progression of angle shown in the arc directly corresponded to the on-screen movement, but visual feedback was shifted upwards to avoid occluding the position of target so that participants always knew where to rotate to. Continuous feedback (in transparent red) was given in the arc segment as the fingers executed the rotational gesture (see Figure 6.4c). A 1-1 mapping was used to translate the amount of angle rotated to the progression of the feedback. For both rotational directions, the initial touch down co-ordinates of each digit had to be within their starting positions to avoid physical stresses on the user's fingers when performing the largest rotational distances. If this did not occur, the gesture area turned red and both digits had to lift off the screen and accurately reselect the starting points again.

Four main design features were carefully considered for the rotational targeting task. Firstly, the key objective was to examine rotations independently, therefore rotational angle was measured. Participants could vary the gap between the digits (when rotating) without affecting the angle, as occurs in standard touchscreen rotations (and was done by Hoggan *et al.* [29]). Secondly, the maximum rotational angle in both directions was constrained to 110° from the starting angle (the difference in angle from the horizontal x-axis) that was calculated between the index finger and the thumb when both digits initially touch the screen. During pilot testing, this range was the physical limit before clutching was required. Thirdly, all rotating gestures were performed in one single action, which is in accordance with a Fitts' Law style targeting task. And fourthly, both digits had to move in the rotational direction required to complete the gesture.

Each trial ended once either digit lifted off the screen. Movement time was calculated from the touch down event of the second digit to the first touch up event of either digit. Target width (W) was defined as the green target area in the feedback arc while target distance (A) was the amount of rotational angle required from the initial angle (calculated from the start touch down positions) to the centre of the target. Both target distance and target width were measured in degrees (see Figure 6.5). There were 3 target widths (6°, 12° and 18°), 3 rotational distances (30°, 60° and 90°) and ID ranged from 1.4 – 4.0 bits. The target widths and distances were selected to cover a range of rotational precision that users may encounter with touchscreen mobile phones (e.g. rotating maps or turning virtual dials).

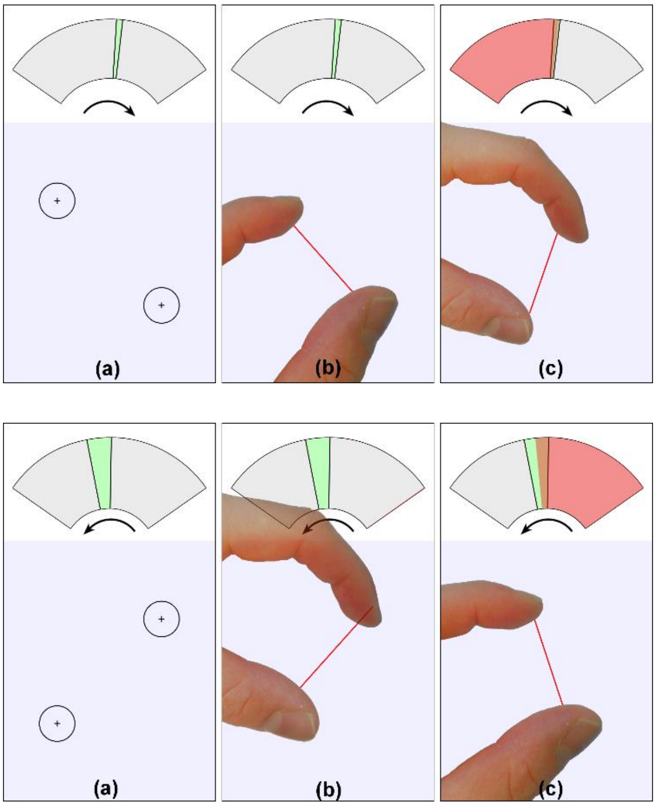

Figure 6.4: The task (a) used for *rotating* clockwise (top) and anticlockwise (bottom). Once both digits have been detected (b), the fingers rotated towards the green target (c). The images are not drawn to scale.

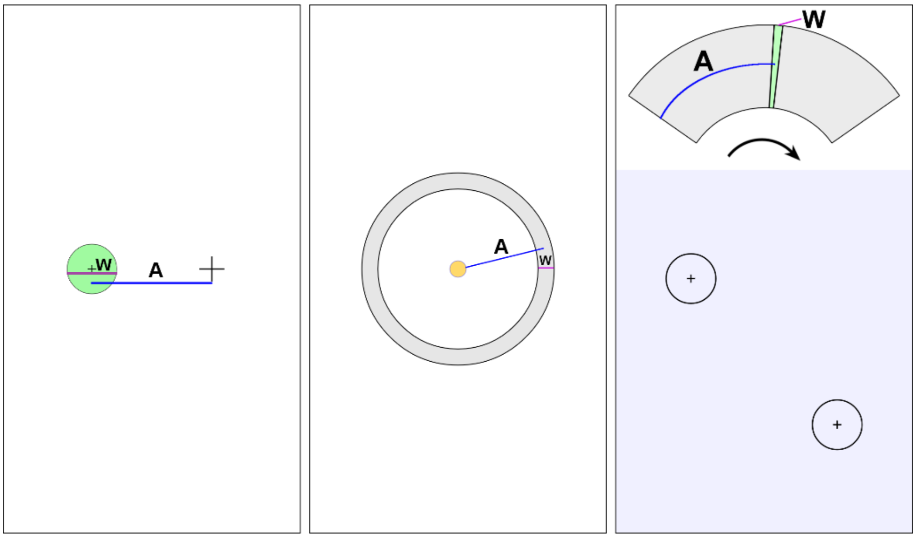

Figure 6.5: The definition of distance (A) and target width (W) for the Fitts' targeting tasks used to measure: *tapping* & *dragging* (left), *spreading* & *pinching* (middle) and *rotating* (right). The images are not drawn to scale.

6.3.3 Encumbrance Scenario

Encumbrance Scenarios 1C was evaluated in *User Study 4*, so a shopping bag was held in each hand to examine the effects of encumbering both hands while performing the different types of on-screen gestures. Each bag measured approximately 330 *x* 480mm and weighed 1.6kg. Figure 6.6 shows the bags selected for the study and the way the objects were held during two-handed input.

Figure 6.6: The encumbrance scenario (a bag held in each hand) evaluated in *User Study 4*.

6.3.4 Walking Environment

The same mobile evaluation methodology as the one applied in *User Study 3* was used. Therefore, a 20m long and 1.5m wide oval path was created in a spacious and quiet room. For each experimental condition, the participants maintained their PWS&I (Preferred Walking Speed during Interaction) by walking side-by-side with a pacesetter. Each individual PWS&I was measured before the main experiment began as the participants walked around the route while completing the abstract targeting task on a touchscreen mobile phone (as described in Section 5.2.2). The pacesetter walked at the calculated PWS&I for each participant by using a metronome application that ran on a mobile phone. Participants were instructed to avoid drifting out of the boundaries of the path during the experiment and were also told to keep up or slow down if they failed to walk in-step with the pacesetter.

As explained previously, controlling the participants' walking speed meant any changes observed on input performance are more likely to be caused by encumbrance since the two physical effects were isolated. In the real world, users are likely to walk slower when interacting with mobile devices and encumbered (as shown in Chapter 3) but controlling walking speed meant a fairer comparison could be made between unencumbered and encumbered input while on the move.

6.3.5 Experimental Design

Twenty students (15 males, 5 females) aged between 21 - 38 years (mean = 26.15, SD = 4.09) recruited from the University took part in the study. Three of the male participants preferred using their left hand for input while all female participants were right-handed. All participants used a touchscreen mobile phone or device on a daily basis. The experiment took around 90 minutes to complete and each participant was paid £8 for taking part. To reduce fatigue, sufficient resting periods were given between conditions and as required by the participants. A training phase for each type of gesture was given at the start of the experiment to familiarise with the different tasks.

There were six gesture conditions: *tapping, dragging, spreading, pinching, rotating clockwise and rotating anti-clockwise*. For *tapping* and *dragging*, each block of trials consisted of the 90 target width/distance combinations. For *spreading* and *pinching*, each block of trials had 45 target selections since each of the nine unique target width/distance combinations was repeated five times. Likewise, there were 45 target selections for each block of *rotating clockwise* and *anticlockwise* gestures as each of the nine unique target width/distance combinations was repeated five times. The order of the trials within each block was randomised. A random delay between 500 - 1500ms was placed between each trial to reduce the chance of any rhythm forming between the user's walking and input behaviour [15]. There were two blocks of trials per condition therefore each participant completed 12 blocks (720 trials) during the experiment.

Each gesture type was completed while walking and either unencumbered or holding a bag in both hands, which resulted in a total of 12 conditions. The conditions were counterbalanced by type of encumbrance and the order of the gestures was further randomised to reduce learning and order effects as much as possible. The Independent Variables were **Type of Gesture** (6 levels), **Type of Encumbrance** (2 levels) and **Target Width-Distance** (either 9 or 12 levels depending on the task). The Dependent Variables were: target accuracy (%) and movement time (milliseconds). A target was selected accurately if the final position was within the target borders. The main hypotheses were:

H1: For each type of gesture, accuracy will be significantly lower when carrying the bags than unencumbered;

H2: For each type of gesture, movement time will require significantly longer when carrying the bags than unencumbered;

H3: *Dragging* will have significantly higher accuracy than *tapping* but significantly slower movement time (due to the findings of Cockburn *et al.* [12]);

H4: *Pinching* will be performed significantly faster than *spreading* (based on Hoggan *et al.*'s results [28]);

H5: *Rotating anticlockwise* will be performed significantly faster than *rotating clockwise* (due to the findings of Hoggan *et al.* [29]).

6.4 Results

A total of 28,800 trials were recorded for the entire experiment. Potential outliers were removed by following the method described by Mackenzie and Isokoski [46] if 1) the measured movement was less than half of the distance to the target (*A*) (only applicable for *tapping* and *dragging*) or 2) the endpoint error was greater than two target widths (*2W*) from the centre of the current target (used for all types of gestures). As a result, 494 trials (1.7%) were deemed as outliers and were removed from the final data analysis. The overall mean target accuracy and movement time for each condition are shown in Figure 6.7 and Figure 6.8 respectively.

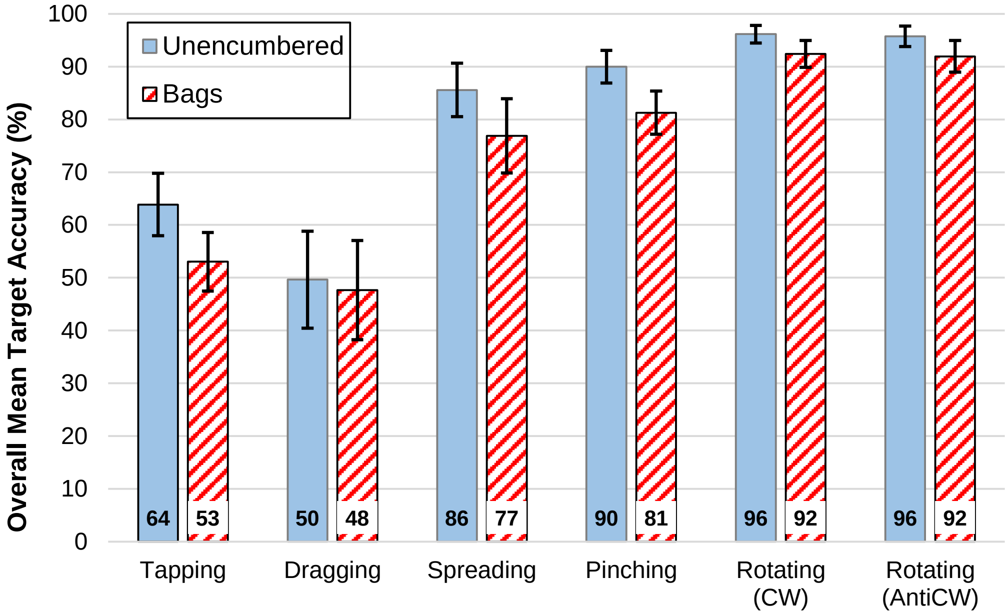

Figure 6.7: The overall mean target accuracy (%) for each condition in *User Study 4*. The solid blue and striped red bars represent the unencumbered and carrying the bags conditions respectively. Error bars denote Confidence Interval (95%).

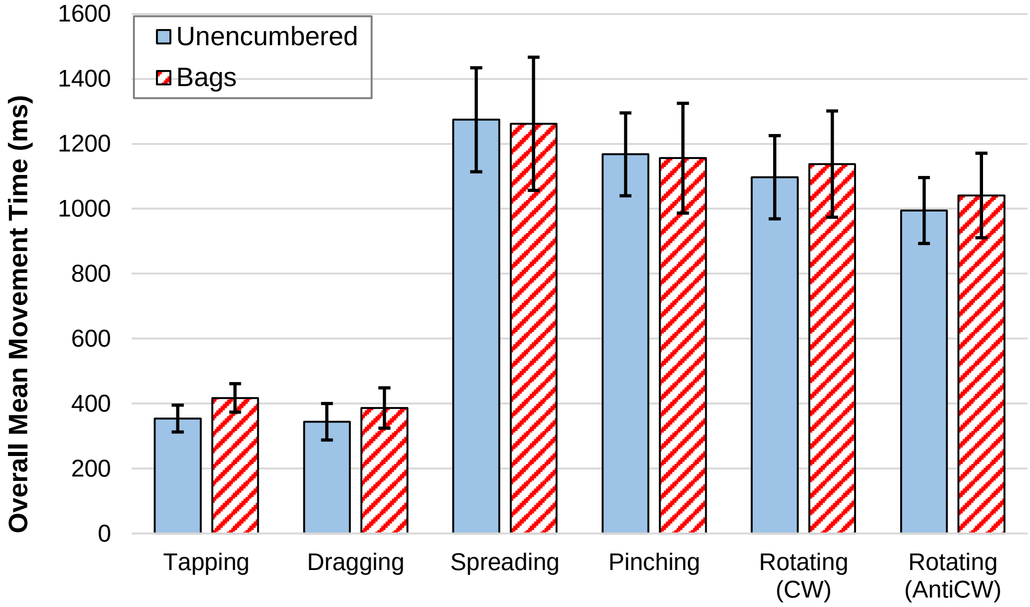

Figure 6.8: The overall mean movement time (mm) for each condition in *User Study 4*. The solid blue and striped red bars represented the unencumbered and carrying the bags conditions respectively. Error bars denote Confidence Interval (95%).

6.4.1 Target Accuracy (Tapping and Dragging)

Shapiro-Wilk tests were conducted to assess the distribution of the data recorded for *tapping* and *dragging* target accuracy. The results are shown in Table 6.1 and all tests were not significant. A three-factor (**Type of Gesture**, **Type of Encumbrance** and **Target Width-Distance**) repeated-measures ANOVA was conducted for target accuracy to compare *tapping* and *dragging*. Greenhouse-Geisser adjustments were used to correct the degrees of freedom since Mauchly's test for sphericity was significant ($\varepsilon < 0.75$). The mean target accuracy for each target width-distance combination when *tapping* and *dragging* is shown in Figure 6.9 and Figure 6.10 respectively.

Type of gesture	Type of encumbrance	W Statistic	Sig.
Tapping	No object	0.94428	0.288472
Tapping	Bags	0.97031	0.761403
Dragging	No object	0.97617	0.875745
Dragging	Bags	0.95984	0.54066

Table 6.1: Shapiro-Wilk normality tests performed on target accuracy for the *tapping* and *dragging* conditions in *User Study 4*.

The ANOVA conducted to compare *tapping* and *dragging* accuracy showed a significant main effect for **Type of Gesture**, $F(1, 19) = 6.03$, $p < 0.05$. Accuracy was significantly higher for *tapping* than *dragging* by 9.8%. There was a significant main effect for **Type of Encumbrance**, $F(1, 19) = 17.69$, $p < 0.01$. The number of correct target selections was significantly higher when unencumbered than holding the bags, a difference of 6.4%. A significant main effect was also observed for **Target Width-Distance**, $F(11, 209) = 105.74$, $p < 0.01$. *Post hoc* pairwise comparisons with Bonferroni corrections showed that for each target distance, increasing target width significantly increased target accuracy.

The interaction between **Type of Gesture** and **Type of Encumbrance** was significant, $F(1, 19) = 14.37$, $p < 0.01$. Encumbrance caused a greater decrease in accuracy when *tapping* than *dragging*. Target accuracy was low for *dragging* whether participants were encumbered or not. The interaction between **Type of Gesture** and **Target Width-Distance** combination was significant, $F(5.2, 98.1) = 11.94$, $p < 0.01$. Accuracy for all target widths at the largest distance 96mm was higher when *dragging* than *tapping*. The accuracy for all other target combinations was higher for *tapping* than *dragging*. The interaction between **Type of Encumbrance** and **Target Width-Distance** was not significant, $F(6.6, 125.8) = 0.99$ $p > 0.05$. The interaction between all three factors for accuracy was also not significant, $F(6.3, 120.2) = 1.17$ $p > 0.05$. The results for the interactions between factors are not relevant to supporting or rejecting the hypotheses stated in the Section 6.3.5.

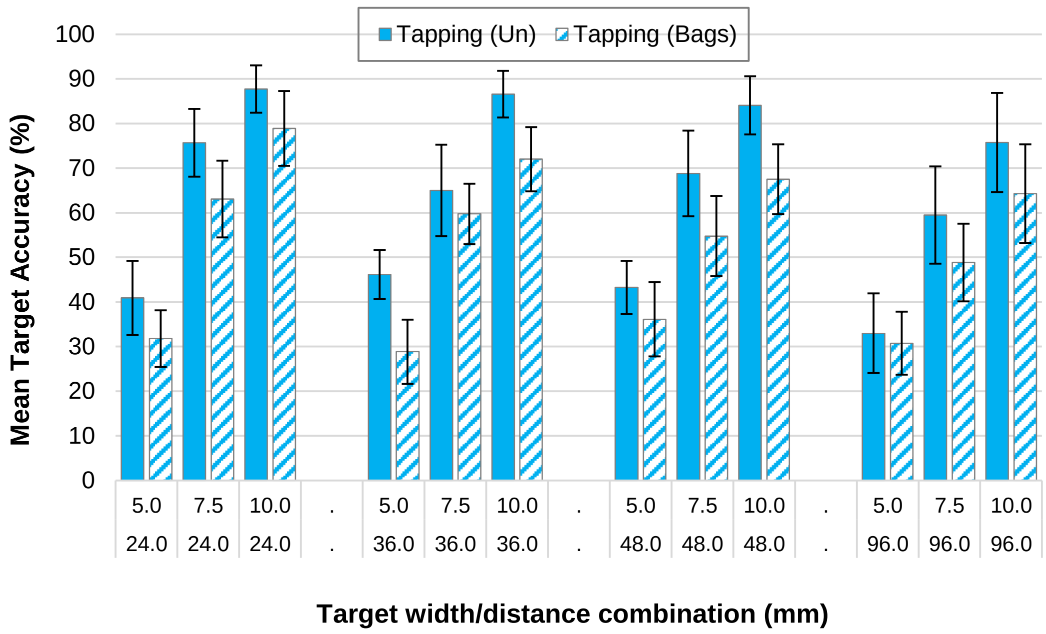

Figure 6.9: The mean target accuracy (%) for each target width-distance combination for *tapping*. The solid and striped bars represent the unencumbered (Un) and carrying the bags (Bags) conditions respectively. Error bars denote Confidence Interval (95%).

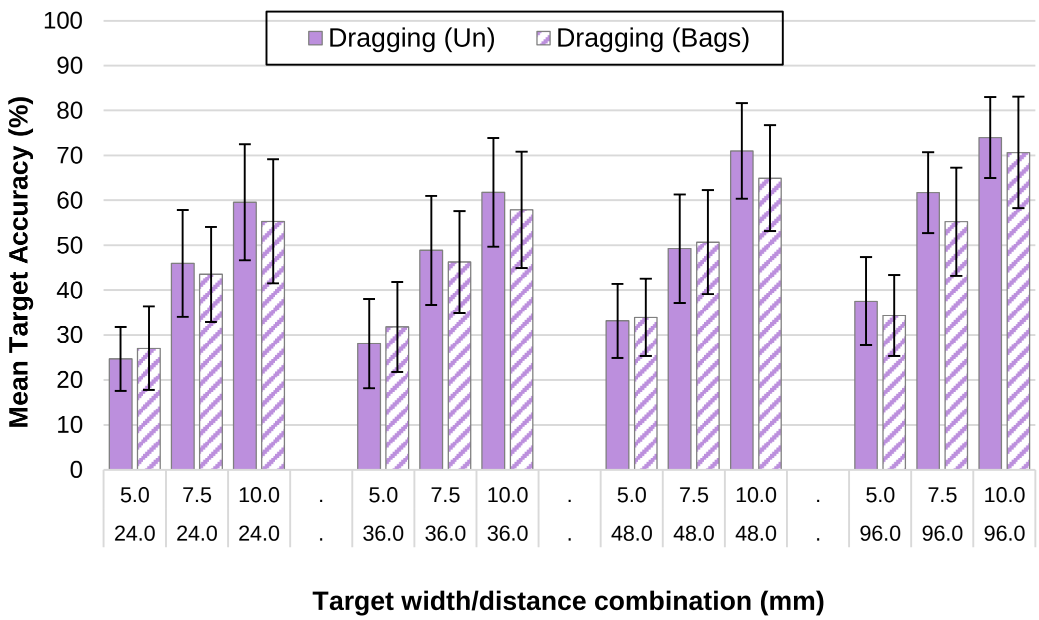

Figure 6.10: The mean target accuracy (%) for each target width-distance combination for *dragging*. The solid and striped bars represent the unencumbered (Un) and carrying the bags (Bags) conditions respectively. Error bars denote Confidence Interval (95%).

6.4.2 Target Accuracy (Spreading and Pinching)

Shapiro-Wilk tests were performed to examine the normality of the data recorded for *spreading* and *pinching* target accuracy. The results are shown in Table 6.2 and all tests were not significant. A three-factor (**Type of Gesture**, **Type of Encumbrance** and **Target Width-Distance**) repeated-measures ANOVA was conducted for target accuracy to compare *spreading* and *pinching*. Greenhouse-Geisser adjustments were used to correct the degrees of freedom since Mauchly's test for sphericity was significant ($\varepsilon < 0.75$). The mean target accuracy for each target width-distance combination when *spreading* and *pinching* is shown in Figure 6.11 and Figure 6.12 respectively.

Type of gesture	Type of encumbrance	W Statistic	Sig.
Spreading	No object	0.935636	0.198081
Spreading	Bags	0.910286	0.06454
Pinching	No object	0.925181	0.12469
Pinching	Bags	0.910085	0.063973

Table 6.2: Shapiro-Wilk normality tests performed on target accuracy for the *spreading* and *pinching* conditions in *User Study 4*.

The ANOVA for *spreading* and *pinching* accuracy showed no significant main effect for **Type of Gesture**, $F(1, 19) = 4.21$, $p > 0.05$. There was a significant main effect for **Type of Encumbrance**, $F(1, 19) = 50.59$, $p < 0.01$. Target accuracy was significantly higher when unencumbered than holding the bags, a mean difference of 8.7%. A significant main effect for **Target Width-Distance** was also observed, $F(3.9, 74.5) = 62.17$, $p < 0.01$. *Post hoc* pairwise comparisons with Bonferroni corrections showed that accuracy significantly increased as target width increased at each gesture distance.

No significant effect was observed for the interaction between **Type of Gesture** and **Type of Encumbrance**, $F(1, 19) = 0.996$, $p > 0.05$. The interaction between **Type of Gesture** and **Target Width-Distance** was significant, $F(3.3, 63.4) = 4.2$, $p < 0.01$. Accuracy was lower for the 1.6mm and 3.2mm target widths at the largest gesture distance of 24mm when *pinching* than *spreading*. The accuracy for all other target combinations was higher when *pinching* than *spreading*. The interaction between **Type of Encumbrance** and **Target Width-Distance** was not significant, $F(4.09, 77.69) = 0.71$, $p > 0.05$. No significant effect was found between all three factors for accuracy, $F(4.07, 77.29) = 0.40$, $p > 0.05$. The interactions between factors are not required to support or disprove the experiment hypotheses.

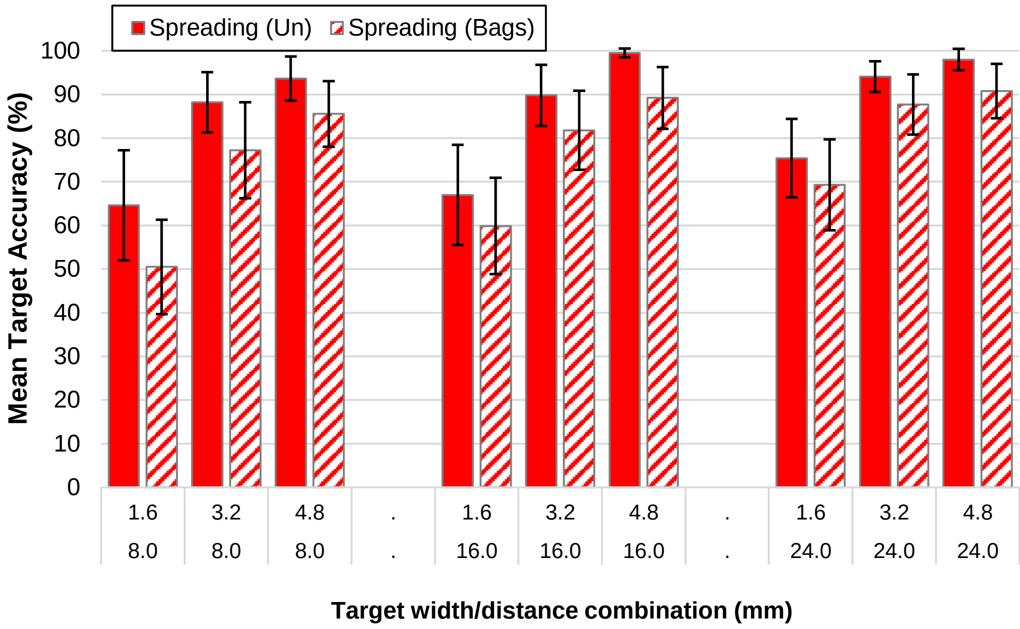

Figure 6.11: The mean target accuracy (%) for each target width-distance combination for *spreading*. The solid and striped bars represent the unencumbered (Un) and carrying the bags (Bags) conditions respectively. Error bars denote Confidence Interval (95%).

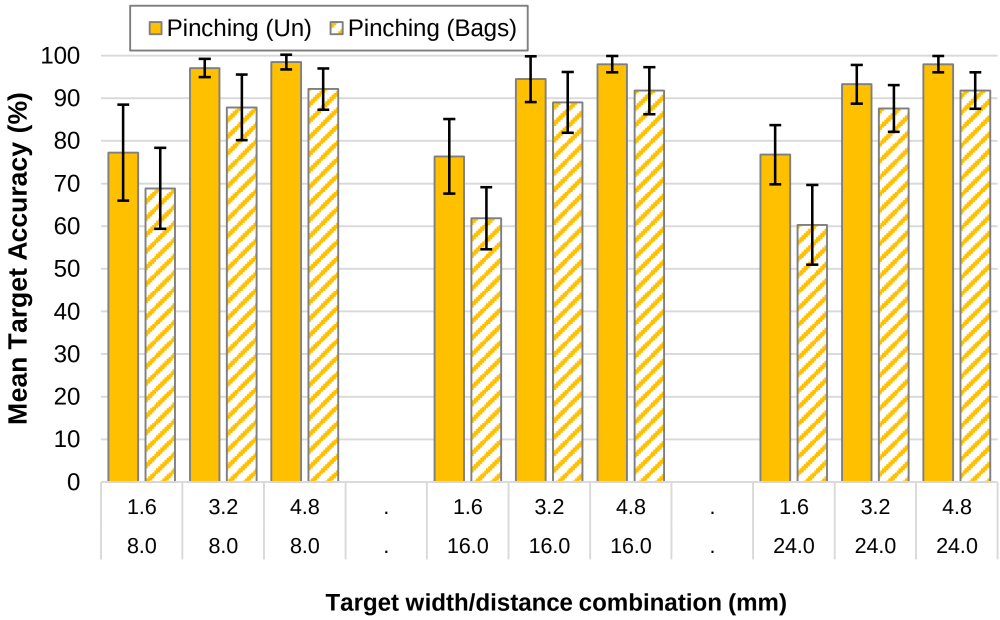

Figure 6.12: The mean target accuracy (%) for each target width-distance combination for *pinching*. The solid and striped bars represent the unencumbered (Un) and carrying the bags (Bags) conditions respectively. Error bars denote Confidence Interval (95%).

6.4.3 Target Accuracy (Rotating Clockwise and Rotating Anticlockwise)

Table 6.3 shows the results from conducting Shapiro-Wilk tests to examine the normality of target accuracy for *rotating clockwise* and *rotating anticlockwise*. Three out of four tests were significant and therefore deviates from a normal distribution. Thus, ART [76] was used to transform the data prior to conducting a three factor (**Type of Gesture**, **Type of Encumbrance** and **Target Width-Distance**) repeated-measures ANOVA to analysis target accuracy between *rotating clockwise* and *rotating anticlockwise*. The mean target accuracy for each target width-distance combination when *rotating clockwise* and *anticlockwise* is shown in Figure 6.13 and Figure 6.14 respectively.

Type of gesture	Type of encumbrance	W Statistic	Sig.
Rotating CW	No object	0.860874	0.008153*
Rotating CW	Bags	0.933184	0.177781
Rotating AntiCW	No object	0.803017	0.000959*
Rotating AntiCW	Bags	0.874734	0.014245*

Table 6.3: Shapiro-Wilk normality tests performed on target accuracy for the *rotating clockwise* (CW) and *rotating anticlockwise* (AntiCW) conditions in *User Study 4*. Significant results are shaded in grey and highlighted with '*'.

The results showed no significant main effect for **Type of Gesture**, $F(1, 665) = 0.329$, $p > 0.05$. There was a significant main effect for **Type of Encumbrance**, $F(1, 665) = 84.77$, $p < 0.01$. Accuracy was significantly higher when unencumbered than holding the bags. A significant main effect was also found for **Target Width-Distance**, $F(8, 665) = 39.3$, $p < 0.01$. *Post hoc* pairwise comparisons showed that accuracy increased as target width also increased at each rotational distance. The interaction between **Type of Gesture** and **Type of Encumbrance** was significant, $F(1, 665) = 5.389$, $p < 0.05$. *Post hoc* Tukey HSD tests showed all pairwise comparisons were not significant. A significant interaction was observed between **Type of Gesture** and **Target Width-Distance**, $F(8, 665) = 6.219$, $p < 0.05$. The interaction between **Type of Encumbrance** and **Target Width-Distance** was significant, $F(8, 655) = 12.728$, $p < 0.01$. The interaction between all three factors was significant, $F(8, 665) = 3.004$, $p < 0.01$. Due to the large number of comparisons and that the interaction effects between the factors are not relevant to the experiment hypotheses, further analyses were not conducted.

