An Analysis of Interaction in the Context of Wearable Computers

by

Alan Chamberlain

A Doctoral Thesis

Submitted in partial fulfilment of the requirements for the award of

Doctor of Philosophy of Loughborough University

September 2004

Abstract

The focus of this thesis is on the evaluation of input modalities for generic input tasks, such inputting text and pointer based interaction. In particular, input systems that can be used within a wearable computing system are examined in terms of human-wearable computer interaction. The literature identified a lack of empirical research into the use of input devices for text input and pointing, when used as part of a wearable computing system.

The research carried out within this thesis took an approach that acknowledged the movement condition of the user of a wearable system, and evaluated the wearable input devices while the participants were mobile and stationary. Each experiment was based on the user's time on task, their accuracy, and a NASA TLX assessment which provided the participant's subjective workload. The input devices assessed were 'off the shelf' systems. These were chosen as they are readily available to a wider range of users than bespoke input systems. Text based input was examined first. The text input systems evaluated were: a keyboard, an on-screen keyboard, a handwriting recognition system, a voice recognition system and a wrist-keyboard (*sometimes known as a wrist-worn keyboard*). It was found that the most appropriate text input system to use overall, was the handwriting recognition system. (*This is further explored in the discussion of Chapters three and seven.*)

The text input evaluations were followed by a series of four experiments that examined pointing devices, and assessed their appropriateness as part of a wearable computing system. The devices were; an off-table mouse, a speech recognition system, a stylus and a track-pad. These were assessed in relation

to the following generic pointing tasks: target acquisition, dragging and dropping, and trajectory-based interaction. Overall the stylus was found to be the most appropriate input device for use with a wearable system, when used as a pointing device. (*This is further covered in Chapters four to six.*)

By completing this series of experiments, evidence has been scientifically established that can support both a wearable computer designer and a wearable user's choice of input device. These choices can be made in regard to generic interface task activities such as: inputting text, target acquisition, dragging and dropping and trajectory-based interaction.

TABLE OF CONTENTS

ACKNOWLEDGMENTS

Thanks to Chloe for supporting me both emotionally and financially throughout the duration of this work. Without her this work would never have been completed. Thanks to my supervisor Roy Kalawsky, my friends for sharing their PhD experiences, my dog Sniff for keeping me company during those long days spent writing up and Mike for suggesting that I, "just get it handed in".

A Special thanks to Stephen Brewster and John Connolly for taking the time and effort involved in marking this thesis.

GLOSSARY

AR - Augmented Reality

CPS - Characters Per Second

GUI - Graphical User Interface

HCI – Human Computer Interaction

HMD - Head Mounted Display

PC – Personal Computer

PPWS – Preferred Percentage Walking Speed

VDU – Visual Display Unit

VR – Virtual Reality

WIMP – Windows Icons Menus Pointers

WPM – Words Per Minute

WWL – Weighted Workload

Chapter 1: Thesis Introduction

1.1 Introduction

Within our lives we use a whole range of objects. Some of these objects are simple, easy and 'natural' to use. The way we interact with and use these many objects comes from our socialisation, cognitive experience and cultural acceptance of the way objects should be and are used within our society.

Wearable computers are 'body-worn' computing systems. They a can be worn in a variety of different ways, from belt-based to vest-based configurations, or even integrated into the user's clothes. The advantages of using such systems are: they allow the user to interact while they are mobile, wearing a computer enables the user to use computer-based technologies in any locale and they can provide continual access to information, as they are always with the user.

The aim of the research reported in this thesis is to give an insight into the way that wearable computers can exist as an extension of the user and their knowledge. Computing, cognition, interaction and culture are the four of the main areas of study that impact upon our understanding of any wearable computing system. One way to initially understand the user's interaction with a wearable computer may be to examine wearable computers in terms of their input mechanisms.

1.1.1 The Specific Focus of This Thesis

The specific focus of this thesis is the evaluation and use of input systems. In particular the thesis aims to examine the input modalities that are used within a wearable computing system and examine these in terms of human-wearable computer interaction. This will be done through adapting a critical evaluation of the existing research in the field and also through a series of scientific experiments.

1.1.2 The Contribution of This Thesis

Wearable computing is a subset of mobile computing and has become an area of academic research in its own right (Bass *Et al.* 1999). Wearable computers in their current form consist of a modular that can be worn. They consist of a power supply and CPU/Data-storage/Memory unit which is able to be connected to input/output peripherals such as: monocular displays, wrist keyboards, off-table mice, microphones and a whole range of input/output devices, that will be further discussed in the later chapters of the thesis. Wearable computers are distinctly different from other computer-based technology as they are designed to be worn and therefore used in a variety of contexts and environments as a viable on body (body worn) computing system.

The aim of this thesis is to research and evaluate human wearable computer interaction paradigms that will enable the user of a wearable computing system to have a more usable relationship with the system they use. This thesis will primarily focus upon the input modalities that are available to wearable computers users.

The key aims and contributions of this thesis are as follows:

- To examine and present techniques for the evaluation of input systems for use with wearable computers.
- To present a set of empirical results that can aid in the design and choice of input systems for use with wearable computers.
- To provide a research platform on which other Wearable Computing-focused Human Interaction studies can be based.
- To provide a set of metrics based on input speeds and error rates through scientific empirical investigation that relate to the use of input systems for wearable computers, while the user is stationary and mobile.

1.1.3 The Thesis Structure

In this section a general outline of the thesis is provided for the reader. The research literature relating to the experiments outlined can be found within Chapter two. Each experimental chapter provides an introduction to the experiment and related research. A graphical representation of the thesis can be found in figure 1.1. This provides a 'road-map' of the research conducted within the thesis.

Chapter 1. This chapter introduces the field of wearable computing in a historical manner, and places this thesis in context as a scientific piece of academic research. It provides an introduction to the literature review and explains the layout of the thesis.

Chapter 2. This chapter begins by outlining the key characteristics of a wearable computing system. The literature review then goes on to discuss the concept of the wearable computing paradigm. It further discusses the application of wearable computers and goes on to look at their common characteristics. This is then expanded upon to focus on input systems that can be used with wearable computers, for text input and as pointing devices. Also explored is the concept of the audio-centric wearable computer and evaluation strategies for investigating input devices. One of the issues raised by the literature review relates to the lack of empirical research into input devices for generic tasks such as text input and pointing in regard to wearable computing systems. This steers the thesis to further explore input devices for wearable computers in two particular areas: text input and pointer-based input.

Chapter 3. This chapter took four off-the-shelf text input systems: speech recognition, handwriting recognition, a wrist-worn keyboard and a QWERTY style virtual keyboard. The four input systems were chosen because the research literature had previously demonstrated that they were systems that had been used with wearable and mobile computer systems. The text input systems were evaluated in terms of the time taken to complete the task, the errors made while completing that task and the participants' subjective workload. The text input systems were evaluated while that participant was wearing a computer whilst in stationary and mobile environments.

Chapter 4. This chapter expands the experimental framework found in Appendix 5 by measuring the errors made during the performed task, and by having the participant select five target sizes instead of one. Each participant selected 1000 targets throughout the course of the study. The targets were selected while the participant was standing stationary and while mobile, using the vest-based wearable compute. It was found that the fastest device when used for target selection was the stylus, while both standing stationary and mobile.

Chapter 5. This chapter takes the three input devices used in Chapter five, and further evaluates their appropriateness as pointing devices for dragging and dropping when used as part of a wearable computer system. This study expands upon the studies in the literature and evaluates the pointing devices in terms of their use with a wearable computing system, while the users were mobile. The elements used to evaluate the devices were taken from the framework laid down in chapter 3. These were the time taken to complete the task, the errors made during the task and the subjective workload of the participants

Chapter 6. This chapter is the last chapter in the series of pointer-based evaluations for wearable computers and focuses on the use of wearable input devices for trajectory-based interaction. It furthers the evaluation of the three input devices in chapters five and six to provide a broader overview, relating to the way that many wearable users use input devices for a variety of pointer based interactions; from target selection and dragging and dropping, to trajectory-based interaction for steering through menus. As in previous experiments, the participants wore a Xybernaut vest-based wearable computer and did the task while stationary and mobile.

Chapter 7. This chapter reviews the findings from chapters three, four, five and six. It presents an overview of the findings, and presents them in a more 'useable' format. The findings are presented in a tabular manner. This allows the reader to see which input devices performed the best in relation to their input speed, accuracy and cognitive workload. This allows the users/designers of wearable systems to see what may be the most appropriate input device for the task they are to accomplish. For example, the task may require the user to input alphanumeric data quickly, while mobile.

Chapter 8. This chapter concludes by focusing on the original contributions to knowledge that this thesis has made. It then goes on to examine the future directions in which the work that has been carried out could be further expanded and offers a series of research routes. The pages following this conclude the thesis and end with a personal reflection.

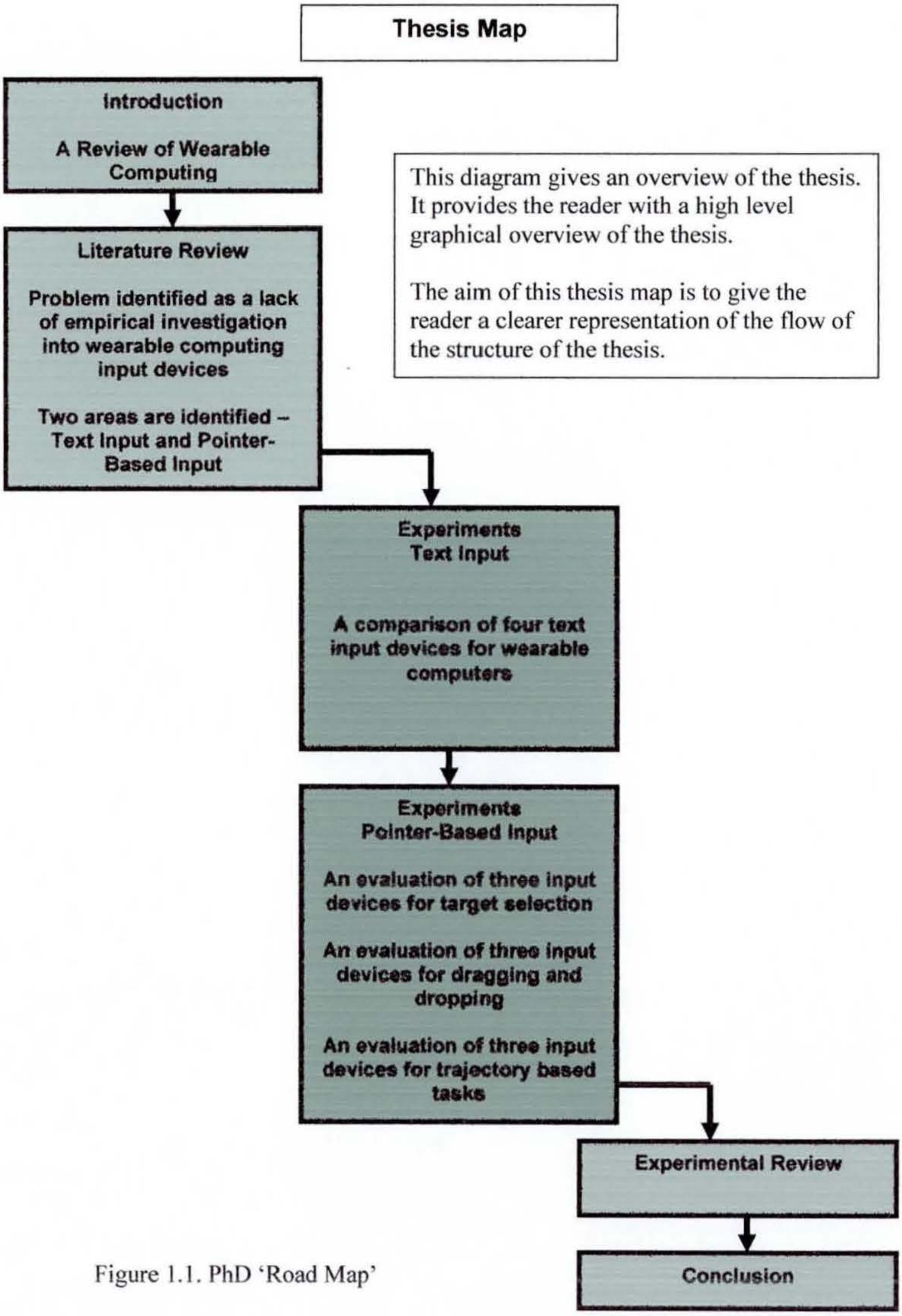


Figure 1.1. PhD 'Road Map'

1.1.4 Wearing a Computer

Different tasks call for users to wear their computers in different configurations. This is sometimes referred to in the literature as wearability - how a user wears their wearable computing system (Gamperele *Et al.* 1998; Watkins and Dunne 2003). It is no use wearing a full vest-mounted system if all the user needs is simple GPS information, but on the other hand it would be inappropriate to give a user such as a soldier a PDA if they constantly need to be sending and receiving AR-based tactical information, maps, weapons reports; photographs and wounded reports, the hardware would most likely be held in vest based configuration such as the Land Warrior System (Murray 2000).

A whole range of different systems have been developed and adopted by different users to carry wearable computers, from vest-based systems (Kortuem *Et al.* 1998; Woolley *Et al.* 2002), belt-based (Starner *Et al.* 1995; Mills & Beliveau 1998) through to digital body piercings (Eves *Et al.* 2000) that can be worn, and illuminated to tell the user when they have an incoming mobile phone message. The actual definition of what a wearable computer is, is further 'clouded' by the use of such terminology as: body-worn (Bass, Siewiorek, Smailagic and Stivoric 1995; Randall & Muller, 2000, 2002; Mann 1997), on-body (Starner 2004) and body-wearable (NASA 2000). These terminologies have been used to define a variety of wearable computing systems. Many definitions, such as vest-worn or belt-worn, refer to the way that the system is worn (the placement on the body). This discussion is further explored in regard to the literature in Chapter 2.

Billinghurst (2002) furthers the debate on the definition of wearable computing by arguing that cell phones could be thought of as wearable computers. He writes;

"*While it may be debatable that a cell phone is a wearable computer, the distinction becomes more uncertain when you consider the combination of devices that a person may be carrying, such as PDA, digital camera, cell phone, heart rate monitor and GPS receiver. These devices are chosen based on their individual attributes yet together they may provide more functionality than a general-purpose wearable machine.*" (Billinghurst 2002).

In light of the proliferation of designs that enable cell phone users to wear their phones by using: neck-straps, pouches and cell phone pockets in jackets; Billinghurst's discourse on the cell phone as a wearable computer is valuable.

A variety of rucksack/backpack 'wearable' systems have been developed. The benefit of using these systems is in their application as a prototyping tool such as the "Touring Machine", a mobile augmented reality system (Fiener *Et al.* 1997), the "Tinmith" project (Piekarski and Thomas 2002) and the orientation and way-finding system developed for visually impaired users (Ross and Blasch 2000). Using a backpack-worn wearable computer may be preferable for some disabled users; such as the visually impaired, as it is less obvious that they are wearing a computer, which could lead to some level of social exclusion.

To further understand wearable computers, the concept of computerised-clothing should also be explored. This type of wearable computing is more akin to the everyday clothes that we already wear as it attempts to merge and integrate everyday wear (clothes) with computer-based technologies. Randall (Randall & Muller, 2000, 2002; Randall 2001) coined this Bristol Fashion, but a more thorough investigation into the literature suggests a parallel evolution in the design of such systems by Philips (Eves *Et al.* 2000), Rekimoto/Sony (2001) and the e-SUIT (Toney, Mulley, Thomas and Piekarski 2002).

Philips (Eves *Et al.* 2002) integrates a level of computerised technology into clothing and explores the use of smart fabrics and their relation to fashion and society. Several novel concepts have been developed: from a suit with an integrated smart fabric keyboard to electronic sportswear that can monitor the user's pulse, temperature and blood pressure and even a jacket for snowboarding that can warn the user of dangerous conditions, which contains heating elements and a radio link. Other examples of Philips designs can be seen below.

Wearable Mp3 Player | Smart Fabric Controls | Play Jacket

Figure 1.2: Clothing-based Wearable Computers.

For a full breakdown of the wearable computing systems developed in figure 1.2 see - http://www.design.philips.com/about/design/section-13526/

These systems are very different from the previous vest-based and wearable backpack designs that we have just discussed. They are different, because computers have been integrated into the 'fabric' of everyday clothing, but there are several problems that have to be noted when examining this branch of wearable computing:

1: Clothes have to be washed so the technology has to be removed (Karvonen & Parkkinen 1994); if the wearable is waterproof it requires the user to have another garment with the same functionality.
2: The systems explored here are on the whole are designed to perform a dedicated task, and as such are designed for task-specific applications, e.g. an MP3 player.
3: Clothes do eventually wear out, become unfashionable and some users such as children grow out of their clothes. This means that the technology needs to be transferable to other garments.
4: Weather conditions often affect the clothes we wear. How do we design for this?
5: How can the designers of such garments apply a physical metric to the placing of controls on the garment, or is it a matter of using the "one size fits all" design approach?
6. Are there technological issues, such as the size and heat of computing technology that can affect the 'wearability' of these garments?

Although there are problems with fabric-based systems, there are also benefits. One of the benefits is the elimination of heavy cabling, as Orth (1998) suggests, "Using sewn fabric sensors and circuits allowed us to eliminate uncomfortable and heavy wires, connectors and electronics."

One interesting slant on wearing a computer is what can be called the wearable thin client paradigm. Systems such as the Panasonic Toughbook CF07LZ5ZY (Panasonic 2002) allow the user to leave the 'core' system containing the CPU, memory and hard drive up to 300 feet away from the screen, which means the user only has to carry the wireless touch screen, which equates to no cabling weight. The Toughbook can be worn on the wrist, placed in a drop down pocket case or simply clipped to a clipboard.

The advantage that this system has over the other vest and belt-based systems is its ability to remain socially and physically unobtrusive, which the literature suggests may be beneficial to the user (Rekimoto 2001). In some situations the unobtrusive nature of wearable computing technology can be important, where time is an issue for example. While conducting a contextual interview with a user of a wearable system he said that "I used to waste up to an hour showing other people the system I was using; as soon as they saw the HMD they wanted to have a go". There are implications related to the novelty of new technologies, and these can impact upon the user's task when used in a social rather than an individual setting (Ockerman & Pritchett 1998).

1.1.5 Conclusion

This introduction has introduced the concept of wearable computing and discussed some of the discourses that exist around the definition of these systems. It outlined the key contributions that the thesis will make and shows the outline of each chapter. Figure 1.1 graphically represents the structure of the thesis. The next chapter further examines the literature and focuses more specifically on input for wearable computing systems.

Chapter 2: Literature Review

2. 1 A Review of Wearable Computing

2.1.1 Introduction

This chapter explores the concept of wearable computing and the associated technologies that come together to form this field of research. It characterises wearable computing in a more formal way so as to set an agenda for the style of the thesis, which will be reflected throughout the rest of this thesis.

2.2 A Characterisation of Wearable Computing

It is important to see the field of wearable computing as a multi-faceted field of computer science that when dissected includes the following elements: Mobile Computing, Ubiquitous Computing, HCI, Ergonomics and Networking.Wearable computers have been classified in several ways (Bass 1997; Rhodes 1997; Mann 1998). Rhodes (1997) characterises wearable computers in the following ways:

Portable while operational: The most distinguishing feature of a wearable is that it can be used while walking or otherwise moving around. This distinguishes wearables from both desktop and laptop computers.

Hands-free use: Military and industrial applications for wearables especially emphasize their hands-free aspect, and concentrate on speech input and heads-up display or voice output. Other wearables might also use chording-keyboards, dials and/or joysticks to minimize the tying up of a user's hands.

Sensors: In addition to user-inputs, a wearable should have sensors for the physical environment. Such sensors might include wireless communications, GPS, cameras or microphones.

"Proactive": A wearable should be able to convey information to its user even when not actively being used. For example, if your computer wants to let you know you have new email and who it's from, it should be able to communicate this information to you immediately.

Always on, always running: By default a wearable is always on and working, sensing, and acting. This is opposed to the normal use of pen-based PDAs, which normally sit in one's pocket and are only woken up when a task needs to be done.

Bass (1997) defines the properties of wearable computers in a more concise manner, but also argues that wearable computers should 'exist within the corporeal envelope of the user', what we might think of as the users 'space', both physically and mentally:

"They may be used while the wearer is in motion
They may be used while one or both hands are free, or occupied with other tasks
They exist within the corporeal envelope of the user, i.e., it should be not merely attached to the body but becomes an integral part of the person's clothing
They must allow the user to maintain control
They must exhibit constancy, in the sense that they should be constantly available." Bass (1997)

Bass (1997), Rhodes (1997) and Mann (1998) use a common descriptive set of features to describe the properties of wearable computing systems. Mann's (1998) keynote address of the 1998 International Conference on Wearable Computing offers a more descriptive outline of what a wearable computer is. This shares common ground with both Bass' and Rhodes' descriptions *(below)*

"A wearable computer is a computer that is subsumed into the personal space of the user, controlled by the user, and has both operational and interactional constancy, i.e. is always on and always accessible...it is a device that is always with the user, and into which the user can always enter commands and execute a set of such entered commands, and in which the user can do so while walking around or doing other activities" (Mann 1998)

Mann (1998) then extends his definition to stress that the main difference between other wearable technology-based devices, is that they are not able to be programmed (re-configurable), whereas wearable computers are.

"what sets the wearable computer apart from other wearable devices such as wristwatches, regular eyeglasses, wearable radios, etc... these other wearable devices that are not programmable (re-configurable), the wearable computer is as re-configurable as the familiar desktop" (Mann 1998)

What is interesting is that in the years since Mann's (1998) paper was given that there are now many devices that can be programmed and are highly re-configurable. Mobile phones are now java enabled and run java-based programs, while some watches have the Palm Operating System, but the key difference is that these devices have a sole purpose and are developments of other technologies. The mobile phone is an extension of the normal land-line based telephone, while the Palm watch is an extension of the wristwatch.

After examining the three definitions it is clear to see that there are commonalities between these definitions:

1. Wearable computers should be able to be used while the user is mobile.

2. The user should be able to enter data and control the system.

3. The system should allow the user to use the system in a hands-free or reduced handed mode.

4. The system needs to be worn.

Within the literature we have identified that there are some shared ways that wearable computers have been defined by Mann (1998), Bass (1997) and Rhodes (1997), but to further understand wearable computing we must examine the systems themselves and see how they are worn, as this is the main feature that separates them from other forms of computing technology.

2.2.1 Wearable Computing Tasks

This section aims to give an overview of some of the application areas and tasks that wearable computers are used for. One of the key differences between desktop computers and wearable computers is the ability to wear and use wearable computers in a variety of different locations. When the user wears their wearable computer the computer goes wherever the user goes. This means that wearable computing technology is ideal for in-field tasks (Matias, MacKenzie & Buxton 1996; Smith *Et al.* 1995; McConnell 2004; Nusser *Et al.* 2001) and environments where a desktop or laptop computer is not able to be used in an appropriate manner (Schmidt *Et al.* 2000).

"A wearable is well suited to work "in the field." This includes data collection and retrieval in out-of-doors situations, such as: field service, inventory, surveying, etc. As well, highly mobile in-door work can also benefit from having a wearable. For example, inspectors performing on-site evaluations of factories can take notes and call-up information." (Matias, MacKenzie & Buxton 1996).

As wearable computers can be used in-field, it means that they are used in a varied range of environments from underwater scenarios (WetPC 2001), archaeological sites (Cross *Et al.* 2002), through to systems designed for use in space (Matias, MacKenzie & Buxton 1996; Carr *Et al.* 2002). The WearSAT (Carr *Et al.* 2002) system allowed its users to monitor oxygen levels, the user's orientation, call up maps and diagrams, and examine daily tasks through a menu-based system.

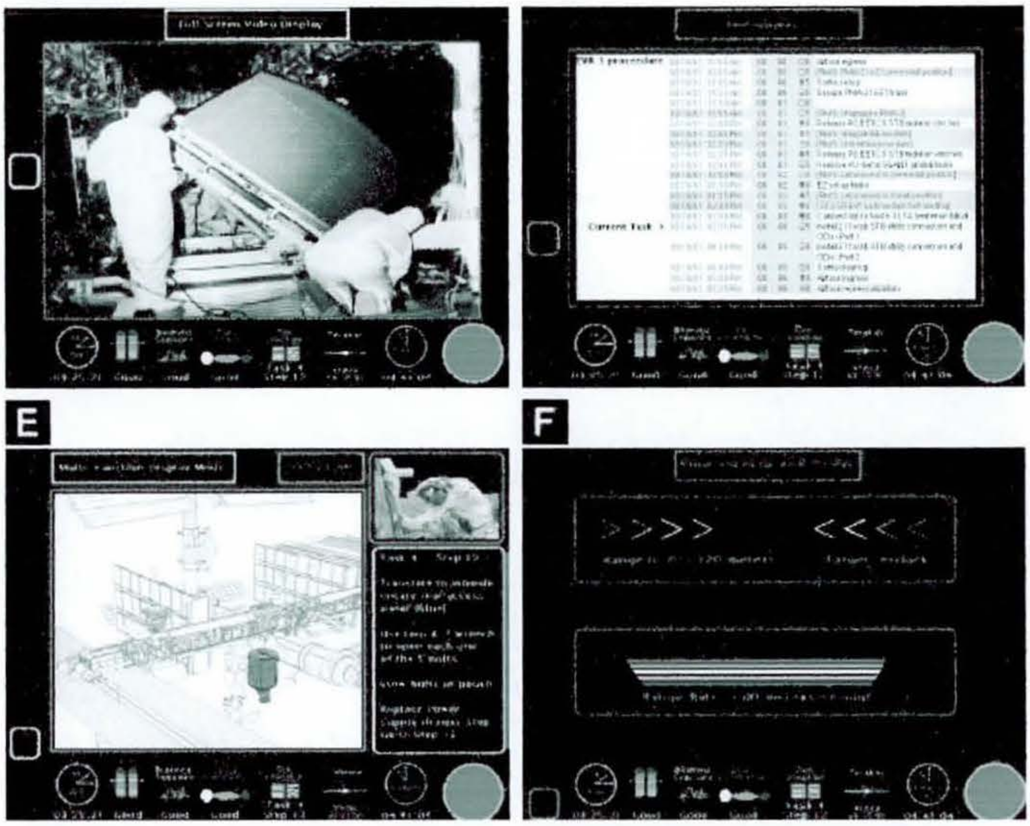


Figure 1.1: Information Management Screens from WearSAT (Carr *Et al.* 2002)

Other systems have been designed for on-site aviation maintenance. These systems were primarily designed so faults could be identified, recorded and the parts ordered (Smith *Et al.* 1995). As McConell writes, the use of wearable computers for aviation-based 'breakdowns' provides a "highly convenient means for performing non-routine maintenance tasks with immediate, real-time data entry" (McConnell 2004). Other wearable computing systems such as the WetPC (2002) allow the user to effectively gather data, access and edit textual material when in underwater environments and performing work such as logging data positions in coral reefs. The interface for the WetPC system is shown below in figure 1.2. This interface consists of a series of buttons for dedicated selection, and a text box, for inputting descriptive information.

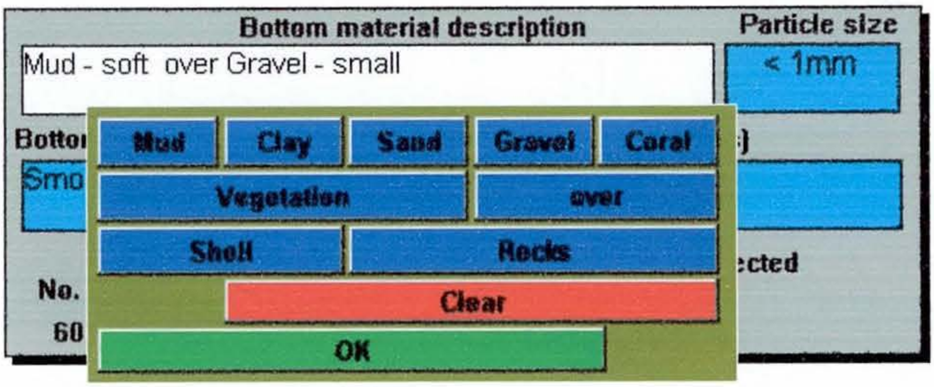


Figure 1.2: The WetPC interface for entering data on seabed (WetPc 2002).

For a full overview of the WetPc system see -
http://wetpc.com.au/html/technology/wearable.htm#diverdiagram

Similarly Nusser (*Et al.* 2001) state that the main task that many wearable computer users perform is collecting field data; as we have seen this field data can be collected in a variety of places. An interesting use of a wearable computing system was the Threat Response system (Rensing *Et al.* 2002), an in-field system that allowed its user to capture images of human faces and identifying them as a threat or friend. The interface is for this system is seen below figure 1.3. It allows the user to point and click (target selection) a series of buttons that perform dedicated functions, such as search, back and main menu.

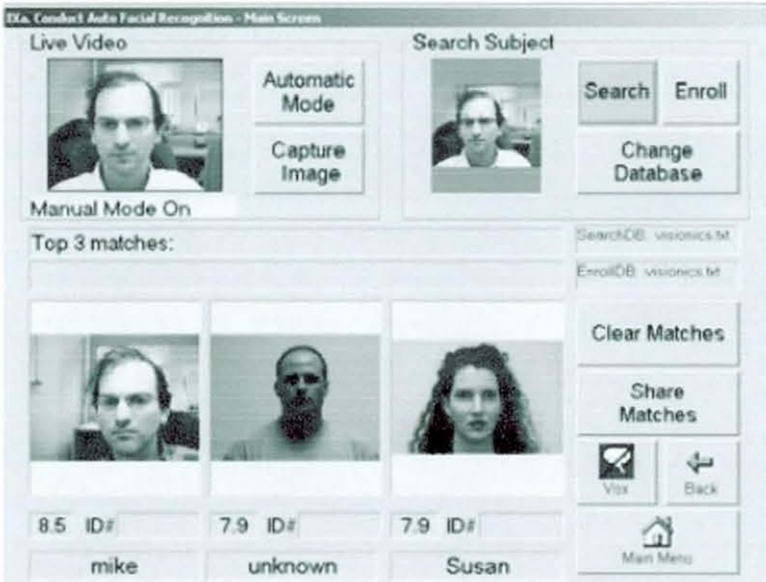


Figure 1.3: Threat Response Wearable Interface. (Rensing *Et al.* 2002).

Bürgy (*Et al.* 2002) report on a system designed for the inspection of bridges and buildings. The key tasks of this system were data gathering and access to data. Bürgy (*Et al.* 2002) further expand upon their system by defining the tasks performed by their system in three main ways that are outlined below.

Primary Task: No interaction with the device; i.e. no IT support is needed and applied

Support Task: Sole interaction with the device and the device supports the user in providing information or accepting information; i.e. 'productive' steps are done; e.g. reading a manual or inputting inspection results.

Control Task: Sole interaction with the device, but the task only involves navigating through the software; i.e. no 'productive' steps are done – scrolling down a page, opening a file or putting a file in a folder. (Bürgy *Et al.* 2002)

What Bürgy (*Et al.* 2002) starts to address is the relationship between the task requirements, user requirements and the interface, input and output modalities that are appropriate to allow the user to complete their task in a useable manner. As we have previously seen, many wearable systems are used in 'the field'. Because of this the user must be able to input data and receive data in an appropriate manner.

The graphical user interfaces that we have shown in this section have been specifically designed for wearable users to input, manipulate and access information; and although they are not desktop-based they do share elements with desktop based interfaces such as the ability to input data, buttons for selecting functions and menu-based interfaces (Basko *Et al.* 2002). Although there are similarities between wearable and desktop interfaces, interaction with a wearable can be done while the wearer is mobile and in a variety of locations, while on the other hand, the desktop user is always stationary. Because interaction with wearable computers can happen while the user is mobile, there is a need for metrics in regard to wearable input devices (Bürgy *Et al.* 2002). By empirically evaluating wearable input devices we may be able to establish what the most appropriate input devices/systems are for use with wearable computing systems.

2.2.2 Summary

In the review of the literature relating to this, we have found that the most common way of wearing a computer is in a vest-based system, although there have been no studies that have reviewed how many systems have been designed to be worn in that manner. We have also seen that the style of wearing a computer depends upon the way it is to be used; a simple MP3 player could be integrated into clothing, but a system previously mentioned such as the Land Warrior (Murray, 2000) requires the user to carry many other components, such as a backpack or compass. By wearing a computer in a vest-based system the user can distribute the weight comfortably over their torso. The system seen below is the commercially available Xybernaut MA V, worn in a vest. These systems are widely available as an off-the-shelf system.

Figure 1.4. A Vest Worn Xybernaut Wearable Computer

For a full overview of the wearable computing products available from Xybernaut see - http://www.xybernaut.com/home.asp

The architecture of this wearable system is modular in design. The CPU is the core unit which allows the connection of any appropriate output devices, such as a Head Mounted Display and touch screen (flat panel display in the diagram above). This system also allows the user to connect a whole series of input peripherals from microphones and touch-pads to wrist-mounted keyboards and stylus for interacting with the touch screen. The wearable computer is very adaptable in its design, as it allows the architecture of the computer to be configured in a way that supports the user's needs and the task that needs to be accomplished. The task and user requirements can impact upon the design of the software, the input and output modalities, the input and output technologies and the configuration of the wearable system.

2.3 Introduction – Text Input

Earlier in this chapter it was noted that the literature that defined wearable computers shared four common characteristics.

1. Wearable computers should be able to be used while the user is mobile.

2. The user should be able to enter data and control the system.

3. The system should allow the user to use the system in a hands-free or reduced handed mode.

4. The system needs to be worn.

In the previous section (1.1.4) the 4^{th} definition was examined; in this section we aim to address the other three. As 2 and 3 both relate to input we shall follow this theme through and relate them to *1. Wearable computers should be able to be used while the user is mobile,* examining the literature to see if any research upon data input has been carried out while the user has been mobile. This is dependent on the task that the user needs to accomplish.

Wearable computing user input paradigms have had to adapt to the contexts where they are used, the task and context are reflected in the way that the wearable computing system is designed and used. One of the concerns for any wearable user is the choice of input device or input system as this is an integral part of interfacing with the application.

Using existing desktop input devices is not appropriate for wearable users for several reasons. The desktop interaction paradigm involves a user sitting in front of a monitor and using a mouse and keyboard to interact with the computer. The act of being mobile and wearing a computer places different constraints upon the user, that the desktop does not. Desktop keyboards are, "difficult to use in particular settings and circumstances, such as conducting inventory on the shop floor or a geology survey in the wilderness...using a keyboard is cumbersome" (NIST 2001) . The same can be said of the mouse. If we are using a computer while mobile, using a desktop mouse as a pointing device is not a viable pointing solution, making the user stop and find a flat surface every time they needed to move the cursor is not appropriate (Rhodes 1997). Desktop devices are designed to be used on flat surfaces, hence the term 'desktop'; weight is not an issue as devices such as keyboards and mice do not have to be carried around for sustained periods. For a wearable user we not only have to consider the weight of the device, but also the weight of the cabling. The device has to be small enough to be worn 'on body' and possibly to be stored on the wearer too. The user may also be required to input with one hand or be working in a hands-free environment (Noble *Et al.* 1997). The wearable computer user may be stationary or mobile when they use an input device, so any input devices must have a high level of usability while the user is mobile.

Integrating existing desktop devices is not a satisfactory design solution, but focusing on the use of desktop devices in a mobile environment has pointed out six very important factors for us to consider when designing and evaluating an input system for mobile use. The task requirements for a mobile user can dictate: the weight of a device and connectors, its size, where to store the input device when not in use, mobile usability, wearing a device and how does the user input in a non-traditional desktop manner?

Wearable computing input models can be defined as two groups; physical and non-physical devices (Dix *Et al.* 2004). Physical refers to many types of input devices such as keyboards, joysticks, stylus/light pens, touch-pads and mice, while voice/speech recognition and eye gaze can be defined as non-physical.

A range of devices has been developed and adopted for use with wearable computers and users on the move. These range from off-the-shelf text input devices such as the wrist keyboard, speech recognition software and finger mice, through to custom devices that have been developed, such as chording gloves, half-keyboards (Matias *Et al.* 1994, 1996), twiddlers (Lyons *Et al.* 2003), smart fabrics (Rantanen *Et al.* 2000, 2001) and virtual keyboards (Howard & Howard 2001).

It is a difficult and abstract task to think of other ways of inputting text, especially after the full size QWERTY keyboard input method has been dominant as a text interaction style for over a hundred years (1868) and the mouse has become a standard pointing device for many desktop users.

How can we move the cursor without a mouse while the user is mobile? How can we input textual information if we are mobile? And are there already existing systems that can be used or adapted to be used as part of a wearable computing system? It is questions such as these, and the associated problems and issues that arise from such questioning, that have given rise to a generation of wearable and mobile computing research. This research has focused upon the development of new and interesting interaction modalities and paradigms for non-desktop computers.

One of the key issues that has not yet been fully investigated is the usability of wearable interaction paradigms and the users' physical and cognitive workload while they are mobile and wearing a computer.

This section of the dissertation examines some of these new input modalities and gives examples of already existing text input technologies as a comparative exercise. It explores their use and looks at the problems and issues associated with these technologies and their use. We start with an examination of text input; as Mackenzie and Soukoreff (2002) write "Although text entry is by no means new in mobile computing, there has been a burst of research on the topic in recent years. There are several reasons for this heightened interest: First, mobile computing is on the rise and has spawned new application domains such as wearable computing".

2.3.1 Text Input

The first part of this section will critically approach the subject of text input in relation to wearable computing-based input. When referring to text input we are examining the use of technologies that are used as part of a wearable input paradigm/system for the input of alphanumeric data. There are many different types of text input devices that are available commercially and custom-made for use with wearable computers. An approach will be taken in this review that encompasses a wide range of text input systems; from common input devices such as the keyboard, through to more 'obscure' devices such as the Twiddler (Lyons *Et al.* 2003), the ISO/IEC 9995-8 (1994) keypad, and half-keyboard designs (Matias *Et al.* 1993,1994,1996) that have been used as part of a wearable computing system.

2.3.2 Numeric Input

Text input has been the primarily focus of this investigation so far, but there are other important studies that have focused upon numeric input. This is important because there are many systems that rely on numeric input. The use of such systems may range from telephoning a friend to collecting stock data in a warehouse. It is often the case that numeric keyboards can be used to input textual data, such as the ISO/IEC 9995- 8 1994 Keyboard/Keypad and TNT (The Number Typer) (Ingmarsson *Et al.* 2004). The TNT system was developed to take advantage of a television remote control and allow the user to input a variety of information using its ten-key (1 to 0) layout. In its simplest form, the user can use the 1 key to input the letter *a*. Its design was focused around inputting information, such as: the time, day and programme title into home multi-media terminals and interactive TV.

Other studies have focused on sonically-enhanced numeric buttons using PDA's (Brewster 2002). This study found that by sonically-enhancing buttons on the screen of the PDA, the usability was improved overall. When the button sizes were reduced from 5mm to 2.5mm there was little loss in relation to the user's quantitative performance when doing data input tasks, although the user's workload with these smaller buttons was significantly raised. The study proved that the usability of mobile devices using small onscreen-buttons is raised if a level of sonification is added to the buttons.

2.3.3 Keyboard Input

Keyboard input for a majority of desktop computer users is the most used mode of text-based input. The traditional 'off the shelf 'computer system invariably comprises of a base unit, monitor mouse and keyboard. It is important to examine the keyboard as it may be a way of allowing the user to transfer the skills they have acquired using this input device to a wearable computing based environment. The QWERTY keyboard has existed for over a hundred years, and was originally designed in an attempt to stop the frequent type-bar jams of mechanical typewriters and was much later adopted as the layout for the computer keyboard. The layout design in 1868 by Christopher Latham Sholes worked by separating the most frequently used letter pairings in an attempt to stop the mechanism jamming. However, as Matias, MacKenzie and Buxton (1996) write, "The QWERTY keyboard has been much maligned over the years. It has been called, by various authors: "less than efficient" (Noyes, 1983, p. 269), "drastically suboptimal" (Gould, 1987, p. 16), "one of the worst possible arrangement[s] for touch typing" (Noyes, 1983, p. 267), "the wrong standard" (Gould, 1987, p. 23), and a "technological dinosaur" (Gopher & Raij, 1988, p. 601), it still remains the predominant method of text input for users today.