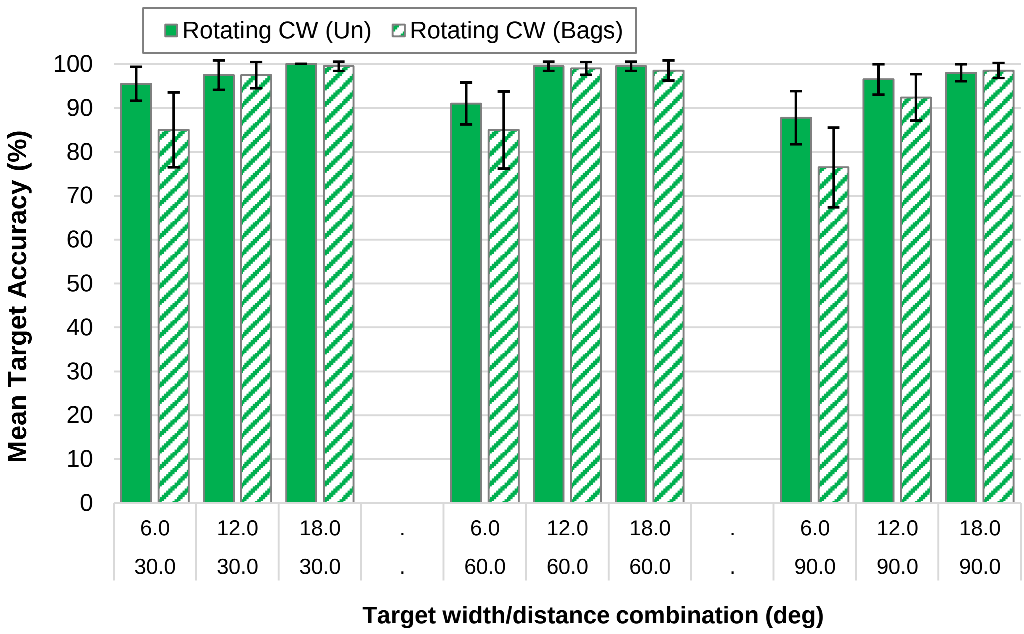

Figure 6.13: The mean target accuracy (%) for each target width-distance combination for *rotating clockwise* (CW). The solid and striped bars represent the unencumbered (Un) and carrying the bags (Bags) conditions respectively. Error bars denote Confidence Interval (95%).

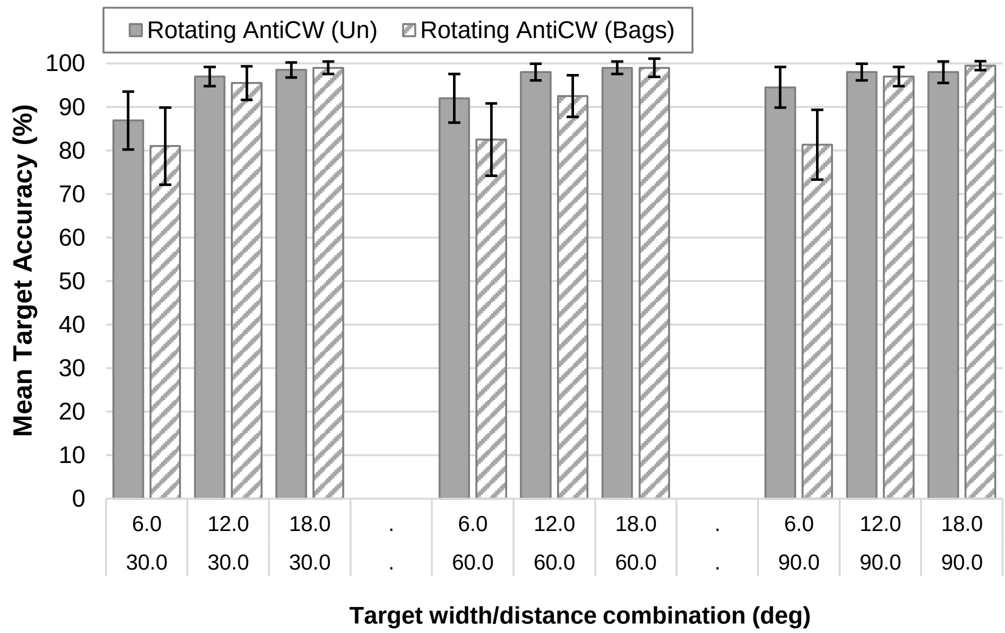

Figure 6.14: The mean target accuracy (%) for each target width-distance combination for *rotating anticlockwise* (AntiCW). The solid and striped bars represent the unencumbered (Un) and carrying the bags (Bags) conditions respectively. Error bars denote Confidence Interval (95%).

6.4.4 Movement Time (Tapping and Dragging)

Shapiro-Wilk tests were conducted to assess the normality of movement time for *tapping* and *dragging*. The results are shown in Table 6.4 and all tests were not significant. A three-factor (**Type of Gesture**, **Type of Encumbrance** and **Target Width-Distance**) repeated-measures ANOVA was conducted for movement time to compare *tapping* and *dragging*. Greenhouse-Geisser adjustments were used to correct the degrees of freedom since Mauchly's test for sphericity was significant ($\varepsilon < 0.75$). The mean movement time for each target width/distance combination for *tapping* and *dragging* is shown in Figure 6.15 and Figure 6.16 respectively.

Type of gesture	Type of encumbrance	W Statistic	Sig.
Tapping	No object	0.981033	0.946732
Tapping	Bags	0.938435	0.223959
Dragging	No object	0.98049	0.940269
Dragging	Bags	0.943179	0.275156

Table 6.4: Shapiro-Wilk normality tests performed on movement time for the *tapping* and *dragging* conditions in *User Study 4*.

The ANOVA conducted to analyse movement time between *tapping* and *dragging* showed no significant main effect for **Type of Gesture**, $F(1, 19) = 0.78$, $p > 0.05$. A significant main effect was found for **Type of Encumbrance**, $F(1, 19) = 12.95$, $p < 0.01$. Movement time took significantly longer when holding the bags than unencumbered (a mean difference of 52.8ms). A significant main effect was also observed for **Target Width-Distance**, $F(1.6, 30.8) = 219.84$, $p < 0.01$. *Post hoc* pairwise comparisons with Bonferroni corrections showed that increasing target width did not have a significant effect on movement time at each target distance. However, increasing target distance significantly increased movement time for each target width.

The interaction between **Type of Gesture** and **Type of Encumbrance** was not significant, $F(1, 19) = 1.52$, $p > 0.05$. The interaction between **Type of Gesture** and **Target Width-Distance** was significant, $F(2.7, 51.7) = 7.19$, $p < 0.01$. Movement time for all target widths at the greatest distance of 96mm took longer when *dragging* than *tapping*. However, movement time for the other nine target combinations was faster when *dragging* than *tapping*. The interaction between **Type of Encumbrance** and **Target Width-Distance** was significant, $F(3.0, 57.6) = 5.59$, $p < 0.01$. Carrying the bags caused longer movement time for all target combinations than unencumbered. The interaction between all three factors for movement time was not significant, $F(4.87, 92.44) = 0.60$, $p > 0.05$. The interaction effects between the factors are not required to support or reject the hypotheses on movement time.

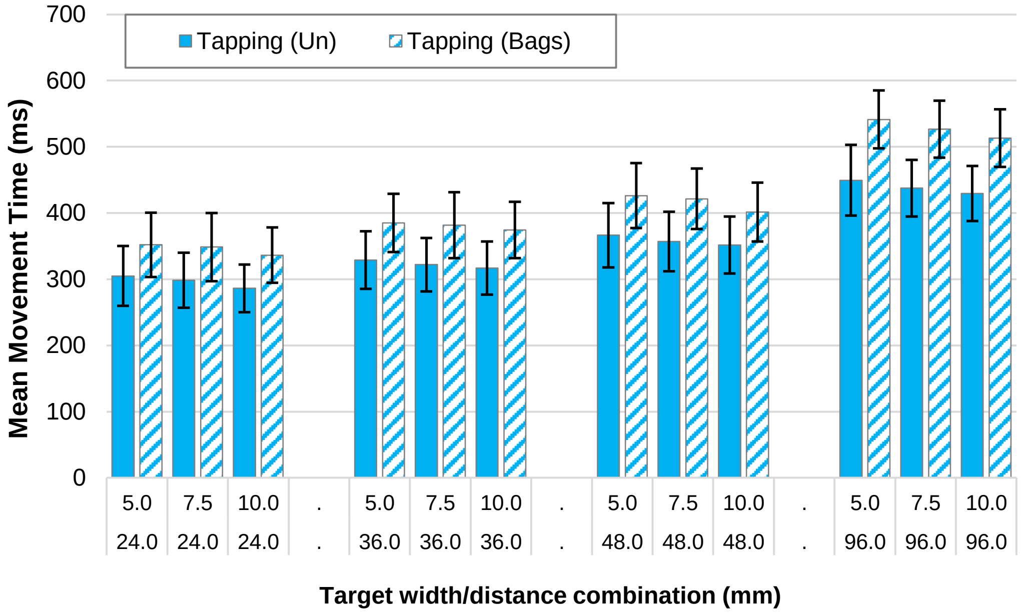

Figure 6.15: The mean movement time (ms) for each target width-distance combination for *tapping*. The solid and striped bars represent the unencumbered (Un) and carrying the bags (Bags) conditions respectively. Error bars denote Confidence Interval (95%).

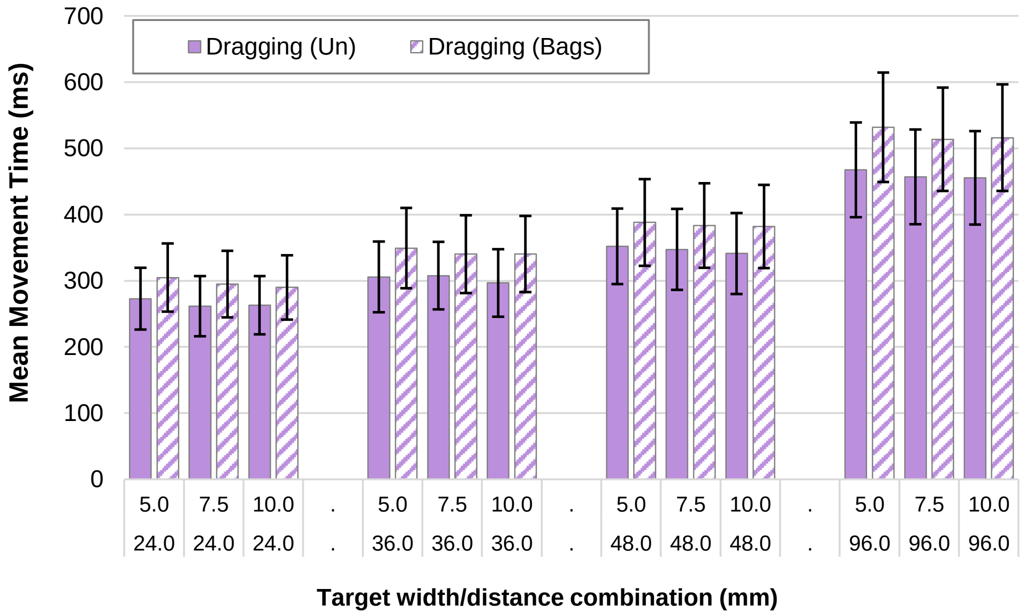

Figure 6.16: The mean movement time (ms) for each target width-distance combination for *dragging*. The solid and striped bars represent the unencumbered (Un) and carrying the bags (Bags) conditions respectively. Error bars denote Confidence Interval (95%).

6.4.5 Movement Time (Spreading and Pinching)

Shapiro-Wilk tests were performed to assess the distribution of movement time for *spreading* and *pinching*. The results are shown in Table 6.5 and all tests were not significant, therefore the data conforms to a normal distribution. A three-factor (**Type of Gesture**, **Type of Encumbrance** and **Target Width-Distance**) repeated-measures ANOVA was conducted for movement time to compare *spreading* and *pinching*. Greenhouse-Geisser adjustments were used to correct the degrees of freedom since Mauchly's test for sphericity was significant ($\varepsilon < 0.75$). The mean movement time for each target width/distance combination for *spreading* and *pinching* is shown in Figure 6.17 and Figure 6.18 respectively.

Type of gesture	**Type of encumbrance**	**W Statistic**	**Sig.**
Spreading	No object	0.948685	0.347575
Spreading	Bags	0.954304	0.437166
Pinching	No object	0.939717	0.236849
Pinching	Bags	0.931063	0.161861

Table 6.5: Shapiro-Wilk normality tests performed on movement time for the *spreading* and *pinching* conditions in *User Study 4*.

The ANOVA for movement time between *spreading* and *pinching* showed a significant main effect for **Type of Gesture**, $F(1, 19) = 7.57$, $p < 0.05$. Movement time was significantly faster when *pinching* than *spreading*. No significant main effect was observed for **Type of Encumbrance**, $F(1, 19) = 0.09$, $p > 0.05$. There was a significant main effect for **Target Width-Distance**, $F(1.6, 30.3) = 56.76$, $p < 0.01$. *Post hoc* pairwise comparisons with Bonferroni corrections showed that movement time was significantly faster as target width increased at each gesture distance. Also, movement time took significantly longer as gesture distance increased for each target width.

The interaction between **Type of Gesture** and **Type of Encumbrance** was not significant, $F(1, 19) = 0.00$, $p > 0.05$. A significant effect was found for the interaction between **Type of Gesture** and **Target Width-Distance**, $F(3.5, 66.0) = 4.36$, $p < 0.01$. With the exception of target combination 3.2/8.0mm, the other eight combinations were selected quicker when *pinching* than *spreading*. However, the difference in movement time was small. The interaction between **Type of Encumbrance** and **Target Width/Distance** combination was not significant, $F(2.04, 38.79)$, $p > 0.05$. The interaction between all three factors for movement time was also not significant, $F(3.88, 73.72) = 0.64$, $p > 0.05$. The results for the interactions between the factors are not required to support or reject the experiment hypotheses.

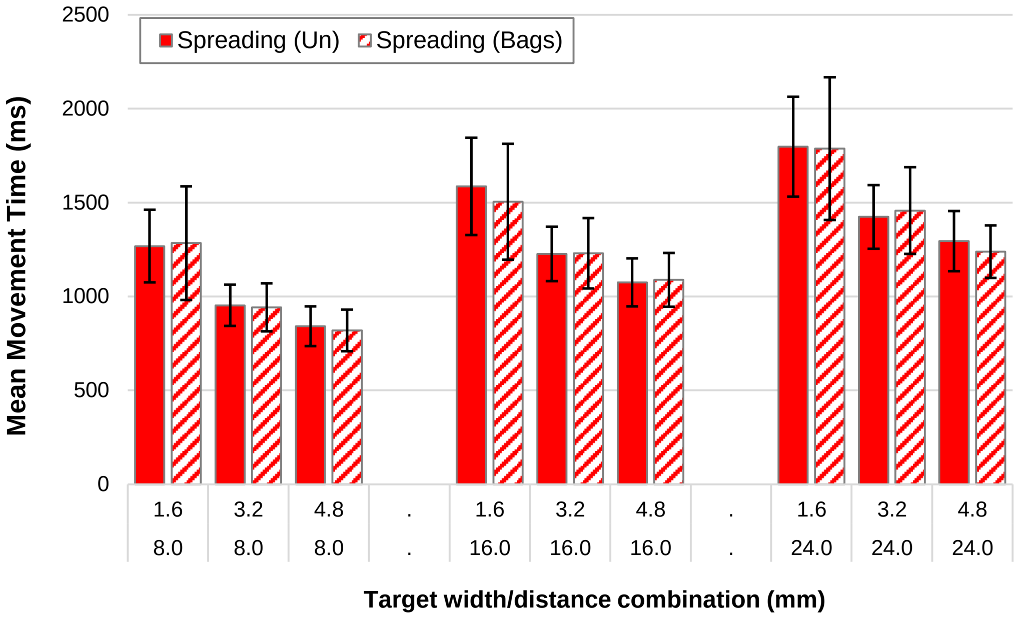

Figure 6.17: The mean movement time (ms) for each target width-distance combination for *spreading*. The solid and striped bars represent the unencumbered (Un) and carrying the bags (Bags) conditions respectively. Error bars denote Confidence Interval (95%).

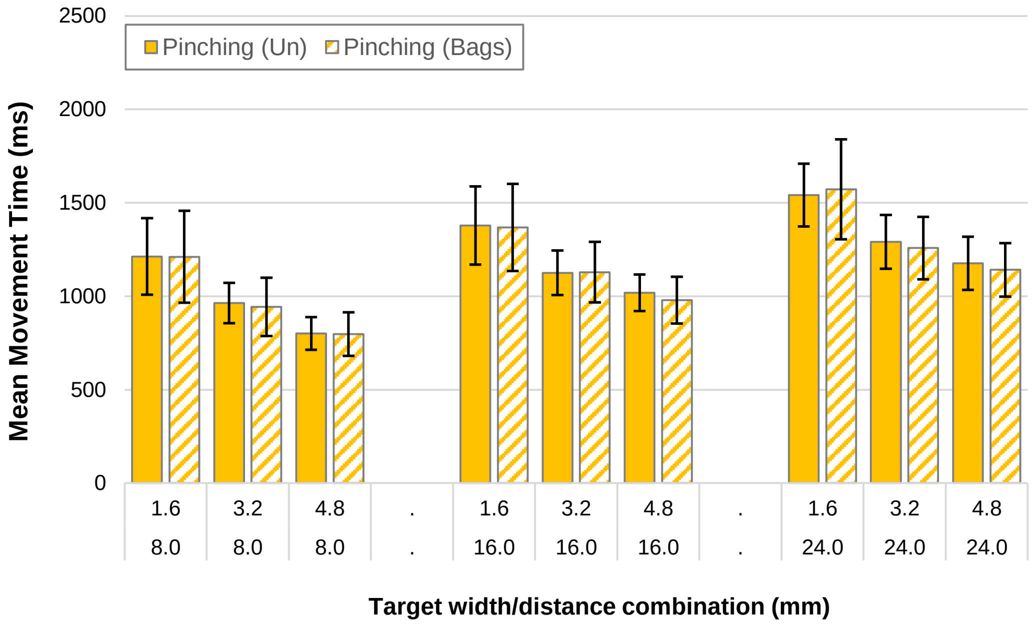

Figure 6.18: The mean movement time (ms) for each target width-distance combination for *pinching*. The solid and striped bars represent the unencumbered (Un) and carrying the bags (Bags) conditions respectively. Error bars denote Confidence Interval (95%).

6.4.6 Movement Time (Rotating Clockwise and Rotating Anticlockwise)

Shapiro-Wilk tests were conducted to examine the distribution of movement time for *rotating clockwise* and *rotating anticlockwise*. The results are shown in Table 6.6 and all tests were not significant, therefore the data conforms to a normal distribution. A three-factor (**Type of Gesture**, **Type of Encumbrance** and **Target Width-Distance**) repeated-measures ANOVA was conducted for movement time to compare *rotating clockwise* and *rotating anticlockwise*. Greenhouse-Geisser adjustments were used to correct the degrees of freedom since Mauchly's test for sphericity was significant ($\varepsilon < 0.75$). The mean movement time for each target width/distance combination for *rotating clockwise* and *rotating anticlockwise* is shown in Figure 6.19 and Figure 6.20 respectively.

Type of gesture	Type of encumbrance	W Statistic	Sig.
Rotating CW	No object	0.94571	0.306611
Rotating CW	Bags	0.906832	0.056617
Rotating AntiCW	No object	0.981033	0.946732
Rotating AntiCW	Bags	0.938435	0.223959

Table 6.6: Shapiro-Wilk normality tests performed on movement time for the *rotating clockwise* (CW) and *rotating anticlockwise* (AntiCW) conditions in *User Study 4*.

The ANOVA conducted to compare movement time for *rotating* showed a significant main effect for **Type of Gesture**, $F(1, 19) = 9.54$, $p < 0.01$. Movement time for *rotating anticlockwise* was significantly quicker than *rotating clockwise*, a mean difference of 99ms. There was no significant main effect for **Type of Encumbrance**, $F(1, 19) = 1.93$ $p > 0.05$. A significant main effect was observed for **Target Width-Distance**, $F(1.6, 29.5) = 86.28$, $p < 0.01$. *Post hoc* pairwise comparisons with Bonferroni corrections indicated that movement time was significantly faster as target width increased at each rotational distance. Movement time also took significantly longer as rotational distance increased for each target width.

The interaction between **Type of Gesture** and **Type of Encumbrance** was not significant, $F(1, 19) = 0.04$, $p > 0.05$. A significant interaction was observed between **Type of Gesture** and **Target Width-Distance**, $F(1.8, 33.3) = 4.55$, $p < 0.05$. Clockwise rotations took longer to perform than anticlockwise rotations for each target width/distance combination. The interaction between **Type of Encumbrance** and **Target Width-Distance** combination was significant, $F(4.1, 78.4) = 3.08$, $p < 0.05$. The movement time for target 6°-30° was significantly quicker when holding the bags than unencumbered. The other eight unique target combinations took significantly longer to select when encumbered. The interaction between all three factors for movement time was not significant, $F(2.7, 51.5) = 1.21$, $p > 0.05$. The interaction effects between the factors are not required to support or reject the experiment hypotheses.

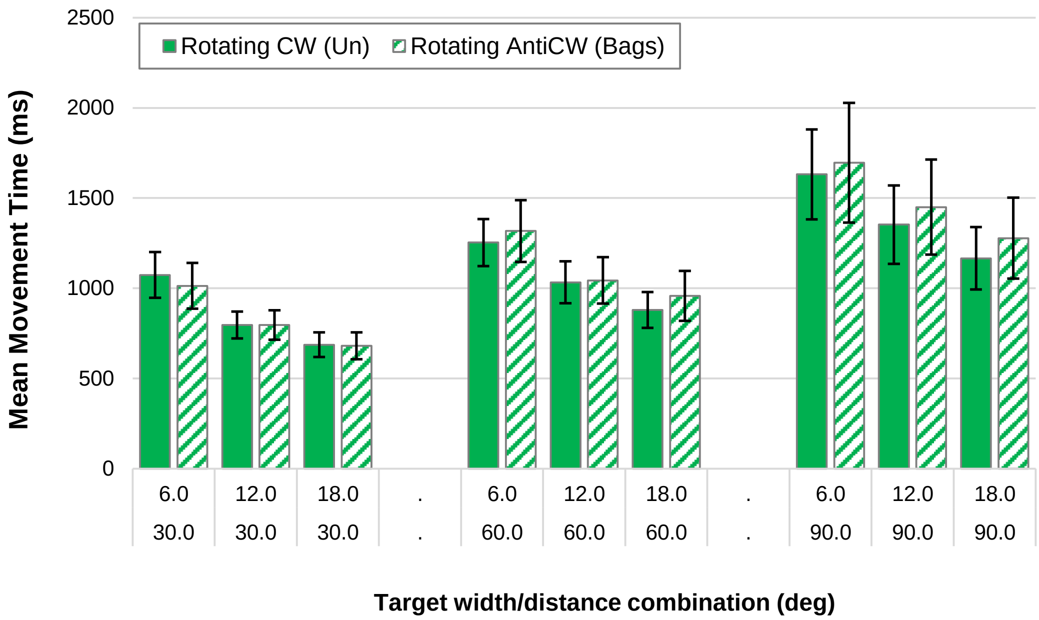

Figure 6.19: The mean movement time (ms) for each target width-distance combination for *rotating clockwise* (CW). The solid and striped bars represent the unencumbered (Un) and carrying the bags (Bags) conditions respectively. Error bars denote Confidence Interval (95%).

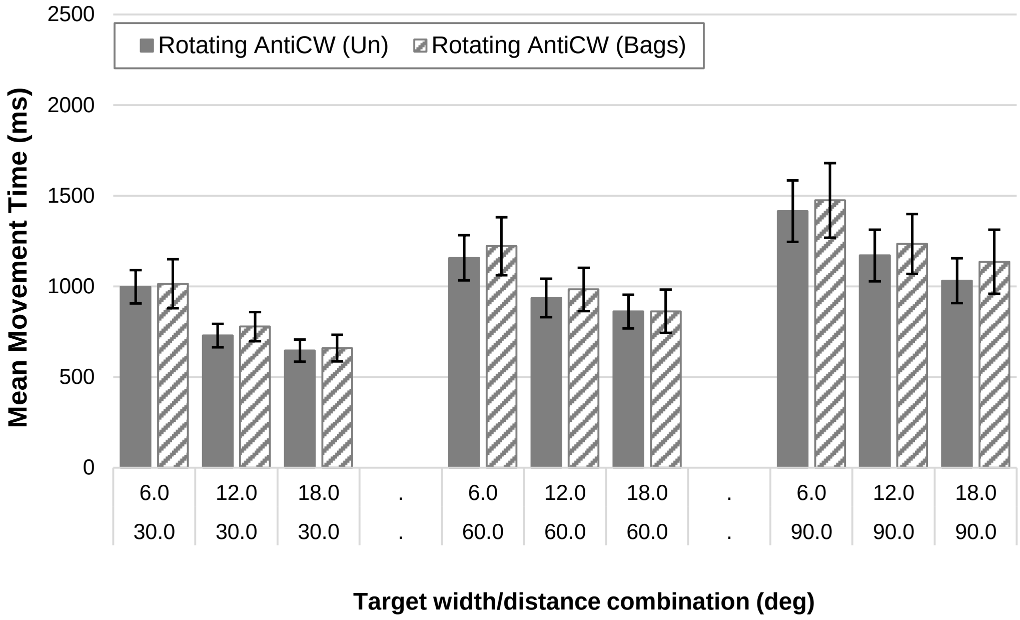

Figure 6.20: The mean movement time (ms) for each target width-distance combination for *rotating anticlockwise* (AntiCW). The solid and striped bars represent the unencumbered (Un) and carrying the bags (Bags) conditions respectively. Error bars denote Confidence Interval (95%).

6.4.7 Fitts' Law Analysis

Figure 6.21 shows a plot of Movement Time (*MT*) against the effective Index of Difficulty (ID_e) for each condition. Table 6.7 shows the values for constants ***a*** and ***b***, the correlation coefficient (*R*), the coefficient of determination (R^2) and Throughput (*TP*) for each condition. Strong correlations ($R^2 > 0.9$) were found for both *tapping* and *dragging* in each encumbrance scenario. This suggests that Fitts' Law is a suitable model to approximate the performance of *tapping* and *dragging* in walking and encumbered situations. The results are also similar to the strong correlation reported in related Fitts' Law *tapping* and *dragging* studies (e.g. [12,48]).

The R^2 values for *spreading* and *pinching* in both encumbrance scenarios indicate a moderate linear relationship (R^2 between 0.6 – 0.8). The *spreading* and *pinching* tasks used in the study were the same ones as Tran *et al.* [71], who reported strong correlations for both types of gestures ($R^2 > 0.9$ for the mobile phone). However, Tran *et al*. controlled error rate since targets had to be selected correctly and participants were seated in their experiment. Since error rate was not controlled in the *spreading* and *pinching* tasks used in

User Study 4, the effective target width was calculated which took into account under- and over- shoots. These results suggest Fitts' Law might not be as useful to model *spreading* and *pinching* performance when users are encumbered and walking. Strong correlations (R^2 > 0.9) were found for *rotating anticlockwise* when unencumbered and carrying bags while walking. Weaker linear relationships (R^2 between 0.75 – 0.89) were determined for *rotating clockwise*. These results show early promise with using Fitts' Law to estimate two-finger rotational performance on touchscreens. However, there might be other reasons to explain why some gestures are a better fit of the Fitts' Law model than others. For example, *tapping* and *dragging* gestures are more frequently used (to select icons and scrolling webpages, for example) than the two-finger gestures, so perhaps users are more skilled with those actions than *pinching* and *spreading* gestures. Further work is required to investigate this and to confirm that Fitts' Law is an appropriate method to model the performance of two-finger gestures.

The throughput results showed a higher rate of information transfer when unencumbered than carrying the bags for each gesture type. Whether users were unencumbered or holding a bag in both hands, the throughput for *tapping* was higher than *dragging*. *Pinching* had higher throughput values than *spreading* while *rotating anticlockwise* had a higher information transfer rate than clockwise rotations.

Gesture	**Encumbrance**	***a***	***b***	***R***	***R²***	***TP***
Tapping	No object	107.2	123.8	0.99	0.97	5.58
Tapping	Bags	151.1	147.4	0.99	0.99	4.27
Dragging	No object	149.9	115.2	1.00	1.00	4.82
Dragging	Bags	155.6	142.0	1.00	0.99	4.16
Spreading	No object	372.7	373.8	0.79	0.62	1.91
Spreading	Bags	310.1	470.5	0.89	0.79	1.60
Pinching	No object	335.7	336.7	0.85	0.72	2.12
Pinching	Bags	458.6	333.4	0.78	0.62	1.82
Rotating CW	No object	122.7	340.7	0.93	0.86	2.60
Rotating CW	Bags	86.5	401.3	0.88	0.77	2.32
Rotating AntiCW	No object	225.0	262.4	0.99	0.98	2.91
Rotating AntiCW	Bags	174.5	324.4	0.99	0.97	2.54

Table 6.7: The values of *a, b, r, R2* and *TP* (Throughput) for each condition. Un = unencumbered, Bags = carrying the bags, CW = clockwise, AntiCW = anticlockwise.

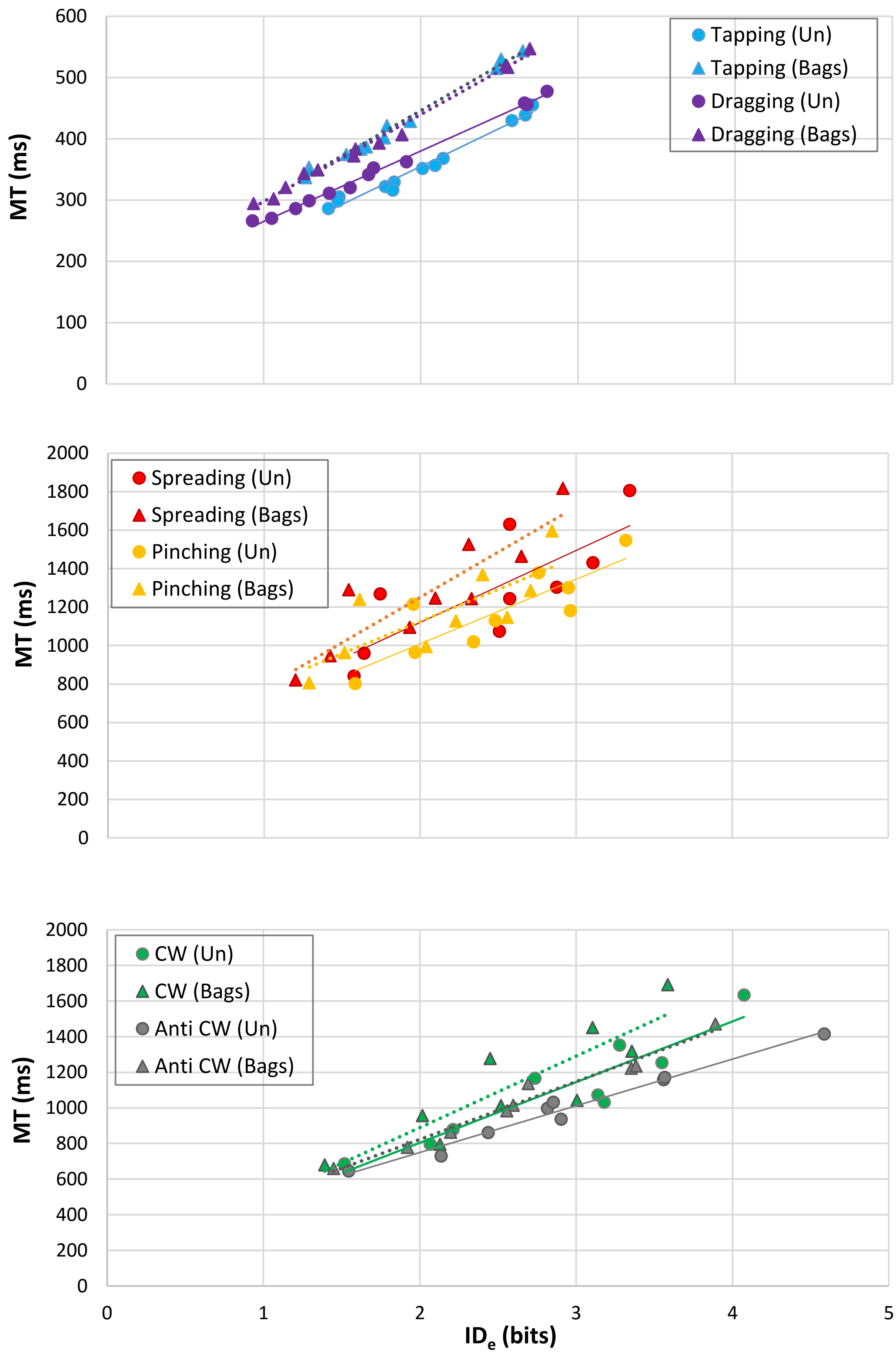

Figure 6.21: Plot of *MT* vs. *IDe* for *tapping* & *dragging* (top), *spreading* & *pinching* (middle) and *rotating clockwise* & *anticlockwise* (bottom). Note the different y-axis units in the top graph.

6.4.8 Distanced Walked and Change in PWS

The estimated mean distance walked per participant over the whole experiment was 3,168.7m (SD = 671.6), approximately 158 laps of the route. Each participant took approximately 41.2 minutes to complete all 12 blocks of targeting tasks. This duration does not include resting periods. The mean baseline PWS and PWS&I measured before the targeting tasks began were 4.57km/h (SD = 0.27) and 3.61km/h (SD = 0.63) respectively. Paired t-test analysis showed a significant difference between the mean walking speeds; $t(19) = 7.197$, $p < 0.05$. Walking speed dropped by 21% while interacting with a mobile phone when compared to walking alone.

6.5 Discussion

The results for target accuracy showed that the number of correct target selections significantly decreased when encumbered for each type of gesture. Therefore, hypothesis *H1* is supported. For *tapping*, the overall mean accuracy while walking and unencumbered was 65% and dropped to 53% when the bags were held. Encumbrance also caused a decrease in accuracy for all target width/distance combinations. As expected, accuracy increased as target width increased at each distance. On the other hand, increasing movement distance did not cause a uniform decrease in accuracy for all target widths. Accuracy gradually decreased as distance between the targets increased when selecting the 10mm targets while the results for 5.0mm and 7.5mm targets were more varied.

Dragging gestures were performed poorly whether users were encumbered or not while walking. The overall mean accuracies for unencumbered and carrying the bags while walking were 50% and 48% respectively. Similar to *tapping*, encumbrance caused a decline in selection accuracy for all twelve target width/distance combinations. Accuracy increased as target width grew at each distance. Unexpectedly, accuracy for each target width was either similar or improved as the movement distance increased regardless of whether users were encumbered or not. In fact, accuracy for all three target widths at the largest distance of 96mm were higher when *dragging* than *tapping*. This suggests that short drags are more difficult to perform and *dragging* might be better than *tapping* for tasks requiring long movements.

Comparing target accuracy between *tapping* and *dragging* showed that taps were performed better than drags. Therefore, hypothesis *H3* is rejected. The poor selection accuracy for *dragging* could have been caused by occlusion issues when the input finger approaches the target. Users might have thought that the target was selected as visually, the finger covered part of the target, but the actual contact point on the touchscreen is not within the target border. Vogel and Baudisch [73] describes this visual issue as the "fat finger" problem where selection is ambiguous due to obstruction by the input finger. Furthermore, in Cockburn *et al.*'s Fitts' Law study [12], an offset cursor was used in their *dragging* task, which meant there were no problems seeing if the target was selected or not. However, Cockburn *et al.* used a one-dimensional task while a two-dimensional *dragging* task was applied in *User Study 4*. Because offset cursors are normally placed vertically above the current finger position, issues occur when the finger approaches the top edge of the touchscreen as it would in a two-dimensional task. In addition, a fair comparison between *tapping* and *dragging* using the same task could not have been possible since selecting targets would have been indirect when *dragging*. Vogel and Baudisch [73] also reported that users tend to compensate the initial touch down position when offset cursors are used. Therefore, to avoid adding noise to the data, an offset cursor was not used when performing *dragging* gestures in *User Study 4*.

S*preading* accuracy dropped from 86% when unencumbered to 77% when both hands were carrying bags. In addition, accuracy of each target width/distance combination was reduced when encumbered. As predicted, the number of correct target selections increased as target width enlarged at each movement distance. Like *dragging* however, target accuracy progressively increased as movement distance also increased. While both digits could touch anywhere on the touchscreen at the beginning of a *spreading* gesture (the distance between the index finger and the thumb was not fixed), it was observed that users tapped near to the control circle. Therefore, both digits were close together which could have made short *spreading* actions more difficult to perform while larger movement distances meant users had more control as the digits expanded away from each other.

For *pinching*, encumbrance caused accuracy to decrease from 90% when no bags were held during input to 81% when encumbered, a similar 9% difference as *spreading*. Encumbrance also caused a decline in selection accuracy for all target width/distance combinations. At each movement distance, *pinching* accuracy increased as target width expanded. However, increasing distance did not have a clear effect on accuracy for each target width. For both unencumbered and carrying the bags, *pinching* accuracy for each target width was very similar at each movement distance. This suggests that the participants had less problems performing large *pinching* movements across the touchscreen when appropriate target sizes are used.

Rotating gestures were performed very well as the overall target accuracies when unencumbered and carrying the bags were 96% and 92% respectively for both rotational directions. For both *rotating clockwise* and *rotating anticlockwise*, accuracy increased as target width also increased at each rotational distance. The smallest target width of 6° caused the most inaccurate target selections while target widths of 12° and 18° at each rotational distance had accuracies greater than 95% regardless of holding the bags or not. As expected, selecting the smallest target width of 6° at the largest rotational distance of 90° (highest *ID*) caused the lowest accuracy for both clockwise and anticlockwise rotations when encumbered.

For *tapping* and *dragging*, movement time took significantly longer when the bags were held than unencumbered. Speed of input for *spreading* and *pinching* was fractionally quicker when encumbered although the results were not significant. Movement time increased when the bags were held in both hands for *rotating clockwise* and *anticlockwise* but the results were not significant. Therefore, hypothesis *H2* cannot be fully supported and is rejected.

The overall mean movement time increased by 17.9% when encumbered compared to holding no objects for *tapping*. Encumbrance also caused an increase in movement time for each target width/distance combination when taps were used. Increasing target width did not have an effect on movement time at each target distance. As expected, selecting each target width with taps took longer as the distance increased. Encumbrance caused an increase in overall movement time by 12.3% when compared to holding no bags for *dragging*. Like *tapping*, each target width/distance combination required more time to execute with drags when encumbered. The time to select each target width with *dragging* gestures was similar at each movement distance. As predicted, long *dragging* distances caused movement time to increase for each target width. Comparing movement time between *tapping* and *dragging* showed that input speed was marginally faster when performing drags but the results were not significant. Since overall accuracy was lower for *dragging* than *tapping*, the participants might have compromised target accuracy for quicker input speed.

Carrying the bags did not significantly increase overall movement time when compared to unencumbered for both *spreading* and *pinching* as differences were less than 1%. For both *spreading* and *pinching*, movement time was faster as target width increased at each gesture distance. Movement time took longer as gesture distance increased for each target width. *Pinching* gestures were performed significantly faster than *spreading* actions and therefore hypothesis *H4* is supported. The overall mean movement time for *spreading* was 1267.7ms, which decreased by 9.1% to 1161.7ms for *pinching*. For comparison, the Fitts' Law tasks used to measure *spreading* and *pinching* were based on the work of Tran *et al.* [71]. They reported overall execution times of 1090ms and 1020ms for *pinching* and *spreading* when the gestures were performed on a mobile phone and users unencumbered and seated. As expected, carrying bags and walking caused slower input speed on the two-finger lateral movements.

Movement time was not significantly affected by encumbrance for rotations in both directions. For *rotating clockwise* and *anticlockwise*, movement time was quicker at each rotational distance as target width increase. Movement time was slower as rotational distance increased for each target width. These results are in accordance to Fitts' Law. The movement time for *rotating anticlockwise* was significantly quicker than *rotating clockwise* therefore hypothesis *H5* is supported.

The strong correlations between movement time and index of difficulty for both *tapping* and *dragging* suggests the applicability of Fitts' Law to approximate the performance of the one-finger gestures in encumbered and mobile contexts. Weaker correlations were found for *spreading* and *pinching* which imply that Fitts' Law is less suited to model two-finger lateral movements when users are walking and carrying bags. Tran *et al.* [71] reported stronger correlations (R^2 > 0.9) for both *spreading* and *pinching* on a mobile phone from their Fitts' analysis. The Fitts' style targeting task designed to quantity the performance of touch-based rotations showed a strong linear relationship between movement time and index of difficulty for *rotating anticlockwise* (R^2 for unencumbered and holding the bags were 0.98 and 0.97 respectively) while weaker correlations for rotating clockwise (R^2 for unencumbered and holding the bags were 0.86 and 0.77 respectively). These results show early promise in using Fitts' Law to model and predict the performance of two-finger rotations on touchscreen devices.

The results for throughput showed that carrying the bags caused a lower information transfer rate for each gesture type than when unencumbered. As expected, a higher difficulty level occurred when bags were held which restricted arm and hand dexterity.

Throughput was affected the most by encumbrance when *tapping* as the information communication rate was reduced by 23.5%. Encumbrance caused a decrease in throughput between 10 - 17% across the other types of gestures. Throughput was higher for *tapping* than *dragging* in each encumbrance scenario.

The information transfer rate was higher for *pinching* than *spreading* for both unencumbered and carrying the bags, which suggests more difficulties with two-finger expanding movements. As to why pinches performed better than spreads, Hoggan *et al.* [28] stated "*the average rotation amplitude of the index finger interphalangeal joint is lower for contraction than expansion* [44]". They recommended using *pinching* whenever possible because *spreading* gestures took longer to execute and were more ergonomically difficult to perform. Even when walking and encumbered, *pinching* gestures were completed marginally better than *spreading* actions.

Throughput was higher for *rotating anticlockwise* than *rotating clockwise*. Hoggan *et al.* [29] explained that "*clockwise rotations by right-handed users are known to generate higher wrist extensor and dominant deltoid muscle activity than anti-clockwise rotations* [13]". They also found that *rotating clockwise* gestures took longer to perform and caused more ergonomic failures than *rotating anticlockwise*. Clockwise rotations maintained its performance advantage over anticlockwise rotations in mobile and encumbered situations but the differences were small. Because three of the participants were left-handed, each participant's performance for each individual gesture was examined prior to conducting statistical tests to observe any potential input differences between left- and right- handed users, especially for executing rotational actions. Observing the data suggested that there was no great disparity in performance across all conditions regardless of the touch-based gestures completed using the left or right hand. The participants were also asked if they found the rotational gestures easier to perform in a particular direction. A majority of the participants commented that there was no preferred rotational direction only that carrying the bags made input subjectively more physically challenging to perform.

To answer research questions *Q2.1 (How do encumbrance and mobility affect tapping performance?)* and *Q2.2* (*How do encumbrance and mobility affect the performance of other standard touch-based gestures on touchscreen mobile phones?*), the study presented in this chapter examined the performance of *tapping, dragging, spreading* & *pinching, rotating clockwise & anticlockwise* while users were walking and encumbered. Using a set of Fitts' Law targeting tasks, the results showed that encumbrance caused a reduction in target accuracy for all types of gestures while movement time took significantly longer for *tapping* and *dragging*.

Although different targeting tasks were used for each gesture type and a direct comparison is not completely possible, in general, the one-finger gestures (*tapping* and *dragging*) were performed poorly in terms of low accuracy, with drags performed the worse when encumbered. Target selections were much more accurate when the two-finger gestures were performed as the overall accuracies for *spreading*, *pinching* and *rotating* in both directions were all over 75% while walking and encumbered. In particular, *rotating* in both directions were executed very well as encumbrance caused a drop in accuracy of only 8%, compared to 52% for *dragging*. The extra digit from the same hand could have made input more stable and therefore target selections were easier to perform accurately. However, overall movement times for the two-finger gestures were at least 990ms while one-finger *tapping* and *dragging* were performed considerably faster as no overall movement times exceeded 420ms.

6.6 Conclusions

To conclude, this chapter has answered research questions *Q2.1* and *Q2.2* by presenting a user study that examined the performance of four main touch-based gestures: *tapping*, *dragging*, *spreading & pinching* and *rotating clockwise & anticlockwise*, while walking and when both hands were encumbered. These gesture types are regularly used on touchscreens and are necessary to interact with certain services such as browsing map applications. Previous work has studied the performance of the mentioned gestures but no research has evaluated how well users can execute these one- and two- finger actions in more physically demanding walking and encumbered contexts.

User Study 4 filled this gap in the literature and the results showed that one-finger taps and drags were performed poorly in terms of target accuracy while the two-finger gestures of *spreading*, *pinching* and particularly *rotating* in both directions were executed surprisingly well, even when walking and encumbered. The expectation was that the two-finger gestures should have been more difficult to perform, especially when encumbered, since they are biomechanically more complex to execute than simpler tap and drag actions. Perhaps the extra finger of the same hand allowed users to stabilise their input and counteract some of the extraneous movements caused when walking and encumbered, therefore targets were selected more accurately. Movement time of the two-finger gestures took substantially longer than one-finger taps and drags but in walking and encumbered situations, incorrect selections are likely to take even longer to recover.

However, there might have been a speed vs. accuracy trade-off since this was not controlled in the experiment. No visual or audio feedback was used to inform the participant if the target was selected correctly or not. Furthermore, incorrect trials were not repeated, due to time constraints, as some studies (such as [46]) repeated trials when error rate is high to control target accuracy. Throughput (which combines accuracy and speed into one metric) for *tapping* and *dragging*, whether unencumbered or carrying the bags, was higher than all two-finger gestures.

User Study 4 also makes a contribution by describing a set of Fitts' Law targeting tasks that other researchers can use to measure the performance of the gestures in a range of contexts. Furthermore, the design of the rotational targeting task is novel and was developed to examine abstract *rotating* performance since no clear method is described in the literature that can be used on small touchscreen mobile devices. The results also showed promise in using Fitts' Law to model two-finger rotational movements in various contexts.

So far, the user studies presented in this and the previous two chapters have collectively evaluated a range of different encumbrance scenarios and showed how touch-based target selections is affected when users are walking and carrying cumbersome objects. The next chapter switches the attention of this thesis towards examining appropriate evaluation approaches for walking and encumbered studies. In the process, two main methodologies were compared, tapping performance was measured in a range of encumbrance scenarios and PWS was varied to see if changing walking speed had any effects on targeting performance while users were encumbered.

Chapter 7

User Study 5: Comparing Evaluation Approaches for Walking and Encumbered Studies

7.1 Introduction

This chapter answers research question *Q3* (*How to evaluate the effects of encumbrance and mobility in controlled user studies?*) by presenting a user study that compared two main evaluation approaches: walking on the ground and walking on a treadmill, to evaluate their effectiveness for mobile and encumbered experiments. The two approaches were selected after reviewing previous walking studies shown in Table 7.1. The methods used in those studies can be categorised into either treadmill or ground walking. Some studies that used the ground walking method also deployed a pacesetter to avoid walking speed having an impact on input performance. Other ground walking experiments did not control walking speed so users could have adjusted their pace and traded walking speed with input performance. This is not an issue when the treadmill approach is used since users always walk at a constant speed.

In *User Study 2* (Chapter 4), which measured tapping performance while carrying different types of bags and boxes, walking speed was not controlled. While the results showed that target accuracy decreased and users walked slower than normal when objects were held during input, it is difficult to be certain if the effects on targeting performance were caused by mobility or encumbrance. Therefore, in *User Study 3* (Chapter 5) and *User Study 4* (Chapter 6), walking speed was controlled to isolate the effects of mobility from encumbrance. This meant that any changes observed to input performance were caused by the effects of encumbrance.

To control walking speed, the participants in *User Study 3* and *User Study 4* walked alongside a human pacesetter. While ecological validity is questionable, human pacesetters (e.g. [24,35,56]) and virtual pacesetters [53] have been commonly used in ground walking studies so that walking speed does not become a dependent variable and add noise to the data collected. However, there are two issues with the pacesetter approach to control walking speed. Firstly, the pacesetter requires training to accurately walk at a range of potential walking speeds, which is costly in terms of time and effort. Secondly, it is difficult for the pacesetter to consistently maintain the same pace for each experimental condition and across all participants. Furthermore, there is no assurance that the participants will walk perfectly in-step with the pacesetter. Consequently, some related walking studies (e.g. [5,7,54]) have used treadmills to address the limitations of using pacesetters but sacrifice more realistic ground walking behaviours.

To reduced experimental complexity and provide guidelines for future walking and encumbered experiments, *User Study 5* was carried out to examine the suitability of using the pacesetter and treadmill approaches to control walking speed. There were three mains objectives for conducting the study. Firstly, it allowed a comparison between treadmill and ground walking which meant any differences between the two evaluation methodologies could be observed. Secondly, an abstract targeting task was used to measure tapping performance while walking and encumbered to answer research question *Q2.1* (*How do encumbrance and mobility affect tapping performance on touchscreen mobile phones?*), which allowed a comparison to the previous results from the studies discussed earlier in this thesis. And thirdly, different levels of PWS were tested to see if walking slower or faster while encumbered had any effects on tapping performance.

Study	Approach	PWS controlled?
Barnard *et al.* [5]	Ground	No
Brewster [8]	Ground	No
Clawson *et al.* [11]	Ground	No
Crossan *et al.* [15]	Ground	No
Lim and Feria [42]	Ground	No
Lin *et al.* [43]	Ground	No
Mackay *et al.* [45]	Ground	No
Mizobuchi *et al.* [50]	Ground	No
Mustonen *et al.* [54]	Ground	No
Pirhonen *et al.* [62]	Ground	No
Schildbach and Rukzio [66]	Ground	No
Yatani and Truong [80]	Ground	No
Goel *et al.* [22]	Ground	Pacesetter
Kane *et al.* [35]	Ground	Pacesetter
Nicolau and Jorge [56]	Ground	Pacesetter
Musić and Murray-Smith [53]	Ground	Pacesetter
Barnard *et al.* [5]	Treadmill	Yes
Bergstrom-Lehtovirta *et al.* [7]	Treadmill	Yes
Lin *et al.* [43]	Treadmill	Yes
Mustonen *et al.*[54]	Treadmill	Yes

Table 7.1: The evaluation approaches used in related walking studies. The last column states if walking speed was controlled. Note that some studies evaluated both treadmill and ground walking approaches.

7.2 Method

7.2.1 Encumbrance Scenarios

Four different encumbrance scenarios were evaluated in *User Study 5*. The participants either held a bag in the non-dominant or dominant hand (*Encumbrance Scenarios 1A and 1B*) or carried a box under the non-dominant or dominant arm (*Encumbrance Scenarios 2A and 2B*). The size of the bag measured 450 *x* 550mm (*w* x *h*) while the dimensions of the box were 370 *x* 300 *x* 150mm (*l* x *w* x *d*). Each object weighed 3kg for each encumbrance scenario and the way the objects were held is shown in Figure 7.1.

Figure 7.1: The encumbrance scenarios evaluated in *User Study 5* (from left to right): holding the bag in non-dominant hand, holding the bag in the dominant hand, carrying the box under the non-dominant arm and carrying the box under the dominant arm.

7.2.2 Task

The same targeting task was used as the one described in *User Study 3* (Section 5.2.2). The participants selected a sequence of targets one at a time on a touchscreen mobile phone as quickly and as accurately as possible. There were nine target positions aligned in a 3 *x* 3 grid. The centre and one of the outer targets were selected in an alternate order. Every second selection was an outer target and the order of the outer targets were randomized for each block of trials. Each outer target was selected ten times which meant there were 160 target selections per block. One block of targets was completed for each condition per participant.

Similar to Crossan *et al.* [15], a random delay from 500 to 1500ms was placed between a selection and the next target shown on-screen to reduce any rhythm created between the user's walking and tapping behaviour. Each target (w *x* h) was 5 *x* 8mm with the central crosshair measuring 2.5mm in both directions. The dimensions of each target were the same size as a key on the standard keyboard for this phone. A Samsung Galaxy S3 mobile phone with a touchscreen resolution of 720 x 1280 pixels (~12 pixels/mm) was used. The mobile device was held in portrait orientation and the two-handed index finger input posture was used across all participants to select the targets. The participants were given a short training phase at the start of the experiment to familiarise them with the targeting task. Figure 7.2 illustrates the target selection task and the two-handed input posture used.

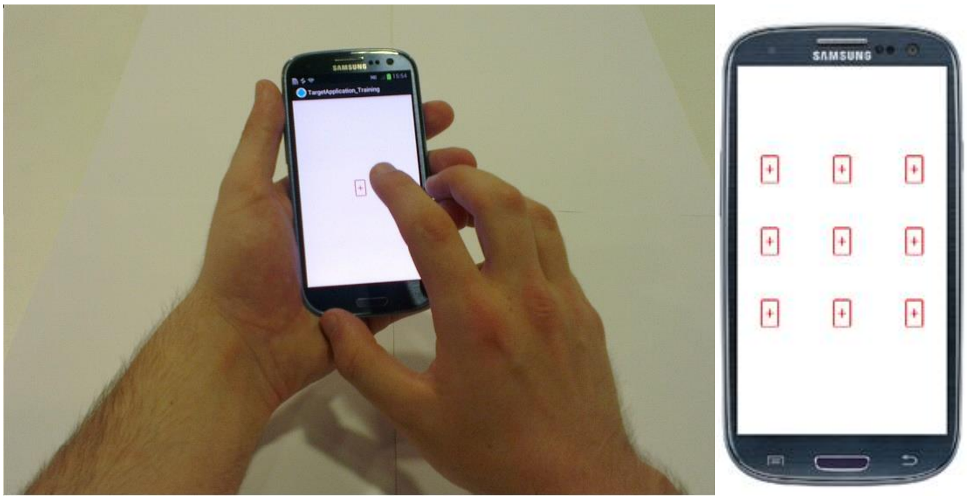

Figure 7.2: The two-handed index finger input posture used to select the targets (left). The nine target positions on a Samsung Galaxy S3 mobile phone (right).

7.2.3 Walking Approaches and Controlling PWS

A calibrated Woodway Bari-Mill treadmill with handrail support (see bottom images in Figure 7.3) 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.* [5] and Bergstrom-Lehtovirta *et al.* [7], participants were asked to think about the pace that they would typical walk while not in a hurry when estimating their 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, the same approach was used as the ones in *User Study 3* and *User Study 4*. Therefore, 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 7.3. The PWS for ground walking was measured by asking the participants to walk the path 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.

The participants walked alongside a pacesetter to control the PWS for the ground walking conditions. The pacesetter used the same metronome application as *User Study 3* to tune the metronome speed for each participant once the PWS was calculated. For each ground walking 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 the targeting task on the mobile phone. Participants were instructed to avoid drifting out of the boundaries of the path during the experiment and were also told to keep up or slow down if they failed to keep in-step with the pacesetter.

Tapping performance while encumbered was also measured at various levels of PWS to simulate situations where the user walked slower (for example, keeping personal distance from other people) and faster (for instance, in a hurry to get to a meeting). This meant that observations could be made to see if the two evaluation approaches were practical at controlling different levels of walking speed and what effects varying PWS would have on targeting performance while encumbered. Based on the findings from Bergstrom-Lehtovirta *et al.* [7], who reported a non-linear drop in target accuracy when normal walking speed was decreased by 20%, the PWS was reduced to 80% in *User Study 5* for the slow walking conditions. The walking speed was increased by the same margin to 120% of PWS for the fast walking conditions.

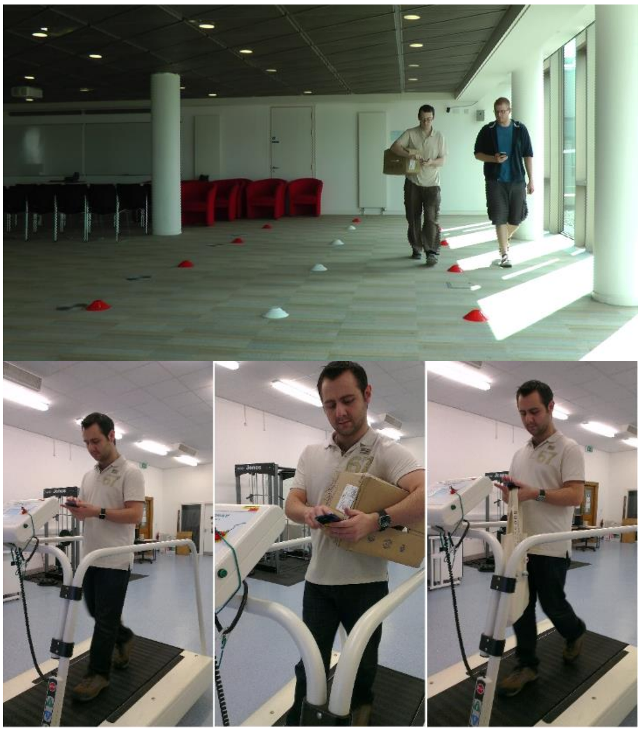

Figure 7.3: Top image illustrates the pre-defined oval route used in the ground walking conditions. The participant (inside) maintained their PWS by walking alongside a pacesetter (outside). The bottom images show a participant performing the task while walking on the treadmill.

7.2.4 Experimental Design

A within-subjects design was used for *User Study 5*. 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 study was split into two sessions and took place on different days to remove any issues of fatigue. The participants were also given sufficient resting periods between each condition and whenever necessary for both sessions. 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 approximately one hour (*introduction* + *training* + *performing the conditions* + *debriefing*) and each participant was paid £12 upon completing both sessions.

The Independent Variables were: **Type of Encumbrance** (5 levels - unencumbered, holding the bag either in the non-dominant or dominant hand and holding the box either under the non-dominant or dominant arm), **Walking Method** (2 levels – walking on the treadmill and walking on the ground around the route) 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 to reduce learning and order effects as much as possible.

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:

H1A: Target accuracy will be significantly lower when encumbered than holding no objects;

H1B: Target error will be significantly higher when encumbered than holding no objects;

H1C: Selection time will be significantly longer when encumbered than holding no objects;

H2A: Target accuracy when the dominant hand/arm is encumbered will be significantly lower than encumbering the non-dominant hand/arm;

H2B: Target error when the dominant hand/arm is encumbered will be significantly higher than encumbering the non-dominant hand/arm;

H2C: Selection time when the dominant hand/arm is encumbered will be significantly longer than encumbering the non-dominant hand/arm;

H3A: Target accuracy will be significantly higher when walking at 80% of PWS compared to walking at 100% of PWS;

H3B: Target error will be significantly lower when walking at 80% of PWS compared to walking at 100% of PWS;

H3C: Selection time will be significantly shorter when walking at 80% of PWS compared to walking at 100% of PWS;

H4A: Target accuracy will be significantly lower when walking at 120% of PWS compared to walking at 100% of PWS;

H4B: Target error will be significantly higher when walking at 120% of PWS compared to walking at 100% of PWS;

H4C: Selection time will be significantly longer when walking at 120% of PWS compared to walking at 100% of PWS;

H5A: Target accuracy will be significantly lower for walking on the ground when compared to walking on the treadmill;

H5B: Target error will be significantly higher for walking on the ground when compared to walking on the treadmill;

H5C: Selection time will be significantly longer 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.

7.3 Results

A total of 96,000 trials (*160 targets x 30 conditions x 20 participants*) was record for the whole study. To filter out unintentional screen taps, targets that took less than 100ms to select were removed from the final data analysis. As a result, 21 trials were eliminated.

7.3.1 Target Accuracy

Shapiro-Wilk tests were carried out to assess the normality of the target accuracy data. The results showed that one condition violated normality: carrying the box under the dominant arm while walking on the ground at 120% PWS. Therefore, ART [76] was used to transform the data before conducting a three-factor (**Type of Encumbrance**, **Walking**

Method and **Walking Speed**) repeated-measures ANOVA to analyse target accuracy. The mean target accuracy for each condition is shown in Figure 7.4.

The ANOVA conducted to examine target accuracy showed a significant main effect for **Walking Method**, $F(1, 551) = 54.0156$, $p < 0.01$. Target selections were significantly more accurate for walking on the treadmill than walking on the ground. The overall mean target accuracy for ground walking was 43.5%, compared 48.3% for treadmill walking.

A significant main effect was observed for **Walking Speed**, $F(2, 551) = 28.0184$, $p < 0.001$. *Post hoc* Tukey HSD comparisons showed that target accuracy was significantly higher when walking at 80% of PWS than both 100% and 120% of PWS. The participants were significantly less accurate at targeting when walking at 120% of PWS than walking at 100% of PWS. The overall mean accuracy for each level of PWS is shown in Figure 7.5.