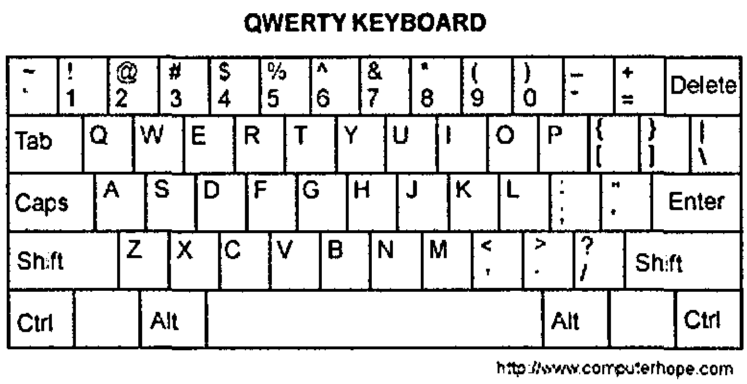


Figure 2: The QWERTY keyboard layout

As the literature suggests (Matias, MacKenzie and Buxton 1996), the QWERTY keyboard is a badly designed for text input, but compared to the original ABC layout that the QWERTY keyboard superseded, the QWERTY keyboard is in some cases faster to use (Noyes 1983). Other keyboard layouts have been designed, most notably the Dvorak keyboard, which places the most commonly used characters together in the touch typists 'home row'. Although finger movement (distance moved to other keys) was reduced by up to 90% there was only a 2%-5% increase in typing speed (Potosnak 1988). This may seem minimal but using a Dvorak keyboard layout gives a reduced finger movement rate of up to 90%, which puts less physical stress upon the users fingers and therefore should equate to fewer instances of RSI (Repetitive strain Injury), an important factor when related to daily use. Yet the Dvorak keyboard is little used in comparison to the QWERTY layout.

2.3.4 Wrist-Worn Keyboards

Wrist-Worn Keyboards conform to the layout of the QWERTY keyboard, but are essentially smaller versions of desktop keyboards, as yet no ABC or Dvorak wrist-worn keyboard are available. A smaller incarnation of their desktop counterparts, the keyboard/keypad has an important role to play as an input intermediary between the user and the computer in a wearable computing system. The wrist-worn/mounted keyboard at first glance would seem to be a 'quick fix' solution as it is easy for users to transfer their already existing cognitive models of the desktop keyboard onto the wrist-worn system, although the wrist-worn keyboard is not without its own problems, as we shall discuss later.

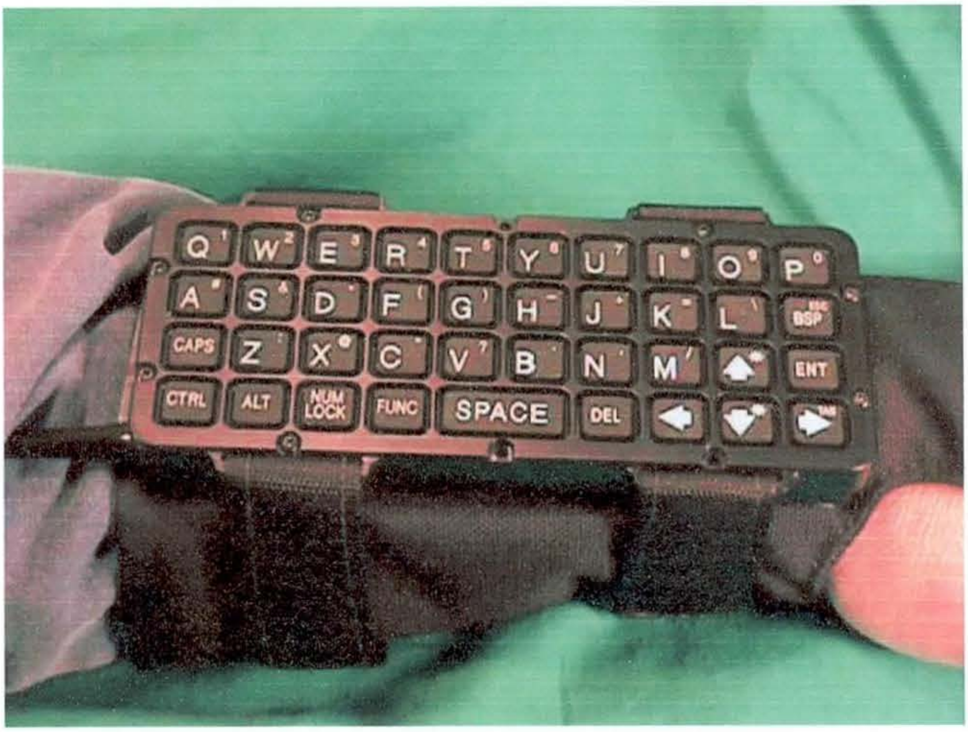


Figure 2.1: WristPC Keyboard *(above)*

Full details and specifications for the keyboard in figure 2.1 can be found at - http://www.l3sys.com/keybd/keybd.html

Mobile and wearable computing devices need to be smaller than conventional desktop computers, because they are carried in pockets, worn in vest systems (Sawhney 2000) and worn by their users in belt based systems. Therefore it is important that they are lightweight, easy to store and enable their users to have them 'on-body' for sustained periods of time. A solution to this problem is for the user to wear the keyboard. These factors have lead to a miniturisation of the associated input peripherals, essentially for the reasons previously explained: size, weight and storage.

Surprisingly, little investigation and research have been done on analysing wrist-worn QWERTY keyboards in terms of their usability, input times and error rates, although they are widely available. In a study by Thomas *Et al.* (1997) it was found that the wrist keyboard (or forearm keyboard as it was known in this study) was much quicker to use than both a belt-worn mouse and a chord keyboard in relation to doing a variety of tasks such as text input, although this may be due to the users' already existing familiarity with the QWERTY layout. The study also did not require the user to use the system whilst mobile, thereby ignoring a whole range of issues that affect wearable user whilst mobile.

There are also underlying issues that relate to the users' physiology in terms of the long term use of wrist keyboards that have yet to be investigated. RSI, although not the focus of this research, could be a factor in using wrist-worn keyboards. One only has to use a wrist keyboard for a while to notice there is some strain placed on the typing wrist elbow, shoulder and back, while the wrist wearing the keyboard has to be held still, involving even more muscle groups, and that is only when the user is stationary. Researching the physical factors influencing the wearing of technologies is a big task (Gemperle, Kasabach and Stivoric, 1999; Knight *Et al.* 2002); although researchers are starting to make inroads into the field it is a difficult area to investigate because of the changing state of the environment and impact upon the users' physical movements.

Other wrist-worn keyboard models do exist that are designed to be used with wearable computers, such as the Matias half-QWERTY keyboard (Matias *Et al.* 1994) Figure 2.2. This was initially designed to help with the transfer of two-handed typing skills to a single-handed keyboard design. The half-QWERTY system works by the user using the space bar to toggle between the two halves of the keyboard; e.g. to type S but the user would press the S key, but if the space was pressed first the S key would change to the L key. Matias (*Et al.* 1993) reported by that experiments had found the half-keyboard to 'exceed the speed of handwriting', and touch typists <10 hours (10 1 hour long sessions) to learn before they could reach speeds of up to 40+wpm (23.8 to 42.8 wpm).

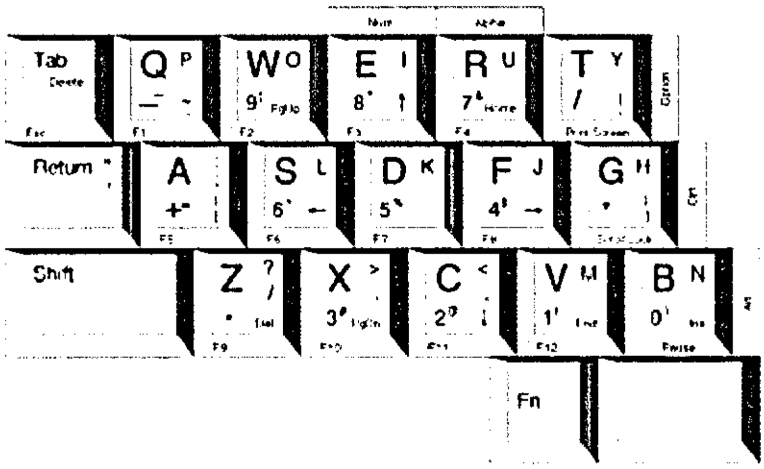


Figure 2.2: Matias half keyboard (Matias *Et al.* 1994)

For more information regarding the half keyboard see - http://half-qwerty.com/

The design of the half-keyboard is problematic for the following reasons: firstly, the user needs substantial training, 10 hours training for a touch typist (as used in the experiment) to reach 40wpm, but how does this training time relate to non touch typists who have to look at the keyboard to type ('hunt and peck typists')? The second problem is that when the user is typing they have to press the space and therefore they are making more keystrokes and more keystrokes can relate to more time and errors. A solution to this may be to build a level of text prediction into the software or auto error correction. Matias (*Et al.* 1994) also reports a 7.44 percent error rate, which he admits is double that of the touch typists' original full size QWERTY input, which means there would have to be a large amount of corrections to make. This system would be of little use for real-time tasks, such as reporting an accident or number plate, because of that error rate. The half-QWERTY was later developed as the main text input modality for a wearable computer for use in microgravity space and other non-desktop environments (Matias *Et al.* 1996). Although it is commercially available, it is not supplied by any of the major wearable computer manufacturers, who prefer to stick with the standard QWERTY-based wrist-worn device.

2.3.5 Summary

The QWERTY wrist worn keyboard uses some of the users' existing knowledge, which is a positive feature of the system, as users instantly recognise the system, but there are negative factors associated with using wrist-worn keyboards in general. Using keyboard commands such as Ctrl+Alt+Delete are difficult for user a user to accomplish using one hand. Key combinations may depend on the task that is being performed; if the Ctrl+Alt+Delete combination was needed frequently, software could be written to accept a different key combination, such as SPACE+Alt. As Goldstein (*Et al.* 1998) state, 'If the QWERTY keyboard is miniaturised, usability suffers both regarding effectiveness and efficiency', so in reducing the keyboard size we may also be reducing its usability. Using a wrist-worn keyboard does not allow the user to be 'hands free', although the user can hold or control something in the other hand whilst they are typing. In physical terms of the hardware configuration on the user's body, it adds more cables, it's another peripheral for the user to consider (uses up a port) and common sense reveals the problems involving clutter that looking at a keyboard through an augmented reality display would cause.

2.4 The ISO/IEC 9995- 8 1994 Keyboard/Keypad

The ISO/IEC 9995-8 1994 12-key Keyboard/Keypad is the keyboard input configuration mostly used on a majority of mobile phones. As mobile phones become part of the fabric of everyday life and the instances of SMS messaging rise the 12-key keyboard becomes an important text input modality in its own right. If we examine the rise in SMS messaging, it has risen to from 16.5 billion in 2002 to 22 billion in 2003 (T-Mobile 2004). Over a billion people own mobile phones worldwide. It is therefore important to examine the 12-key keyboard as a viable form of mobile text entry. There have even been instances of touch typing using this configuration reported; "teens in Finland are learning to touch type with their thumbs," said Marc Retting, a member of HannaHodge, a market research firm in Chicago (Batista, 2001). Yet other critics such as Goldstein (1998) claim 'the 12-key keypad (0-9, *, #) is not adequate for touch-typing', but no reference is made to any study or definition of touch typing. Other mobile phone keyboard layouts do exist, such as Fastap, seen in Figure 2.3. This increases the key density of mobile phones on average by 240%, by adding more keys, placed at the corners of the main keys (Hare 2002).

Figure 2.3 A Fastap Keyboard layout.

Further information on Fastap products can be found at - http://www.digitwireless.com/

Modern mobile phones are capable of allowing the user to input text in 2 different ways, multi-key input and predictive input. Multi-key means that to write a word such as ***on*** the user must press button **6**, three times for the **o** and two times for the **n:** all together five presses for a two-letter word. Predictive text input attempts to match a word to the key pattern that the user selects. So the user simply presses the keys with the coordinating letters upon them. As we can see from the diagram below, to write the word 'how' the user simply has to press buttons 1=h, 6=o and 9=w. This takes three presses instead of six. Although T9 prediction is not perfect since some words have the same key sequences, such as 'cake' and 'calf', and if this situation occurs the user selects the 'next word' function key to select the word required, but these cases are rare (MacKenzie *Et al.* 2001). In a recent report/user survey conducted by Digit Wireless (Hare 2002) it was found that 45% of users use T9 prediction, while 50% still use multi-key (Hare 2002).

What this displays is that T9 input has a large level of acceptance and usage within the mobile market, although this maybe counterbalanced by the fact that Nokia uses T9 as its main form of predictive input; interestingly the survey did not report on any other forms prediction such as Motorola's iTap.

Figure 2.4: Diagram of T9 Predictive Input, see - http://www.t9.com/

2.4.1 Summary

Using an ISO/IEC 9995- 8 1994 Keyboard/Keypad, seems like an ideal text input solution. Keystroke input times are fast and the keyboards are small. This keyboard layout is used by a large amount of people and so skills transfer would be an easy for the user. 12 key keyboard layouts were used because a fully usable QWERTY keyboard layout cannot fit onto a small mobile phone. The 12 key layout is an inherited design from existing telephones, whose original designers did not design it for text input. The typing style used for input is thumb based, so the user holds the handset in their palm and uses their thumbs to type. This could become uncomfortable after sustained periods of input.

2.5 Stylus-Based Input

A stylus is a pen like device that is used scribe, tap and control a cursor. Stylus based input systems fall into four main categories: traditional handwriting recognition, stroke based input, soft keyboards (sometimes referred to as virtual keyboards) and gestural-based input.

2.5.1 Handwriting Recognition

Handwriting recognition is a very desirable input modality as a majority of people can either handwrite or print. Mapping the user's already existing knowledge and motor skills onto another technology is an ideal solution as it minimises the amount of training needed to interact with and use the system (Connell 2000). It could be said that, "data entry using a pen forms a natural, convenient interface" (Connell 2000). In a wearable handwriting interaction paradigm, the paper and the pen which we normally associate with handwriting are replaced with a touch-screen and a stylus; no cables are needed for the stylus and the input source is in the same place as the output source. One advantage of using a stylus with a touch-screen is the ability to augment the user's written material with pictures and diagrams that can be integrated into a document or formalised at a later date using software such as Microsoft Journal.

Figure 2.5: User uses handwriting recognition system.

Vest-based Xyber panel with stylus
http://www.xybernaut.com/Solutions/accessories/accessories_display.asp

The main two types of 'natural' writing input are, cursive (joined handwriting) and printed writing. Soukoreff's *Et al.* study (*Et al.* 1995), examines handwriting input speeds for paper-based systems to provide an index for what we may expect the range of input speeds using a computer-based handwriting recognition system. Printing-based input ranged from 12wpm to 23wpm (Bailey 1987. Card *Et al.* 1983); in comparison cursive handwriting speeds are somewhat higher, ranging from 16wpm (Devoe 1967) to 30wpm (Wilkund and Dumas 1987) for advanced users. What are not compared in these studies, however, are the other factors that are associated with handwriting recognition based systems such as error rate, which may involve a level of auto correction, and error feedback, the amount of time the handwriting takes to be processed by the system, the interface through which the user inputs the text and most importantly the impact that being mobile will have upon the user in terms of input speed and error rate. As we have already noted, the user of any wearable system will certainly be mobile so we must examine the impact of mobility upon the users input performance.

2.5.2 Symbolic/Stroke Based Stylus Input

Symbolic/stroke based input is when a symbol refers to a letter. These include systems such as Unistrokes, shown in Figure 2.6 (Goldberg & Richardson 1993) and Graffiti (Blickenstorfer 1995) (for a full investigation of Graffiti refer to MacKenzie, and Zhang (1997)). Unistrokes, so called because it takes one stroke to scribe each letter, was a system developed by Xerox at its Palo Alto Research Centre (Goldberg & Richardson, 1993). It is a strictly alphabetical system and contains no punctuation, numbers or other special characters. These have to be used by dedicating your own strokes to other symbols Chang *Et al.* (1994) notes that the best form of text entry relies on transference of existing handwriting skills, with reference to studies by Wolf *Et al.* (1991) and McQueen (1994), but Chang also references two other studies by Veniola (*Et al.* 1994) and Goldberg (*Et al.* 1993), noting that stroke-based text input is problematic because the strokes first have to be learnt. This means the user has to learn the strokes before full input functionality can be achieved, but Microsoft's Jot as seen below in Figure 2.7, uses symbols that look the same as standard letters, in an attempt to lower learning times. It also allows users to use up to four different symbols for each character; it also incorporated the Unistrokes alphabet into its pattern recognition system. Goldberg &

Richardson's (1993) original study founds an input rate of 2.8 characters per second, but failed to take into account errors within that rate. Another single stroke system was the minimal device-independent text input method, MDTIM (Isokoski 1999). Isokoski found a 7.5wpm average using this system, but the symbols developed by Isokoski are abstract and need to be learnt by the user before faster input speeds may be obtained.

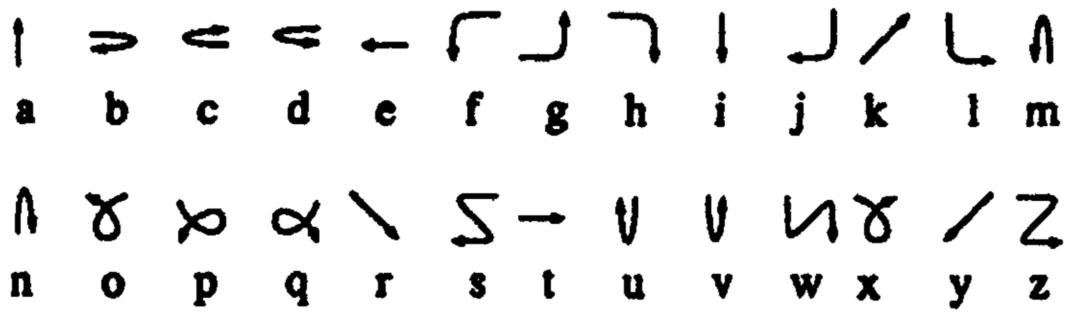


Figure 2.6: The Unistroke alphabet above (MacKenzie and Zhang 1997).

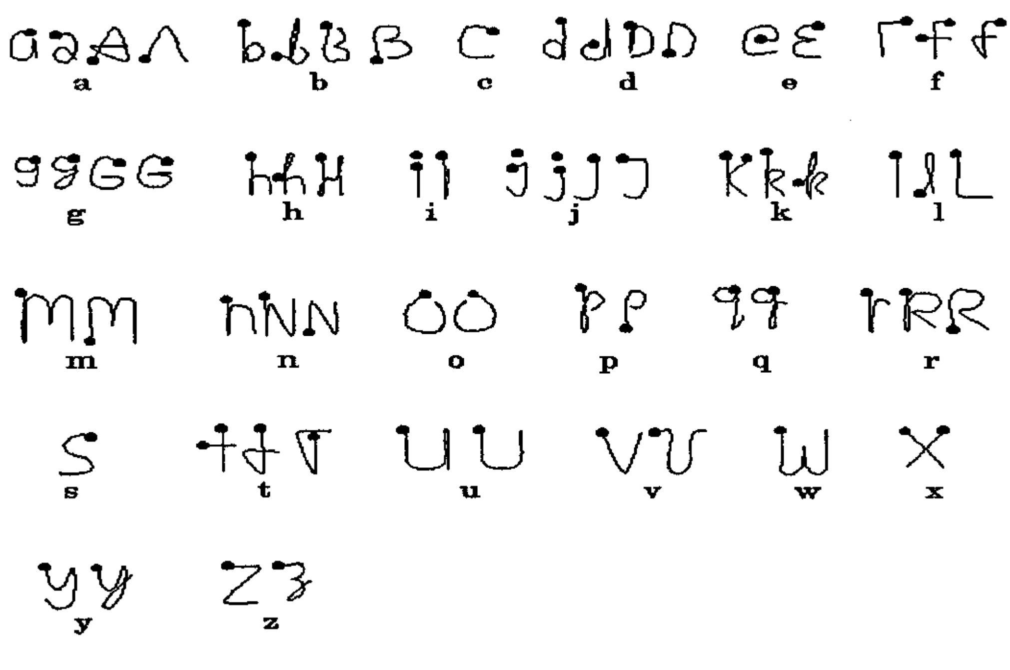


Figure 2.7: Microsoft Jot character recognition set (MacKenzie and Soukoreff 2002).

A vision-based single stroke character recognition system for wearable computers (Faruk-Özer *Et al.* 2001) was developed as an evolution of the stylus touch-pad/screen model. It uses a vision-based system (mounted camera) to recognize each stroke that the user makes. So in theory the user could be in any environment wearing this technology and be able to input textual information anywhere, as long as they had a stylus and the camera was pointed at it. The main problem with this system is that it was never integrated into a wearable computer and no evaluation was carried out while the user was mobile; this could obviously add extra movement, such as the sway of the body and its effect on the limbs that could affect the recognition system (Schmidt, Gellersen and Merz 2000); however it was evaluated using three different lighting conditions to mimic the lighting conditions in different environments that the user may be in. The developers reported a recognition rate of 97% and 10 words per minute (wpm) compared to the 30 wpm reported in other studies on the Graffiti system (MacKenzie and Zhang 1997).

2.5.3 Soft Keyboards

Soft keyboards, short for software keyboards, are sometimes referred to as virtual or on-screen keyboards when they relate to a desktop keyboard layout. They are software based graphical representation of character layouts; the character is chosen mainly by means of a stylus but can be used with mice, arrow keys or fingers. The user does not have to touch the screen if they are using a mouse or the arrow keys, but can instead hover. These keyboards, do not necessarily have to conform to the layout of standard keyboards, but may be circular, such as the Cirrin soft keyboard (Mankoff and Abowd 1989), or rearranged like the Opti (MacKenzie and Zhang 1999) and Metropolis (Hunter *Et al.* 2000) designs.

The soft keyboard is very flexible as it lends itself to a variety of designs, so depending upon the task and users' required layout, the keyboard could be designed accordingly. The ability to have a flexible input interface is especially useful to users who are differently-abled, who may have limited hand, finger or body movement. For example, the Cubon keyboard (Zhai *Et al.* 2000), originally designed for rehabilitation is specifically for users to use one finger or a head based pointer.

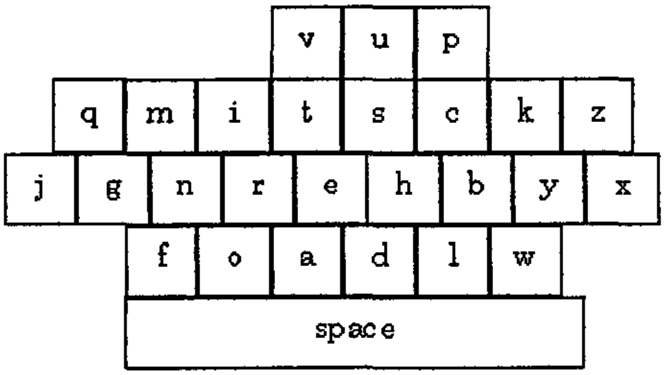


Figure 2.8: The Cubon Keyboard (Zhai *Et al.* 2000).

The soft keyboard could be highly adaptable as part of pervasive or context-aware wearable computing system. Depending on the environment of the user and who the user is, the soft keyboard interface could adapt to a user's needs accordingly. If low light levels are sensed the output could become brighter; if a user who is visually impaired is known to be using the system, the keys could instantly become larger brighter and the system could offer audio feedback, such as the repeating of characters and words that have been input. Although, this would be a function of an whole system incorporating a soft keyboard, rather than an actual property of the soft keyboard itself.

Unlike wrist keyboards and chord keyboards the soft keyboard focuses the user's gaze in a single location, the screen, so the input and output are displayed in the same area. This is an advantage when learning as the novice user does not have to keep looking for feedback in another on-body location. The downside to this is that the 'keyboard' takes up valuable screen real estate, which is problematic, because devices that often use soft keyboards are mobile and wearable computers such as Palm Pilots, Windows CE devices and Xybernaut tablets have small screens with the result that the keyboard is often reduced to the smallest usable size. Studies have shown that larger the key size does equate to a higher level of words per minute and a lower task load and it also helps if audio is used as a feedback method (Brewster 1999), but sonically enhancing buttons can have its drawbacks as some users can find this annoying (Potosnak 1988). In systems such as Microsoft's onscreen keyboard, visual feedback is used, but if the user is using a stylus to input, the feedback may also be physical; when the user touches the screen they are feeling the pressure that the stylus exerts upon their fingers. This tells us that there are two levels of feedback that may be needed in such a system: one that informs the user that they are doing the task, such as the audio click of a key, and one that shows that the task has been done, for instance the word appearing upon the screen.

Using any kind of soft keyboard is an ocular-centric task; it relies upon the user's visual attention being focused in one place, both for input and output (non-audio). The user's attention is partially focused upon the GUI; this is the opposite of what Brewster (2003) coined an eyes-free system.

Soft Keyboard WPM – Type - Study Table			
Key	WPM = Words Per Minute	p = predicted rate	a = actual rate

Study	**Type**	**Novice** WPM	**Average** WPM	**Expert** WPM
sears Et al. 1993	QWERTY	9.9	X	21.1
MacKenzie *Et al.* 1999	QWERTY	8.9p	20.2a	X	43.2p
	Dvorak	8.7p	8.5a	X	38.7p
	ABC	9.0p	10.6a	X	40.9p
	Telephone	9.1 p	8.1a	X	43.5p
	Just type	9.8p	7.3a	X	44.2p
Bohan *Et al.* 1999	QWERTY		26	
	T9		19	
MacKenzie & Zhang 1999	OPTI		17a	X	58.2p	44.3a

Table 2.1: Soft keyboard layouts with their input speeds.

The table above shows different soft keyboard layouts and the times they have recorded and theoretically predicted for the use of soft keyboard input. Both the QWERTY (20.2_a wpm) and ABC (10.6_a wpm) layouts had fast input rates for novice users. This is probably because the users are very familiar with these layout models. In examining MacKenzie's (*Et al.* 1999) upper and lower input limits for the QWERTY soft keyboard model of 8.9_p wpm and 43.2_p wpm, it is clear to see there is a difference between these and the 26 wpm input rate using the QWERTY layout recorded by Bohan (*Et al.* 1999). Interestingly, the OPTI layout offered a lower actual input level of 17 wpm and an expert level of 44.2 wpm, giving it an average of 30.65 wpm average (MacKenzie & Zhang 1999). So in examining these studies, what the evidence illustrates to us is that the QWERTY layout is initially the fastest to use for novice users, but users can become much faster at inputting text using a different keyboard layout such as the OPTI.

As soft keyboards are software based they are able to have any key assigned to a function, such as cut and paste or a macro function. Designs like the POBox (Masui1998), designed to minimize key presses, offer an added level of functionality; if the key **f** is presses a small menu appears offering the user the nine most common words they use beginning with the letter **f**, but this does take up screen real estate, which would be a problem on devices with smaller screens. The POBox could be used in Japanese or English mode.

2.5.4 Summary

Stylus-based systems are especially useful for users of touch screen based systems. These systems are software based so they also reduce cabling and all the user needs is a stylus, which is lightweight, easy to use and economical. Many of these systems currently exist on mobile systems such as Palm Pilots and Win CE devices; these have also been integrated into wearable systems such as the Xybernaut MA V (Comdex 2004) by others (Smailagic and Siewiorek 2002; Kumar *Et al.* 2001; Reitmayr and Schmalstieg 2001). Another recognized problem of using a system like this is that it focuses the user's attention into one place (Brewster and Walker 2000), so this system may not be appropriate for tasks that require the user to look at what they are doing while they record data. Having the keyboard on screen or a space to input on-screen also takes up screen space, which gives less room for other operations to happen on the screen.

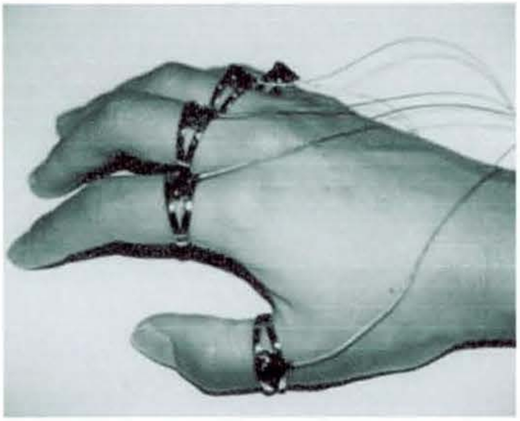
As we previously saw in the literature, the soft keyboard only allows the user to focus on one key at a time, which is akin to typing with one finger, so input speeds are slower than two/one-handed typing systems that were reported on earlier. Although have we have seen there is a wealth of evaluation into the stylus-based input, there were no systems that actually took input speeds for users while they were mobile. This is odd as most of the systems that we have seen are implemented in one type of mobile device or another; indeed some of the papers that were reviewed here claim their system was for use with wearable computers, but then go on to say that they have not yet used the system with a wearable computer. It is plain to see that what this review is bringing to light is a lack of research into text input devices for wearable computers, while the participant is mobile and wearing a computer.

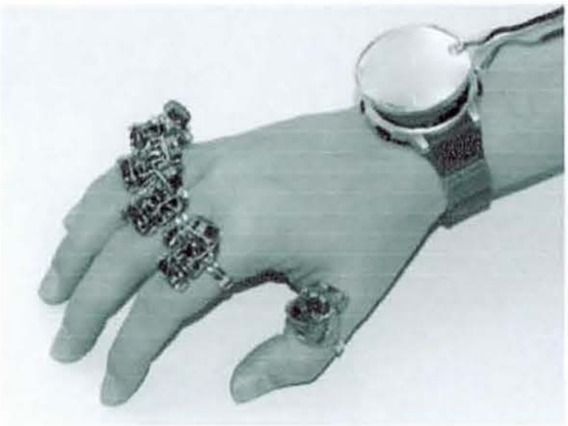
2.6 Alternative Text Input Methods

A whole range of other input devices have been designed with the mobility of the user in mind. Although they do not adhere to the traditional stylus, voice, and keyboard input models, it is important to briefly examine these designs in their contexts of as wearable interfaces for alpha numerical data. A double mouse text input system was developed by Nakamura, Tsukamoto and Nishio (2001). This system worked by using combinations of two mouse trackball movements. Other suggested systems have used 'tilting' technology for text input. This technology uses an accelerometer to judge the direction in which the user is tilting the device in which they are inputting text. By tilting in different directions, different items can be selected. Dunlop (2004) uses this technology in a wristwatch-based system, while Wigdor *Et al.* (2003) employs the use of this technology to input text into a mobile phone. As Sazawal *Et al.* (2002) write, "Tilting has proved to be a successful mechanism to enable selection on tab-like devices, and so we decided to apply a similar approach to text entry. A tilt-based approach to writing offers an alternative to the stylus when screen space is small or nonexistent. Tilt-to-write also offers a one-handed method for text entry." Using a tilt-based system also means the users do not need any additional input devices, such as styli or mice and it is an easy system to initially understand, as it mimics physical actions that occur in the real world.

Another interesting wearable text input system was the finger-joint wearable keypad (Goldstein and Chincholle 1999). This was essentially an evolution of Rosenburg's (1998) chording glove. The system worked by essentially using the finger joints a as a keypad. Different functions could be used by pressing the nails of the hand. Although their research claims that training was 'negligible', no analysis was given in the research. This system was renamed as the joint-gesture palm-keypad (Goldstein *Et al.* 2000).

A similar system was produced by Fukumoto *Et al.* (1994), previous to Rosenburg's glove based design, but this did not have to rely upon the thumb selecting each character. The Finger-Ring placed a set of rings around the user's fingers of one hand and allowed the user to input text or commands though making a finger-tip typing action. This action could be made against any surface, making the system portable but not hands-free. The finger-ring was later developed into a wireless system where a full evaluation was done on its input speeds and a specialized training system was developed (Fukumoto *Et al.* 1997).





Figures 2.9 & 2.10 Finger Ring (Fukumoto *Et al.* 1994). (Fukumoto *Et al.* 1997).

Although an impressive input rate of 210 letters per minute (average) was reached, the user needed specialized training that was based around the piano keyboard layout (although the training is not thoroughly explained in this paper). This input rate is representative of lab-based trials, but different tasks may give different input rates.

Using a hand/glove-based input may seem like an ideal solution, but in many real life situations both hands are needed to complete real world tasks (Rekimoto 2001); Reikimoto, goes on to say "Many current wearable input devices look unusual, or too "hi-tech", and often prevent normal social activities, such as shaking hands. Even though it is functional for the purpose of demonstration, it often becomes troublesome, even for the user, to wear it long-term in everyday situations. In contrast, traditional wearable accessories, such as wristwatches, eyeglasses, jewellery, and clothes are a vital part of our lifestyle and are designed for comfortable long-time use. Wearable devices should be as acceptable as today's accessories, or they should be a part of these things." (Rekimoto 2001). Instead of obtrusive physical devices, Rekimoto developed two systems: the wrist gesture which was built into a digital wristwatch and gesture pad, a gesture recognition system that could be embedded into the users clothing.

The GestureWrist (2001) system, seen in Figure 2.11 is based around a wristwatch and is small and unobtrusive. The basic idea is that the user can learn different hand signals that they can use while wearing a computer in any location. Rekimoto also embedded the recognition technology into a suit jacket. Gesture pads are placed throughout the jacket and the user makes the previously learnt gestures on top of the pads. What is not answered by this research is the question, "why the pads are placed where they are, how many gestures can the system can recognize and what are the gestures based on?" Although the idea of using gestures in this case is to make the user's interaction with wearable technology less socially obtrusive, it may prove that using gestures in public is not socially unobtrusive and could put off some users. So although the system may be small or hidden from other people that we may interact with, the method of input is certainly not.

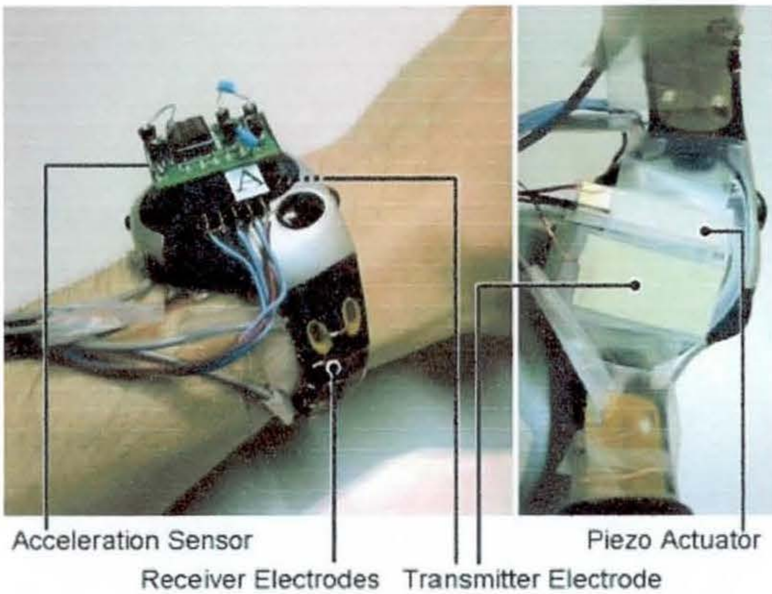


Figure 2.11 GestureWrist (Rekimoto 2001).
See - http://www.csl.sony.co.jp/person/rekimoto/gwrist/

There has been other research that has place input systems into fabric, but these have merely been existing designs such as a QWERTY keyboard sewn into a sleeve and while novel in material usage they were not in design terms. There are also problems with embedding technology into clothes (Lehikoinen and Roykee 1994) as we has previously discussed previously in the thesis.

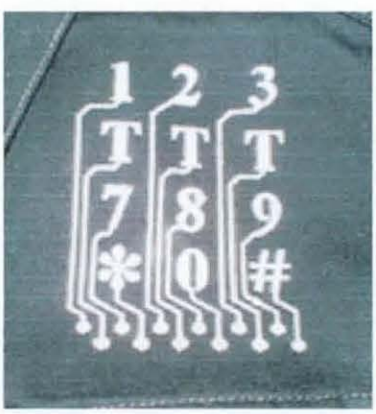


Figure 2.12: Embroided Keypad. (Orth, Post and Cooper 1998)

2.6.1 Summary

Although interesting, the use of the specialized custom input devices as we have examined here often only relate to the project that conceived them, and as such are often one-off projects. Real world users do not always have the access, time or knowledge to develop such technologies, while they do have the ability to buy 'off the shelf' systems such as voice input technologies, wrist keyboards and stylus-based devices. These specialised systems are important in the realms of research and design. Systems such as the gestural systems mentioned in this chapter and Orth's (*Et al.* 1998) fabric-based input systems may reduce unwanted cable weight and may lead to a greater level of social acceptance of wearable devices. Still, there needs to be more research done on currently existing input systems, as there has been little empirical research into the integration of this type of technology with wearable computers.

2.7 Speech and Sound

This section explores speech recognition, both in the areas of text input and as a command-based input. In a wearable computing system where the task analysis has identified a hands-free requirement, speech input/control can be an ideal solution for text input (Smith, Bass and Siegel 1995; Smailagic 1997; Buergy 2001; Smailagic and Siewiorek 2002). It is desirable because, besides its hands-free (Rhodes and Crabtree 1998), eyes-free qualities (Brewster 2000), it is lightweight and therefore causes less physical stresses to the wearable systems user (Gemperle, Kasabach and Stivoric, 1999), it is fast; indeed claims have been made that some software such as Dragon's Naturally Speaking 6 offers up to 160 words per minute (2.6 words a second) (Scansoft 2002). The literature (Isokoski 1999) also discovered that speech recognition is over sixteen times faster than other input devices, when related to studies such as Isokoski (1999) that compared words per minute input, using the trackball, mouse, joystick and track-pad for text input. There are many off-the-shelf packages and speech engines that are available for development. One of the main problems with such systems is the accuracy of word or command recognition. This problem was identified at the Boeing Workshop on wearable computing 1997.

"The quality of recognition is a function of the size of the vocabulary, the acoustic characteristics of the environment and the microphone, and the quality of the recognition algorithms. A recognition rate of 90% still means that one word in ten is incorrectly recognized. Current speech systems operating in ideal circumstances have a recognition rate of over 95% (one error in twenty words). " (Boeing 1997)

The recognition rate can lead to further problems and time delay/usage when the unrecognised words have to be corrected. If error correction is a problem, we must ask ourselves; *"In a hands free situation how can we correct errors in voice recognition systems?"* (Boeing 1997). Can speech recognition systems be relied upon to correct their own errors?

In systems such as IBM's Via Voice, Dragon's Naturally Speaking and Philips' Free Speech the mouse is used to correct mistakes, but this is not a viable solution in a hands free context, as the user is not able to use the mouse. Trying to correct mistakes by voice can be done, but can also be extremely problematic; for instance, "how would the user correct a mistake by voice? And what would happen if the user of the speech recognition system encountered a mistake while trying to solve a mistake?" This sort of scenario could put the user under a large amount of stress and raise the user's subjective task load when using the system. It is easy to see how using a simple sentence to control a spoken system could end up taking far longer than the user would expect. Human speech patterns are intricate, complicated and are mostly based around human-to-human interaction. There may be a whole range of contextual, geographical and psycho-socio impacts that affect the way we communicate with each other.

Although we hear speech (spoken language) as a collection of words, each separate yet placed together in such a way as to form a sentence, the sentence when spoken has no spaces between the words (known as coarticulation); it is only our knowledge of the language and the way that we psychologically interpret speech that enables us to hear each word separately. The biggest problem that computers have is deciphering these words as separate entities (Michie and Johnson 1985).

Using natural language sentence structures can also be problematic in two other ways. The first is the size of the vocabulary dictionary used, although contemporary speech recognition systems such as Via Voice may have a vocabulary of 80,000 words, this does not take into account the large amount of slang and jargon terms used in everyday life. Vocabulary size, if too small, can be a hindrance, while on the other hand problems can occur if the vocabulary is too large. Many words and phrases can sound the same or similar, yet have completely different semantic meanings, for example: 'bye' and 'buy', 'son' and 'sun', 'hi' and 'high' all sound the same. If the speech recognition system is used to control the interface, a large vocabulary may prove difficult for the user to remember.

In mobile devices where small amounts of memory are used the vocabulary is even smaller and can be embedded within the hardware (IBM 2004). In Via Voice's latest embedded mobile incarnation the vocabulary size is a mere 50 words and it can run on Palm and Pocket PC platforms. 50 words may prove inadequate for full scale text input, depending upon the task, but may be adequate if used for command entry, for example: 'open file', 'exit', 'view' and 'save' are all commands that one might want to use for speech command control, as opposed to 'full' natural language based text entry .

A way to resolve some of the recognition problems is through training, and by implementing what is known as the 'scaffolding' technique (Rosson, Carroll and Bellamy 1990). New users are given a scaled down, simple, easy-to-learn version of a full application. This provides basic functionality and is quicker to learn than the full package. As the user progresses and gathers more experience and expertise in the application, another version of the interface can be used to give the user more options, flexibility and therefore functionality.