A significant main effect was also found for **Type of Encumbrance**, $F(4, 551) = 137.1244$, $p < 0.01$. *Post hoc* Tukey HSD comparisons showed that target accuracy was significantly higher when unencumbered compared to holding the objects. Target accuracy while carrying the bag in the dominant hand was significantly lower than carrying the bag in the non-dominant hand ($t = 5.240$, $p < 0.001$). Target accuracy while holding the box under the non-dominant arm was significantly higher than holding the box under the dominant arm ($t = 3.453$, $p < 0.001$). The overall mean accuracy for Type of Encumbrance is shown in Figure 7.6.

The interaction between **Walking Method** and **Walking Speed** was not significant, $F(2, 551) = 1.2709$, $p > 0.05$. No significant effect was observed for the interaction between **Walking Method** and **Type of Encumbrance**, $F(4, 551) = 1,4361$, $p > 0.05$. The interaction between **Walking Speed** and **Type of Encumbrance** was not significant, $F(8, 551) = 0.4071$, $p > 0.05$. The interaction between all three factors was also not significant, $F(8, 551) = 0.1390$, $p > 0.05$. The interaction effects are not required to support or reject the experiment hypotheses on target accuracy.

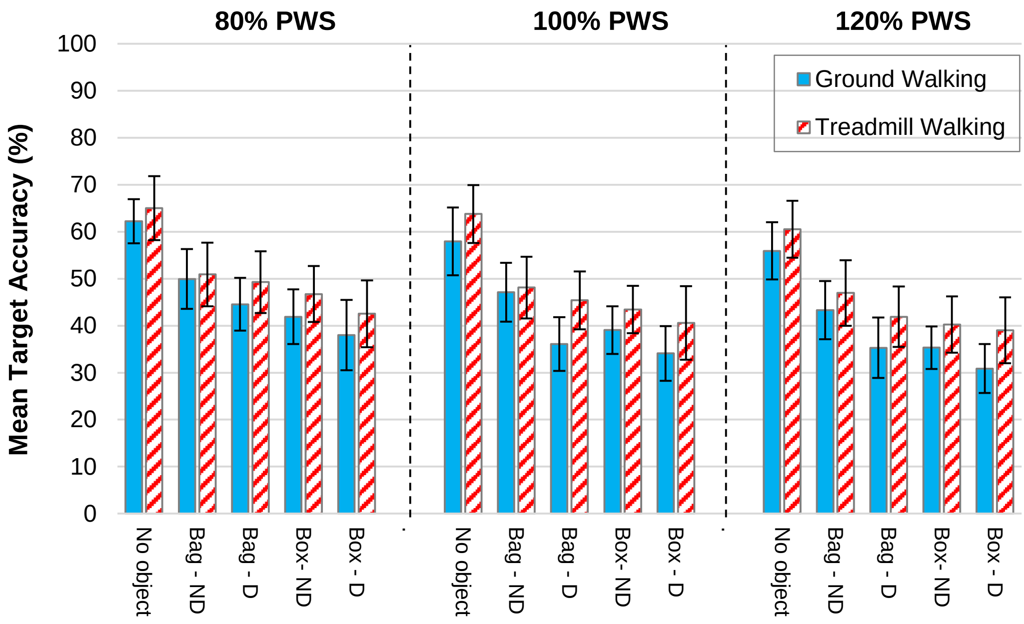

Figure 7.4: The mean target accuracy (%) for each walking condition in *User Study 5* (grouped by Walking Speed). The solid blue and striped red bars represent the ground and treadmill walking conditions respectively. Error bars denote Confidence Interval (95%).

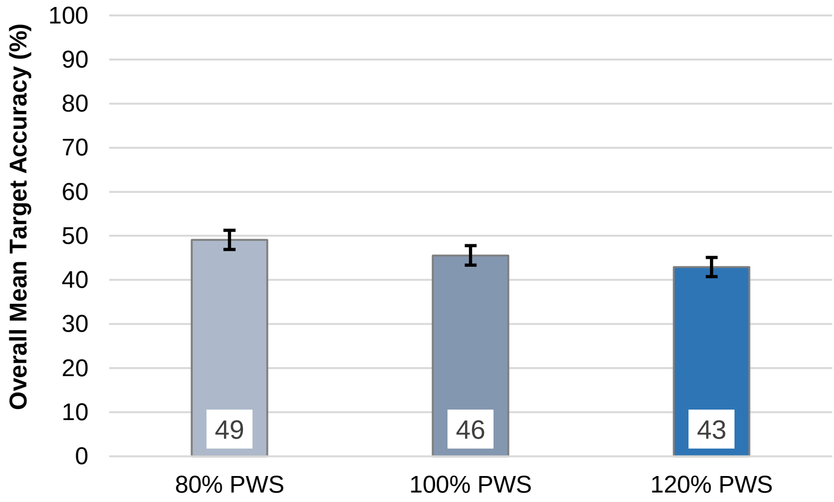

Figure 7.5: The overall mean target accuracy (%) for each level of PWS in *User Study 5*. Error bars denote Confidence Interval (95%).

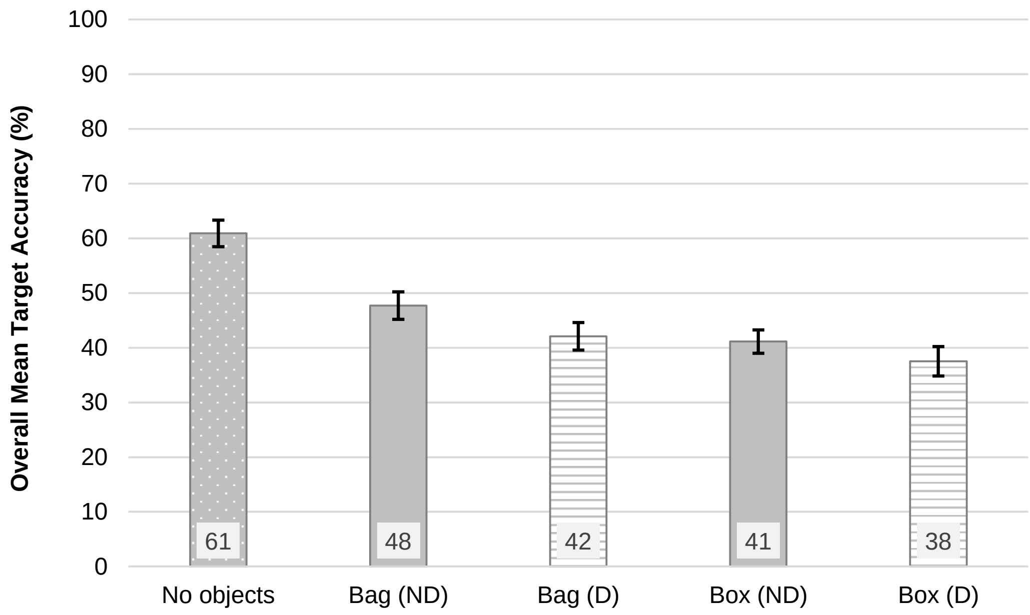

Figure 7.6: The overall mean target accuracy (%) for each Type of Encumbrance in *User Study 5*. ND = non-dominant (solid) and D = dominant (horizontal stripes). Error bars denote Confidence Interval (95%).

7.3.2 Target Error

Shapiro-Wilk tests were carried out to assess the normality of target error. The results showed that three conditions violated normality: (1) carrying the bag in the dominant hand while walking on the ground at 100% PWS, (2) carrying the bag in the non-dominant arm while walking on the treadmill at 120% PWS and (3) carrying the box under the dominant arm while walking on the treadmill at 80% PWS. Therefore, ART [76] was used to transform the data before conducting a three-factor (**Type of Encumbrance**, **Walking Method** and **Walking Speed**) repeated-measures ANOVA on target error. The mean target error for each condition is shown in Figure 7.7.

The ANOVA for target error showed a significant main effect for **Walking Method**, $F(1, 551) = 48.4032$, $p < 0.05$. Target error was significantly higher when walking on the ground than walking on the treadmill, although the difference was small. The mean errors for ground and treadmill walking were 4.6mm and 4.3mm respectively. A significant main effect was observed for **Walking Speed**, $F(2, 551) = 32.9826$, $p < 0.01$. *Post hoc* Tukey HSD comparisons showed that target error was significantly lower when walking at 80% of PWS than 100% of PWS ($t = 4.032$, $p < 0.001$) and 120% of PWS ($t = 8.122$, $p <$

0.001), both small mean differences of 0.3mm and 0.5mm respectively. Target error when walking at 100% of PWS was significantly lower than walking at 120% of PWS (t = 4.090, $p < 0.001$), a small mean difference of 0.3mm. The overall mean error for each level of PWS is shown in Figure 7.8.

A main effect was found for **Type of Encumbrance**, $F(4, 551) = 129.6910$, $p < 0.001$). *Post hoc* Tukey HSD comparisons showed that target error was significantly higher when holding the objects than unencumbered input. Target error was significantly higher when the bag was held in the dominant hand than when the bag was held the non-dominant hand (t = 8.313, $p < 0.001$). Likewise, holding the box under the dominant arm resulted in a significant increase in error compared to holding the box in the non-dominant arm (t = 3.740, $p < 0.001$). The overall mean error for each **Type of Encumbrance** is shown in Figure 7.9.

The interaction between **Walking Method** and **Walking Speed** was significant, $F(2, 551) = 4.2811$, $p < 0.05$. *Post hoc* Tukey HSD tests showed that no comparisons were significant. The interaction between **Walking Method** and **Type of Encumbrance** was significant, $F(4, 551) = 4.1023$, $p < 0.05$. Due to the large number of comparisons and that the interaction effect is not relevant to the experiment hypotheses, further analysis was not conducted. No significant effect was found for the interaction between **Walking Speed** and **Type of Encumbrance**, $F(8, 551) = 1.2622$, $p > 0.05$. The interaction between all three factors was not significant, $F(8, 551)$ = , $p > 0.05$. The interaction effects between the factors are not required to support or reject the hypotheses on target error.

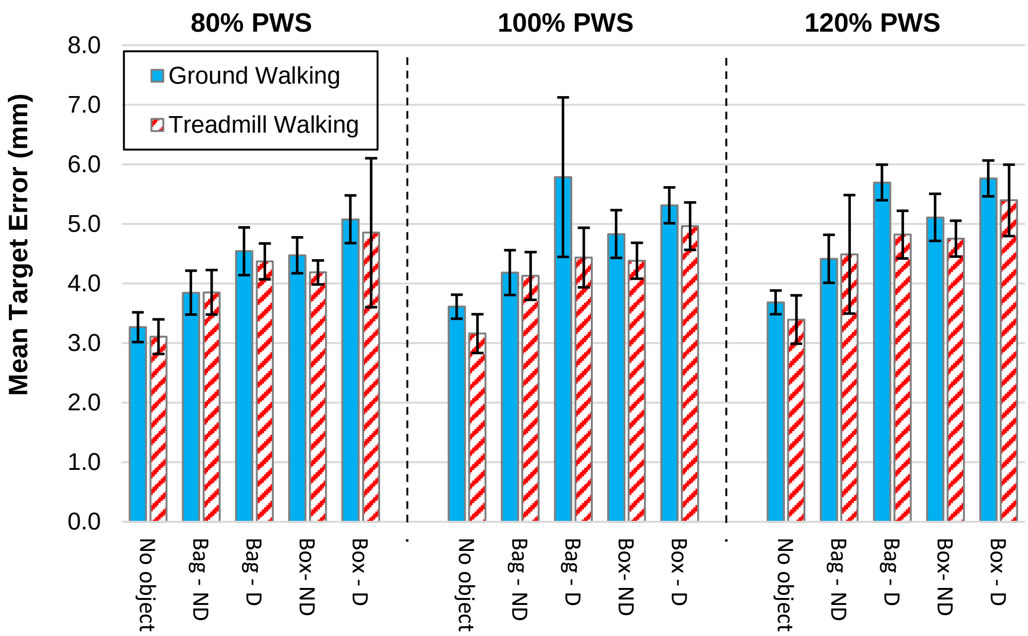

Figure 7.7: The mean target error (mm) for each walking condition in *User Study 5* (grouped by Level of PWS). The solid blue and striped red bars represent the ground and treadmill walking conditions respectively. Error bars denote Confidence Interval (95%).

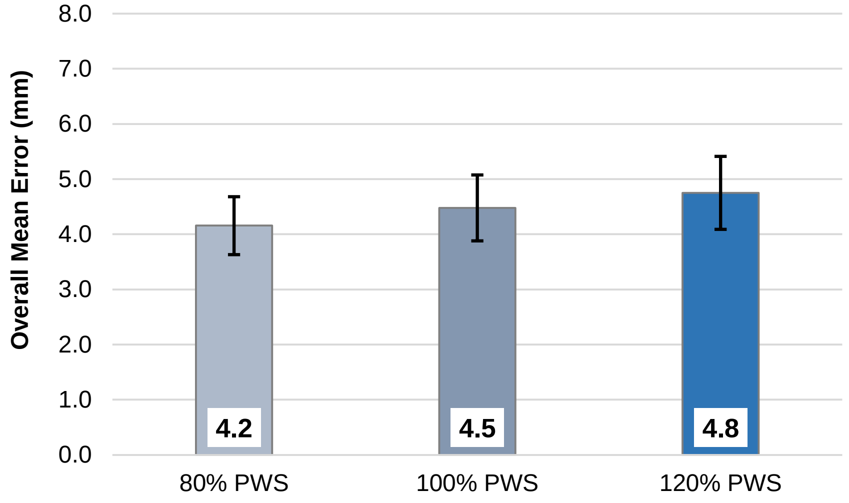

Figure 7.8: The overall mean error (mm) for each level of PWS in *User Study 5*. Error bars denote Confidence Interval (95%).

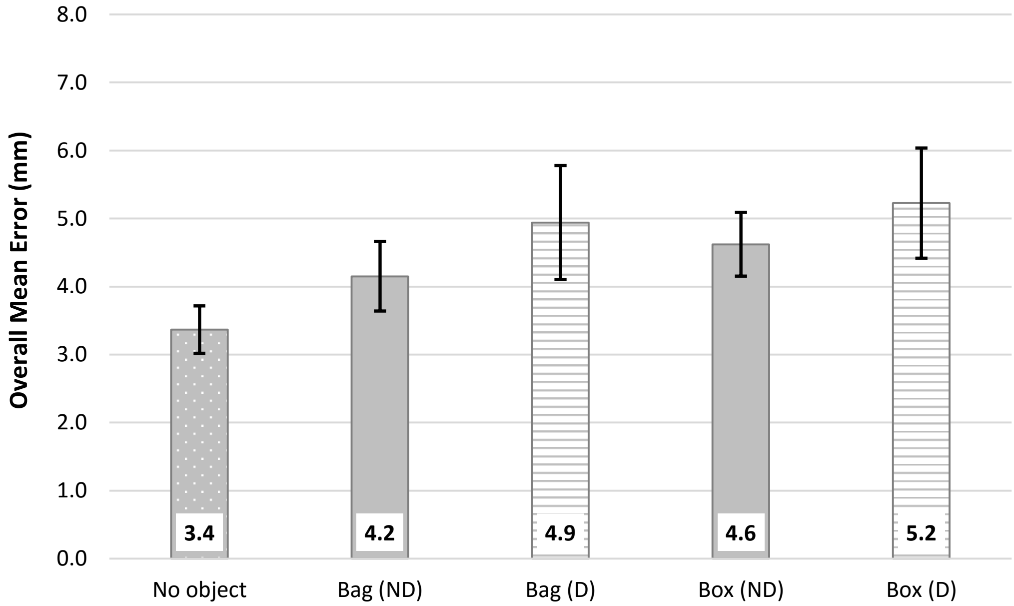

Figure 7.9: The overall mean error (mm) for each Type of Encumbrance in *User Study 5*. ND = non-dominant (solid) and D = dominant (horizontal stripes). Error bars denote Confidence Interval (95%).

7.3.3 Selection Time

Shapiro-Wilk tests were conducted to examine the distribution of selection time. The results showed that three conditions deviated from a normal distribution: (1) carrying the box under the non-dominant arm while walking on the ground at 100% of PWS, (2) carrying the bag in the dominant hand while walking on the ground at 120% of PWS and (3) carrying the box under the dominant arm while walking on the treadmill at 120% PWS. Therefore, the data was transformed using ART [76] before conducting a three-factor (**Type of Encumbrance**, **Walking Method** and **Walking Speed**) repeated-measures ANOVA. The mean selection time for each condition is shown in Figure 7.10.

The ANOVA for selection time showed a significant main effect for **Walking Method**, $F(1, 551) = 87.4845$, $p < 0.001$. Selection time was significantly quicker for walking on the treadmill than walking on the ground. The mean selection times for ground and treadmill walking were 518ms and 492ms respectively, small difference of 26ms.

A significant main effect was observed for **Walking Speed**, $F(2, 551) = 6.3288$, $p < 0.005$. *Post hoc* Tukey HSD comparisons showed that selection time was not significantly quicker for walking at 80% of PWS than 100% of PWS ($t = 1.621$, $p > 0.05$). There was also no significant difference for selection time between walking at 100% and 120% of PWS ($t = 1.932$, $p > 0.05$). However, selection time was significantly quicker when walking at 120% of PWS than 80% of PWS ($t = 3.553$, $p < 0.01$), a small mean difference of 14.3ms. The overall mean selection time for each level of PWS is shown in Figure 7.11.

A significant main effect was found for **Type of Encumbrance**, $F(4, 551) = 59.0126$, $p < 0.01$. *Post hoc* Tukey HSD comparisons showed that target selections were significantly quicker when unencumbered compared to carrying the objects, except for holding the bag in the non-dominant hand. The participants were significantly quicker at selecting the targets when the bag was held in the non-dominant hand than holding the bag in the dominant inputting hand ($t = 8.474$, $p < 0.05$). Selection time was significantly quicker when the box was held under the non-dominant arm than the dominant arm ($t = 6.435$, $p < 0.01$). The overall mean selection time for each **Type of Encumbrance** is shown in Figure 7.12.

The interaction between **Walking Method** and **Walking Speed** was not significant, $F(2, 551) = 0.6656$, $p > 0.05$. The interaction between **Walking Method** and **Type of Encumbrance** was also not significant, $F(4, 551) = 1.1264$, $p > 0.05$. No significant effect was found for the interaction between **Type of Encumbrance** and Walking Speed, *F*(8, 551) = 0.345, *p* > 0.05. The interaction between all three factors was not significant, *F*(8, 551) = 0.703, *p* > 0.05. The interaction effects between the factors are not required to support or reject the hypotheses on selection time.

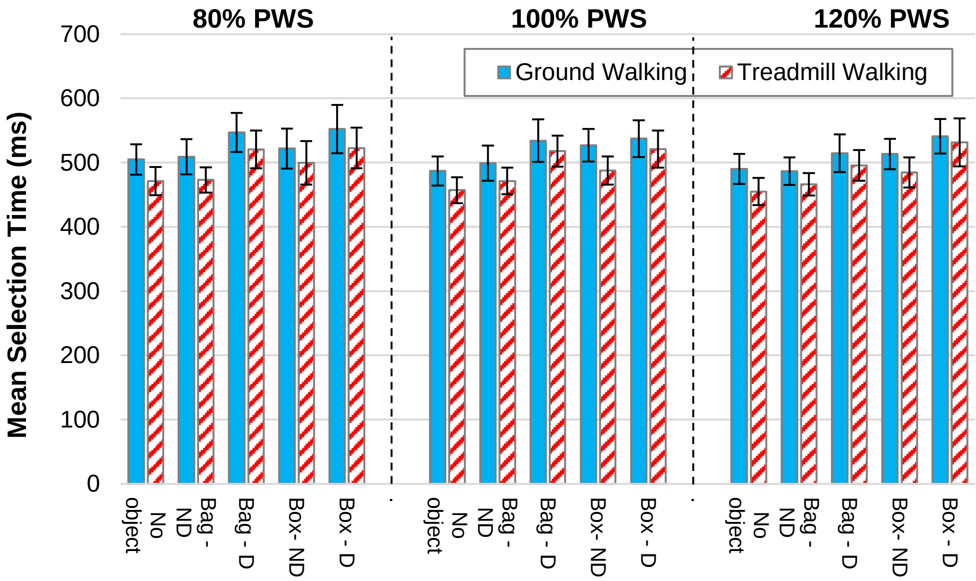

Figure 7.10: The mean selection time (ms) for each walking condition in *User Study 5* (grouped by Level of PWS). The solid blue and striped red bars represent the ground and treadmill walking conditions respectively. Error bars denote Confidence Interval (95%).

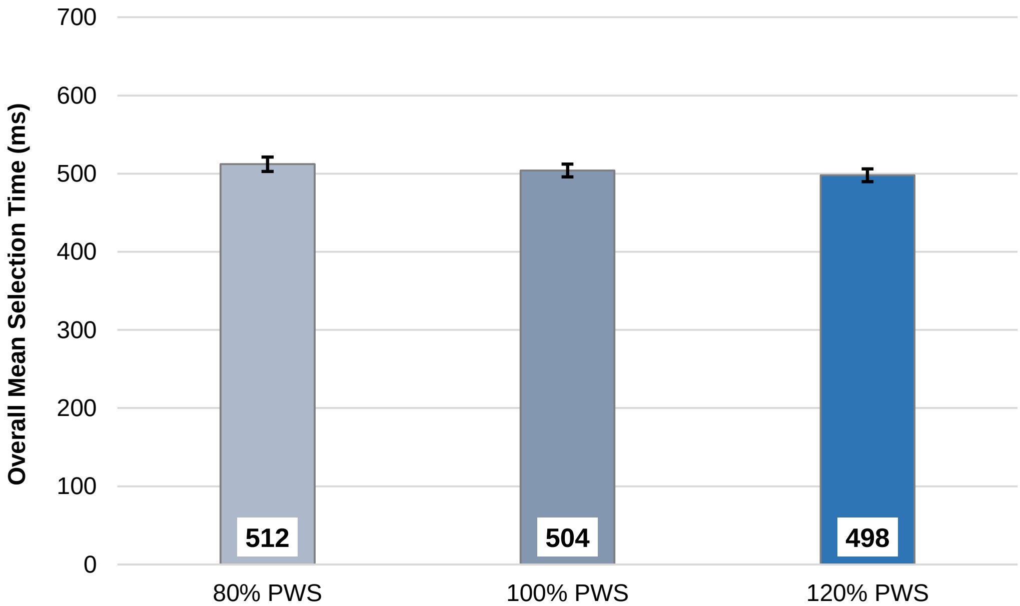

Figure 7.11: The overall mean selection time (ms) for each level of PWS in *User Study 5*. Error bars denote Confidence Interval (95%).

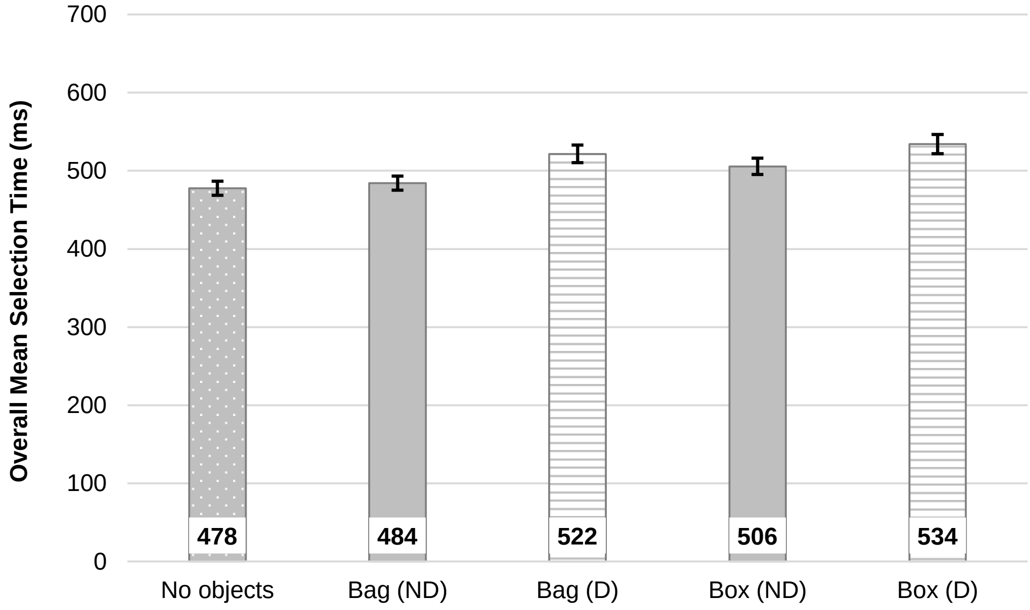

Figure 7.12: The overall mean selection time (ms) for each type of encumbrance in *User Study 5*. ND = non-dominant (solid) and D = dominant (horizontal stripes). Error bars denote Confidence Interval (95%).

7.3.4 Comparison of PWS and Distance Walked

Shapiro-Wilk normality tests were conducted on the walking speeds recorded for ground and treadmill walking. No significant results were found therefore the data conforms to a normal distribution. 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 difference in walking speed of 26.8%. Furthermore, each participant's PWS on the treadmill was slower than walking on the ground. Table 7.2 shows the estimated 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.

	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 7.2: The mean walking distance and total interaction time for each walking method.

7.4 Discussion

The results from *User Study 5* showed a decrease in target accuracy when encumbered compared to holding no objects during input. Therefore, hypothesis *H1A* is supported. Holding the bag in the dominant hand caused more inaccurate target selections than holding the bag in the non-dominant hand. Likewise, target accuracy was significantly lower when the box was carried under the dominant arm than carrying the box under the non-dominant arm. Therefore, hypothesis *H2A* is supported. The lowest overall mean accuracy of 37.5% occurred when carrying the box under the dominant arm. For comparison, the overall mean accuracy of the same encumbrance from *User Study 2* but selecting slightly smaller 4 *x* 6 targets was 40%. Holding the bag in-hand had higher overall accuracies than carrying the box underarm. When the bag was held in the dominant hand, an overall accuracy of 42.1% occurred and improved to 48% when held in the non-dominant hand. As expected, the participants made more accurate target selections when unencumbered as the overall mean accuracy was 60.9%.

Comparing target accuracy between the different levels of PWS showed that the number of inaccurate selections was lower when walking 80% of PWS than 100% of PWS. Therefore, hypothesis *H3A* is supported. The overall mean accuracy when walking 100% of PWS was 45.6% and improved marginally to 49.1% when walking at 80% of PWS, a small difference of 3.5%. This suggests that walking speed needs to be reduced further in order to substantially improve tapping accuracy. For comparison, Bergstrom-Lehtovirta *et al.* [7] reported an accuracy of ~82.5% when users walked at 80% of PWS in their unencumbered treadmill walking study. Musić and Murray-Smith [53] found that median error rates were higher when walking slower than 100% of PWS when users tapped on similarly sized targets. As expected, walking faster at 120% of PWS increased the number of incorrect target selections when compared to walking at 100% of PWS. Therefore, hypothesis *H4A* is supported. Again, the difference in accuracy between the two speeds was small at 2.6%. Perhaps this suggests that input has reached a limit of poor targeting performance and walking faster does not substantially decrease accuracy further. For comparison, Bergstrom-Lehtovirta *et al.* found that targeting accuracy dropped by ~8% between 100% and 120% of PWS.

As predicted, target accuracy for walking on the treadmill was significantly higher than walking on the ground. Therefore, hypothesis *H5A* is supported. However, it is worth noting that the recorded PWS across all participants for walking on the treadmill was slower than walking on the ground. This might have been one reason why the overall target accuracy was higher when walking on the treadmill than on the ground. Furthermore, the difference in accuracy between the two walking methods was small - 4.8%. The overall mean target accuracies for ground and treadmill walking were 43.5% and 48.3% respectively, which illustrates the negative impact mobility has on tapping performance. In *User Study 2*, the overall mean accuracy when walking on the ground was even lower at 37%, although participants selected smaller targets than the ones in this study.

The results for targeting error showed that all four encumbrance scenarios increased tapping imprecision when compared to unencumbered. Therefore, hypothesis *H1B* is supported. Carrying the bag in the dominant hand caused an increase in error compared to holding the bag in the non-dominant hand. Similarly, error was higher when the box was held under the dominant arm than the non-dominant arm. Therefore, hypothesis *H2B* is supported. However, it is worth noting that the difference in error between the dominant and non-dominant hands when holding the bag was small at 0.7mm. Likewise, the difference between carrying the box under the dominant and non-dominant arms was only 0.6mm. The highest overall mean error of 5.2mm occurred when the box was carried under the dominant arm, an increase of 52.9% when compared unencumbered tapping.

Walking slower at 80% of PWS resulted in a decrease in targeting error when compared to walking at 100% of PWS. Therefore, hypothesis *H3B* is supported. However, the difference in error between the two walking speeds was very small at 0.3mm. As predicted, walking faster at 120% increased error when compared to walking normal at 100% of PWS. Therefore, hypothesis *H4B* is supported. Again, the difference in error between 100% and 120% of PWS was marginal at 0.3mm. Comparing targeting error between the two walking methods showed that the participants were more precise when walking on the treadmill than walking on the ground. Therefore, hypothesis *H5B* is supported. The overall mean error for treadmill walking was 4.6mm, an increase of 6.5% from 4.3mm when walking on the ground so the difference is small.

The results for target selection time showed that tapping speed was quicker when no object was held than encumbered with the exception of carrying the bag in the non-dominant hand. Therefore, hypothesis *H1C* cannot be fully supported and is rejected. Selection time was quicker when the bag was held in the non-dominant hand than the dominant hand. Similarly, targeting speed was quicker when the box was held under the non-dominant arm than carrying the box under the dominant arm. Therefore, hypothesis *H2C* is supported. Carrying the box under the dominant arm resulted in the slowest overall mean selection time of 534ms. Although the results were significant, selection time only reduced by 5.3% when the box was held under the non-dominant arm. Likewise, there was a small difference of 7.2% in overall mean selection time between the non-dominant and dominant hand when the bag, despite the results being significant. These results show that targeting performance when encumbered between the non-dominant and dominant sides were evenly matched since the difference in accuracy was also marginal.

Walking slower at 80% of PWS did not significant increase targeting speed when compared to walking 100% of PWS. Therefore, hypothesis *H3C* is rejected. Selection time did not take significantly longer when walking at 120% of PWS than 100% of PWS. Therefore, hypothesis *H4C* is also rejected. The consistency in overall mean selection times between the three levels of PWS showed that walking speed did not have a major effect on tapping speed. Target selection time was significantly faster when walking on the treadmill than walking on the ground. Therefore, hypothesis *H5C* is supported. However, the difference in overall mean selection time between the two walking methods was very small (26ms). It was anticipated that target selection would have been much slower when walking on the ground since full visual attention could not be used for input and the participants had to divide its resources for navigation as well as for interaction.