These different iterations of the user interface allow the user to seamlessly graduate from a novice to an expert user of the package at their own pace. This 'scaffolding' technique could be applied to the use of speech recognition systems. First giving the user a limited vocabulary that allows them to get user; to the system and have limited functionality, then allowing them to develop their skills and adapt to using more spoken commands. This may prove more comfortable to the user, instead of giving the user a large cognitive workload, we are creating a steady learning curve with a small amount of workload at every step of the learning curve. The start of the process could even be as simple as learning to turn the system on and off, which is a very important first step to learn in hands-free environment and also directly relates to the user interface design guidelines for wearable computer speech recognition applications (Najjar *Et al.* 1998). It is argued in the literature that speech input applications are not useable in environments where there is a high level of background noise (Siegel and Bauer 1997); it could be argued that a throat microphone (Figure 2.13) may be of more use than a conventional microphone when used with acceptable levels of background noise (Rhodes 1998; Randall and Muller 2000) or a directional noise cancelling microphone (Peckham 1994), so only audio input originating from the direction of the users mouth may be picked up.

Figure 2.13 Throat Microphone.
See - http://www.rahq.com/throat_microphone.htm

If the background noise levels are too high it is far better to use other input devices, such as wrist-keyboard (Ditlea 2000). Background noise can interfere with the recognition-rate of a speech-recognition system, as Yankelovich *Et al.* (1995) notes, "Background noise, especially words spoken by passers by, can be mistaken for the user's voice." This may cause problems for the users of speech recognition systems, such as input errors and a degree of variability in terms of the recognition rate. If the users are in environments where there is a level of background noise that will impact upon the recognition rate of their system, the system may be rendered un-useable. As we have seen, this could be from other speech acts from passers by, traffic, and environmental factors such as wind-noise on a microphone.

There have been attempts to minimize the effects of background noise using a combination of speech recognition and visual tracking of the lips (Bregler and Omohundro 1995), the use of noise cancelling microphones (Gauvin *Et al.* 1996) and a range of different filtering algorithms developed to solve the problem of background noise (Gales and Young 1995). Yet, in settings with high background noise levels, recognition still remains a problem. As Gales and Young (1995) state, "Increasingly, as they are applied in real world applications, speech recognition systems must operate in situations where it is not possible to control the acoustic environment. This may result in a serious mismatch between the training and test conditions, which often causes a dramatic degradation in the performance of these systems."

If hands-free input is required it may be possible to combine two forms of input, or design a combined input system that can be used with a voice recognition system, such as the CMU (Carnegie Melon University) hypertext dial (Smith 1995) which allows its users to jog through menus then select the required option directly by voice; although this involves its users learning to use two input devices as oppose to one, this was beneficial in the long term in contexts where noise levels fluctuated dramatically.

Difficulties such as frustration, or the user getting 'lost' and repeating tasks can occur when feedback is not continuous (Shneiderman 1993). Speech recognition is not instantaneous; it takes a small amount of time before the text appears on the screen. This can sometimes leave users wondering where they are when dictating text. *"You'd have to wait* 10 *seconds after some voice commands because of processing time,"* said Bruce Knaack who, as manager of licensing and solutions for IBM's Personal Systems Group, oversaw the pilot tests of the Wearable PC prototypes (Ditlea 2000). In the literature it is also reported that off-the-peg speech recognition packages *(Dragon Naturally Speaking* 6

2002) have got faster, but there is still a noticeable lag while using them (Greenburg 2002).

A set of wearable computing speech recognition design principles (Najjar *Et al.* 1998) were produced to be used by designers/developers using speech recognition as part of a wearable computing system. The fifteen guidelines are categorised under the following three headings: General, Software and Hardware, and are expanded upon below. Although the guidelines provide a platform for the use of speech recognition as part of a wearable computing system, some important principles that have not been fully considered that impact upon the integration and use of speech recognition as an input modality for wearable computers. The following are the original guidelines plus an extension of the guidelines:-

Wearable Computing Speech Recognition Design Principles (Najjar *Et al.* 1998)
Use speech when the user's eyes and hands are busy or when the user is moving.
Train the speech recognition system in the user's work environment.
Iteratively evaluate and re-design the speech recognition application.
Keep small the number of words in the speech recognition vocabulary.
Keep short each speech input.
Use speech inputs that sound distinctly different from each other.
Provide immediate feedback for every speech input.
Keep the user interface simple.
Make error correction obvious.
Avoid Modes.
Don't use speech to position objects.
Use a command-based user interface.
Allow users to quickly and easily turn off and on the speech recogniser.
Consider using headphones or an earphone for auditory feedback.
Use full duplex audio.
Use a highly directional, noise-cancelling microphone.
Consider providing a back-up input technique to speech.

Table 2.1: Table of Wearable Computing Speech Recognition Design Principles (Najjar *Et al.* 1998).

After examining the guidelines it was obvious that these could be developed further in relation to the literature. Table 2.1: (above) shows the extension of the guidelines.

Table 2.2: Extended Audio Input/Output Guidelines (below)

Extended Guidelines based on Previous Literature
Give the user a list of words they can use to instantly control the system.
Use the users' already existing knowledge to develop the vocabulary used, e.g. back, forwards, and knowledge of existing software such as Internet Explorer etc.
Take into account the scaffolding effect
Do not allow audio feedback to occur at the same time as audio input. Confusion occurs.
Make all audio hardware (microphone) as lightweight as possible to avoid physical stresses. If long-term wear of equipment is part of the working life of the hardware it is important to use light weight equipment; it may also be important to make equipment more rugged depending on the physical stresses and strains it is placed under (Baber and Noyes 1996; Kalawsy 2000).

2.7.1 Audio Application Areas

So far we have addressed the problems and issues that are associated with the use of speech recognition in wearable computing systems, but we have not yet looked at the application areas where speech recognition as part of a wearable computing system could be beneficial to its user. Speech recognition is used in a huge variety of professions and real life situations: medical, legal, business, commercial/warehouse, universal access handheld devices, toys and education and automobile applications, according to Weinschenk and Barker (2000).

The main area of research for speech recognition applications as part of a wearable computing system is often called the 'hands-free environment' (Bass *Et al.* 1995, Gloria *Et al.* 1998). This includes scenarios where the user is on the move, using their hands and eyes for tasks other than input, or needs to input two different sets of information into two different systems (Helander, 1993; Jones, Frankish, and Hapeshi, 1992; Peckham, 1994; Simpson *Et al.* 1985). Although speech recognition, as an input modality, seems to be an 'obvious' way to interact with a computer (Freed, 1997; Newsome, 1997; Slater 1997) the key problem is that talking to a computer is not the same as talking to a human being. Users have to learn speech recognition systems.

2.7.2 Audio Based Wearable Computing

We have seen that there are problems associated with speech as an input modality, but speech (audio) can be beneficial as an output modality too. Many tasks require the user to look at a keyboard or screen to input data, and as such it is difficult for users to focus all of their attention on the interface, because they may be engaged in other non interface based tasks while mobile (Brewster *Et al.* 2003; Brewster 2002). Using the wearable system may be a secondary task and as such the user will not focus their attention constantly on the display. It should be noted that when users are using desktop systems, many of them are using them for a primary task, such as typing, using other 'office' based applications, developing or browsing, whether it's the internet, databases or other documents. When using a wearable or mobile system the task may not involve using the wearable system at all.

The nature of many wearable computing systems relies heavily on the user's visual channel, and as such is problematic in terms of focusing the user's attention constantly 'on-screen' or on the 'input device'; this places users in a non eyes-free situation which could be problematic if the user has to constantly check the visual interface; this is especially true if the user is involved in a task that requires their full visual attention. As Rhodes (1998) explains, "Because audio does not distract the user in the same way as a screen or display interface, audio output is especially useful where the user is driving, involved in delicate operations, or may be visually impaired" (Rhodes 1998), but this may depend upon the level of background noise and the type of information that is to be conveyed. Within our own existence, most of us do not live in a purely visual world. We rely on a combination of sound, vision, touch, smell and taste to interpret the world we inhabit; even hearing-impaired users can use sound waves/vibrations that add another dimension to their understanding of their environment. Adding the ability to use sound as or part of a wearable interface may give support to the visual elements of the interface and add "synergies and further informational dimensions" to the system (Blattner & Dannenberg, 1992).

The output may be given to the user by means of an earphone-based system or through speakers that may be built into the environment or hardware. The system's output may be music-based, a variety of sounds or speech. The type of audio output used may alter the levels of interaction with a wearable system. A system may be fully audio based, which may be designed for a visually impaired user or may just have the odd audio indicator, to say for example that mail has arrived or on a scroll bar; experiments have shown an auditory enhanced scroll bar to be quicker and have a lower error rate than conventional scroll bars (Brewster, Wright & Edwards, 1994). Audio output can be an excellent way of providing the user with feedback in respect of the state of the system. In our everyday life we use audio to speak and convey information, to alert us, and to hear the state of various systems from microwaves to telephones. Each intrinsic sound, whether it is speech-based or not, contains a level of semantic meaning/s, which the user can learn, understand and interpret.

2.8.3 Audio Centric Wearable Computing

A majority of the research has been aimed at visual and multimodal/media-based wearable computers, as was seen in chapter one. Within the literature there are many references to the 'Nomadic Radio' project (Sawney and Schmidt 1999). The Nomadic Radio is an audio-centric wearable computer that has no visual output metaphors. In an attempt to unify the interface the researchers brought together a wide range of audio based components, such as voice mail, email, diary dates, calendar, updates and news/media broadcasts. As we can see from the illustration below the input consisted of a voice recognition system, while the output could be synthetic speech, audio cues and for spatialised audio.

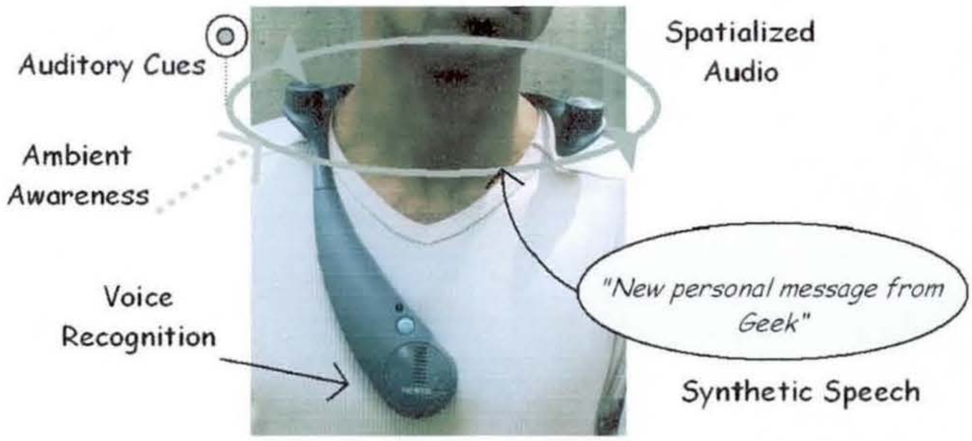


Figure 2.14: Nomad 1, Nomadic Radio. From (Sawhney and Schmandt 1998).

The Nomadic Radio project attempted to take the audio input/output interface metaphor to a pure audio only level and integrated it within a wearable computing paradigm. The Nomadic Radio project was discontinued in 1999 and there were no plans to put the system into production or develop it further. One of the benefits of using auditory techniques on a wearable device is that it provides hands-free access and navigation as well as lightweight and expressive notification (Sawhney and Schmandt 1998). Leaving the user's hands free allows for a greater flexibility, not only in terms of the user's ability to accomplish a wider range of tasks but also in allowing the user the freedom to express themselves in terms of their body language and to use gesture-based signals, which may be an important factor for someone using sign language to sign language interpreter.

2.8.4 Summary

Using speech recognition systems to control and input data into a wearable computing system seems like the ideal solution to the problem of inputting whist mobile. There are many problems that are associated with the use of speech recognition: low and variable recognition rates, variability of background audio levels impacting upon the recognition and having to learn a spoken series of commands to control a system. There are also issues relating to the comfort levels of users interacting with (speaking to) computers. However, there are benefits to using audio-based input and output, such as allowing the user to be eyes-free (Brewster *Et al.* 2003) and therefore not focusing their attention in one place and also by providing a means that allows the user to interact with their system in a hands-free manner.

There are entire wearable systems, such as the Nomadic Radio (Sawney and Schmidt 1999) that have been developed. These rely purely upon the user's speech for input and have audio-only output. Both audio input and output are important areas that need to be considered when developing any wearable system.

2.9 Pointer-Based Input

One way of interacting with wearable computers is through the use of pointing devices. These devices may be touchpads (track-pads), off-table mice, the stylus or speech to select interface items. This section explores the devices and literature by relating to them and their use in terms of their interaction with the wearable graphical user interface MacKenzie, Kauppinen and Silfverberg 2001). The interaction techniques that have been specified are: trajectory-based interaction (Accot and Zhai, 1997 and 1999), target selection (Smith, Ho, Ark and Zhai, 2000; Thomas, Grimmer, Zucco and Milanese 2002; Curry, Hobs and Toube 1996) and dragging and dropping (MacKenzie, Sellen and Buxton 1991; Inkpen 2001). This is not a technology review, but instead offers an insight into wearable pointer-based interaction.

2.9.1 Gestural Input

Gesture recognition allows the user to interact with a computer by using body movements. These movements may be hand-based (Tsukada and Yasumura 2002; Bowden *Et al* 2004), use arm and upper body movement (Kang, Lee and Jung 2004), and involve head-based gesturing (Brewster, Lumsden, Bell, Hall and Tasker 2003). Unlike recognition systems that use the arms and hands as the source of the gestures to be recognised, head-based interaction is hands-free. This means that the user can use their hands to accomplish physical tasks and still use head-based gestures to interface with a computer.

Other benefits of using gesture recognition systems are that the user can interact with computer-based systems without the need for physical input devices such as mice, joysticks and keyboards. Gestures that are already learnt, such as nodding the head to say yes, or shaking the head to mean no, may be also be used and may slow down the time it takes to learn how to use a gesture-based system.

Systems such as Ubi-Finger (Tsukada and Yasumura 2002) a gestural input device for mobile and ubiquitous environments have been developed. This system allowed the user to be mobile in a ubiquitous environment and control elements within that environment, by using gestures. At its simplest, the user could make a gesture, such as pointing to switch on and off electric lights. One of the drawbacks of using this system was that the user needed to wear a cabled device on their finger, which attached them to the recognition system. The Touch-Player (Pirhonen, Brewster and Holguin 2002) allowed the user to make gestures on a PDA in order to interact with an MP3 player. This system allowed the user to use 'natural' gestures to select different functions. The user used their finger to make gestures on the screen of the PDA, these were:

- Sweep across screen left side → right side = next track (this could be reversed for left - handed users)
- Sweep across screen right side → left side = previous track
- Single tap = start/stop
- Sweep from bottom → top of screen = volume up
- Sweep from top → bottom of screen = volume down

(Pirhonen, Brewster and Holguin 2002)

Kang, Lee and Jung (2004) developed a gestural method for playing the computer game Quake (see - http://www.idsoftware.com/). Their system consisted of a series of ten gestures that allowed the player to run, fight and move from side to side. In this paper they outlined the problems that they had come across during the development of the gestural system. Unintentional movements by the user, such as a scratch of the head, were interpreted by the system as a command. Continual gestures (co-articulation) were sometimes not recognised, as it was difficult for the system to see where one gesture finished and another started, and the user had to first learn the gestures. Their study found that any gesture-based system should be: fast, reliable, economically viable and easy to use and learn.

2.9.2 Target Selection

Target selection is when the user uses an input system to point at an item on the screen. These items may be buttons on a web page, selection boxes on an order form, folders or icons on a desktop or radio buttons. It is therefore important to investigate this area, because of the wide range of GUI-based 'targets' that can be selected. Studies have been carried out on touch screens (Sears and Shneiderman 1991), pen based systems (McClintock and Hoiem 1993), eye-based interaction (Bates 1999) and desktop systems, and with a variety of devices.

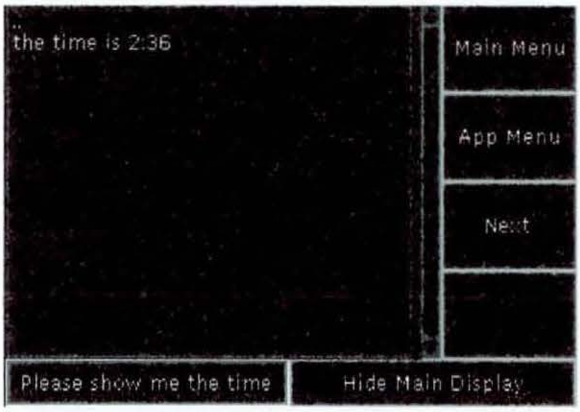


Figure 2.15: The Gili AR wearable interface. A five button point and click interface.
See - http://wearables.essex.ac.uk/sulawesi/io/gili.html

Point and click based interfaces do exist on wearable systems (Schmidt, Gellersen and Beigl 2000, Gili 1998), yet little empirical research has been carried out on the range of input devices that could be used in conjunction with wearable interfaces that use this type of interfacing strategy.

2.9.3 Trajectory Based Interaction

A trajectory is the path of a moving body through space. In terms of trajectory-based interaction, it is the physical path that a user takes to accomplish tasks such as; writing, drawing, selecting command lines (see figure 2.14) and navigating through menu structures (see figure 2.13), when using a graphical user interface. As Acott an Zhai (1997) state, "Trajectory-based interactions, such as navigating through nested-menus, drawing curves, and moving in 3D worlds, are becoming common tasks in modern computer interfaces." If we examine applications that exist today such as browsers, word processing software and desktop publishing applications it is difficult to find a piece of software that does not use drop down menus to offer options, or scroll bars that allow the user to move a screen up and down. Even selecting text on a command line requires a user to use a degree of trajectory-based interaction to select text on that line.

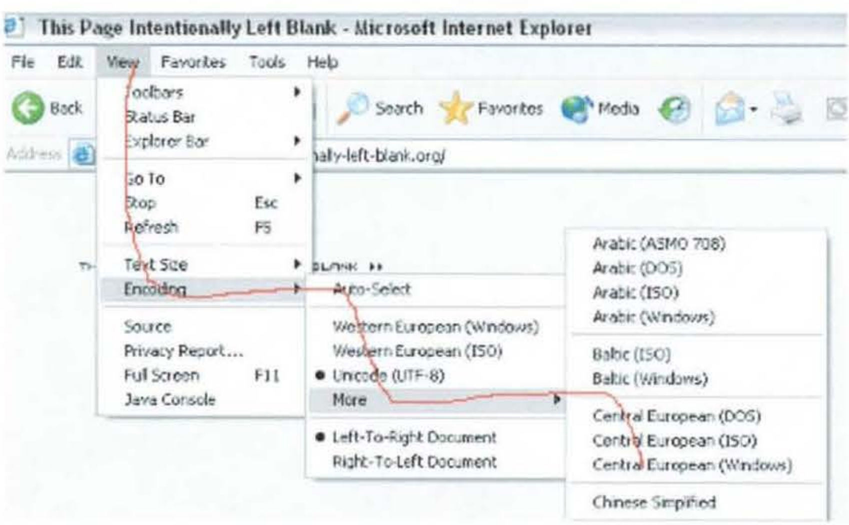
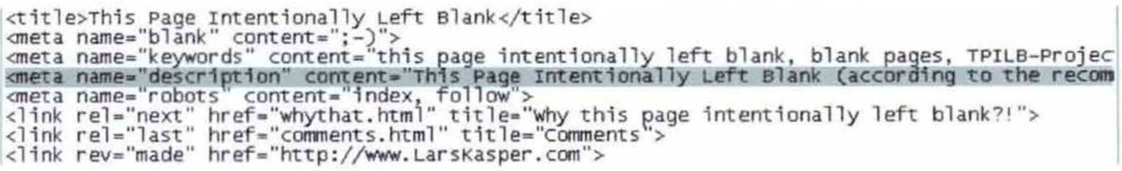


Figure 2.16: A typical trajectory based task. Steering thought the menus in Internet Explorer. The screenshot was taken from a 'traced-trajectory' menu route using Microsoft's Internet Explorer.



```
<title>This Page Intentionally Left Blank</title>
<meta name="blank" content=";-)">
<meta name="keywords" content="this page intentionally left blank, blank pages, TPILB-Projec
<meta name="description" content="This Page Intentionally Left Blank (according to the recom
<meta name="robots" content="index, follow">
<link rel="next" href="whythat.html" title="Why this page intentionally left blank?!">
<link rel="last" href="comments.html" title="Comments">
<link rev="made" href="http://www.LarsKasper.com">
```

Figure 2.17: Steering through a command line interface, to select a line of text.
The screenshot was taken from the text editing screen of Microsoft's Front Page.

Other wearable menu-based structures, such as pie menus (Schmidt *Et al.* 2000), as seen below in Figure 2.18, also require the user steer to different parts of the menu, although these are not as prevalent as drop down and on-screen linear type menus.

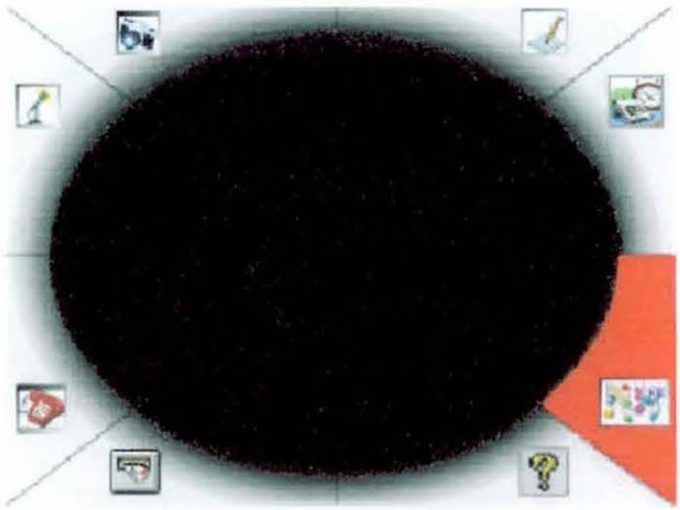


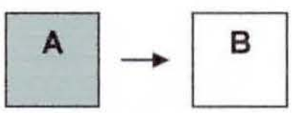
Figure 2.18: wGUI. (Schmidt *Et al.* 2000)

Other systems that have been primarily used for this type of interaction with wearable computers have been the Blasko and Feiner's (2002) synaptic touch pad, which split a touchpad into three sections each relating to a menu and was belt worn. Similarly the QBIC, a Wearable Computer integrated in a Belt, designed and built by the wearable computing lab at Zurich (Wearable Computing Lab 2004) also use their touchpad (track-pad) system belt mounted. However, it has been found that the most appropriate place for users to have a touchpad attached is the upper thigh (Thomas, Grimmer, Zucco and Milanese 2002).

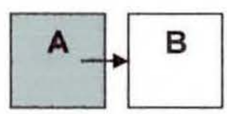
2.9.4 Dragging and Dropping

This section examines direct manipulation in relation to the use of pointing devices as part of a wearable computing system. As we have previously seen, there are two other methods of using a pointing device within a WIMP based interface: target selection and trajectory-based interaction.

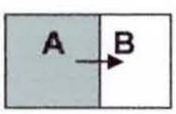
Dragging and dropping (Dix, Finlay, Abowd and Beal 2004) is a metaphor used to explain the action of placing items within the structure of the graphical user interface. It is most commonly associated with WIMP based systems and the movement of files, folders and shortcuts within this environment. This system is used within a wide variety of contexts word processing packages, email clients and using desktop publishing packages. We can see from the diagram below that the process is a direct manipulation of one file onto another that uses a real world metaphor. The folder (or file) is first dragged over to the folder with an input device and then dropped (placed) into the folder. Often, as the file is dragged over the folder, the folder will be highlighted, to show that the drop function can now be performed. The diagram below graphically represents this process.



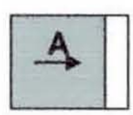
1) Folder A & folder B



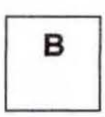
2) Folder A is dragged over to B



3) Folder A is dragged over to B



4) Folder A is dropped into/onto B



5) Folder A is inside folder B

Figure 2.19 Dragging and Dropping.

One novel use of the 'drag and drop' metaphor was developed by Rekimoto (1997, 2001). Called the 'Pick and Drop' (Rekimoto 1997) system, the wearable user was able to select a target, in this case an image on the screen of a device, and drag it across to the visual display of another device.

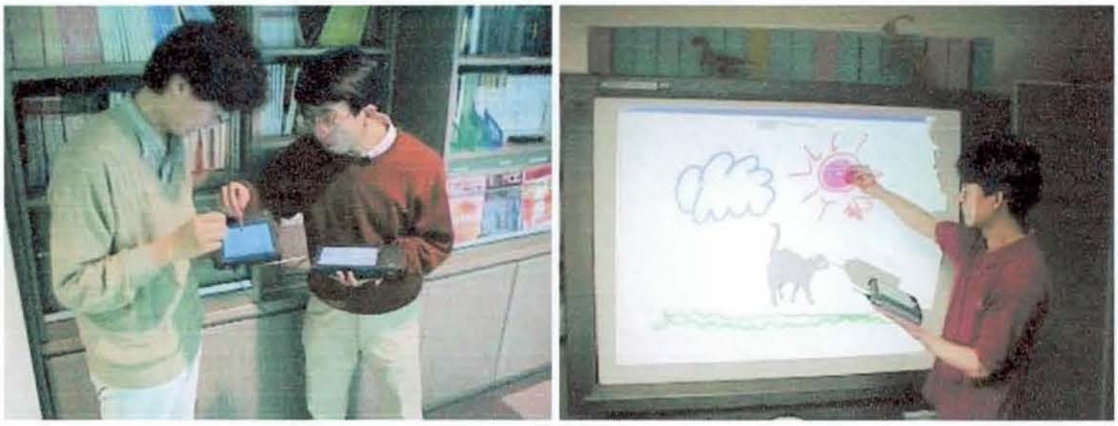


Figure 2.20: Pick and Drop. (Rekimoto 1997).

Folders can be dragged from one device to another. They can be dragged from one mobile device and dropped to another, or from a large screen to a mobile device.

This action was coined hyperdragging, and was used with a variety of devices from mobile phones (Kohno and Rekimoto 2002) through to mobile computers and large screen displays (Rekimoto and Saitoh 1999). Other wearable drag and drop systems have included FieldNotes (www.xybernaut.com), a system designed for automated field data collection and mapping, produced by the wearable computing manufacturer Xybernaut. This system relies heavily on manipulation by dragging and dropping data. The Graylevel Visual-Glove (Iannizotto *Et al.* 2001) was a vision based hand-tracking system, which recognized different gestures made by the user's hand; before use the user was required to learn a set of hand movements that related to different tasks to be performed by the system, such as dragging a file into a folder. Although the system worked within a desktop environment its developers did not port it to a wearable platform. Other tracking systems have been developed (Kurata 2002) but they are not yet sophisticated enough to take into account the differing light levels that can occur whilst mobile and the low power consumption that many wearable users often need for extended in-field use. Other systems (Brewster 1998) have evaluated the effects of sonically enhancing the drag and drop process. The experimenter found that sonic-highlighting in relation to a dragging and dropping task significantly lowered workload, time on task and significantly increased system usability. Although these results were not specifically aimed at wearable computer users, nevertheless providing multimodal cues to tell a user if they have correctly completed a task is a very powerful feedback-based interactive technique. This is important as it allows the user to use a visual or audio-centric "eyes-free" system.

2.9.5 Summary

The literature defines pointer-based interaction as three distinct categories: target selection (Smith, Ho, Ark and Zhai, 2000; Thomas, Grimmer, Zucco and Milanese 2002; Curry, Hobs and Toube 1996), trajectory-based interaction (Acott and Zhai 1997;1999) and dragging and dropping (Accot and Zhai, 1997 and 1999). As we have seen, within the literature there are a variety of pointer-based interaction methodologies that allow the user to use: touch screens (Sears and Shneiderman 1991), pen based systems (McClintock and Hoiem 1993), eye-based interaction (Bates 1999) and gestural interaction (Brewster, Lumsden, Bell, Hall and Tasker 2003; Pirhonen, Brewster and Holguin 2002). Within these studies there has been research into the impact of walking and using gestural menu selection (Brewster, Lumsden, Bell, Hall and Tasker 2003) and gesture for interface-based manipulation (Pirhonen, Brewster and Holguin 2002). Importantly, both of these studies offer an insight into the effects of movement upon a user within a dual-task environment. Yet, what needs to be examined further is interaction with widely available 'off-the-shelf' technologies, such as off-table mice and stylus-based interaction.

2. 10 Evaluation of Input Devices for Wearable Users

Many pointing device studies (Accot and Zhai 1997; MacKenzie, Kauppinen and Silfverberg 2001; Card, English and Burr 1978) have been focused upon desktop interaction. These studies, although useful, have not fully explored the use of pointing devices for use with wearable computing systems. Indeed, Curry *Et al.* (1996) attempt to relate their findings to the use of wearable computers, but this research was carried out using a desktop system and can only be seen as a indication that more empirical work with 'actual' wearable computer systems is needed. A commonly used assessment method for comparing pointing devices is the application of Fitts' law (MacKenzie, Sellen and Buxton 1991; MacKenzie, Kauppinen and Silfverberg 2001). However, problems occur when Fitts' law is applied to evaluate and assess wearable pointing devices; while the user is mobile the 'targets' the user selects are not stationary. Therefore, Fitts' law cannot be related to users that are mobile, because the 'target' sizes and distances involved in a Fitts' law-based evaluation need to be constant. This is because the display may be moving with the motion of the body, which obviously impacts upon the movement of the objects to be displayed.

Other systems have been used in an attempt to assess the user whilst mobile, such as the PPWS (percentage preferred walking speed) of the user (Brewster, Lumsden, Bell, Hall, and Tasker 2003; Lumsden and Brewster 2003). This compares a user's 'normal' walking speed to the speed when the user is performing a task whilst mobile. Unfortunately, the origins of this assessment method lie in the evaluation of visually impaired users and their stress levels in relation to different navigational tasks (Clark-Carter *Et al.* 1986) and there are no associated usability metrics relating to this method. Brewster's (*Et al.* 2003) results relating to frustration scale (scores)in a NASA TLX workload assessment could have been correlated to the PPWS to see if there is indeed a relationship between stress and walking speeds for non-visually impaired users, this work has not yet been done.

2.10.1 Workload

Workload can be defined "...as the effort invested by the human operator into task performance; workload arises from the interaction between a particular task and the performer" (Hart and Wickens 1990). The concept of workload arises from the theoretical 'viewpoint' that humans have a set of limited cognitive abilities that can be 'invested' into the performance of a variety of tasks. It relates directly to research into attention, processing capacity, dual-task performance, and allocation of mental resources (Schvaneveldt *Et al.* 1998). The use of workload is especially prevalent within studies relating to the evaluation of mobile devices (Brewster 1999 (HCI), 1999 (CHI); Lumsden & Brewster 2003). This is because it allows evaluators to compare similar tasks that are done in a mobile and a stationary manner. In effect, it enables the evaluators to further examine the effect of movement upon the user's workload, by comparing a stationary user completing a task in comparison to a user dual-tasking; walking while completing the same task. This is important, as it can highlight trade-offs that may occur while the user is mobile, such as a rise in mental workload. For a discussion on the debates relating to the definition of workload see Braarud (2001) and Schvaneveldt *Et al.* (1998).

There are problems associated with the use of workload when correlating between workload and objective measures of performance such as: heart-rate, error-rate and time on task. "Objective measures of performance sometimes correlate strongly with perceived workload or effort, sometimes not. However, this dissociation cannot be taken as an indication of measurement problems when it comes to workload *per se*. Performance and workload measures are sensitive to different task factors." (Svensson *Et al.* 1997). For example, one impacting factor may be the location of the task. It could be theorised that the workload may differ for the same task in different environments depending on the user's relation to that environment. Users may have to put more mental effort into conducting similar tasks in unfamiliar settings as opposed to familiar settings. This is a point that an evaluator may need to be aware of when conducting workload assessments that relate to the use of real world settings. The use of real world settings may be more variable than a lab based environment and therefore different environmental factors such as noise, lots of people, traffic and lighting/weather could all impact upon a user's workload.

The workload experienced by mobile users has been assessed traditionally using a workload assessment exercise called NASA TLX (Task Load Index). This system has been used to find the workload of users in many wearable computing studies (Brewster *Et al.* 2003; Tang *Et al.* 2003; Billinghurst *Et al.* 1998). This system can be paper-based or a computerised version is available that outputs a workload score at the end of the assessment. It is a multi-dimensional assessment method that is based on the average score of six sub-scalar elements: Mental Demands, Physical Demands, Temporal Demands, Own Performance, Effort and Frustration. These combined elements make up the user's final workload (or WWL, weighted workload) score.

See Appendix 6 for the user instructions that come with the computerised version of NASA TLX. (To obtain a computerised version of NASA TLX contact, The Navy Center for Applied Research in Artificial Intelligence (NCARAI)) –
(http://www.nrl.navy.mil/aic/).

2.10.2 The Dual-Task Paradigm

Many users have the ability to multi-task. They are able to walk and talk, drive and listen to the radio, or text a message to a friend using a mobile phone while walking. When users perform two tasks simultaneously in this manner, it is known as the dual-task paradigm. There is often a primary task, such as changing the tracks on an MP3 player and a secondary task such as walking (Pirhonen, Brewster and Holguin 2002).The dual-task paradigm is especially pertinent to evaluators working within the field of mobile-HCI, as mobile users can use their systems while they are moving (mobile). This means they can perform a primary task, such as inputting data into a PDA for example and a secondary task such as navigating around obstacles (people) on a busy street (Kjeldskov & Skov 2003).

As Hommel (1998) writes, "When people perform more than one task at a time, their performance often decreases more or less dramatically, even if the tasks are quite simple. According to an assumption of Welford (1952), this is so because the architecture of the human information-processing system does not allow for the transformation of more than one stimulus into a response at a time: somewhere on the way from sensory coding to muscle contraction there must be a bottleneck that renders human information processing a single-channel system." (Hommel 1998)

It is often the case that, when a user is dual-tasking there are trade-offs in relation to the user's ability to successfully perform both tasks. A good example of these trade-offs comes in the form of studies that have shown that drivers are not able to give their full attention to the task of driving, while they are having a mobile phone conversation (Strayer & Johnston 2001; Strayer, Drews & Johnston 2003). Their driving is hindered by attempting to dual-task, when placing their attention and trying to focus in two places concurrently. Interestingly, other studies (Schumacher et al 2001; Hazeltine,Teague & Ivry 2002) have shown that with practice and when an equal emphasis is placed on each task, the dual-task 'costs' are minimal.

Goodman, Brewster and Gray (2004) discuss a range of factors that can be measured to evaluate users while they are mobile. Besides the time taken to complete a task and the errors made, there are also specific measurements that relate directly to a user's movement. The distance a user walks, how fast they walk and the route they take are all valuable indicators of the impact an interface, or different input and output systems can have upon the user.

Studies by Petrie *Et al.* (1998), Pirhonen, Brewster and Holguin (2002) and Brewster, Lumsden, Bell, Hall and Tasker (2003) have all employed the dual-task paradigm to evaluate users while mobile and using mobile and wearable computing technologies. This is because the act of being mobile places the mobile users firmly within the dual-task paradigm. They are attempting to do two tasks simultaneously, which, as we have seen, leads to trade-offs when the user's attention is shared between two tasks. Using dual-task theory to evaluate users is not without its flaws: "Despite its intuitive appeal, the commonly held assumption that there is some general limitation on dual-task performance has been shown to be seriously flawed...Central to this has been the inability to measure the attentional demands of tasks, without which there is no way to determine whether their joint demands exceed the hypothetical general limit" (Bourke 1997).

Petrie *Et al.* (1997) evaluated a wearable navigation aid for the visually impaired. To evaluate this system she employed several tests that ranged from a NASA TLX workload assessment, through to the time taken to walk a set route and analysed the participant's preferred percentage walking speed (PPWS). PPWS (Clark-Carter *Et al.* 1986) is an assessment method that takes the participant's average walking speed around a set distance, and uses this as a metric against which the participant's walking speed while doing other tasks can be correlated. In Petrie's (*Et al.*1997) evaluation the visually impaired participants navigated using their primary navigation system (a long cane or a guide dog), while in the other condition the participants also used a wearable navigation aid, to help them find their way to a set location. Interestingly, the participants preferred using their primary navigation system combined with the use of the wearable navigation aid, this condition also had a lower workload.

We also find these techniques taken forwards to analyse gestural interaction by Pirhonen, Brewster and Holguin (2002). They found that using gestures to control the Touch-Player, when compared to the standard Media-Player (MP3 players): reduced the participants' subjective workload, allowed them to complete the task while walking at a pace closer to their PPWS and complete the experiment more quickly. By analysing the PPWS of the participants it can be hypothesised that there may be a relationship between walking speeds and usability, when the participant is engaged in a dual-task scenario. The more usable a system is, the nearer their pace is to their PPWS (the slower they are, the less usable the system). Other studies have also used walking speeds to assess the usability of mobile computer-based systems (Brewster, Lumsden, Bell, Hall, and Tasker 2003; Lumsden and Brewster 2003).

From this review of the literature there was evidence that many mobile computer users are involved in dual-tasking. Mobile users may be inputting text while walking, using gestures to interact with software system, or using navigation aids. What this can suggest is that there are trade-offs in relation to the performance of the user. It may be suggested therefore that a user will perform better when doing two single tasks, such as walking around an obstacle course and then using a gesture-based system. By doing the task simultaneously there may be a performance loss until, as the literature suggests, the user is practised at doing the tasks at the same time. This performance loss may be in the form of: a slowing down of the participant's walking speed, a higher subjective workload or raised stress levels.

2.10.3 Mobile and Stationary Interaction

When we examine wearable computer based interaction with pointing devices, rather than desktop-based interaction, we have to remember that the user may be on the move, there may be no flat surfaces available to use with a conventional mouse, and there could be a high level of variable background noise that may impact upon the use of voice recognition systems if used (Najjar, Ockerman and Thompson 1998). Therefore the mobile user has a limited amount of devices that they can use for input. The user may also be in a situation where they have to dual task. The task of pointing may also be a sub-task supporting the user's primary task. The user's environment and movement can be continually changing, this research intends to investigate two different movement conditions, standing and walking, these may affect the user's performance when using input devices.

We also need to take into account the type of display system that the user is wearing: many PDA's have small screens, the system may be an output-only screen integrated into clothing, a touch screen that the user can also use as an input device, a head mounted display system or an audio-centric system such as Sawhney *Et al.* (2000), or the Carnegie Mellon TIA Series (Smailagic 1997). Each of these separate output systems has qualities that may affect the pointing system we choose for a wearable computer user to use. For example, would it be impossible to use a stylus for pointing with a HMD (Head Mounted Display) or an auditory feedback system? There may also be requirements placed upon the user that mean that a method of hands-free input and gestural-interaction are the only viable input methods available. One key, and often neglected issue, is the use of devices while mobile. Wearable computers are able to go wherever the user goes and because of this they can be used while the user is mobile. It may also be the case that the user is continually interacting with the system, or the system is monitoring them. An important factor in this research may be to assess the user in both stationary and mobile conditions, as its importance has been noted in other studies (Brewster, Lumsden, Bell, Hall and Tasker 2003; Lumsden and Brewster 2003).

Although there have been studies relating wearable computing and text-based input, as we have already seen earlier in the chapter, no empirical results were reported that relate to the user while wearing a computer. Studies relating to input with wearable computers have found that it is important to recognise the movement condition of the wearable user, whether it was walking, running or standing still, and that this movement impacts upon the task performed by the user. Some studies have taken into account the movement of the user when using mobile devices, both Brewster (2002) in relation to audio feedback on PDAs and importantly Petri *Et al.* (1998) in their examination of wearable computers and visually impaired users, and both have stressed the need for evaluation in real world envirionments.

However, it is important to focus on the real world, in real world situations there is much more variablitiy that relates to the user's environment such as the weather, lighting conditions, background noise, the amount of other users in that environment, how dangerous the environment is, as well as a vast range of factors that one may find in the user's natural day to day environment. Although this is certailny true, it is also true that every user's use environment is different and therefore it must be noted that controlling a natural environment so that every participant will have a uniform experience is near impossible. So for any experimental investigation it is necessary that we take into acount that the user may be mobile while using their wearable ssyem. It is therefore important that a mobile condition be implemented, but this must initially be within a controlled environment to create a uniform set of user experiences so as not to negatively affect any resulting data.

2.10.4 Overall Conclusion

It has been found that wearable computers share similar features, but they have a disparate range of uses: from underwater systems, such as the WetPC, military-based wearable computers like the Land Warrior and generic 'off-the-shelf' systems such as those made by Xybernaut. Interfaces have been designed for AR, there have been audio-centric systems and even gesture-based systems, but many of these have been special one-off systems made in academia. If a user is to obtain a wearable system, they can either build their own, such as Steve Mann's wearcam (often known as 'homebrew' systems), or buy a ready built system, such as the Xybernaut MA series wearable computers.