The comparison in PWS between the walking methods showed that the participants walked significantly faster on the ground than on the treadmill. Thus, hypothesis *H6* is rejected. It was expected that PWS for walking on the ground would be slower than walking 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 participant dropped their PWS by as much as 71% when walking on the ground. The difference in PWS between the two walking methods could have been one reason why accuracy was marginally better for walking on the treadmill than on the ground.

To find out why there was a difference in walking speed between the two evaluation methods, at the end of the study, each participant was asked to walk at their measured PWS 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 clear reference point. 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 [1,52] associated with treadmill-based evaluations that cause participants to walk differently. Furthermore, people walk on the ground regularly and therefore have more experience with ground walking than treadmill walking. The participants in the experiment were given time to familiarise with walking on the treadmill before the tasks were carried out. The observations from *User Study 5* suggest the treadmill approach should be used cautiously for assessing the effects of walking on mobile interactions. The ground walking method gives a better approximation of PWS than using treadmills and should be used if natural walking speed is an important factor in future mobile studies.

Despite the inconsistency in PWS between the two evaluation techniques, both methods are suitable to use to examine the effects of walking and encumbrance if extra care is taken when planning user studies. The treadmill approach is appropriate if limited space is available to setup a walking route indoors. Also, no additional effort is required from an experimenter to act as a pacesetter to control each participant's walking speed. The participants walk at a consistent pace during the experiment without variation in walking speed. On the other hand, the ground walking method requires an experimenter to act as a pacesetter to control each participant's walking speed. It is a challenging task for the pacesetter to walk at the required walking speed consistently for each participant across all the conditions. Furthermore, training is required for the pacesetter to walk at a range of walking speeds, which substantially increased experimental time and effort. The participants might also struggle to keep in-step with the pacesetter. The participants in *User Study 5* were able to maintain walking speed with the pacesetter, except for a few minor instances where the participant slowed down to avoid drifting out of the path.

In terms of using the evaluation approaches to examine the effects of encumbrance while walking, a potential issue with the treadmill approach is the restricted space due to the safety sidebars. In *User Study 5*, it was ensured that carrying the bag and the box while walking on the treadmill would not cause the participants any unnecessary inputting 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 objects (like carrying bags in both hands in *User Study 3* and *User Study 4*) and new encumbrance scenarios that require more complex movements such as pushing objects (e.g. a pushchair). There is no such problem with the ground walking approach as the user is not restricted in upper body movements and has more space to carrying the objects and interact with the device at the same time.

To answer research question *Q2.1* (*How do encumbrance and mobility affect tapping performance on touchscreen mobile phones?*), the results from *User Study 5* supports the findings from *User Study 2* and *User Study 3*, which also measured tapping performance while users were walking and encumbered. The results from the user study presented in this chapter showed that carrying the objects reduced tapping accuracy while encumbrance had less of an effect on targeting speed. Reducing PWS by 20% only resulted in a small increase in target accuracy when compared to walking 100% of PWS but no significant effect was observed for selection time. Walking faster at 120% of PWS marginally reduced accuracy when compared to walking normally at 100% of PWS and again, no significant difference in target selection time was found. In general, walking and carrying the objects caused poor targeting performance.

To answer research question *Q3* (*How to evaluate the effects of encumbrance and mobility?*), the results showed that there was little difference in terms of targeting performance between the two mobile evaluation methods. However, the treadmill approach has two main limitations. Firstly, it was more difficult for the participants to accurately judge their normal PWS on the treadmill than walking on the ground. And secondly, the restricted space with treadmills limits the types of encumbrance scenarios that could be evaluated. Whenever possible, it is more suitable to use the ground walking approach to avoid these issues. In addition, walking on the ground maintains a certain level of ecological validity since visual attention is required for navigating and keeping within the path. With the treadmill approach, full attention is used for interaction, which is too ideal when compared to real world situations. If the ground walking method is used, it is important to control the participant's walking speed to isolate the effects of mobility from encumbrance and avoid a trade-off between walking speed and interaction, which will give a more accurate comparison of input performance between walking unencumbered and walking while carrying objects.

7.5 Conclusions

To conclude, Chapter 7 has answered research questions *Q2.1* and *Q3* by presenting a user study that examined tapping performance in a range of encumbrance scenarios, which required users to carry either a bag in-hand or a box underarm and compared the appropriateness of two main mobile evaluation approaches: treadmill and ground walking. The results from *User Study 5* showed that in general, target accuracy was poor as no mean accuracy when walking and encumbered was greater than 51%. Reducing PWS by 20% only improved accuracy by a small margin, which suggests walking speed needs to be greatly reduced further to allow users to maintain a standard level of targeting performance.

The results also showed that the participants performed marginally better in terms of higher target accuracy and quicker selection times when walking on the treadmill than when walking on the ground. However, the participants walked slower on the treadmill than on the ground and perhaps this was one reason why there was a difference in performance between the two approaches. *User Study 5* has also highlighted the issues with both walking methods but recommends using the ground walking evaluation approach for future studies looking to examine the effects of encumbrance and mobility.

The research discussed in this and the previous three chapters have all examined targeting performance in a range of encumbrance scenarios and the results have given a comprehensive understanding of how touch input is affected in these physically demanding contexts. The next chapter discusses ways to improve usability when users are walking and encumbered by defining an appropriate target size that is likely to increase selection accuracy and evaluate the effectiveness of non-standard input techniques.

Chapter 8

User Study 6 and 7: Improving Input Performance While Walking and Encumbered

8.1 Introduction

The user studies presented in the previous four experimental chapters of this thesis have all examined abstract targeting performance in a range of different encumbrance scenarios. The results from these studies have shown how input deteriorates on touchscreen mobile phones in terms of poor selection accuracy, especially for one-finger tapping and dragging actions, when bags and boxes were held while on the move. Therefore, more effective user interfaces and interaction techniques are required to improve input performance and usability with handheld devices in walking and encumbered situations. This chapter answers research questions *Q4.1* (*What are the appropriate target sizes and target placements for encumbered and walking interactions?*) and *Q4.2* (*Can pointing techniques improve targeting performance while walking and encumbered?*) by discussing two main approaches.

The first approach defines a target size, based on the results from *User Study 2 - 5*, to minimise inaccurate tapping selections in walking and encumbered situations. Previous research (e.g. [35,59,66]) has reported an improvement in accuracy as target size increase. However, further design considerations are required when altering user interfaces on small touchscreens such as those found on mobile phones. Therefore, the implications of using large on-screen targets are discussed, with a use-case frequently performed on mobile devices as an example.

The second approach examines the effectiveness of one particular pointing technique in walking and encumbered contexts. Increasing target size to enhance usability might not always be possible (e.g. when many targets are placed close to each other), so alternative interaction techniques have been developed to increase targeting accuracy when tapping on small on-screen elements. However, there is no empirical work to suggest whether established input techniques can maintain their performance gains in real world contexts, such as when walking and encumbered, since the usefulness of these methods have only been tested in ideal non-mobile settings. Therefore, one pointing technique was examined to see how well it performed when users experience extraneous movements caused by walking and carrying objects. Later, an explanation is given as to why *Shift* [73] was the pointing technique evaluated in the first user study presented in this chapter.

This chapter presents the results and discusses the findings from two user studies. The first study (*User Study 6*) had three purposes, which were to measure: (1) the results from increasing target size, (2) the performance of an alternative user interface when compared to a standard comparable interface and (3) the effectiveness of *Shift*, while users were walking and carrying bags in both hands. The second study (*User Study 7*) was carried out as a follow-up experiment to understand and explain an anomaly in the results from *User Study 6*.

8.2 Defining an Appropriate Target Size

The four user studies discussed in Chapters 4 - 7 have all measured input performance and showed poor target accuracy when users were walking and holding bags and boxes. In *User Study 2* (Chapter 4), walking and carrying a bag in the dominant hand caused the lowest mean accuracy of 29.8% when selecting 4 *x* 6mm targets. The targeting task was altered for *User Study 3* (Chapter 5) and *User Study 5* (Chapter 7), where each target measured 5 *x* 8mm. In *User Study 3* when participants were walking and carrying a bag in both hands, the lowest mean accuracy of 48% occurred in the two-handed index finger input posture. In *User Study 5*, walking at 100% of PWS and carrying a bag in the dominant hand caused a mean accuracy of 36%. These three studies have revealed the low tapping accuracy when selecting those particular target sizes. A question then arises, *what is an effective target size to improve accuracy in walking and encumbered contexts?*

In *User Study 4* (Chapter 6), three target widths (5.0, 7.5 and 10.0 mm) were examined in the tapping Fitts' Law task. The number of inaccurate selections decreased as target width increased, as illustrated in Figure 8.1. The largest target width of 10mm had a mean accuracy of 70.7%, which shows wider targets are required to improve accuracy further. Linear regression showed a strong correlation between increasing target width and accuracy (R^2 = 0.98) and approximates a target width of 13.5mm to achieve 100% accuracy. A study could have been carried out to examine whether (1) the approximated value does indeed achieve perfect accuracy or (2) systematically examine a range of target sizes to find the width in which the number of incorrect selections level off when users held objects while on the move. However, since numerous studies presented in this thesis have already examined tapping performance while walking and encumbered, the *Prediction Interval* [69] method was used to define a target width by using the targeting error results.

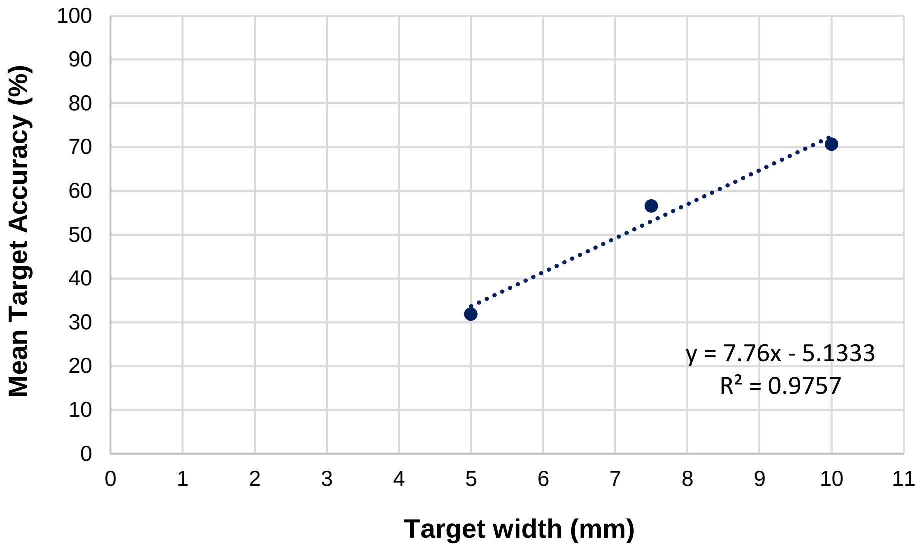

Figure 8.1: The mean target accuracy (%) for the three target widths (5, 7.5, 10mm) tested in the Fitts' Law tapping task in *User Study 4*.

Prediction Interval (*PI*) is a statistical analysis method used to estimate the range in which a future value will fall, based on a chosen probability and a sample already obtained. *PI* is calculated as follows:

$$\boldsymbol{\overline{X} \pm t_{(\alpha/2,n-1)}\, s \sqrt{1+\frac{1}{n}}}$$

Where,

$\boldsymbol{n}$ = number of samples

$\boldsymbol{\overline{X}}$ = sample mean

$\boldsymbol{s}$ = sample standard deviation

$\boldsymbol{t}$ = t distribution value

Targeting error (the absolute distance from the centre of the target to the touch up position) measured in *User Study 2, 3 and 5* gave an indication as to how imprecise input became when participants were walking and either carrying bags or boxes. By using the errors from the walking and encumbered conditions gathered in *User Study 2, 3* and *5* as the sample to calculate the range in which tapping is likely to lie within, a suitable target width can be created to increase selection accuracy. The data from the Fitts' tapping task in *User Study 4* was not included in the calculation since *User Study 6* precedes it. In addition, only the errors measured in the two-handed index finger posture in *User Study 3* was included in the calculation since error was heavily biased by difficult to reach targets when the one-handed preferred thumb posture was used. A logarithmic transformation was applied to normalise the errors prior to calculating the interval. The results gave a *PI* (95%) range of 2.6 - 11.2mm. Using the upper limit of this interval, a target width of 22.4mm was defined in which interface elements based on this size would be accurately selected with a probability of 95% when users are walking and encumbered. The next section discusses the implications of implementing larger target sizes on standard user interfaces for mobile phones.

8.3 The Effects of Increasing Target Size

The previous section defined a large target size of 22.4 x 22.4mm to improve selection accuracy while users are walking and encumbered. Related work (e.g. e.g. [35,59,66]) has shown the benefits of increasing target size to improve tapping accuracy and given suggestions on more effective target dimensions. While these recommendations are valid on an abstract level, increasing the size of on-screen elements, such as buttons and icons, create other design challenges that are often not discussed when developing user interfaces for small touchscreen devices. Therefore, this section discusses the design considerations of increasing the size of on-screen targets with one particular use-case example.

The design challenge undertaken is to do with the way mobile applications (app) are arranged and organised on mobile phones. Apps have become a new means and a necessity to access services on mobile devices (such as fitness trackers, games and sharing information on social media). The popularity and diverse range of apps allow users to have a great number of functionalities stored on their mobile devices, so the arrangement and layout design of apps are important. The way apps are accessed and arranged on mobile devices depend on their operating system. With Apple devices running iOS8[13], existing apps can be accessed directly on the homescreen while new apps (once downloaded) are ordered by installation date and placed to the next available position on the last screen. Mobile devices running the latest version of the standard non-customised Google Android[14] require an Application/Task Launcher to access the current list of apps, which are normally ordered alphabetically. Currently, both mobile operating systems visually arrange the apps in a 4 *x* 5 grid, excluding the default quick access apps at the bottom of the screen in iOS, as shown in Figure 8.2.

Figure 8.2: The arrangement of apps in standard Google Android (left) and Apple iOS (right).

The design of both arrangements has one common problem: finding particular apps can become a time consuming task as the number of apps increases, which also means that more screens are required to hold the apps and creates an additional cost in time of switching between them. This issue is likely to increase further in walking situations when full visual attention is not available for input. An interesting question then arises: *can alternative app arrangements that reduce task time while improving selection accuracy by using larger on-screen targets be achieved?* The following systematic approach tries to address this design challenge to improve input for walking and encumbered contexts. For consistency with the user studies discussed so far in this thesis, the design of the alternative app arrangement was implemented in Google Android. The arrangement ran on a

[13] Access date January 2015
[14] Android Lollipop - January 2015

Samsung Galaxy S3, which has a touchscreen resolution of 720 *x* 1280px (59.7 *x* 106.2mm), approximately 12.05px/mm.

Consider the example of the standard app arrangement on the S3 as shown in Figure 8.3. The apps are aligned in 4 *x* 5 grid and take up appropriately 720 *x* 1000px (59.7 x 83mm) of the screen. The target area for each app is therefore 180 *x* 200px (14.9 *x* 16.6mm). If the area for each app increases to the proposed dimensions of 270 *x* 270px (22.4 *x* 22.4 mm) as discussed previously, each screen can no longer hold apps in the 4 *x* 5 layout. A maximum of six 22.4 x 22.4mm target areas can be placed within the arrangement (two horizontally and three vertically). As a result, four screens are required to hold 20 apps compared to the standard layout which only needs one screen. To reduce the amount of screen switching in the alternative arrangement, the remaining unused screen area was utilised more effectively.

Figure 8.3: The app arrangement in the Samsung S3 Android mobile phone. The dimensions taken up by the arrangement are shown by the red outline.

Consider two 22.4 x 22.4mm target areas placed beside each other on a particular row as shown in Figure 8.4. Since no more large targets areas can be horizontally placed on the row, there is an unused area of 180 *x* 270px (14.9 *x* 22.4mm). To fully utilise the remaining space, the highest common denominator of the width and height of the unused area was calculated. As a result, smaller target areas of 90 *x* 90px (7.5 *x* 7.5mm) were created and exactly six of them could be placed in a 2 *x* 3 grid within the remaining space, as shown in Figure 8.5. *User Study 2 - 5* showed that it was difficult to select small targets accurately when walking and encumbered, so a pointing technique was implemented to help reduce the number of errors. The selected pointing technique will be discussed in more detail later in this chapter.

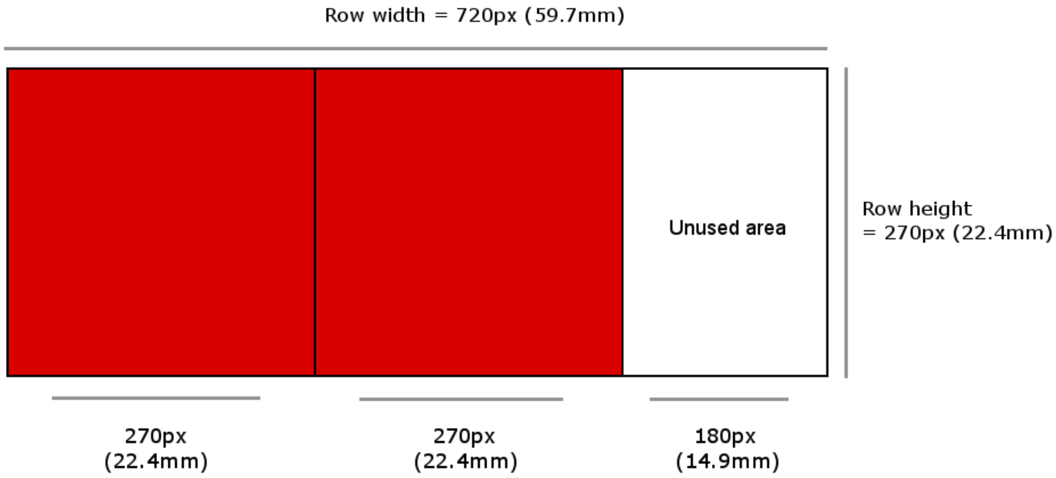

Figure 8.4: An example row of two 22.4 x 22.4mm target areas (red) placed beside each other.

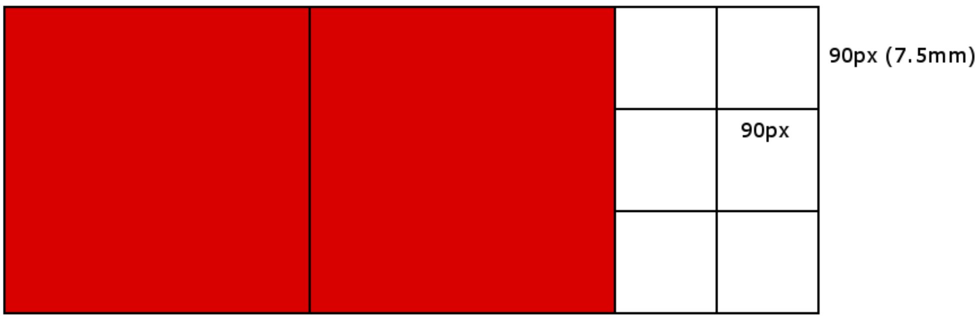

Figure 8.5: The remaining space on a row with 22.4 x 22.4mm target areas utilised with 7.5 x 7.5mm targets.

As previously mentioned, the arrangement can afford three rows containing 22.4 x 22.4mm target areas which meant there is an unused region of 720 *x* 190px (59.7 *x* 15.8mm), as shown in Figure 8.6. To avoid creating a third target size, the smaller 90 *x* 90px (7.7 *x* 7.5mm) targets were used to exploit the remaining space available in the arrangement. As a result, eight small target areas can be placed in a row and two rows can be located within the remaining region as illustrated Figure 8.7. An area of 720 *x* 10px (59.7 *x* 0.8mm) was left unused at the bottom of the layout. The example described here illustrates one possible arrangement, which meant the rows of smaller targets could be placed at the top of the layout or in between rows that contain the larger 22.4 x 22.4mm targets.

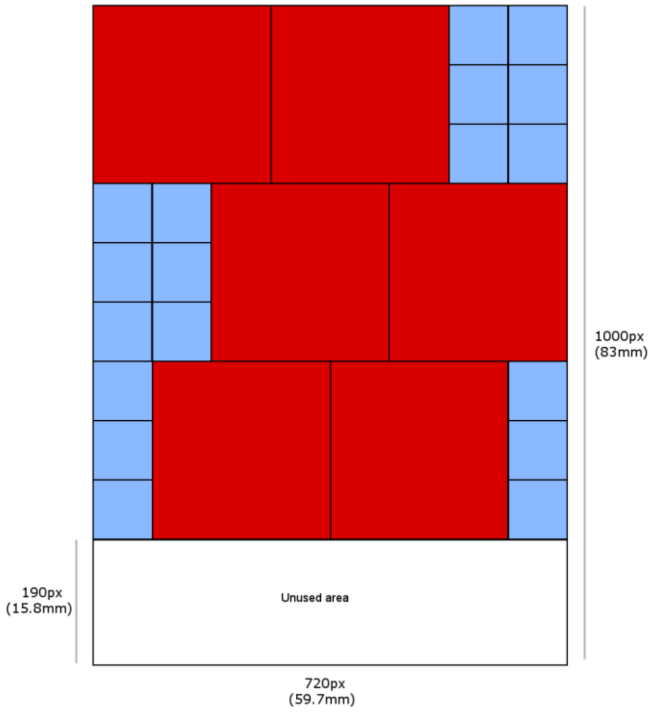

Figure 8.6: One example of arranging a maximum of six 22.4 x 22.4mm target areas. The unused space at the bottom of the arrangement.

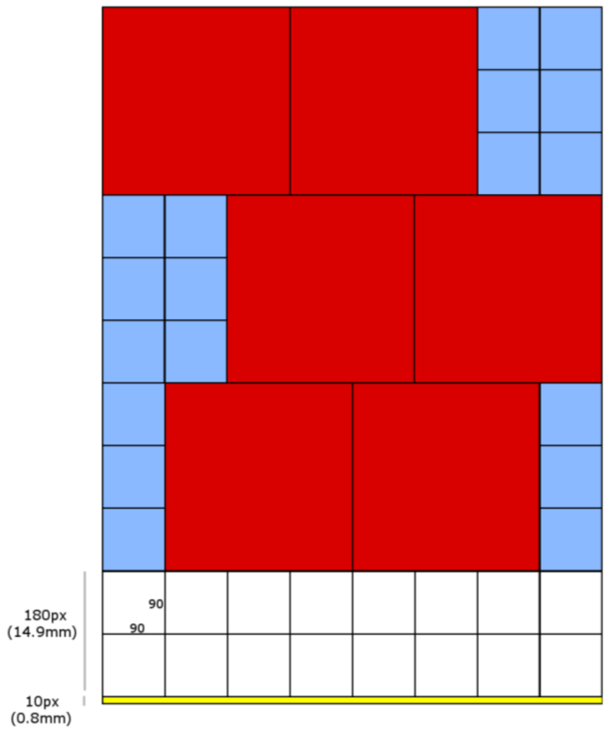

Figure 8.7: One example of a full alternative arrangement. The yellow area remained unused.

The methodical design of the alternative arrangement meant six 22.4 x 22.4mm targets and 34 smaller 7.5 x 7.5mm targets can be placed within the layout. The arrangement can therefore hold a total of 40 apps per screen, double the amount of apps held in the standard layout. An interesting comparison can be made in selection time between placing more apps per screen against spreading them uniformly across several screens when users are walking and encumbered. Before describing the user study which was conducted to make this comparison and evaluate the effectiveness of the alternative arrangement, two issues require further discussion: (1) the practicality of the non-standard layout and (2) the selection of the small 7.5 x 7.5mm targets.

One of the main goals of the asymmetric arrangement is to improve selection accuracy using larger target areas while designing a more efficient interface layout that utilises the available screen space as much as possible. Each screen can hold six 22.4 x 22.4mm targets at most, so the question then arises as to which apps should be held in those areas. Without going into detail about behavioural patterns with app organisation on mobile phones, which is beyond the scope of this thesis, one way to choose could simply be down to the user's personal preference. For example, there is likely to be particular apps that are used most often, so those could be placed in the larger target areas. Similarly, an automated approach could be implemented where the apps that are frequently opened are automatically enlarged while those apps that not regularly used are held in the smaller target areas. There are limitations to the alternative arrangement, which are discussed later in the chapter. But one of the main purposes of the layout is to show a concept, which examines the trade-offs between implementing larger on-screen elements and its effects on user interface design.

The second issue is concerned with selecting apps placed in the smaller 7.5 x 7.5mm targets. In *User Study 4*, the mean accuracy for selecting the 7.5mm target widths in the Fitts' Law tapping task was 57% when users were walking and carrying bags in both hands. To reduce tapping errors on the smaller target areas, a benchmark pointing technique was implemented to assist the user's input while encumbered and on the move. The next section explains why *Shift* [73] was the pointing technique chosen.

8.4 Pointing Techniques

Previous work has examined different ways to enhance target selections with different input modalities. For example, Grossman and Balakrishnan's *Bubble Cursor* [25] showed an improvement over standard cursor control via mouse input. Later, Mott and

Wobbrock's *Bubble Lens* [51] addressed the limitations of *Bubble Cursor* when selecting very small targets (<10px) and reported an improvement in target accuracy and selection time. While the mentioned input techniques were originally designed to enhance pointing via mouse input, they can be translated for finger input on touchscreens as shown by Au *et al.* [2], who compared the performance of their enhanced target selection technique *LinearDragger* to those of *Bubble Cursor* and other pointing techniques such as *Escape* [79] and *Shift* [73]. *LinearDragger* attempts to improve the selection of small, densely packed targets on touchscreens. Instead of directly tapping on targets, users simply drag in specific directions to select targets without worrying about precise touch control. Yatani *et al.*'s touch-based input technique *Escape* [79] operates similarly to *LinearDragger* but visual colour-coded icon cues were used to aid the gesture direction to select targets.

While both Au *et al.* and Yatani *et al.* showed that their input techniques perform better than Vogel and Baudisch's *Shift* [73] in terms of target selection time while accuracy was similar between all three methods, *Shift* was the method used to improve the selection of the smaller targets in the alternative arrangement as described above. *Shift* was selected over *Escape*, *LinearDragger* and other input techniques despite slower selection times because tapping behaviour is direct and unaltered. Only the area occluded by the input finger is visually placed to a non-occluded area of the screen by using a circular callout. Furthermore, the selection of individual pixel elements is unaffected with *Shift*, whereas it becomes a problem for indirect techniques such as *LinearDragger*. A cancellation mechanism is also unnecessary for *Shift* since on-screen adjustments can be made to correct the initial inaccurate selection (if required), which is ideal for walking and encumbered contexts.

The implementation of *Shift* was kept similar to the original design by Vogel and Baudisch [73] with a few alternations after pilot tests. The diameter of the circular callout was marginally increased to 29mm from the original size of 26mm. Whenever possible, the callout was placed 22mm above the initial touch position, the same as the original design. One observation during pilot testing was the difference in angle of the input finger between standing and walking when *Shift* was used, as shown in Figure 8.8. The input finger was almost perpendicular to the vertical axis of the phone when walking, compared to a more parallel position when stationary. This can cause visual problems of occluding the callout by the input finger when manipulating the top region of the touchscreen, as shown in Figure 8.9(a). To reduce this issue as much as possible, the callout was placed below the touch position when interacting with the top part of the screen, as shown in Figure 8.9(b).

The distance from the touch position to the centre of the callout was doubled to 44mm. As the finger moved towards the left edge of the screen, the callout gradually moved to the right while keeping vertical position. The opposition callout movements occurred as the finger approached the right edge of the screen.

A cursor was also placed at the centre of the callout to aid target selection. The design of cursor depended on the size of the target areas, as shown in Figure 8.10. For the small 7.5 *x* 7.5mm targets, a small red circle (2mm in diameter) showed the current on-screen position and a green border highlighted the app currently selected. For the large 22.4 *x* 22.4mm target areas, a green crosshair (2.5mm in length) replaced the red circle and no border was used to illustrate the current app. In the standard arrangement, the same cursor was used as selecting the 7.5 *x* 7.5mm target areas in the alternative arrangement, except a larger red cross-hair (2.5mm in length) replaced the circle.

Figure 8.8: The observed angle of the input finger when standing (left) and walking (right).

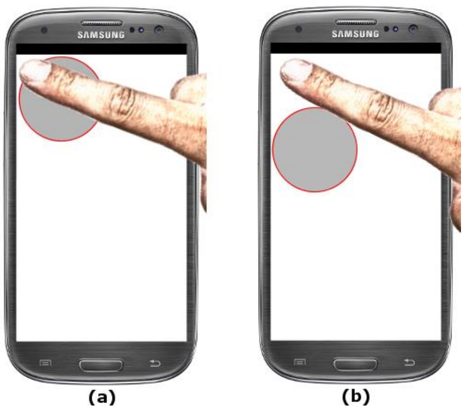

Figure 8.9: The occlusion of the callout due to the input angle of a right-handed user (a). The callout was therefore placed lower down when manipulating the top region of the screen to reduce this issue as much as possible (b).

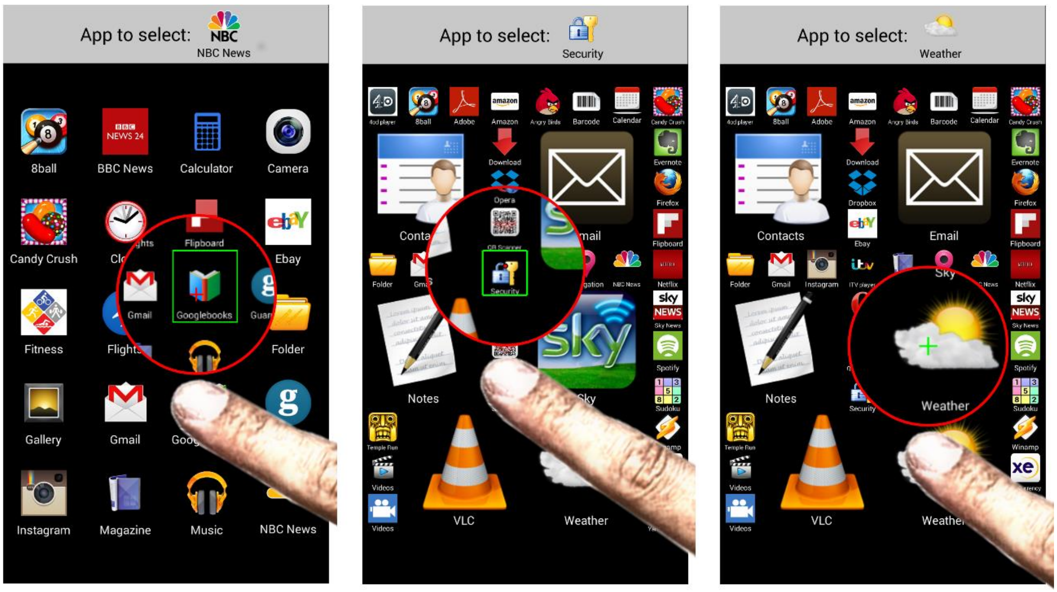

Figure 8.10: The different types of cursors used in the standard arrangement (left) and selecting the small 7.5 x 7.5mm targets (middle) and large 22.4 x 22.4mm targets (right) in the alternative arrangement.