What this thesis aims do is to examine some of the 'off-the-shelf' input devices for text input and as pointing devices for wearable computers. These devices will be chosen because they are easily available to a wide range of users, and therefore have more relevance to 'real life' usage. The wearable computer that will be used will be a Xybernaut MAIV, a system that is widely available and comes with a Windows based operating system.

In order to carry out these evaluations, an experimental approach will be taken. This will encompass a range of evaluation techniques that have been used in the literature. These will be time on task, error rate and an examination of the user's subjective workload. The literature suggests that the user will use their device while mobile and stationary and therefore it will be imperative to build a level of mobile use into these experiments. It was discovered that many text input and pointer based tasks are generic tasks, and as such the experiments are not aimed at a particular product or application, but are based on previous 'generic' input-based studies.

Chapter 3: An Evaluation of Four Text Input Devices for Wearable Computers

3. 1 Introduction

It was noted within the literature that one of the primary concerns for any wearable user is the choice of input system/device, as this is an integral part of interfacing with the wearable system for control and information input (Mann 1998; Bass *Et al.* 1997; Rhodes 1998). This is a major problem when designing software for the users of wearable computers as there is little empirical evidence to support a designer's, developer's or wearer's choice of interface in regard to simple generic tasks, such as text input, a commonly noted task performed by many wearable computer users.

Using existing desktop input devices is not appropriate for wearable users for several reasons; the desktop interaction paradigm involves a user sitting in front of a monitor and using a mouse and keyboard to interact with the computer, whereas the wearable interaction paradigm allows the user to be mobile and a range of environmental factors affect the user's choice of modality (Schmidt *Et al.* 2000). The act of being mobile and wearing a computer places constraints upon the user that the desktop does not (Gemperle *Et al.* 1998). Desktop keyboards are "difficult to use in particular settings and circumstances, such as conducting inventory on the shop floor or a geology survey in the wilderness...using a keyboard is cumbersome" (NIST 2001). The same can be said of other input devices such as the mouse, if we are using a computer while mobile; using a desktop mouse as a pointing device is not a viable pointing solution, since making the user stop and find a flat surface every time they needed to move the cursor is not appropriate (Rhodes 1998). Desktop devices are designed for use on flat surfaces, hence the term 'desktop'; they are essentially a range of devices that are designed for use on a desk.

For a wearable user we not only have to consider the weight of the device, but also the size of the device; it has to be small and light enough to be worn 'on body' and possibly to be stored on the wearer too (Thomas *Et al.* 202). The user may also be required to input with one hand (Matias *Et al.* 1993), working in a hands-free (Noble *Et al.* 1997) or eyes free environment (Brewster *Et al.* 2003) The user may be stationary or mobile when they use an input device (Randell and Muller 2000), so any input devices must have a high level of usability while the user is mobile. Therefore the task that the user is completing impacts upon the input and output modalities that are available and can support the user's needs. This study aims to find the best text input system for use by a wearable computer user. This criterion for judging this will be based on input times, error rates and the participant's subjective workload.

3.2 Rationale

Thomas *Et al.* (1997) carried out some initial investigation into text input devices for wearable computers. The study did not require the user to wear a computer, instead the user interacted with a desktop system and no experimental data was recorded while the user was mobile and the user did not wear a computer. After reviewing the literature it became apparent that a different experimental methodology from Thomas *Et al.* (1997) was needed to evaluate text input devices for wearable computers. This was done for several reasons: evaluations with wearable computers need the user to wear a wearable computing system; there are put systems that have been specifically designed or have been previously used with a wearable systems and the study needs to show the differences that mobility has upon the user.

Some studies have taken into account the movement of the user when using mobile devices, both Brewster (2002) in relation to audio feedback on PDAs and Petrie *Et al.* (1998) in their examination of wearable computers and visually impaired users, and Johnson (1998) has stressed the need for evaluation in real world envirionments.

In real world situations there is much more variablitiy that relates to the user's environment, such as the weather and a whole other range of factors that one may find in the user's natural day-to-day environment. Although this is certainly true, it is also true that every user's natural environments are different and therefore it must be noted that controlling a natural environment in order to give the participant a uniform experience is nearly impossible. So for these trials we have taken into acount that the user may be mobile and wearing a computer, and implemented a movement level within the experimental conditions that examines the participant whilst mobile.

The empirical research is formed around three strands of data. These related to the time taken to do the task and the amount of errors produced by the participant (any deviation from the text is classed as an error) and a task load rating. Unlike previous research that has been desktop based (Thomas *Et al.* 1997), we take into account two movement conditions: the participants were assessed while stationary and mobile (walking around stationary cones). This evaluation strategy relates more closely to the mobile conditions of a wearable computer user and therefore provides a more valid insight. It was important to use a wearable computer for evaluative purposes as some studies (Faruk-Özer *Et al.* 2001; Thomas *Et al.* 1997) fail to evaluate their wearable systems on wearable computers. After the completion of each part of the experiment the user is then asked to take a NASA TLX assessment. NASA TLX is a system for assessing the participant's subjective workload. It derives a weighted workload (WWL) from the way the participants rate six subscales: mental demands, physical demands, temporal demands, participant's performance, effort, and frustration. The use of NASA TLX as a task load assessment tool can be found throughout the research literature (Hart and Wickens 1990; Brewster and Cryer 1999).

3.2.1 Hypotheses

The null hypotheses are as follows:

1. The input device type does not effect the time taken to complete the text input tasks.

2. The input device type does not have an effect on the error rate of the text input tasks.

3. The movement condition of the participant standing and walking has no effect on the time taken to complete the task or error rate.

4. The input device type does not affect the subjective workload of the participant.

3.3 Method

3.3.1 Apparatus

The wearable computer used in all of the experiments was a Xybernaut MA IV. This computer weighed exactly 900g and measured 19 x 6.2 x 12 cm size. The Xybernaut MA IV had an Intel Pentium MMX 233 MHz CPU and a 128 MB of SDRAM installed. The size of the hard-drive was 4.3GB. The display unit used in the experiments was a Xybernaut FPD (flat-panel display) touch-screen. The screen had a diagonal display and weighed 520g. The operating system was Windows 2000. The touch screen was used instead of an augmented reality system as a previous pilot study that we had conducted had shown binocular rivalry (Larame and Ware 2002) to be a factor that could affect the user's performance, because of feelings of nausea, dizziness and watering eyes. This happens when a very different image is seen by each eye, as can be the case when using monocular displays. Instead of the user seeing one 'whole' image, their brain flicks between both images, continually seeing one image and then the other. This causes the physical side-effects previously mentioned. The apparatus listed here was used in all of the experiments in the thesis.

Microsoft Notepad was used as the base text input software. The following were used in the trial to input text: a Xybernaut wrist worn keyboard model kbd 00400 (figure 4.1), the Windows 2000 on-screen keyboard and stylus, MyScript (for Windows 2000/NT) handwriting recognition system by Vision Objects, with stylus. MyScript was chosen as it needed very little user training to recognise hand-printed (non-cursive) standard printed writing. MyScript is a widely used handwriting recognition system as it is the underlying recognition system used by the tablet pc version of Windows XP.

The speech recognition system used was Naturally Speaking Professional six, voice recognition software, by Dragon Software, along with a Labtec single ear piece microphone-headset. The speech recognition system required the users to take part in an initial training exercise. The exercise was part of the speech recognition package and required the participants to read sections from 'Alice in Wonderland'. Six road cones were placed at one metre intervals for the movement section of the experiment. This was found to be a satisfactory distance for the participants to walk around. Error correction was not used.

Figure 3.1: Touch screen – vest-based.

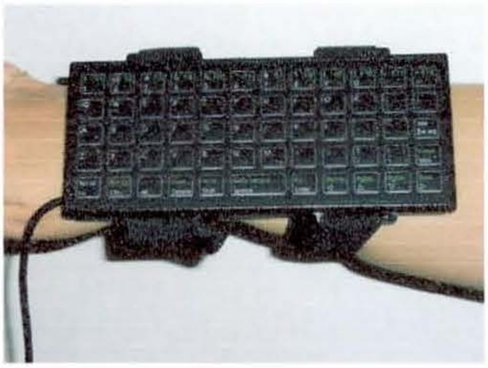


Figure 3.2: Xybernaut Wrist Keyboard.

For more information on figures 3.1 and 3.2 see -

Figure 3.1: http://www.xybernaut.com/Solutions/accessories/accessories_display.asp

Figure 3.2: http://www.xybernaut.com/Solutions/accessories/accessories_additional.asp#Key

The principles of the input system to be used were shown to the participants. Each participant was allowed to write a small amount of text to see if they had understood how to use the input system. The participants wrote *Hello World*, and their name. If the participants had any questions they were answered by the experimenter. No auto-correcting was used in any of the input systems used in this experiment.

3.3.2 The Text Used

Two sets of text were used in the trials (seen in Table 3). These were used because they contained an equal amount of letters, a large amount of the same letters and an equal amount of spaces (the same keystrokes), this allowed the experiment to be fully counterbalanced. They were designed specifically for this experiment to remove any carry-over effects. The two texts were of equal size: characters without spaces 119, characters with spaces 143. It was designed to take into account the issues that have been outlined in other text input evaluations (MacKenzie and Soukoreff 2002) such as the lack of numbers and punctuation.

Text 1	Text 2
Pack my box with five dozen liquor jugs. 9231558470. AIU ! He relies on everything listing round to me. H659 EJF. mirage super slime tribesman	Lazy dogs jump over the quick brown fox. 1879452630. EOU ? Here lies one very thin silting ground tome. K327 NTO. sugar empire miles brainstem

Table 3: The two sets of texts used in the experiment.

The two sets of text were checked against the speech recognition dictionary of the software and any words that did not previously exist there were added. Ideally real world users in 'day-to-day' situations would be used to evaluate text entry, but there are obvious problems in matching what the user wanted to say with what the user actually says. In this experiment we used the two texts that have been designed for evaluation purposes. There are other phrase sets available, such as those developed by MacKenzie and Soukoreff (2003), but these did not take into account the use of numerical and punctuation-based input in written language.

3.3.3 Design

The experiment consisted of a 2 by 4 mixed design. The 2 independent variables (*IV*) were the movement condition and input device condition. The 2 levels of the movement condition *IV* were standing and walking. The 4 levels of text input *IV* types were wrist keyboard, on-screen keyboard, speech recognition and handwriting recognition. These input devices were chosen as they are readily available 'off the shelf' and therefore have pre-existing drivers and support available.

There were three dependent variables measured: the time to input the text, the errors in the inputted text and a task load rating. The movement condition (standing and walking) and the input text (1 or 2) were counterbalanced using a Latin-Square design.

This experimental design is used in the three experiments that follow this one. The two independent variables (*IV*) are the movement condition and input device condition and the two levels of the movement condition *IV* are standing and walking. The walking section of the experiment as explained in the procedure (*below*) was also carried through to these experiments. In these experiments there are three levels of pointer input *IV* types. These are stylus, off-table mouse and a track-pad. As previously mentioned there are three dependent variables that are measured: the time on task, the errors made during the task and the task load rating. Ten participants were assigned to each input type. (This section will be referred to in the following experiments to avoid repetition).

3.3.4 Procedure

The participants then put on the vest-based wearable computing system; depending on the handedness of the participant the touch screen and the wrist keyboard could be moved to either the left or right had side of the participant. They were then asked to walk in a straight line for 10 metres, to see if the vest was comfortable; if not, the vest could be adjusted to fit more comfortably. Each of the 40 participants spent a session training/understanding the principles of their assigned input system. The participants were requested to input text whilst walking and to input text whilst standing stationary, the order of this being done in relation to the counterbalancing. The text (*text1* or *text2*) was displayed on the screen in Notepad for the participants to read. During the speech condition, the participants read out each word and spoke the non-lexical letter strings; digit strings were read out character by character; the letter strings were read out in an alphabetical manner. The participant was asked to walk around six stationary road cones placed one metre apart in a straight line while taking part in the walking part of the experiment, as seen in Figure 3.2.

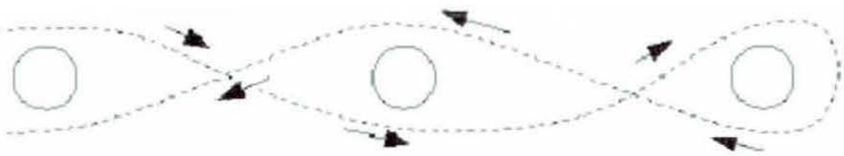


Figure 3.2: An example of participants walking around obstacles used in this experiment. From Pirhonen, Brewster and Holguin (2002).

There was no background noise during the experiments as this might have conflicted with the voice recognition software. After finishing each movement condition of the experiment, the participant completed a NASA task load assessment.

3.4 Results

A mixed factorial ANOVA was carried out upon each data set. Tukey HSD *post-hoc* tests were used to find where the actual significance differences between the devices lay (Field 2000; Howell 2002; Gravetter and Wallnau 1985). Tukey HSD *post-hoc* tests were used, as they allow a pairwise comparison of all the treatment groups, and minimise the probability of a type 1 error occurring. This would occur if the null hypothesis was rejected when it was true

3.4.1 Time

The results based on the time on task are displayed in Table 3.1. The data from this experiment can be found in Appendix 1. Table 3.1 displays the average time along with the standard deviations. This is based on the mean, speech recognition was the fastest method of input in both movement conditions, while the on-screen keyboard was the slowest in both movement conditions. All of the results relating to this experiment can be found in Appendix 1.

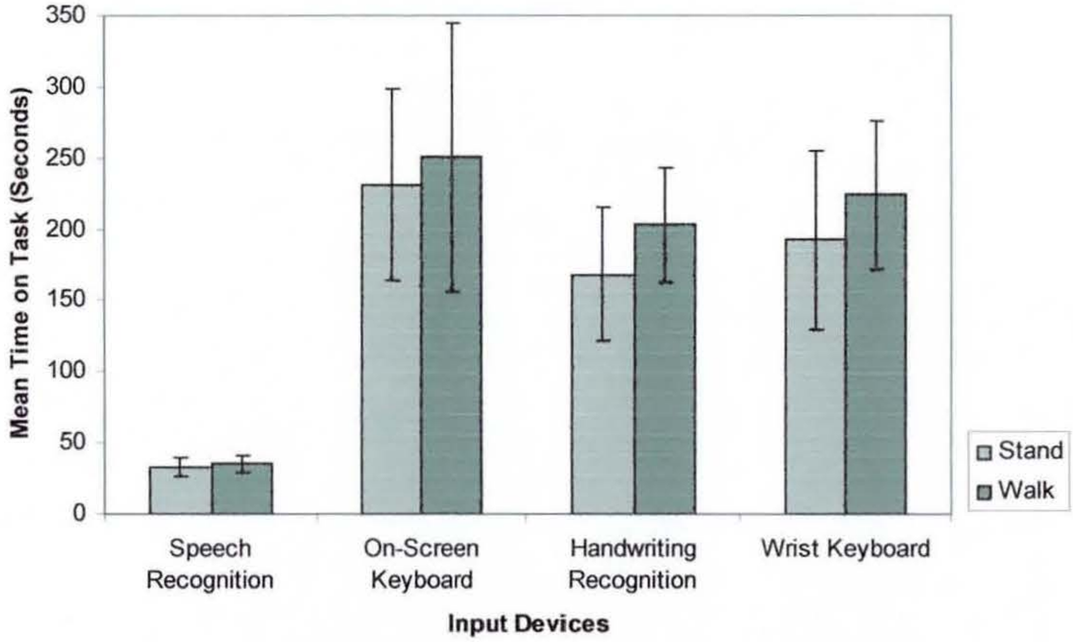
Descriptive Statistics - Time (*Seconds*)			
Movement	Device Type	Mean	Standard Dev.
Standing	Speech Recognition	33.054	6.385
	On-Screen Keyboard	230.847	67.723
	Handwriting Recognition	168.287	46.693
	Wrist Keyboard	192.341	62.628
Walking	Speech Recognition	35.039	5.992
	On-Screen Keyboard	250.410	94.519
	Handwriting Recognition	203.033	39.981
	Wrist Keyboard	224.076	52.424

Table 3.1: Wearable text input table (Time - seconds).
The table shows the participant's mean input times and standard deviations for both movement conditions.

A mixed factor ANOVA was used to statistically analyse the results using SPSS. There was a highly significant main effect of input device type ($F(3,36)$, = 30.938, $p<0.001$), suggesting that the time taken to input the text was significantly affected by the input device type.

Tukey HSD *post-hoc* tests revealed highly significant differences between speech recognition and the other input devices only. The largest difference occurred between speech recognition and the on-screen keyboard ($p<0.001$) with a mean difference of 206.582 seconds. This was followed by speech recognition and the wrist-keyboard ($p<0.001$) with a mean difference of 174.162 seconds and finally between speech recognition and handwriting recognition ($p<0.001$) with a mean difference of 151.613 seconds. These differences are represented in Graph 3.1. The *post-hoc* tests found no significance differences between the on-screen keyboard, handwriting recognition and wrist-worn keyboard input devices.

There was a highly significant main effect of movement ($F(1,36)$, = 16.320, $p<0.001$), with walking being 22.007 seconds slower than standing. There was no significant interaction between the input device type and movement condition



Graph 3.1: Wearable text input graph (Time - seconds).
This graph shows the participant's mean input times while standing and walking. They are shown here with standard deviation error bars.

3.4.2 Error

Any deviation from the initial texts was counted as an error. These were counted and recorded by the experimenter. A mixed factor ANOVA showed that the errors made were significantly affected by the input device type ($F(1,36)$, = 8.906, $p<0.01$). The mean errors and standard deviations can be seen in Table 3.2.

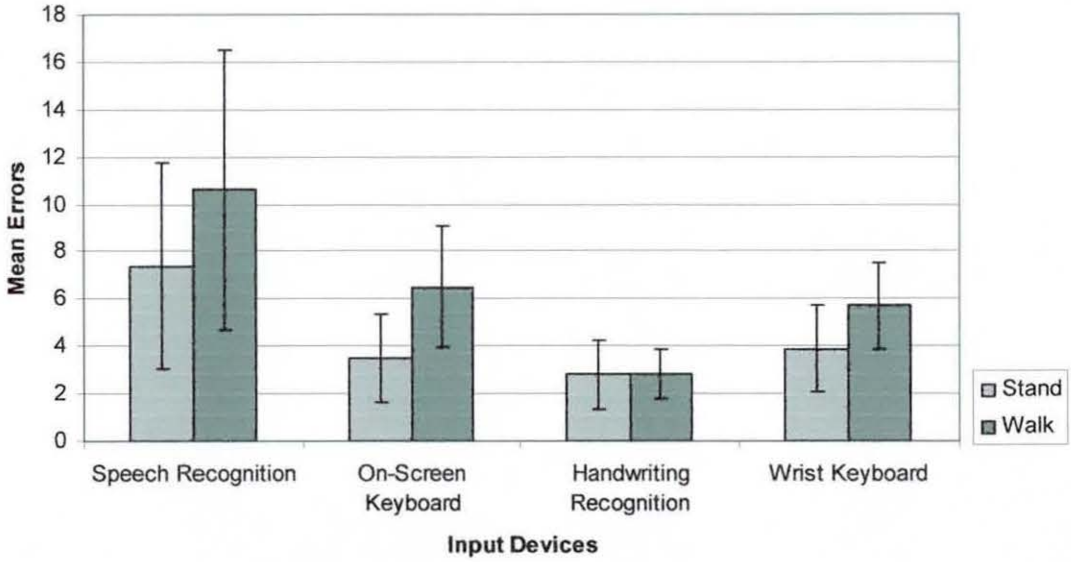
Descriptive Statistics - Errors			
Movement	Device Type	Mean	Standard Dev
Standing	Speech Recognition	7.40	4.377
	On-Screen Keyboard	3.50	1.841
	Handwriting Recognition	2.80	1.476
	Wrist Keyboard	3.90	1.792
Walking	Speech Recognition	10.60	5.948
	On-Screen Keyboard	6.50	2.593
	Handwriting Recognition	2.80	1.033
	Wrist Keyboard	5.70	1.829

Table 3.2: Wearable text input (Errors made on task).
The table shows the mean amount of errors made while standing and walking. They are shown here with their associated standard deviations.

Graph 3.2 gives a more intuitive representation of the mean amount of errors from each device while standing and walking. The errors were higher in the walking level of the experiment, except for handwriting recognition which interestingly had a consistent mean error rate of 2.8 per 143 characters in both movement levels. This is examined in further detail in the discussion section.