8.5 User Study 6

This section discusses *User Study 6*, the first of two studies presented in this chapter, which compared the effectiveness of the alternative app arrangement to the baseline standard layout and examined the performance of *Shift* while users were walking and encumbered.

8.5.1 Method

8.5.1.1 Encumbrance Scenario and Walking Environment

Similar to *User Study 6*, *Encumbrance Scenario 1C* (see Figure 6.6) was evaluated to replicate situations where both hands are encumbered during input. The participants held a 1.6kg shopping bag, which measured 330 *x* 480mm (w *x* d), in each hand.

Like *User Study 3*, *4 and 5*, a 20m long and 2m wide oval route was setup in a spacious room for the participants to walk. Each participant maintained their Preferred Walking

Speed during Interaction (PWS&I) by walking alongside a pacesetter, as described in Section 5.2.3. PWS&I was measured before the main experiment began as each participant walked around the path while performing the abstract targeting task on a Samsung Galaxy S3 phone, as described in Section 5.2.2. All participants performed the task in the two-handed index finger input posture. As explained in previous chapters that also used the pacesetter approach, walking speed was controlled to isolate the effects of mobility from encumbrance. Therefore, the changes observed in tapping performance while walking are most likely due to carrying the bags.

8.5.1.2 Task

The standard layout on the S3 was used as the baseline comparison to examine the effectiveness of the alternative app arrangement. The apps on each screen in the standard layout were uniformly aligned in a 4 *x* 5 grid, so each target area measured 180 *x* 200px (14.9 *x* 16.6mm). The small and large target areas in the alternative arrangement were 90 *x* 90px (7.5 *x* 7.5mm) and 270 *x* 270px (22.4 *x* 22.4mm) respectively, as described above in Section 8.2. The design of the alternative arrangement meant 40 apps can be placed on each screen, which is double the amount that can be held in the standard layout. Therefore, 40 apps were placed on one screen in the alternative arrangement while 20 apps were placed over two screens in the standard arrangement.

The apps placed in each arrangement were selected from the Google Play Store[15]. A set of 80 apps was pseudo-randomly chosen based on popularity, most downloaded and top rated criteria. For each participant and each experimental condition, the apps were arbitrarily chosen from the set to reduce learning effects. Once the apps have been selected, they were placed in each arrangement in alphabetical order from top left to bottom right of each screen-page (see Figure 8.12). The design of the alternative arrangement meant there are numerous ways in which the small and large target areas can be organised. To keep the number of experimental conditions to a feasible amount, one arrangement was arbitrarily selected, as shown in the middle and right images of Figure 8.10.

For each trial, the screen first displayed the app (icon and name) to select. Once the participant tapped on the screen, the trial started and the participant was asked to find and

[15] Google Play: https://play.google.com/store

select the current app as quickly and accurately as possible. The current app to select was also displayed at the top of the screen (above the arrangement area) once the trial had started as a reminder. The task ran on a Samsung Galaxy S3 phone and all participants held the device in the non-dominant hand while only using the index finger of the dominant hand to select the targets.

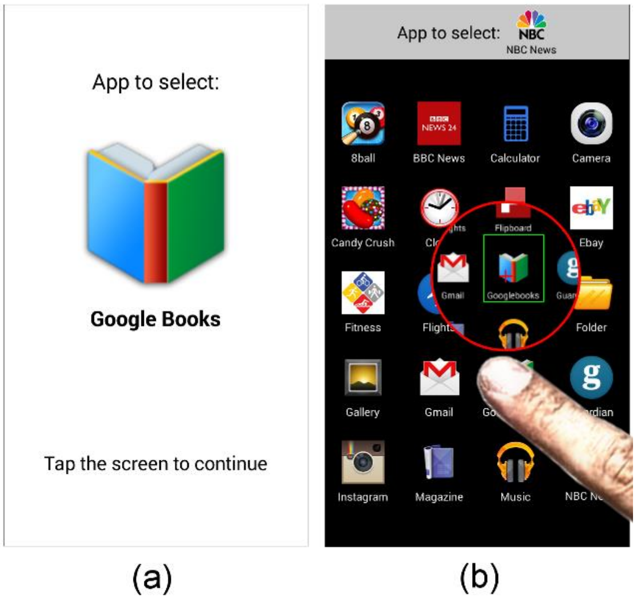

Figure 8.11: The task completed in *User Study 6*. Each trial began by instructing which app to select (a). Once the screen was tapped, the arrangement was presented to the user to find and select the app (b).

8.5.1.3 Experimental Design

Eighteen students (12 males, 6 females) aged between 20 - 44 years (mean = 26.44, SD = 0.06) recruited from the University took part in the study. All participants owned and used a touchscreen mobile phone on a daily basis and preferred using their right hand for input. The study took approximately 60 minutes to complete and each participant was paid £6 for taking part. Resting periods were given between conditions and when required to reduce fatigue.

There were two types of arrangements: standard (*STA*) and alternative (*ALT*) and two input techniques: unassisted touch (*Touch*) and *Shift*. As a result, there were four conditions in total and each condition was completed while walking at the measured PWS&I and carrying a bag in each hand. A within-subject design was used and the conditions were counter-balanced by type of arrangement and further randomised by input technique to reduce learning and order effects as much as possible. The number of trials for each condition is explained as follows:

STA_{Touch}: In the standard arrangement on the S3, the apps are aligned in a 4 *x* 5 grid. Therefore, each screen can hold a maximum of 20 apps. In this condition, two screens were created, giving a total of 40 unique apps. Each app was selected twice, giving a total of 80 target selections per condition.

STA_{Shift}: To keep selection time comparable with the other conditions, *Shift* was activated immediately once a touch down event was detected after the trial began. However, this design choice meant that swiping between screens was not possible. Therefore, the performance of *Shift* was only measured on one screen in the standard arrangement. Each of the 20 apps was selected twice, resulting in 40 target selections per condition.

ALT_{Touch}: As explained above in Section 8.3, the design of the alternative arrangement meant each screen can hold a maximum of 40 apps. For this condition, one screen was created and each app was selected twice, which resulted in a total of 80 target selections per condition.

ALT_{Shift}: Same procedure as ALT_{Touch}, but *Shift* was used to select the targets instead of standard touch input. Like STA_{Shift}, *Shift* was activated immediately after a touch down event was detected.

The Independent Variables were **Target Size** (three levels: 7.5 *x* 7.5, 14.9 *x* 16.6 and 22.4 x 22.4 mm) and **Input Technique** (two levels: *Touch* and *Shift*). The three target sizes will be referred to as "*Small*" (7.5 *x* 7.5mm), "*Medium*" (14.9 *x* 16.6mm) and "*Large*" (22.4 x 22.4 mm). The Dependent Variables were accuracy (%) and selection time (ms). An app was successfully acquired if the touch up position was within its target area. Selection time was the duration between the first touch up event (to start the trial) and second touch up event (to selected the app). The hypotheses were:

H1: *Small* targets will be selected significantly less accurately than *Medium* targets;

H2: *Small* targets will be selected significantly slower to select than *Medium* targets (on the first screen only in *STA*);

H3: *Large* targets will be selected significantly more accurately than *Medium* targets;

H4: *Large* targets will be selected significantly quicker than *Medium* targets (on the first screen only in *STA*);

H5: *Small* targets will be selected significantly less accurately than *Large* targets;

H6: *Small* targets will be selected significantly slower to select than *Large* targets;

H7: Target accuracy for *Shift* will be significantly higher than *Touch*.

H8: Selection time for *Shift* will be significantly slower than *Touch*.

H9: Target accuracy will be significantly increased when selecting the *Small* targets using *Shift* than unassisted touch input.

H10: Using unassisted touch input, selecting the targets on the alternative layout will be significantly faster than selecting the targets placed on the second screen in the standard layout.

8.5.2 Results

Each participant completed two blocks of target selections per condition, which meant a total of 10,080 trials (*280 selections x 2 blocks x 18 participants*) was recorded for *User Study 6*. To remove trials that were selected accidentally, selections that took less than 300ms were not included in the final data analysis. As a result, nine trials were removed. The error bars in the figures shown in this section represents confidence interval (95%) of the mean.

8.5.2.1 Target Accuracy

Shapiro-Wilk tests were performed to examine the distribution of target accuracy for each condition and the results are shown in Table 8.1. Only one condition was not significant: unassisted tapping on targets placed on the first screen in the standard arrangement. Therefore, target accuracy was transformed using ART [76] before conducting a two-factor (**Target Size** and **Input Technique**) repeated-measure ANOVA. Only the accuracies of targets on the first screen in STA_{Touch} were included in the test. The mean accuracy for each condition is shown in Figure 8.13.

Target Size	Screen Page	Input Technique	W Statistic	Sig.
Standard	Page1	Touch	0.909358	0.083822
Standard	Page2	Touch	0.868014	0.016504*
Standard	Page1	Shift	0.536804	0.000002*
Small	Page1	Touch	0.790768	0.001128*
Small	Page1	Shift	0.805752	0.001833*
Big	Page1	Touch	0.613631	0.000009*
Big	Page1	Shift	0.456961	0*

Table 8.1: Shapiro-Wilk normality tests performed on target accuracy for each condition in *User Study 6*. Significant results are shaded in grey and highlighted with '*'.

A significant main effect was observed for **Target Size**, $F(2, 85) = 11.795$, $p < 0.01$. *Post hoc* Tukey HSD comparisons showed that accuracy for the *Small* targets was significantly lower than both the *Medium* targets ($t = 2.537$, $p < 0.05$) and *Large* targets ($t = 4.855$, $p < 0.001$). There was no significant difference in accuracy between the *Medium* and *Large* targets ($t = 2.318$, $p > 0.05$). A significant main effect was found for **Input Technique**, $F(1, 85) = 24.754$, $p < 0.01$. Target accuracy for *Shift* was significantly higher than *Touch*. The interaction between the two factors was significant, $F(2, 85) = 4.589$, $p < 0.05$. *Post hoc* Tukey HSD tests showed that no pairwise comparisons were significant. The interaction effect is not required to support or reject the hypotheses on target accuracy.

A Wilcoxon signed-rank test was conducted to compare target accuracy between the two input techniques on the *Small* targets. There was a significant difference in accuracy between *Touch* and *Shift* ($Z = -2.485$, $p < 0.05$). Selecting the *Small* targets with *Shift* was significantly more accurate than *Touch*.

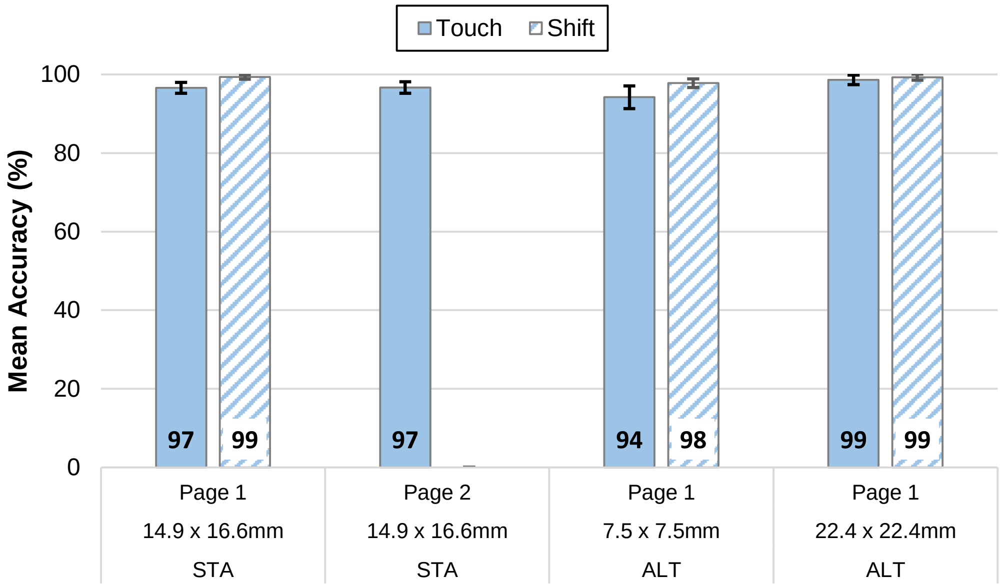

Figure 8.12: The mean selection accuracy (%) for each condition in *User Study 6*. The solid and striped bars represent *Touch* and *Shift* respectively. Note - accuracy is separated between the first and second screens in STA_{Touch}. Error bars denote Confidence Interval (95%).

8.5.2.2 Selection Time

Shapiro-Wilk tests were conducted to assess the distribution of the selection time data for each condition is *User Study 6*. The results are shown in Table 8.2 and no condition was significant therefore normality was not violated. The mean selection time for each condition is shown in Figure 8.14. A two-factor (**Target Size** and **Input Technique**) repeated-measure ANOVA was conducted for selection time. Only the selection times of targets on the first screen in STA_{Touch} were included in the ANOVA test. Huynh-Feldt adjustments were used to correct the degrees of freedom since Mauchly's test for sphericity was significant ($\varepsilon > 0.75$).

A significant main effect was observed for **Target Size**, $F(1.8, 30.7) = 22.087$, $p < 0.05$. *Post hoc* pairwise comparisons with Bonferroni corrections show that selection time took significantly longer for the *Small* targets than the *Medium* targets, a mean difference of 523.1ms. The *Medium* targets were selected significantly quicker than the *Large* targets, a mean difference of 322.3ms. There was no difference in selection time between the *Small* and *Large* targets. There was a significant main effect for **Input Technique**, $F(1, 17) =$

6.582, $p < 0.05$. Selection time for *Shift* was significantly slower than *Touch*, a mean difference of 235.8ms. The interaction between the factors was not significant, $F(1.2, 19.7) = 0.408$, $p > 0.05$. The interaction effect between the factors is not necessary to support or reject the hypotheses on selection time.

Target Size	Screen Page	Input Technique	W Statistic	Sig.
Standard	Page 1	Touch	0.93653	0.252528
Standard	Page 2	Touch	0.903701	0.066678
Standard	Page 1	Shift	0.979955	0.949543
Small	Page 1	Touch	0.896189	0.049334
Small	Page 1	Shift	0.966185	0.723664
Big	Page 1	Touch	0.944996	0.351977
Big	Page 1	Shift	0.919917	0.128835

Table 8.2: Shapiro-Wilk normality tests performed on selection time for each condition in *User Study 6*.

To compare the selection time between placing more apps per screen in the alternative arrangement and spreading the apps over two screens in the standard arrangement, a one-factor, repeated-measures ANOVA was conducted. There were four levels: STA_{Touch} on the first screen (*Medium*), STA_{Touch} on the second screen (*Medium*), ALT_{Touch} (*Small*) and ALT_{Touch} (*Large*). Greenhouse-Geisser adjustments were used to correct the degrees of freedom since Mauchly's test for sphericity was significant ($\varepsilon < 0.75$).

A significant main effect was observed, $F(1.7, 28.7) = 17.085$, $p < 0.05$. *Post hoc* pairwise comparisons with Bonferroni corrections showed that selection time for the *Medium* targets on the second screen in STA_{Touch} was significantly slower than both the *Small* targets (mean difference of 598ms) and *Large* targets (mean difference of 709.7ms) in ALT_{Touch}. The *Medium* targets on the first screen in STA_{Touch} were selected significantly faster than the same sized targets placed on the second screen, a mean difference of 1075.7ms.

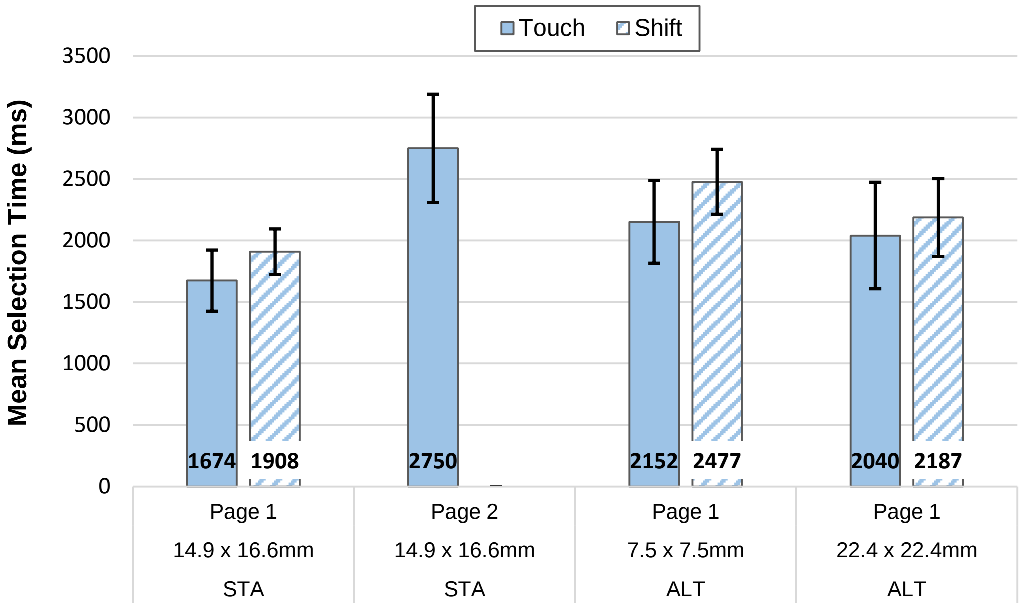

Figure 8.13: The mean selection time (ms) for each condition in *User Study 6*. The solid and striped bars represent *Touch* and *Shift* respectively. Note - selection time is separated between the first and second screens in STA_{Touch}. Error bars denote Confidence Interval (95%).

8.5.2.3 Walking Speed, Distance Covered and Completion Times

A paired t-test was conducted to compare the mean walking speed (km/h) between PWS (no interaction and unencumbered) and PWS&I. There was a significant difference between the two types of walking speeds; $t(17) = 8.947$, $p < 0.05$. PWS&I (mean = 3.60, SD = 0.46) was significantly slower than PWS (mean = 4.39, SD = 0.43), a decrease of 17.3%. The approximated distance walked and completion time for each condition is shown in Table 8.3.

	STA_{Touch}	STA_{Shift}	ALT_{Touch}	ALT_{Shift}	**Total**
Estimated mean distance walked (metres)	484.0	213.1	444.9	489.4	**1631.4**
Mean condition completion time (mins)	8.12	3.54	7.44	8.14	**27.3**

Table 8.3: The estimated mean distance walked (metres) and mean task completion time (minutes) for each condition in *User Study 6*. The condition completion times do not include resting periods. Note - distance walked and completion time of STA_{Shift} is less than the other conditions because there were less trials to complete.

8.5.3 Discussion

The results from *User Study 6* showed that accuracy for the *Small* targets were significantly lower than the *Medium* targets. Therefore, hypothesis *H1* is supported. The overall mean target accuracy between the *Small* and *Medium* targets was only 3.2%, as participants selected the *Small* targets much better than expected. As expected, the number of accurate selections was significantly less for the *Small* targets than the *Large* targets. Therefore, hypothesis *H5* is supported. However, the difference in overall mean target accuracy between the *Small* and *Large* targets was again small at 4.2%. No statistical difference was found for target accuracy between the *Medium* and *Large* targets, therefore hypothesis *H3* is rejected. Target selection was significantly more accurate when *Shift* was used for input than *Touch*. Therefore, hypothesis *H7* is supported. The comparison between the two input techniques on the *Small* targets showed than *Shift* was significantly more accurate than *Touch*. Thus, hypothesis *H9* is supported.

One unexpected result from the study was the surprisingly high accuracy of 94.2% for selecting the *Small* targets in ALT_{Touch} when participants were walking and carrying bags. *User Study 2, 3* and *5* have all shown poor target accuracy and in the Fitts' Law tapping task of *User Study 4*, the overall accuracy for selecting the 7.5mm target width was much lower at 57%, where participants were encumbered in the exact same way as *User Study 6*. To find out why the participants were targeting exceptionally well on the *Small* targets when walking and encumbered, a follow-up study was conducted, which will be presented in the next section.

As predicted, accuracy was very high when selecting the *Large* targets in the alternative arrangement, whether or not *Shift* was used. The number of inaccurate selections when choosing apps placed in the *Large* target areas were less than 1% for both *Touch* and *Shift*. In the standard arrangement, the same mean accuracy of 97% was recorded for the *Medium* targets placed on both the first and second screens. Selecting the apps with *Shift* improved accuracy to 99% for targets placed on the first screen in the standard arrangement. These results show the benefits of implementing target sizes greater than 14.9mm to improve accuracy when users are walking and encumbered.

The results showed that *Shift* is still effective in terms of high input accuracy in walking and encumbered situations. Accuracy was greater than 98% across all three target sizes when *Shift* was used, although, its performance on the *Small* targets was underwhelming due to *Touch* accuracy was much higher than predicted. *Shift* was also well received as feedback from the participants suggested a preference for *Shift* over *Touch* for input. Participants commented that they felt less hurried to select the target accurately on touch down since correction can be made if the initial selection was incorrect. Furthermore, *Shift* gave users more assurance as to which target was acquired, especially when selecting the *Small* targets, due the clear visual feedback from the callout.

The results for selection time showed that the *Small* targets in the alternative arrangement were acquired significantly longer than the *Medium* targets placed on the first screen in the standard arrangement. Therefore, hypothesis *H2* is supported. However, the *Large* targets in the alternative arrangement were not selected significantly quicker than the *Medium* targets on the first screen in the standard arrangement. Therefore, hypothesis *H4* is rejected. Perhaps the *Medium* targets were big enough for the participants to search for a particular app at the same rate as the *Large* targets and therefore resulted in no significant difference in selection time. The difference in selection time between the *Small* and *Large* targets was not significant, thus, hypothesis *H6* is rejected. Hypothesis *H8* is supported since input speed for *Shift* was significantly slower than *Touch*. The *Small* and *Large* targets in the alternative arrangement were both selected significantly quicker than the *Medium* targets on the second screen in the standard layout. Therefore, hypothesis *H10* is supported.

The *Medium* targets on the first screen in STA_{Touch} took on average 1674ms to select, which was the fastest of all the conditions. The mean selection time for the *Medium* targets on the second screen in STA_{Touch} was 2750ms, just over one second more than the apps targets located on the first screen. Apps placed in the *Small* and *Large* targets in ALT_{Touch} were both selected slower than the *Medium* targets on the first screen in STA_{Touch}. As expected, the *Small* targets required more time to acquire since there were more targets in the alternative arrangement to search. The selection time of the *Large* targets was only 112ms faster than the *Small* targets in ALT_{Touch}, although the results were not significantly different. Comments from the participants indicated that there were difficulties of finding apps placed in the *Large* target areas, despite its size advantage over the *Small* targets. The apps in the alternative arrangement were placed close to each other and therefore made the interface more visually "busy".

The advantage in selection time of the *Medium* targets in the standard arrangement starts to diminish when apps are not placed on the first screen. On average, apps on the second screen took just over one second more to select than those placed on the first. The same margin in selection time likely for subsequent screens as the number of apps increases.

The *Small* and *Large* targets in the alternative arrangement were selected quicker than the *Medium* targets on the second screen of the standard arrangement, which suggests promise in reducing task time by assigning more apps within each screen while maintaining high accuracy, despite participants commenting that the alternative arrangement was subjective more visually demanding to search for specific apps than the standard layout.

Selection time between the two input techniques showed that overall, *Shift* was only 235.8ms slower than *Touch*. A greater difference in task time was expected but observations and feedback from the participants suggested that with *Touch*, the participants took more time to locate and aim carefully at the screen to select the apps accurately. Perhaps this is one reason why the accuracy of selecting the *Small* targets was much better than expected.

Conversely, there was no need to be as accurate as *Touch* when *Shift* was used since participants can recover their initial input if incorrect and therefore spent less time precisely aiming at the apps and more effort fine-tuning the final selection. The advantage of *Shift* was also exemplified when used on the *Small* targets in alternative arrangement, as selection time was quicker and accuracy was higher than selecting the *Medium* targets on the second screen using *Touch*. These results illustrate the benefits of *Shift* to assist users to target more effectively on small touchscreen devices in walking and encumbered situations.

The participants, while walking and both hands encumbered, selected the *Small* 7.5 x 7.5mm targets better than expected as accuracy reached 94% even without using *Shift*. In *User Study 4*, where the same encumbrance scenario was evaluated, a much lower mean accuracy of 57% was recorded for selecting the 7.5mm target width in the Fitts' Law tapping task. To find out why the results of *User Study 6* are not in line with the findings of *User Study 2 - 5*, a follow-up experiment was conducted.

8.6 User Study 7

A subsequent study was carried out to investigate an anomaly in the results of *User Study 6*. Without the assistance of *Shift*, basic touch accuracy for selecting the *Small* 7.5 x 7.5mm targets in the alternative arrangement reached 94%, even when the users were walking and both hands encumbered with bags. This finding did not support those from earlier user studies discussed in this thesis as a much lower target accuracy was shown when users are encumbered and on the move. One possible reason for the inconsistency in the results could be due to the different targeting tasks used across the studies.

The abstract targeting tasks in *User Study 2 - 5* measured fundamental input performance whereas in *User Study 6*, a more realistic and common tapping activity was evaluated. Arguably, the app selection task in *User Study 6* was more mentally demanding than the abstract tapping tasks in *User Study 2, 3, 4 and 5* since participants had to spend extra effort to search for the current target to select. In contrast, the abstract tasks were more straightforward to complete since the current target to select was always clearly displayed on-screen. While this should have made the app selection task in *User Study 6* more difficult to complete, especially when selecting the *Small* targets, the extra time spent searching for the current app to acquire could have meant more effort was also used to input more accurately. Comments from the participants also suggested that they were more careful at selecting the current target accurately due to other targets placed nearby.

Another possible reason as to why there was a discrepancy in the results between the tasks might have been caused by the visual design and interpretation of the different target types. In the abstract tasks, each target was represented by a border and a cross-hair placed at the centre, which meant that a selection was correct as long as the touch up position was within the limits of the target. Conversely, no border was used to outline the target areas in the app selection task (like in a standard Android arrangement) of *User Study 6*, so participants might have tapped more precisely to ensure that the target was selected accurately.

To see if (1) *the visual presentation of targets* and (2) *when multiple targets are placed close to each other* changed the user's input behaviour, variations of the abstract targeting task (used in *User Study 3* and *User Study 5*) and the app selection task of *User Study 6* were created and compared in *User Study 7*. The remainder of this section describes the approach and discusses the main findings from *User Study 7*.

8.6.1 Method

8.6.1.1 Tasks

Two variations of the abstract targeting task (denoted as *Task 1* and *Task 2*) used in *User Study 3* and *User Study 5* were created to see if altering the visual appearance of the targets made any differences to targeting performance. Six variants of the app selection task (denoted as *Task 3 - 8*) used in *User Study 6* were created to see if gradually placing more targets on-screen resulted in any changes in tapping behaviour. As a result, there were eight targeting tasks all performed while walking and encumbered. The tasks are described as follows:

Task 1: The exact same abstract targeting task as the one used in *User Study 3* and *User Study 5*. Each target measured 5 *x* 8mm and was presented on-screen by a border and cross-hair (2.5mm in length), as shown in Figure 8.14. There were nine target positions aligned in a 3 *x* 3 grid and 160 selections per block. See Section 5.2.2 for more details.

Task 2: Same as *Task 1* but each target's border and cross-hair was replaced by an app icon, as illustrated in Figure 8.15. For each trial, a randomly selected app icon was presented on-screen at the target position to select. The apps were selected from a set of 80 apps, which were chosen from the Google Play Store as described above in Section 8.5.1.3. The icons were scaled to fit within the 5 *x* 8 mm target area.

Task 3: Similar to the app selection task (described above in Section 8.5.1.2) when selecting targets in the alternative arrangement, but instead of apps placed in all 40 target areas, only two *Small* targets were presented on-screen for each trial. This allowed tapping performance to be measured when the target to select is placed among distractors (one distractor in this case). As a reminder, there were 34 *Small* targets (7.5 *x* 7.5 mm) and 6 *Large* targets (22.4 *x* 22.4 mm) in the alternative arrangement. No *Large* targets were selected in *Task 3* and each of the 34 *Small* target positions was selected twice, which meant there were 68 trials per block. The one distractor target position was randomly chosen for each trial. The apps placed in the *Small* target areas were pre-selected and stayed constant for each block of trials. The way the task operated was the same as described above in Section 8.5.1. *Task 3* is illustrated in Figure 8.16.

Task 4: Same as *Task 3* but with three distractors, therefore four *Small* targets displayed on-screen for each trial, as shown in Figure 8.17.

Task 5: Same as *Task 3* but with seven distractors, therefore eight targets shown on-screen for each trial (see Figure 8.18).

Task 6: Same as *Task 3* but with 15 distractors, therefore 16 targets shown on-screen for each trial as illustrated in Figure 8.19.

Task 7: Same as *Task 3* but all 34 *Small* targets were displayed on-screen for each trial, as shown in Figure 8.20.

Task 8: Same as *Task 3* but both *Small* and *Large* targets were shown i.e. the full alternative arrangement, as shown in see Figure 8.21. The *Large* targets only acted as distractors and were not selected.

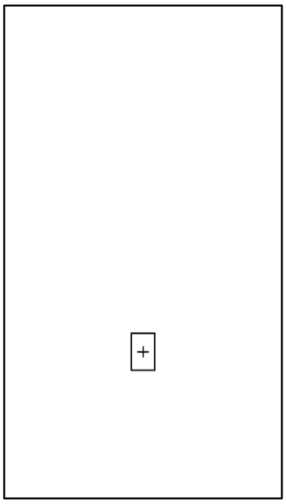

Figure 8.14: Screenshot of Task 1 tested in *User Study 7*.

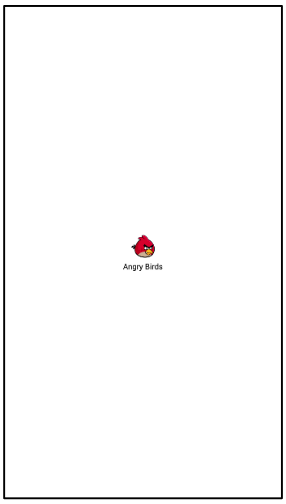

Figure 8.15: Screenshot of Task 2 tested in *User Study 7*.

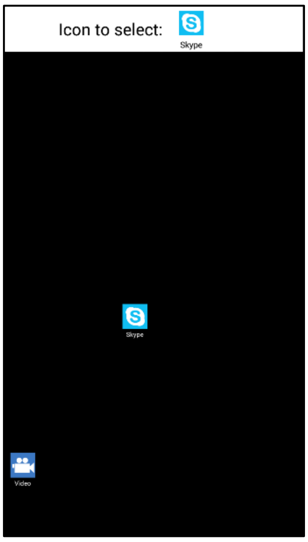

Figure 8.16: Screenshot of Task 3 tested in *User Study 7*.

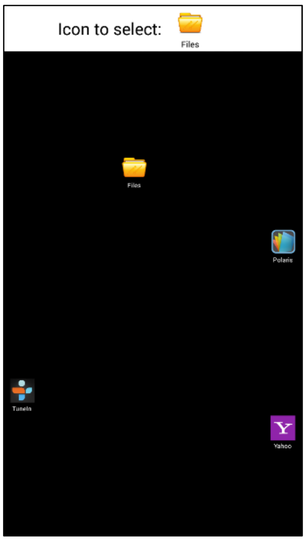

Figure 8.17: Screenshot of Task 4 tested in *User Study 7*.

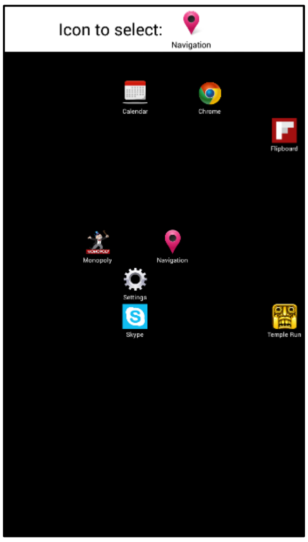

Figure 8.18: Screenshot of Task 5 tested in *User Study 7*.

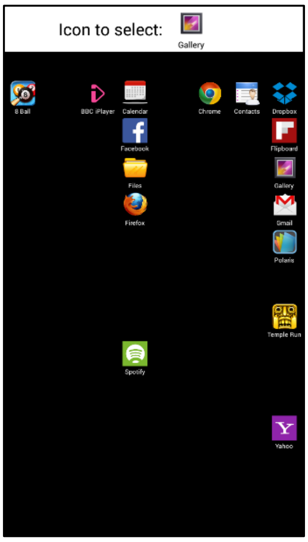

Figure 8.19: Screenshot of Task 6 tested in *User Study 7*.

Figure 8.20: Screenshot of Task 7 tested in *User Study 7*.

Figure 8.21: Screenshot of Task 8 tested in *User Study 7*.

All eight tasks ran on a Samsung Galaxy S3 phone (~12.05px/mm) and each participant held the device in the non-dominant hand while only the index finger of the dominant hand was used to select the targets. Participants were instructed to select each trial as quickly and as accurately as possible.

8.6.1.2 Experimental Design

The tasks were completed while participants were walking and encumbered. The same encumbrance scenario and walking evaluation approach as *User Study 6* was used. Therefore, a 1.6kg shopping bag, which measured 330 *x* 480mm, was held in each hand during input. Each participant walked around the same oval route as *User Study 6* and maintained their PWS&I by walking alongside a pacesetter.

Twenty right-handed students (12 males, 8 females) aged between 21 to 45 years (mean = 27.6, SD = 7.0) were recruited from the University to take part. All participants used touchscreen mobile phones and devices on a daily basis. The study took approximately one hour to complete and £6 was paid to each participant. Each targeting task was completed while carrying the bags and walking, giving a total of eight experimental conditions. A within-subject design was used and the conditions were randomised to reduce order and learning effects as much as possible.

The Independent Variable was **Type of Targeting Task** (eight levels) and the Dependent Variables were accuracy (%) and selection time (ms). For all tasks, a target was selected accurately if the touch up position was within the target border. For *Task 1* and *Task 2*, selection time was the duration from the display of the current target to a touch up event recorded. For *Task 3 - 8*, selection time was recorded as the duration between the first (to start the trial) and second (to select the target) touch up events. The hypotheses were:

H1: Target accuracy of *Task 1* (border and cross-hair) will be significantly lower than *Task 2* (app icon);

H2: Selection time of *Task 1* (border and cross-hair) will be significantly quicker than *Task 2* (app icon);

H3: Target accuracy of *Task 1* will be significantly lower than *Task 8*;

H4: Selection time of *Task 1* will be significantly quicker than *Task 8*.

8.6.2 Results

Two blocks of target selections were completed for each condition. Therefore, each participant performed 1,456 target selections and a total of 29,120 trials was recorded for *User Study 7*. For *Task 1* and *Task 2*, unintentional taps were filtered out if target selections took less than 100 milliseconds, similar to the method used in *User Study 3* and *User Study 5*. As a result, 11 trials were removed. For *Task 3 - 8*, the same method described above in Section 8.5.2 was used to remove accidental selections. Therefore, 17 trials were removed.

8.6.2.1 Target Accuracy

Shapiro-Wilk tests were conducted to assess the normality of the target accuracy data for each condition. The results are shown in Table 8.4 and no condition was significant therefore normality was not violated. A one-factor, repeated-measures ANOVA was conducted for target accuracy. Greenhouse-Geisser adjustments were used to correct the degrees of freedom since Mauchly's test for sphericity was significant ($\varepsilon < 0.75$). The mean target accuracy for each condition is shown in Figure 8.22.

Condition	W Statistic	Sig.
Task 1	0.990663	0.998826
Task 2	0.974017	0.836361
Task 3	0.91362	0.074735
Task 4	0.958514	0.514602
Task 5	0.947502	0.330758
Task 6	0.919192	0.095599
Task 7	0.924916	0.123233
Task 8	0.93495	0.192182

Table 8.4: Shapiro-Wilks normality tests performed on target accuracy for each condition in *User Study 7*.

The ANOVA for accuracy showed that a significant main effect was observed, $F(3.3, 70) = 79.426$, $p < 0.01$. *Post hoc* pairwise comparisons with Bonferroni corrections showed that there was a significant difference in accuracy between *Task 1* and *Task 2*. Changing the targets visually in the abstract targeting tasks did not have a significant impact on tapping accuracy. Target accuracy of *Task 1* was significantly lower than *Task 3 - 8* (varying the number of targets displayed on-screen in the alternative arrangement). Likewise, the number of accurate selections for *Task 2* was significantly less than *Task 3 - 8*.

There was no difference in accuracy between *Task 3* (1 distractor) and *Task 4* (3 distractors), or between *Task 3* and *Task 5* (7 distractors). However, accuracy for *Task 3* was significantly lower than *Task 6* (15 distractors), *Task 7* (all 34 *Small* targets were displayed) and *Task 8* (full arrangement). The number of accurate selections for *Task 4* was not significantly different from *Task 5*. Target accuracy for *Task 6*, *Task 7* and *Task 8* were all significantly higher than *Task 4*. The accuracy between *Task 5* and *Task 6* was not significant but *Task 5* was significantly less accurate than both *Task 7* and *Task 8*. The participants performed *Task 6* significantly less accurate than *Task 7* and *Task 8*. No difference in accuracy was observed between *Task 7* and *Task 8*. Table 8.5 shows all the target accuracy pairwise comparisons.

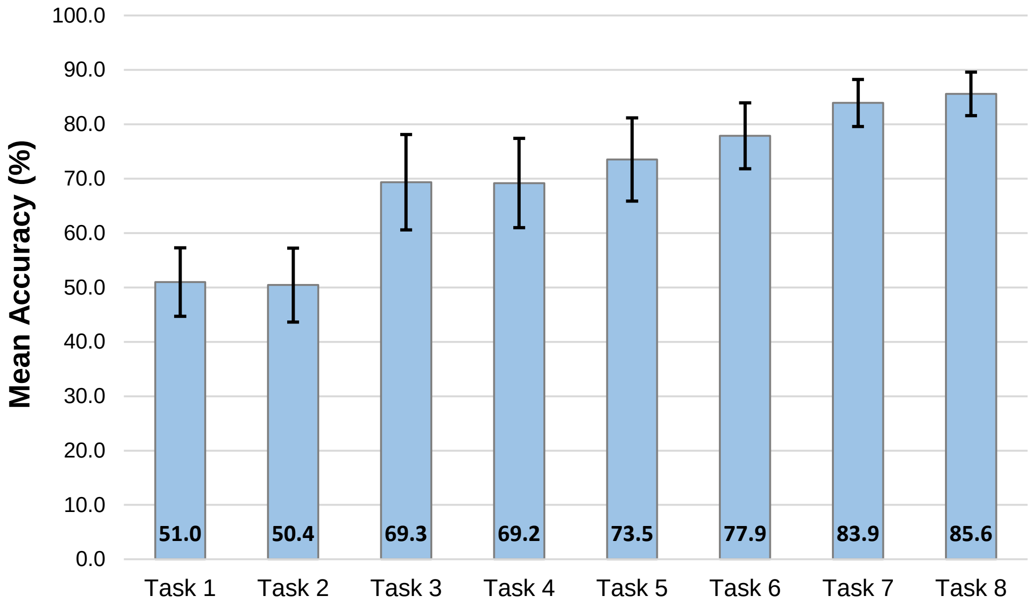

Figure 8.22: The mean accuracy (%) for each task in *User Study 7*. Error bars denote Confidence Interval (95%).

Task Comparison		Mean Difference (%)	Std. Error	Sig
1	2	0.547	1.683	1.000
1	3	-18.354	2.389	0.000*
1	4	-18.207	2.272	0.000*
1	5	-22.545	1.903	0.000*
1	6	-26.92	1.704	0.000*
1	7	-32.95	2.022	0.000*
1	8	-34.604	1.986	0.000*
2	3	-18.901	2.424	0.000*
2	4	-18.754	2.316	0.000*
2	5	-23.092	1.913	0.000*
2	6	-27.467	2.097	0.000*
2	7	-33.496	2.29	0.000*
2	8	-35.151	2.578	0.000*
3	4	0.147	1.983	1.000
3	5	-4.191	1.433	0.244
3	6	-8.566	2.185	0.026*
3	7	-14.596	2.786	0.001*
3	8	-16.25	3.022	0.001*
4	5	-4.338	1.213	0.056
4	6	-8.713	1.85	0.004*
4	7	-14.743	2.594	0.000*
4	8	-16.397	2.809	0.000*
5	6	-4.375	1.607	0.378
5	7	-10.404	2.211	0.004*
5	8	-12.059	2.505	0.003*
6	7	-6.029	1.554	0.028*
6	8	-7.684	1.729	0.008*
7	8	-1.654	0.939	1.000

Table 8.5: All target accuracy (%) pairwise comparisons for Type of Targeting Task. The comparisons that were significant are shaded in grey and highlighted with '*'.

8.6.2.2 Selection Time

Shapiro-Wilk tests were conducted to assess the distribution of selection time for each condition. The results are shown in Table 8.6 and no condition was significant, therefore the data for selection time did not deviate from a normal distribution. A one-factor, repeated-measures ANOVA was conducted for selection time. Greenhouse-Geisser adjustments were used to correct the degrees of freedom since Mauchly's test for sphericity was significant ($\varepsilon < 0.75$). The mean selection time for each condition is shown in Figure 8.23.

Task	W Statistic	Sig.
T1	0.936436	0.205166
T2	0.94486	0.295716
T3	0.944024	0.285327
T4	0.910449	0.065004
T5	0.926267	0.130849
T6	0.929196	0.149012
T7	0.949583	0.360801
T8	0.951899	0.396831

Table 8.6: Shapiro-Wilks normality tests performed on selection time for each condition in *User Study 7*.

The ANOVA conducted to analyse selection time showed that there was a significant main effect, $F(2.1, 39.5) = 46.209$, $p < 0.01$. *Post hoc* pairwise comparisons with Bonferroni corrections showed that the targets in *Task 1* were not selected significantly quicker than *Task 2*. Changing the targets' visual appearance in the abstract targeting tasks did not have an impact on input speed. The selection time for *Task 1* was significantly quicker than *Task 3 - 8*. Likewise, the targets in *Task 2* took less time to select than *Task 3 - 8*, which required more time to visually search for the current target to select among distractors.

There was no significant difference in selection time between *Task 3* (selecting one target out of two in the alternative arrangement) and *Task 4* (selecting one target out of four in the alternative arrangement). The targets in *Task 3* were selected significantly quicker than those in *Task 5 - 8*. The selection time of *Task 4* took significantly less time than *Task 5 - 8*. The targets in *Task 5* were selected significantly quicker than *Task 6 - 8*. The selection time of *Task 6* was significantly quicker than *Task 7* but no significant difference was observed between *Task 6* and *Task 8*. The selection time between *Task 7* and T*ask 8* was not significant. All selection time pairwise comparisons are shown in Table 8.7.

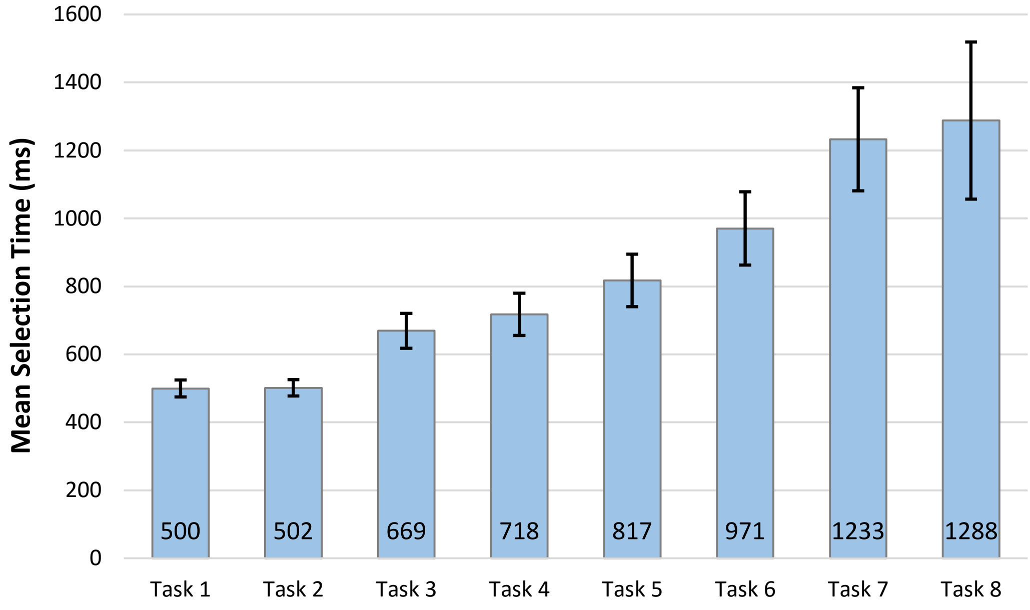

Figure 8.23: The mean selection time (ms) for each task in *User Study 7*. Error bars denote Confidence Interval (95%).

Task Comparisons		Mean Difference (ms)	Std. Error	Sig
1	2	-1.862	6.814	1.000
1	3	-169.671	21.045	0.000*
1	4	-217.856	25.89	0.000*
1	5	-317.455	35.532	0.000*
1	6	-470.786	52.459	0.000*
1	7	-733.415	68.537	0.000*
1	8	-788.199	108.504	0.000*
2	3	-167.808	19.007	0.000*
2	4	-215.994	23.961	0.000*
2	5	-315.592	33.926	0.000*
2	6	-468.923	50.626	0.000*
2	7	-731.553	67.607	0.000*
2	8	-786.336	107.1	0.000*
3	4	-48.185	16.266	0.224
3	5	-147.784	26.301	0.001*
3	6	-301.115	45.855	0.000*
3	7	-563.744	66.199	0.000*
3	8	-618.528	100.984	0.000*
4	5	-99.599	21.108	0.004*
4	6	-252.929	43.518	0.000*
4	7	-515.559	65.852	0.000*
4	8	-570.343	103.768	0.001*
5	6	-153.331	35.57	0.011*
5	7	-415.96	61.721	0.000*
5	8	-470.744	101.029	0.005*
6	7	-262.629	47.168	0.001*
6	8	-317.413	92.241	0.077
7	8	-54.784	92.705	1.000

Table 8.7: All selection time (ms) pairwise comparisons for type of targeting task. The comparisons that were significant are shaded in grey and highlighted with '*'

8.6.2.3 Walking Speed, Distance Walked and Completion Times

A paired t-test was performed to examine the difference in walking speed (km/h) between PWS (no interaction and unencumbered) and PWS&I. A significant difference was found between the two types of walking speeds; $t(19) = 9.345$, $p < 0.05$. PWS&I (mean = 3.61, SD = 0.62) was significantly slower than PWS (mean = 4.75, SD = 0.62), a decrease of 23.9%. The estimated mean distance walked and mean completion time for each condition is shown in Table 8.8. Note the condition completion times do not include resting periods.

	Estimated mean distance walked (m)	**Mean Condition Time (min) (2 blocks)**	**Number of trials (2 blocks)**
Task 1	159.4	2.67	320
Task 2	160.1	2.68	320
Task 3	159.5	2.68	68
Task 4	163.5	2.74	68
Task 5	181.4	3.04	68
Task 6	202	3.41	68
Task 7	241.8	4.07	68
Task 8	245.6	4.16	68
TOTAL	***1513***	***25.4***	***1048***

Table 8.8: The estimated mean distance walked (metres) and completion time (minutes) for each condition in *User Study 7*. The condition completion times do not include resting periods.

8.6.3 Discussion

The results from *User Study 7* showed that in terms of target accuracy, there was no significant difference between *Task 1* and *Task 2*. Therefore, hypothesis *H1* is rejected. The difference in mean accuracy between *Task 1* and *Task 2* was only 0.6%, which suggests that changing the targets visually in the abstract targeting tasks did not have a substantial impact on the user's tapping performance while walking and encumbered. Furthermore, the mean accuracy for *Task 1* was 51%, which was similar to the 48% accuracy recorded when the identical targeting task was performed in the same input posture and encumbrance scenario in *User Study 3*. This shows the replicability of results of the abstract targeting task, even when users are walking and encumbered

The number of accurate targets selections for *Task 1* was also significantly less than selecting the *Small* targets in *Task 8*. The difference in accuracy between *Task 1* and *Task 8* was 34.6%, despite the targets in *Task 1* were only 2.5 mm wider than the *Small* targets in *Task 8*. Therefore, hypothesis *H3* is supported. This result also confirms that the participants did indeed perform well when selecting the *Small* targets in the alternative arrangement while walking and encumbered in *User Study 6*. While the mean accuracy for *Task 8* (85.6%) was not as high as the 94% accuracy for selecting the exact same targets in *User Study 6*, the participants still tapped well considering both hands were encumbered and participants were walking.

Tasks 3 - 8 were designed to see if tapping behaviour changed when more targets were displayed on-screen. The results showed that accuracy gradually increased from 69.3% for *Task 3*, which had two targets shown at each instance to 85.6% for *Task 8*, where the entire alternative arrangement was displayed. This suggests that the participants selected the current target more accurately when placed among other targets since target size stayed the same.

The results for selection time showed that there was no significant difference between *Task 1* and *Task 2*. Therefore, hypothesis *H2* is rejected. The mean selection times for *Task 1* and *Task 2* were 500ms and 502ms respectively, which shows that changing the targets visually from its outline to an app icon scale to fit the target size, did not have a significant impact on tapping speed. For comparison, a mean selection time of 545ms was recorded in *User Study 3* when the same abstract targeting task was conducted in the same input posture and encumbrance scenario.

The selection time of *Task 1* was significant faster than *Task 8*. Therefore, hypothesis *H4* is supported. As expected, the additional time required to visually search for the target to select in the alternative arrangement in *Task 8* meant it took the participants substantially longer to perform the selections than in *Task 1*. The mean selection time to select the *Small* targets in *Task 8* was 1288ms, almost three times greater than *Task 1*. For comparison, the participants in *User Study 6* took 2152ms to select the *Small* targets, almost one second more than the participants to perform the same task (*Task 8*) in *User Study 7*. The participants in *User Study 6* might have taken longer to select the *Small* targets even more accurately.

As anticipated, the participants took longer to select the *Small* targets as the number of targets displayed in the alternative arrangement increased. The selection time for *Task 3* was 669ms and gradually increased to 1288ms for *Task 8*. While not entirely conclusive, users might have taken some of the selection time to steady their input to tap on the targets more accurately since the number of correct target selections incrementally increased from

Task 4 to *Task 8*. Furthermore, the abstract targeting tasks of *Task 1* and *Task 2* required less mental workload since the target to select was always shown without distractors, which meant no visual search time was required and therefore the participants might have compromised accuracy for input speed. *Task 3 - 8* might have made the participants more cautious of selecting the wrong target since many targets were clustered together as the number of apps displayed on-screen increased.

8.7 Conclusions

To summarise, this chapter began by defining a target size of 22.4 *x* 22.4mm, based on the results of previous user studies presented in this thesis, to improve tapping accuracy when users are walking and encumbered. The chapter then discussed the effects of implementing user interfaces with larger target sizes and used app arrangement on touchscreen mobile devices as a use-case example. A systematic approach was used to create an alternative app arrangement, which implemented larger targets to increase accuracy and utilised unexploited screen space with smaller targets. Previous user studies discussed in Chapters 4 - 7 and related work in the literature have reported poor accuracy when selecting small targets using touch input. Therefore, *Shift*, a benchmark interaction technique was implemented to see if it can still maintain its performance gains when used in walking and encumbered situations.

The first user study (*User Study 6*) presented in this chapter evaluated the effectiveness of the alternative app arrangement when compared to a standard layout, and the performance of *Shift* when compared to unassisted touch input. To answer research question *Q4.1* (*What are the appropriate target sizes for encumbered and walking interactions?*), the results from *User Study 6* showed that the large 22.4 *x* 22.4mm targets almost achieved 100% accuracy with standard touch input. However, smaller target sizes could be as effective since accuracy when selecting the 14.9 *x* 16.6mm targets in the standard layout only dropped by 2% to 97% at the expense of limiting the number of apps that can be placed on each screen.

To answer research question *Q4.2* (*Can non-standard input techniques improve input performance while walking and encumbered?*), *Shift* improved accuracy for all target sizes evaluated in *User Study 6*. *Shift* reduced selection uncertainty i.e. the current touch position underneath the finger is always shown in a callout. But more importantly, *Shift* gives users a chance to correct their input (if required) and therefore reduced the need to be as precisely as unassisted touch at the initial stage of a selection since adjustments can be made. The results of *Shift* from *User Study 6* also showed that users could make fine on-screen adjustments while encumbered and on the move, which illustrates the usefulness of pointing techniques (*Shift* in this case) to assist users to input more accurately and effectively in physically demanding contexts.

The accuracy for selecting the 7.5 *x* 7.5mm targets using basic touch input was unexpectedly high in *User Study 6*. To understand why the participants performed so well when walking and encumbered, a follow-up user study (*User Study 7*) was conducted, which compared the performance of the different targeting tasks used in the studies presented in previous chapters. The results from *User Study 7* showed that changing the targets visually in the abstract targeting task used in *User Study 3* and *User Study 5* did not have a significant effect on either target accuracy or selection time. Furthermore, similar results to *User Study 3* were reproduced when the same experimental design was evaluated.

To find out if placing targets close to each other caused users to input differently, the number of on-screen apps shown in the alternative arrangement was varied. The results from *User Study 7* showed that both target accuracy and selection time gradually increased as the number of apps displayed on-screen also increased. While it was anticipated that selection time would progressively take longer to visually search for the current target to select, accuracy also gradually improved, which suggests participants took more care to input more accurately to avoid selecting the wrong target.

The abstract targeting task is useful for understanding how basic tapping performance is affected when users are walking and encumbered. It provides a means to compare different encumbrance scenarios on a measureable level. However, results from *User Study 6* and *User Study 7* suggest caution when designing user interfaces with abstract measurements. Participants performed unexpectedly well when selecting small targets on a realistic task when similarly sized targets were poorly acquired in the abstract tasks. This suggests that when using a standard task for developing novel user interfaces, accuracy should be controlled to avoid a trade-off with input speed, which can be achieved by either 1) placing more targets on-screens as distractors or 2) using some form of feedback so that the participant knew when a target was selected accurately and if not, they could adjust their targeting behaviour to do so.

In terms of improving usability in walking and encumbered contexts, it is best to avoid visually demanding user interfaces. While accuracy was high for a range of target sizes, participants from *User Study 6* and *User Study 7* commented that the tasks were mentally demanding. The alternative arrangement was designed as a concept to show that large targets can increase accuracy while reducing overall task time by placing more targets on-screen and to utilise the screen space more effectively. The main drawback of this design was that targets were placed close to each other, which created a visually 'busy' interface. If this is avoidable, better spacing between targets should be used. However, there will be applications where it is difficult to reduce the number of targets that are tightly packed together (e.g. a map with many points of interests) and in these situations, an effective pointing technique like *Shift* should be available to assist the user's input.

Chapter 9

Conclusions

9.1 Introduction

This chapter first sums up the main findings from each user study presented in this thesis. The chapter then concludes with the limitations of this thesis and a discussion on future research on encumbered and mobile interactions with mobile devices.

9.2 Summary of Work Completed

The portability of mobile phones allows interaction to take place in a range of different daily situations. Therefore, researchers have acknowledged that it is inadequate to assume that mobile phones and devices are operated only when users are stationary and have consequently evaluated the effectiveness of user interfaces and input techniques in mobile situations, such as when users are walking. These studies, which examined how well users interacted with handheld devices while on the move, have reported a decline in input performance since extraneous movements make input physically challenging while visual and mental resources are constantly switched between different activities.

While mobile phones are commonly used in walking situations, interaction also happens when the user's hands are occupied with other activities, such as carrying shopping bags, personal gear and children, for example. This thesis defined these as "*encumbered*" situations and is likely to make mobile input even more problematic. However, the usability issues associated with the effects of encumbrance are often overlooked. Consequently, there is a lack of empirical work in the literature that has examined the interaction difficulties in these physically and mentally demanding situations. The research presented in this thesis contributes by filling this gap in the literature and a series of user studies was conducted to investigate how targeting performance on touchscreen mobile phones is affected in walking and encumbered contexts.

Bags and boxes are commonly held during mobile usage

To examine the effects of encumbrance and mobility, common encumbered situations had to be identified first. *User Study 1* (Chapter 3) was carried out to answer research questions *Q1.1 (What are the typical objects held during interaction with mobile devices?)* and *Q1.2* ***(****How are these typical objects held during interaction with mobile devices?)* to define a set of common encumbrance scenarios. Users in various types of public settings were observed and the results showed that different forms of bags and boxes were the two most frequently held object types. Therefore, a range of encumbrance scenarios, based on carrying bags in-hand and boxes underarm, was simulated in subsequent user studies to examine the effects of encumbrance and walking on interactions with mobile phones.

Tapping performance deteriorates when walking and encumbered.

To investigate tapping performance while walking and encumbered, *User Study 2* (Chapter 4) was conducted to answer research question *Q2.1 (How do encumbrance and mobility affect tapping performance?)*. Users performed an abstract targeting task on a touchscreen mobile phone while walking around an indoor route and carrying either a bag or a box during input. The highlight findings showed that walking while unencumbered caused target accuracy to decrease to 46.2% when selecting 4 *x* 6mm targets. Accuracy decreased further when users were also encumbered while walking, as carrying a medium-sized bag in the dominant hand resulted in the lowest mean accuracy of 29.8%. The number of correct target selections suffered on a similar level when a wide box was held under the dominant arm while walking, as accuracy dropped to 30.6%. Carrying the medium-sized bag in the non-dominant hand caused an accuracy of 37.9%, although, the accuracy between the different object types were very similar when held in the non-dominant side.

Further, selection time was quicker when the non-dominant hand or arm was encumbered than carrying the objects in the dominant side, which illustrates the interaction difficulties in situations where the hand performing the input is hampered. The slowest selection time of 552ms occurred when a small bag was held in the dominant hand while walking, compared to 525ms when unencumbered. Interestingly, the selection times when the non-dominant side was encumbered were quicker than unencumbered targeting, which suggest the participants in *User Study 2* might have traded accuracy for input speed. On average, walking slowed targeting speed by 12.1% when compared to standing, which shows the negative effects of mobility on interactions with touchscreen mobile devices.

Motion capture cameras were also used in *User Study 2* to see how the physical movements caused by walking and encumbrance affected tapping performance. The results could only show that there were more movements in the non-dominant hand holding the mobile phone when the dominant hand or arm was encumbered. The motion of the dominant hand performing the input was also tracked but limitations mainly due to the placement of the markers meant it was difficult to know for certain if surplus movement was caused by the effects of encumbrance or users were making more adjustments to aim and tap on the touchscreen.

Tracking the input finger accurately is a challenging task but perhaps in future walking and encumbered user studies, which look to examine the user's body movements during interactions with handheld devices, should consider a better setup with the motion capture hardware. The fixed position of the cameras in *User Study 2* meant there was less flexibility for adjustments and therefore it was not possible to use smaller markers to attach to the users, which would have made the tracking of the input hand more precise.

Despite the constraints with the setup, the users' walking speed was accurately measured. As anticipated, users walked slower when completing the targeting task, which supports the findings of related work. Walking speed dropped further when the user was also encumbered as carrying the wider box under the non-dominant arm resulted in the biggest dropped in PWS of 41.2% when compared to the baseline (walking unencumbered and no interaction). Despite the slowdown in walking speed, targeting performance did not improve which illustrates the difficulties of input while walking and carrying cumbersome objects.

Encumbrance affects tapping performance regardless of input posture.

Because touchscreen mobile devices can be used in different input postures and users are likely to switch to a more suitable position when walking and encumbered, *User Study 3* was conducted to answer research question *Q2.1.1 (How does the change of input posture affect tapping interactions when walking and encumbered?)*. Therefore, three common input postures were tested to see whether there were any noticeable differences between one- and two- handed input and if there were any advantages of having an extra thumb for input. Furthermore, people often carry multiple objects and may require both hands to carry them, so to evaluate situations were both hands are hindered, participants in *User Study 3* held a typical shopping bag in each hand while performing a target selection task. The results from *User Study 3* showed that target accuracy decreased and selection time was slower when carrying the bags than unencumbered while walking in all three input postures. Despite different overall target accuracy between the three input postures, the results were not statistically significant. The selection time of the one-handed preferred thumb input posture was slower than the two-handed index finger position. However, tapping speed was heavily biased by the longer time taken to select the targets further away from the thumb's optimal reach. These findings suggest that there is no clear input posture that users can switch to in order to improve tapping performance when both hands are encumbered while on the move. For comparison, Musić and Murray-Smith [53] found that target accuracy using the two-handed index finger posture was higher than using thumb-based input when tapping on a touchscreen phone while walking but unencumbered.

A more appropriate evaluation approach was used in *User Study 3* than the one applied in *User Study 2*. Users are likely to walk slower, as shown in *User Study 2* or vary their walking speed when interacting with mobile devices while encumbered. However, one problem with deviations in walking speed when conducting controlled experiments is that the effects of encumbrance are mixed with the effects of mobility. Therefore, it is difficult to know if the changes observed in targeting performance are caused by the effects of walking or carrying the objects. To remove this ambiguity and to be more certain that changes in targeting performance are indeed caused by the effects of encumbrance, a pacesetter was deployed to control the participants' walking speed during each experimental condition. This evaluation approach was also used in all subsequent user studies and should be applied in future encumbered and walking experiments to get a more accurate reflection of the impact encumbrance has on interactions with mobile devices while on the move.

Two-finger on-screen gestures were accurately executed while walking and encumbered.

Touchscreen interfaces afford many different types of touch-based gesture interactions and no related work in the literature has examined their performance while users are walking and/or encumbered. Therefore, four main gesture inputs (tapping, dragging, spreading &

pinching and rotating) commonly used on touchscreen mobile devices were examined while walking and encumbered in *User Study 4* to answer research question *Q2.2 (How do encumbrance and mobility affect the performance of other standard touch-based gestures?).*

A set of novel Fitts' Law targeting tasks was designed to test the four gesture types. The main finding from *User Study 4* showed that despite both hands were carrying bags and users were walking, the accuracy of the two-finger gestures (spreading, pinching and rotating) were much higher than both one-finger tapping and dragging. The rotating gestures were performed particularly well as the overall accuracy was greater than 92%, even when users were walking and encumbered. Arguably, the two-finger gestures are more biomechanically complex to perform than a simple on-screen tap or drag since both fingers have to move in sequence with each other. However, having an extra finger from the dominant hand for input possibly gave users more stability to complete the target selections more accurately while walking and carrying bags in-hand. Furthermore, the design of the two-finger targeting tasks meant visual feedback had to be used, which might have made input visually easier to perform.

One possible reason why the one-finger gestures were much less accurate than the two-finger gestures, especially for dragging which resulted in an overall accuracy of 48% when encumbered, was that visual feedback was not used in those tasks and therefore, it was difficult for the user to know if the target was correctly selected or not, due to occlusion by the input finger. Despite the fact that both tapping and dragging had poor target accuracy, movement time for the one-finger gestures was substantially quicker than for the two-finger gestures. When encumbered, tapping had the lowest overall movement time of 417ms for one-finger input while rotating anticlockwise had the quickest targeting speed of 1041ms out of all the two-finger gesture types. The nature of the two-finger gesture types possibly means spreading, pinching and rotating actions are always going to take longer to execute than more straightforward one-finger taps and drags.

Use ground walking over treadmills and control PWS for mobile and encumbered experiments.

User Study 5 answered research questions *Q3 (How to evaluate the effects of encumbrance and mobility in controlled user studies?)* and *Q2.1* by comparing the ground walking evaluation approach (applied in *User Study 3* and *User Study 4*) and walking on a treadmill to see which method was more appropriate for conducting walking and encumbered experiments. Furthermore, the user's walking speed was varied in *User Study 5* to investigate if walking slower or faster caused any changes to tapping performance while encumbered. The results showed that reducing PWS by 20% only improved overall accuracy by 3.6% while walking faster than normal by 20% decreased overall accuracy by 2.6%. Despite statistical significant results, the selection times for walking at 80%, 100% and 120% of PWS were evenly matched. The effect of encumbrance while walking was again apparent as target accuracy decreased and selection time took longer when users carried either a bag in-hand or box underarm.

The difference in targeting performance between ground walking and treadmill walking was small. However, when measuring PWS, users walked slower on the treadmill than on the ground. Therefore, the ground walking approach is the more suitable method to apply for mobile and encumbered experiments. In addition, the user's PWS should be controlled with the use of a pacesetter to isolate the effects of mobility from encumbrance to ensure that any changes in performance are a consequence of carrying objects.

Large targets, dense target grid and Shift can improve tapping accuracy while walking and encumbered.

In Chapter 8, a target size of 22.4 *x* 22.4mm was defined to improve tapping accuracy while walking and encumbered. The implications of designing user interfaces with large on-screen elements were discussed with an alternative app arrangement created systematically as a use-case example. Small on-screen elements are inevitable so a benchmark pointing technique *Shift*, which has shown to improve tapping accuracy when selecting small targets, was also discussed. *User Study 6* tested the effectiveness of the large target size, the alternative app arrangement and *Shift* collectively while users were walking and carrying bags in both hands to answer research questions *Q4.1 (What are the appropriate target sizes and target placements for walking and encumbered interactions?)* and *Q4.2 (Can pointing techniques improve targeting performance while walking and encumbered?).*

The results from *User Study 6* showed that *Shift* improved accuracy across all target sizes, which suggests that *Shift* is still effective in walking and encumbered contexts. Furthermore, there is less emphasis on aiming accurately at the initial stage of a selection when *Shift* is used since on-screen adjustments can be made if necessary. The 22.4 *x*

22.4mm targets improved accuracy to almost 100% when walking and encumbered but selection time was only slightly faster than tapping on smaller 7.5 *x* 7.5mm targets. Unexpectedly, users when walking and encumbered selected the small 7.5 *x* 7.5mm targets very well as accuracy reached 94% without the assistance of *Shift*.

A follow-up study (*User Study 7*) was carried out to investigate why accuracy for selecting the small targets was much higher than expected by comparing variations of the abstract targeting task (used in *User Study 3* and *Study 5*) and the app selection task of *User Study 6*. The results showed that replacing the target outline with borderless icons in the abstract targeting task did not have a statistical significant effect on tapping performance as both accuracy and selection time were near identical. Furthermore, the results were similar to those in *User Study 3*, where the same task, input posture and encumbrance scenario were evaluated. As the number of apps shown in the alternative arrangement progressively increased, accuracy also gradually improved. Target accuracy when two apps were displayed was 69.3% and reached 85.6% when the full arrangement was shown. Selection time gradually took longer as the number of apps shown increased (from 699ms to 1288ms). This suggests that when numerous targets are placed close to each other, users are likely to take more care to input more accurately in walking and encumbered situations since the time required to recover an incorrect selection is likely to take longer.

9.3 Limitations and Future Work

While this thesis has extensively examined the effects of encumbrance and mobility on input performance with touchscreen mobile phones and results have shown poor target accuracy while walking and carrying bags or boxes, further research questions remain. This section describes several key areas for future work on encumbered interactions.

Evaluate other types of encumbrance scenarios

One main limitation of this thesis is that only encumbrance scenarios based on carrying bags in-hand and boxes underarm were evaluated. The observational study presented in Chapter 3 found that people carried or held a variety of objects while on the move. But due to the scope of this thesis, only the two most common object types were examined.

Future work should investigate the impact of carrying a broader range of objects and different types of encumbrances to see whether the effects on input performance is similar to carrying bags and boxes. For example, holding a cup requires a different type of hand grip to carrying a bag, so it is difficult to predict how input on touchscreen devices is affected.

Evaluate the effectiveness of input techniques

User Study 6 showed that *Shift* [73] improved target accuracy while walking and encumbered and was well received by the users. One drawback of *Shift* is the increase in selection time over basic touch input regardless of the user's context. One limitation of this thesis is that other input techniques such as *LinearDragger* [2] and *Escape* [79], which have shown to improve the selection time of *Shift* while maintaining similar levels of accuracy, were not evaluated. It would be interesting for future work to examine the performance and usability of these interaction techniques in more physically demanding mobile and encumbered situations and therefore develop better tools for users to interact with touchscreen mobile devices more efficiently when input becomes problematic.

Explore multi- finger interactions for encumbered and mobile contexts

User Study 4 showed that two-finger gestures were performed better than one-finger actions in terms of higher target accuracy even when users were walking and both hands encumbered. Perhaps the additional finger from the same targeting hand gave users more stability by reducing the amount of extraneous movements caused when walking and carrying objects, and therefore made input easier to perform. One limitation of this thesis is that multi-finger gestural interaction techniques were not designed to see whether they could be used to replace standard one-finger inputs. For example, instead of a tap to select a button, perhaps a two-finger rotational action will give higher accuracy in walking and encumbered situations. It would be interesting for future work to explore this and design novel multi-finger interaction techniques and test their effectiveness in mobile and encumbered contexts.

Alternative non-touch input modalities for encumbered and mobile interactions

If on-screen touch input is a major challenge for users, an interesting area of research for future work is to examine the effectiveness of alternative input modalities. Previous studies have shown promise in using body gestures as an alternative means to acquire on-screen targets with mobile devices without the need for explicit finger input. For example, custom-made sensor packs were used to design head tilting [14] and wrist rotation [16] gestures to select targets on mobile phones and results showed reasonable accuracy and selection times when these gestures were performed while walking. Perhaps the shift towards more powerful and functional wearable technologies such as smartwatches and augmented reality glasses/headgear could improve this type of touchless interaction and address the limitations of earlier work. Furthermore, the use of wearable technology opens up the possibility of hands-free and eyes-free interactions, without the need to take mobile devices out of pocket for input. Visual attention can then be used to attend situational distractions when users are on the move while both hands are free for other activities such as carrying objects. Wearable technology is currently at an early development stage but its potentials remain both exciting and advantageous to improve the way mobile devices are used in the future.

The possibility of detecting encumbered contexts

Mobile devices now contain various inertial sensors (such as accelerometers, gyroscopes and magnetometers) and researchers in the past have utilised these lightweight, low-cost and energy efficient technologies not only to infer the orientation of the device but also to estimate and classify the user’s current context or activity. The ability to identify the orientation of handheld devices meant they could be used in different ways and allows user interfaces to dynamically adjust to the current context at any given moment. But user interfaces could potentially improve further if the on-board sensors could detect when users are encumbered and provide assistance to interact more effectively when input on touchscreen devices is known to be problematic. Perhaps, when searching for apps, the alternative arrangement (discussed in Chapter 8) could be automatically deployed when encumbrance is detected, without an explicit command from the user.

Previous research has shown promise in detecting encumbrance by using simple accelerometers. For example, Bao and Intille [4] attached five biaxial accelerometers to different parts of the user’s body and showed high accuracy in classifying different activities (e.g. sitting down, walking only and walking while carrying objects). However, deploying many sensors on to various body parts is impractical and a simpler solution is to exploit the sensors already built into mobile devices. Perhaps the development and popularity of wearable devices will make sensors become ubiquitous and address this limitation. Nevertheless, related work (e.g. [21,53,82]) has shown the possibility of detecting different physical contexts, such as when the user is stationary or on the move, by using accelerometers in or attached to mobile phones. Future work should extent this research to see if on-board sensors could detect when users are encumbered, which will open opportunities to improve usability with mobile devices in physically demanding situations.

9.4 Final Comment

Mobile phones have become important tools in our everyday lives as they provide access to functionality and applications to complete daily activities. Despite the usefulness of these handheld devices, the different contexts that they are commonly used in are less well studied. It is important to evaluate how mobile phones and devices are used other than in ideal conditions to understand the interaction difficulties and usability issues in different common everyday situations. This thesis examined one particular context in which mobile phones are frequently used - the effects of *encumbrance* in terms of carrying typical bags and boxes while on the move.

The research presented makes three main contributions. First, the results from a series of user studies have shown the extent in which targeting performance on touchscreen mobile phones declines when users are carrying cumbersome objects while walking. Second, a suitable evaluation approach has been described for future studies looking to examine the effects of encumbrance and mobility on interactions with mobile devices. Third, several approaches have been evaluated to improve targeting performance on mobile phones while encumbered and on the move.

But more work could be done by researchers and designers to help users interact more effectively in physically and mentally demanding situations. This closing chapter has presented several areas for future work on encumbered and mobile interactions. If researchers explore these ideas and design new interaction techniques and applications that are useful in a range of everyday situations, then usability with touchscreen mobile devices will be greatly enhanced.

Bibliography

[1] Alton, F., Baldey, L., Caplan, S., and Morrissey, M.C. A kinematic comparison of overground and treadmill walking. *Journal of Clinical Biomechanics 13,* 6 (1998), 434–440.

[2] Au, O.K.-C., Su, X., and Lau, R.W.H. LinearDragger: a Linear Selector for One-finger Target Acquisition. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '14* (2014), ACM Press, pp. 2607–2616.

[3] Azenkot, S. and Zhai, S. Touch Behavior with Different Postures on Soft Smartphone Keyboards. In *Proceedings of the 14th International Conference on Human-Computer Interaction with Mobile Devices and Services - MobileHCI '12* (2012), ACM Press, pp. 251–260.

[4] Bao, L. and Intille, S.S. Activity Recognition from User-Annotated Acceleration Data. In *Proceeding of the 3rd International Conference on Pervasive Computing - PERVASIVE 2004 (2004)*, Springer Berlin Heidelberg, pp. 390–397.

[5] 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.

[6] Bau, O., Poupyrev, I., Israr, A., and Harrison, C. TeslaTouch: Electrovibration for Touch Surfaces. In *Proceedings of the 23nd Annual ACM Symposium on User Interface Software and Technology - UIST '10* (2010), ACM Press, pp. 283-292.

[7] Bergstrom-Lehtovirta, J., Oulasvirta, A., and Brewster, S. The Effects of Walking Speed on Target Acquisition on a Touchscreen Interface. In *Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices - MobileHCI '11* (2011), ACM Press, pp. 143–146.

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

[9] Buschek, D., Rogers, S., and Murray-Smith, R. User-Specific Touch Models in a Cross-Device Context. In *Proceedings of the 15th International Conference on Human-Computer Interaction with Mobile Devices and Services - MobileHCI '13* (2013), ACM Press, pp. 382–391.

[10] Clark-Carter, D.D., Heyes, a D., and Howarth, C.I. The efficiency and walking

speed of visually impaired people. *Journal of Ergonomics 29*, 6 (1986), 779 – 789.

[11] Clawson, J., Starner, T., Kohlsdorf, D., Quigley, D.P., and Gilliland, S. Texting While Walking: An Evaluation of Mini-QWERTY Text Input while On-the-Go. In *Proceedings of the 16th International Conference on Human-Computer Interaction with Mobile Devices & Services - MobileHCI '14* (2014), ACM Press, pp. 339–348.

[12] Cockburn, A., Ahlström, D., and Gutwin, C. Understanding performance in touch selections: Tap, drag and radial pointing drag with finger, stylus and mouse. *International Journal of Human-Computer Studies 70*, 3 (2012), 218–233.

[13] Cohen, R.G. and Rosenbaum, D. A. Prospective and retrospective effects in human motor control: Planning grasps for object rotation and translation. *Journal of Psychological Research 75*, 4 (2011), 341–349.

[14] Crossan, A., McGill, M., Brewster, S., and Murray-Smith, R. Head Tilting for Interaction in Mobile Contexts. In *Proceedings of the 11th International Conference on Human-Computer Interaction with Mobile Devices and Services - MobileHCI '09* (2009), ACM Press, pp. 1–10.

[15] Crossan, A., Murray-Smith, R., Brewster, S., Kelly, J., and Musizza, B. Gait Phase Effects in Mobile Interaction. In *CHI '05 Extended Abstracts on Human Factors in Computing Systems - CHI EA '05* (2005), ACM Press, pp. 1312–1315.

[16] Crossan, A., Williamson, J., Brewster, S., and Murray-Smith, R. Wrist Rotation for Interaction in Mobile Contexts. In *Proceedings of the 10th International Conference on Human Computer Interaction with Mobile Devices and Services - MobileHCI '08* (2008), ACM Press, pp. 435–438.

[17] Everett, S.P. and Byrne, M.D. Unintended Effects: Varying Icon Spacing Changes Users' Visual Search Strategy. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '04* (2004), ACM Press, pp. 695–702.

[18] Field, A. *Discovering Statistics Using IBM SPSS Statistics*, 4th edition, SAGE Publishing, 2013.

[19] Findlater, L., Froehlich, J.E., Fattal, K., Wobbrock, J.O., and Dastyar, T. Age-Related Differences in Performance with Touchscreens Compared to Traditional Mouse Input. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '13* (2013), ACM Press, pp. 343–346.

[20] Fitts, P.M. The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology, 47*, 6 (1954), 381–

391.

[21] Fitz-Walter, Z. and Tjondronegoro, D. Simple Classification of Walking Activities using Commodity Smart Phones. In *Proceedings of the 21st Annual Conference of the Australian Computer-Human Interaction Special Interest Group on Design: Open 24/7 - OZCHI '09* (2009), ACM Press, pp. 409-412.

[22] Goel, M., Findlater, L., and Wobbrock, J. WalkType: Using Accelerometer Data to Accommodate Situational Impairments in Mobile Touch Screen Text Entry. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '12* (2012), ACM Press, pp. 2687-2696.

[23] Goel, M., Jansen, A., Mandel, T., Patel, S.N., and Wobbrock, J.O. ContextType: Using Hand Posture Information to Improve Mobile Touch Screen Text Entry. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '13* (2013), ACM Press, pp. 2795-2798.

[24] Goel, M., Wobbrock, J., and Patel, S. GripSense: Using Built-In Sensors to Detect Hand Posture and Pressure on Commodity Mobile Phones. In *Proceedings of the 25th annual ACM symposium on User Interface Software and Technology - UIST '12* (2012), ACM Press, pp. 545-554.

[25] Grossman, T. and Balakrishnan, R. The Bubble Cursor: Enhancing Target Acquisition by Dynamic Resizing of the Cursor's Activation Area. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '05* (2005), ACM Press, pp. 281-290.

[26] Hatfield, J. and Murphy, S. The effects of mobile phone use on pedestrian crossing behaviour at signalised and unsignalised intersections. *Journal of Accident Analysis and Prevention 39,* 1 (2007), 197–205.

[27] Hoggan, E., Brewster, S. A., and Johnston, J. Investigating the Effectiveness of Tactile Feedback for Mobile Touchscreens. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '08* (2008), ACM Press, pp. 1573-1582.

[28] Hoggan, E., Nacenta, M., Kristensson, P.O., Williamson, J., Oulasvirta, A., and Lehti, A. Multi-Touch Pinch Gestures: Performance and Ergonomics. In *Proceedings of the 2013 ACM International Conference on Interactive Tabletops and Surfaces - ITS '13* (2013), ACM Press, pp. 219–222.

[29] Hoggan, E., Williamson, J., Oulasvirta, A., Nacenta, M., Kristensson, P.O., and

Lehtiö, A. Multi-Touch Rotation Gestures: Performance and Ergonomics. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI ’13* 2013), ACM Press, pp. 3047-3050.

[30] Holz, C. and Baudisch, P. The Generalized Perceived Input Point Model and How to Double Touch Accuracy by Extracting Fingerprints. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI ’10* (2010), ACM Press, pp. 581–590.

[31] Holz, C. and Baudisch, P. Understanding Touch. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI ’11* (2011), ACM Press, pp. 2501–2510.

[32] Hoober, S. UX Matters: How Do Users Really Hold Mobile Devices? (2013) Retrieved from: http://www.uxmatters.com/mt/archives/2013/02/how-do-users-really-hold-mobile-devices.php. Retrieved on 14/02/15.

[33] Hyman, I.E., Boss, S.M., Wise, B.M., McKenzie, K.E., and Caggiano, J.M. Did you see the unicycling clown? Inattentional blindness while walking and talking on a cell phone. *Journal of Applied Cognitive Psychology 24*, 5 (2010), 597–607.

[34] Jain, S.S.L. The means of mobility. *Journal of Consumer Culture 2*, 3 (2002), 419–438.

[35] 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 *Proceedings of the 10th International Conference on Human Computer Interaction with Mobile Devices and Services - MobileHCI ’08* (2008), ACM Press, pp. 109–118.

[36] Karlson, A., Bederson, B., and Contreras-Vidal, J. Understanding One Handed Use of Mobile Devices. In Handbook of Research on User Interface Design and Evaluation for Mobile Technology, Technical Report (2006), pp. 86-101.

[37] Karlson, A. and Bederson, B. ThumbSpace: Generalized One-Handed Input for Touchscreen-Based Mobile Devices. In *Proceedings of the 11th IFIP TC 13 International Conference on Human-Computer Interaction - INTERACT '07* (2007), Springer-Verlag, pp. 324–338.

[38] Karlson, A.K., Bederson, B.B., and SanGiovanni, J. AppLens and LaunchTile: Two Designs for One-Handed Thumb Use on Small Devices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI ’05* (2005), ACM Press, pp. 201–210.

[39] Karlson, A.K. and Bederson, B.B. One-Handed Touchscreen Input for Legacy Applications. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI ’08* (2008), ACM Press, pp. 1399–1408.

[40] 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.

[41] Kobayashi, M., Hiyama, A., Miura, T., Asakawa, C., Hirose, M., and Ifukube, T. Elderly User Evaluation of Mobile Touchscreen Interactions. In *Proceedings of the 13th IFIP TC 13 International Conference on Human-Computer Interaction - Volume Part I - INTERACT 2011* (2011), Springer-Verlag, pp. 83–99.

[42] Lim, J.J. and Feria, C. Visual Search on a Mobile Device while Walking. In *Proceedings of the 14th International Conference on Human-Computer Interaction with Mobile Devices and Services - MobileHCI ’12* (2012), ACM Press, pp. 295–304.

[43] 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.

[44] Lozano, C., Jindrich, D., and Kahol, K. The Impact on Musculoskeletal System during Multitouch Tablet Interactions.In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI ’11* (2011), pp. 825–828.

[45] 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 *Proceedings of the 7th International Conference on Human Computer Interaction with Mobile Devices & Services - MobileHCI ’05* (2005), ACM Press, pp. 183-190.

[46] MacKenzie, I.S. and Isokoski, P. Fitts’ Throughput and the Speed-Accuracy Tradeoff. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI ’08* (2008), ACM Press, pp. 1633–1636.

[47] MacKenzie, I.S., Kauppinen, T., and Silfverberg, M. Accuracy Measures for Evaluating Computer Pointing Devices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI ’01* (2001), ACM Press, pp. 9–16.

[48] MacKenzie, I.S. Fitts’ Law as a Research and Design Tool in Human-Computer Interaction. *Journal of Human-Computer Interaction 7*, 1 (1992), 91–139.

[49] Mainwaring, S.D., Anderson, K., and Chang, M.F. Living for the Global City: Mobile Kits, Urban Interfaces, and Ubicomp. In *Proceeding of the Seventh International Conference on Ubiquitous Computing - UbiComp ’05* (2005), Springer-Verlag, pp. 269–286.

[50] Mizobuchi, S., Chignell, M., and Newton, D. Mobile Text Entry: Relationship between Walking Speed and Text Input Task Difficulty. In *Proceedings of the 7th International Conference on Human Computer Interaction with Mobile Devices & Services - MobileHCI ’05* (2005), ACM Press, pp. 122-128.

[51] Mott, M.E. and Wobbrock, J.O. Beating the Bubble: Using Kinematic Triggering in the Bubble Lens for Acquiring Small, Dense Targets. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI ’14* (2014), ACM Press, pp. 733–742.

[52] 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.

[53] Musić, J. and Murray-Smith, R. Nomadic Input on Mobile Devices: The Influence of Touch Input Technique and Walking Speed on Performance and Offset Modeling. *Journal of Human–Computer Interaction,* 1–52.

[54] Mustonen, T., Olkkonen, M., and Hakkinen, J. Examining Mobile Phone Text Legibility while Walking. In *Extended Abstracts on Human Factors in Computing Systems - CHI EA ’04* (2004), ACM Press, pp. 1243-1246.

[55] Nasar, J., Hecht, P., and Wener, R. Mobile telephones, distracted attention, and pedestrian safety. *Journal of Accident Analysis and Prevention, 40,* 1 (2008), 69–75.

[56] Nicolau, H. and Jorge, J. Touch Typing using Thumbs: Inderstanding the Effect of Mobility and Hand Posture. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI ’12* (2012), ACM Press, pp. 2683–2686.

[57] Oulasvirta, A. and Bergstrom-Lehtovirta, J. Ease of Juggling: Studying the Effects of Manual Multitasking. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI ’11* (2011), ACM Press, pp. 3103–3112.

[58] Oulasvirta, A., Tamminen, S., Roto, V., and Kuorelahti, J. Interaction in 4-Second Bursts: The Fragmented Nature of Attentional Resources in Mobile HCI. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI ’05 (*2005), ACM Press, pp. 919-928.

[59] Parhi, P., Karlson, A.K., and Bederson, B.B. Target Size Study for One-Handed Thumb Use on Small Touchscreen Devices. In *Proceedings of the 8th Conference on Human-Computer Interaction with Mobile Devices and Services - MobileHCI ’06* (2006), ACM Press, pp. 203–210.

[60] Park, J., Patel, A., Curtis, D., Teller, S., and Ledlie, J. Online Pose Classification and Walking Speed Estimation using Handheld Devices. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing - UbiComp ’12* (2012), ACM Press, pp. 113–122.

[61] Petrie, H., Johnson, V., Furner, S., and Strothotte, T. Design Lifecycles and Wearable Computers for Users with Disabilities. In *Proceedings of the First Workshop on HCI Mobile Devices* (1998).

[62] Pirhonen, A., Brewster, S., and Holguin, C. Gestural and Audio Metaphors as a Means of Control for Mobile Devices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI ’02* (2002), ACM Press, pp. 291–298.

[63] Potter, R.L., Weldon, L.J., and Shneiderman, B. Improving the Accuracy of Touch Screens: An Experimental Evaluation of Three Strategies. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI ’88* (1988), ACM Press, pp. 27–32.

[64] Prasanta Chandra Mahalanobis. On the Generalised Distance in Statistics. In *Proceedings of the National Institute of Sciences of India* (1936), pp. 49–55.

[65] Schedlbauer, M. and Heines, J. Selecting While Walking : An Investigation of Aiming Performance in a Mobile Work Context. In *Proceedings of Americas Conference on Information Systems - AMCIS 2007* (2007), electronic library.

[66] Schildbach, B. and Rukzio, E. Investigating Selection and Reading Performance on a Mobile Phone while Walking. In *Proceedings of the 12th International Conference on Human Computer Interaction with Mobile Devices and Services - MobileHCI 2010* (2010), ACM Press, pp. 93–102.

[67] Sears, A., Lin, M., Jacko, J., and Yang, X. When computers fade: Pervasive computing and situationally induced impairments and disabilities. In *Proceedings of HCI International 2003*, Lawrence Erlbaum Associates, pp. 1298–1302.

[68] Sears, A. and Shneiderman, B. High precision touchscreens: design strategies and comparisons with a mouse. *International Journal of Man-Machine Studies, 34*, 4 (1991), 593–613.

[69] Simonof, J.S. Interval Estimation and Statistical Inference (2011). Retrieved from: www.stern.nyu.edu/~jsimonof/classes/1305/pdf/intestim.pdf. Retrieved on 23/09/14.

[70] Tamminen, S., Oulasvirta, A., Toiskallio, K., and Kankainen, A. Understanding mobile contexts. *Journal of Personal Ubiquitous Computing 8*, 2 (2004), 135–143.

[71] Tran, J.J., Trewin, S., Swart, C., John, B.E., and Thomas, J.C. Exploring Pinch and Spread Gestures on Mobile Devices. In *Proceedings of the 15th International Conference on Human-Computer Interaction with Mobile Devices and Services - MobileHCI '13* (2013), ACM Press, pp. 151–160.

[72] Tseng, Y. and Howes, A. The Adaptation of Visual Search Strategy to Expected Information Gain. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '08* (2008), ACM Press, pp. 1075-1084.

[73] Vogel, D. and Baudisch, P. Shift: A Technique for Operating Pen-Based Interfaces Using Touch. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '07* (2007), ACM Press, pp. 657–666.

[74] Weir, D., Buschek, D., and Rogers, S. Sparse Selection of Training Data for Touch Correction Systems. In *Proceedings of the 15th International Conference on Human-Computer Interaction with Mobile Devices and Services - MobileHCI '07* (2013), pp. 404–407.

[75] Weir, D., Rogers, S., Murray-Smith, R., and Löchtefeld, M. A User-Specific Machine Learning Approach for Improving Touch Accuracy on Mobile Devices. in *Proceedings of the 25th Annual ACM Symposium on User Interface Software and Technology - UIST '12* (2012), ACM Press, pp. 465-475.

[76] Wobbrock, J.O., Findlater, L., Gergle, D., and Higgins, J.J. The Aligned Rank Transform for Nonparametric Factorial Analyses Using Only ANOVA Procedures. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '11*, ACM Press, pp. 143-146.

[77] Wobbrock, J.O., Myers, B. A., and Aung, H.H. The performance of hand postures in front- and back-of-device interaction for mobile computing. *International Journal of Human-Computer Studies 66*, 12 (2008), 857–875.

[78] Wobbrock, J.O., Shinohara, K., and Jansen, A. The Effects of Task Dimensionality, Endpoint Deviation, Throughput Calculation, and Experiment Design on Pointing Measures and Models. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '11* (2011), ACM Press, pp. 1639–1648.

[79] Yatani, K., Partridge, K., Bern M., and Newman, M.W. Escape: A Target Selection Technique Using Visually-cued Gestures. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '08* (2008), ACM Press, pp. 285–294.

[80] Yatani, K. and Truong, K.N. An evaluation of stylus-based text entry methods on handheld devices studied in different user mobility states. *Journal of Pervasive and Mobile Computing 5*, 5 (2009), 496–508.

[81] Yatani, K. and Truong, K.N. SemFeel: A User Interface with Semantic Tactile Feedback for Mobile Touch-screen Devices. In *Proceedings of the 22nd Annual ACM Symposium on User Interface Software and Technology - UIST '09* (2009), ACM Press, pp. 111-120.

[82] Yi, J.S., Choi, Y.S., Jacko, J. A., and Sears, A. Context Awareness via a Single Device-Attached Accelerometer During Mobile Computing. In *Proceedings of the 7th International Conference on Human Computer Interaction with Mobile Devices & Services - MobileHCI '05* (2005), ACM Press, pp. 303–306.

[83] Yu, N.-H., Huang, D.-Y., Hsu, J.-J., and Hung, Y.-P. Rapid Selection of Hard-to-Access Targets by Thumb on Mobile Touch-Screens. In *Proceedings of the 15th International Conference on Human-Computer Interaction with Mobile Devices and Services - MobileHCI '13* (2013), ACM Press, pp. 400–403.

[84] Zhao, J., Soukoreff, R.W., and Balakrishnan, R. Exploring and modeling unimanual object manipulation on multi-touch displays. *International Journal of Human-Computer Studies, 78,* C (2015), 69-80.