There was a highly significant main effect of movement. Overall the mean errors were higher for the walking level of the movement condition ($F\ (1,36), = 23.114, p<0.01$), with the largest mean difference occurring between movement levels on the speech recognition condition, a difference of 3.2 errors per text set. There was also highly significant interaction between the device type and the movement condition ($F\ (3,36), = 8.906, p<0.05$) suggesting a relationship between movement and the type of input device. Interestingly the handwriting recognition system had the same amount of errors in both movement conditions. These results are further expanded upon in the discussion section.

Tukey HSD *post-hoc* tests showed highly significant differences between speech recognition and all of the other input devices; on-screen keyboard ($p=0.013$), wrist keyboard ($p=0.008$) and the largest difference occurred between speech recognition and hand writing recognition at ($p<0.001$). On further examination of the *post-hoc* results it was found that there were no significant differences in the comparative tests between handwriting recognition, wrist keyboard and on-screen keyboard. The significance lay between speech recognition and the other input devices only.



Graph 3.2: Wearable text input (Errors made on task).
Mean errors for each text input device while standing and walking, shown with standard deviation error bars

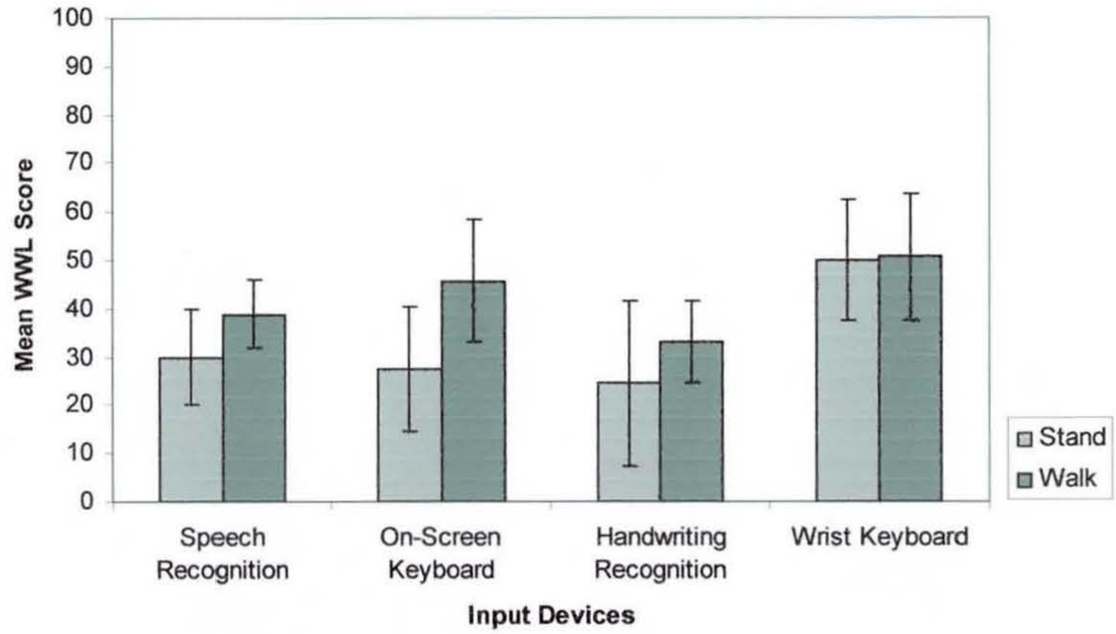
3.4.3 Workload

After each movement level and input type, the participant completed a software-based NASA TLX assessment to assess their subjective workload (NASA TLX v1.0 – Ames Research). The Weighted Workload was calculated by this software package. The mean workload results are shown in Table 3.3 (next page).

Descriptive Statistics - Workload			
Movement	Device Type	Mean	Standard Dev
Standing	Speech Recognition	30.0	9.820
	On-Screen Keyboard	27.4	12.842
	Handwriting Recognition	24.5	17.083
	Wrist Keyboard	50.0	12.337
Walking	Speech Recognition	38.8	6.972
	On-Screen Keyboard	45.7	12.667
	Handwriting Recognition	33.2	8.469
	Wrist Keyboard	50.7	13.182

Table 3.3: Wearable text input (Mean - Workload).
Average WWL (Weighted Workload) score for each text input device while standing and walking, shown with their associated standard deviations.

The results from these assessments can be seen graphically illustrated in Graph 3.3, these were analysed using a mixed factor ANOVA. It was found that there was a significant main effect of device type upon the participants' subjective workload (F (3,36), =12.846, p=0.01). Tukey HSD *post-hoc* tests found highly significant differences, but only between the wrist keyboard and each of the other three devices: speech recognition (p=0.002), on-screen keyboard (p=0.008), with the main difference occurring between the wrist keyboard and handwriting recognition ($p<0.001$), with the wrist-keyboard having a significantly higher workload. There was no significant interaction between the device type and movement levels. A significant main effect of movement was found between device type and movement F (1,36), =702.903, $p<0.001$). This shows that the movement levels had a significant effect upon the input device in terms of workload. Graph 3.3 indicates that the largest difference relates to the on-screen keyboard, while the smallest difference relates to the wrist keyboard.



Graph 3.3: Wearable text input (Mean - Workload).
Average WWL (Weighted Workload) score for each text input device while standing and walking, displayed with standard deviation error bars

.

A further analysis of the six subscales that make up the WWL were carried out in order to examine the effects of the different input devices on the six NASA TLX subscales. In order to do this we carried out a series of mixed factor ANOVAs. There were no significant main effects found for performance or for frustration. The results for the four remaining subscales are given here. The analysis of the mental demands subscale found a highly significant main effect of input device type upon mental demands (F (1,36),= 29.424, p=0.01). The *post-hoc* tests found highly significant differences between the pair-wise comparisons as follows: on-screen keyboard and speech recognition (p=0.01), on-screen keyboard and wrist keyboard (p=0.003), and wrist keyboard compared with handwriting recognition (p=0.004). This suggests that the significance here occurs mainly between the on-screen keyboard and other devices, but interestingly not between on-screen keyboard and handwriting recognition. This may be because they are both 'stylus' based systems, but more research would have to be carried out to answer this. A highly significant main effect of input device on physical demands was shown by the analysis (F (1,36),=14.429, p=0.01). The *post-hoc* tests found highly significant differences between the wrist keyboard and all of the other input devices: speech recognition (p=0.001), on-screen keyboard (p=0.018), handwriting recognition (p=0.003). This difference showed the wrist keyboard to be the most subjectively physically demanding input device. Speech recognition was the least physically demanding input system.

There was a significant main effect of device type upon temporal demands (F (1,36), =8.360, p=0.006). The *post-hoc* tests displayed a highly significant difference between the wrist keyboard and other text input devices: speech recognition ($p<0.001$), on-screen keyboard ($p<0.01$), handwriting recognition ($p<0.001$). It was found that the wrist keyboard had the highest subjective temporal demands, while speech recognition had the lowest, a difference of 62 on the scale. A significant main effect of device type on the effort rating was found (F (1,36),=4.325, $p<0.05$). The only significance in the *post-hoc* tests was displayed between the wrist keyboard and handwriting recognition ($p<0.05$), with the wrist keyboard showing significantly higher effort than handwriting recognition.

3.5 Discussion

This evaluation has shown that out of the four input devices, speech recognition, although the fastest, also had the highest error rate and second highest cognitive workload. This experiment did not allow the participants to correct any mistakes made in the text. Other studies (Karat, Halverson, Horn and Karat 1999) have found that when participants are allowed to correct their input, the keyboard was faster than speech recognition. This was in relation to desktop users. The study in this chapter found that without correction, speech was the fastest method of input, but it also had the highest amount of errors. It could be hypothesised that correcting all of the errors may make speech recognition a slower system to use.

The slowest method of text input was the on-screen keyboard, this may have been because, "QWERTY is a poor choice for stylus keyboarding. The polarizing positions of common English digraphs in QWERTY, mean that the stylus has to move back and forth more frequently and over greater distances than necessary. The key to a good virtual keyboard is exactly opposite to the idea behind QWERTY." (Żhai, Kristensson and Smith 2004). Yet the results have shown that the on-screen keyboard did not have the highest cognitive workload or the highest error rate. It may have been the case that participants took longer to do the task in order to get fewer errors.

While the wrist-keyboard was not the slowest input system used, it was slower than both speech recognition and handwriting recognition. This may have been because of the small key sizes and also because the user is expected to type with one hand. As Zhai *Et al.* (2004) state, "Typing on these keyboards is difficult due to their reduced size that prevents ten finger touch typing." (Zhai, Kristensson and Smith 2004). Interestingly a participant commented on this fact after finishing the task.

The significant results from the analysis meant that we could reject the first null hypothesis, as we have proved that the type of input device does have a significant effect upon the input time. The rejected null hypothesis was: *1. The input device type does not affect the time taken to complete the text input tasks.* We were also able to reject the second null hypothesis which was: *2. The input device type does not have an effect on the error rate of the text input tasks.* The results had previously shown that the type of input device used did significantly affect the error rate. A significant interaction was also found between the input device and error rate.

On further analysis we found that the error rate for the handwriting recognition condition was the same for both the standing and walking levels of the experiment. This had an average of 2.8 errors, made in both of the movement levels. This suggests that the different movement levels had no impact upon the error rate of this input device. This may be because handwriting as a form of text input is practised by many users on a daily basis; therefore many users are used to using this input system in a variety of different settings, such as standing and walking. It also uses the affordances of the pen (Kristensson 2004). Handwriting-recognition was not the fastest method of input, but had the lowest errors and workload. This may be because, as Lemmens *Et al.* (2000) state, "A possible explanation is that participants striving for as few errors as possible take more time in the dual-task situation." (Lemmens *Et al.* 2000). It could also be hypothesized that the text was not difficult or long enough to thoroughly examine the effects of walking upon handwriting recognition. It might be possible to further extend this study by using longer tracts of text such as the phrase sets developed by MacKenzie and Soukoreff (2003) (these can be found at - http://www.yorku.ca/mack/PhraseSets.zip). The use of evaluative software and methods, such as those developed by Soukoreff and MacKenzie (2004) may have also aided in the collection of error based data, as this would have identified any possible corrections that the participants made.

The results also meant that we could partially reject the third null hypothesis: *3. The movement condition of the participant standing and walking has no effect on the time taken to complete the task or error rate.* This was because the movement condition had a significant effect on the error rate, but the movement condition did not have a significant effect on the time it took to complete the task.

If we examine the overall workload we can reject the 4th null hypothesis which was: *4. The input type does not affect the subjective task load of the participant.* There was a significant difference between the input devices in relation to the participants' weighted workload. On closer inspection there were no significant main effects for performance or frustration found in relation to the input device types. So, overall there was a significant main effect of the input devices on workload, but at the lower subscale level, only four of the categories - mental demands, physical demands, temporal demands and effort - showed any significance relating to the input device used.

Although speech recognition had the highest error rate, it is one of the few plausible methods of 'hands free' input, along with body-gestural input systems (Brewster,

Lumsden, Bell, Hall and Tasker 2003). In certain circumstances speech may be the only input modality available to a user. The nature of many wearable computing systems relies heavily on the user's visual channel, and this is problematic in terms of focusing the user's attention constantly 'on-screen' or on 'input device', thus placing users in a non eyes-free situation which can detract from the user's task. However, it must be noted that there are audio-centric wearable systems (Sawhney and Schmandt 2000). It may therefore be appropriate for a level of auto-correction to be implemented into the speech recognition, such as the auto-correct facility found in Microsoft word, or earcons (Blattner *Et al.* 1989) that alert the user of a syntactical mistake or other error in the text, as Rhodes (1998) explains: "Because audio does not distract the user in the same way as a screen or display interface, audio output is especially useful where the user is driving, involved in delicate operations, or may be visually impaired", but this may depend upon the level of background noise and the type of information that is to be conveyed.
Interestingly the input system that fared best overall was the handwriting recognition system; it had the lowest error rate and lowest workload although it had the second fastest input time, and the error rate was not affected by the movement conditions. This suggests that the users may be able to compensate more for physical movement when they are inputting text in this manner. One advantage of using a stylus with a touch screen is the ability to augment the users' written material with pictures and diagrams that can be integrated into a document or formalised at a later date using software such as Microsoft Journal. The stylus also requires no cables; this may be an advantage as it removes the danger of entanglement. Being cable-free reduced the burden of an already computer-laden user, although some studies (Lumsden and Brewster 2003) have found that users who were carrying extra weight seemed oblivious to it. As handwriting recognition is mainly software-based, the interface could also be adapted to a variety of users' preferences. Another advantage of using a stylus is that it can also be used as a pointing device.

As we noted earlier, there was no significant interaction between the device type and the movement conditions in relation to the time on task, yet there was significant interaction between device type and the movement condition in relation to the errors in the text. This may have been in part due to the participants slowing down while in motion in an attempt to avoid errors.

This research has opened up many areas of future research that apply to our work and to the work of others working in this area. The next step in this research is to further examine the devices in relationship to their longer term use, by using task oriented evaluation, and so further evaluate their usability and appropriateness to mobile wearable users. The stylus used in the handwriting condition can also be used as a pointing device; it would be intriguing to see it compared to other pointing devices used with wearable computing systems, such as off-table mice.

During the experiment it was noted that all participants moved slowly around the cones. This is because they were dual-tasking. Their attention was focused in two places simultaneously. One was the text input task, while the other was walking around the obstacles. It would have certainly been advantageous to have measured the distance as this would have provided another set of metrics that could have been used to evaluate the user's performance (Brewster, Lumsden, Bell, Hall, and Tasker 2003; Lumsden and Brewster 2003). Interestingly in Karat, Halverson, Horn and Karat's (1999) study users when questioned, said "they found it harder to speak and think and easier to use a keyboard and think". Yet, the investigation carried out in this chapter found that the wrist keyboard had the highest workload. This may be due to the fact that the user's visual attention was being focused on the keyboard for input, the screen for output (feedback) and the path they needed to take to avoid the obstacles in their way. In this study we can hypothesise that handwriting recognition was the best overall text input device because users have a lot of prior knowledge of the system, as it mimics pen and paper based systems, the user's attention is focused on the screen for input and output (attention may be focused in other places if the user is dual tasking) and it has a small cognitive workload.

3.6 Conclusion

This set of experiments has highlighted that the text input systems used in this experiment have a significant effect on the user in terms of time and error rates and significantly affects the amount of workload that a user experiences.

In relating back to the contributions of the thesis we have: presented techniques for the evaluation of input systems for use with wearable computers from the literature, presented a set of empirical results that can aid in the design and choice of input systems for use with wearable computer and provided a research platform on which other Wearable Computing-focused Human Interaction studies can be based. This research also forms a set of metrics based on input speeds and error rates through scientific empirical investigation that relate to the use of input systems for wearable computers, while the user is stationary and mobile.

This experiment importantly found that overall handwriting recognition had significantly fewer errors and gave the user significantly less workload as compared to the other input systems. Speech recognition was the fastest input system, but also had the highest error rate. It may be the case that the text sets used in this experiment were too constrained to examine the appropriateness of speech recognition as an input system for mobile users. In this experiment the background noise was controlled (it was low level). In many environments there is a continually changing background noise level. There are other factors associated with the users' voice, such as accent, speech patterns and the effects of: stress, background noise levels (*Lombard effect* (Chi and Oh 1996)) and emotion upon the user. They need to be addressed equally. The results showed that there was a significant difference in terms of the users' mobility upon input errors and workload, with walking having a higher workload and more input errors than standing.

We found that although speech recognition was the fastest method of input, it had the highest error rate and second largest workload. It has been noted within the literature that, 'it is very hard to handwrite while on the move.' (Brewster, Lumsden, Bell, Hall and Tasker 2003). The study in this chapter found that handwriting recognition had the lowest workload and error rate and had the second fastest input time; there was no significant difference in relation to the mean errors made between the standing and walking levels of the movement condition (as discussed in the discussion). This however does not suggest that handwriting is easy to use while on the move, but we may hypothesise that it is not as hard (in terms of errors and cognitive workload) as the other three systems evaluated. It may simply be the case that the participants found that they could transfer their existing handwriting skills over to the handwriting recognition system. The wrist keyboard had the longest input time and the highest workload, but there was no significant difference between the amount of errors using the wrist keyboard or the on-screen keyboard. The on-screen keyboard had the third fastest input time and also showed the greatest difference in input errors between standing and walking.

Chapter 4: An Extended Evaluation of Three Wearable Computing Input Devices for Target Selection

4.1 Introduction

This experiment examines the users' performance and compares the input devices used, in terms of their accuracy, speed and cognitive workload when used to complete a target selection task as part of a wearable computing system. This study leads on from the previous experiment in chapter five and is primarily aimed at comparing stylus, track-pad and off-table mouse. The targets selected in a real world scenario may be as diverse as icons, buttons, textboxes, radio buttons or drop down menus. In this experiment the track-pad replaced the speech recognition system. It was added to the experiment as it is a commonly used input device for laptop users and is easily available 'off the shelf' item.

Although there have been studies relating wearable computing and pointer based input, as we have already seen, few report empirical results that relate to the user while wearing a computer. Studies relating to input with wearable computers have found that it is important to recognise the movement condition of the wearable user, whether it was walking, running or standing still, and that this movement impacts upon the task performed by the user. Some studies have taken into account the movement of the user when using mobile devices, both Brewster (2002) in relation to audio feedback on PDAs and importantly Petri *Et al.* (1998) in their examination of wearable computers and visually impaired users. Both have stressed the need for evaluation in real world envirionments. The way in which a user interacts with a pointer based system whilst mobile and using a wearable computer is complex, and dependent upon a whole range of factors that we will further discuss. It is important to focus design and evaluation in the real world, because it can help to guide and focus evaluations in relation to the user's task within their day-to-day environment, although this can be problematic when a uniform set of experiences is needed so as not to bias any experimental results. In real world situations there is much more variablity relating to the user's envirionment such as: the weather, lighting conditions, background noise (Najjar, Ockerman and Thompson, 1998), the number of other users in that environment and how dangerous the environment is, as well as a vast range of factors that one may find in the user's natural day-to-day environment. Every user's envirionment is different, and therefore it must be noted that attempting to control a 'natural' environment so as to create a uniform experience is near impossible; this may also prove contrary to where and how the user actually uses their system. So, for any experimental investigation it is necessary that we take into account certain factors: firstly, all participants in an experiment must have a set of uniform experiences and secondly, that the user may be mobile while using their wearable system. It is therefore important that a mobile condition be implemented within any evaluation structure, but this must initially be within a controlled environment to create a uniform set of user experiences so as not to negatively affect any resulting data.

4.1.1 Rationale

The experiment compared three input methods that can be used for target selection with a wearable computing system. The input methods were stylus, a track-pad and an off-table mouse. These devices were chosen because they are commonly available to many wearable computer users. The experiment contained two movement conditions and used one target size. This first experiment used time and errors as dependent variables and a cognitive workload scoring system (NASA TLX). The use of NASA TLX as a task load assessment tool can be found throughout the research literature (Hart and Wickens 1990; Hart and Staveland 1988). This software also recorded any errors made by the participant, when the participant missed the target.

This study was designed with three goals in mind: to compare the three input methods in terms of speed, errors and workload; to detect if any of the input methods or any experimental techniques used could cause problems in future generations of experiments; and to examine and identify any factors affecting the use of pointing devices whilst the user is mobile and wearing a computer.

In this experiment we excluded voice recognition. The main reason for this was its poor performance in a previous target selection pilot study (Appendix 5). It should be pointed out that long-term usage of speech recognition software may heavily impact upon the performance of the input system and therefore contribute to its accuracy. Speech recognition is also one of the few feasible means of input that could be used in a hands-free environment, but for this study we are comparing three devices that require the user to use their hand.

4.1.2 Hypotheses

The null hypotheses are as follows:

h1. The input device type does not affect the time taken to complete the target selection task.

h2. The input device type does not have an effect on the error rate of the target selection task.

h3. The movement condition of the participant standing and walking has no effect on the time taken to complete the task or on the error rate.

h4. The input device type does not affect the subjective workload of the participant.

h5. The movement condition does not affect the subjective workload of the participant.

4.1.3 Participants

30 participants took part in the study. The participants were volunteers from the staff and student body of Loughborough University. The participants that took part in the trials were naïve to the purpose of the experiment and had no previous experience using the target selection devices used in the trial or of using wearable computers. The participants were put into groups of ten at random. This was also the same in the two experiments following this one.

4.2 Method

4.2.1 Apparatus

The three input devices used are shown below. The consisted of: a stylus and touch-screen (used in Chapter 3), an off-table mouse and a track-pad. These devices will also be used in the pointing experiments in Chapters five and six. The wearable computing system used was the vest-based Xybernaut MA IV as used in Chapter three.

Figure 4.1: Stylus

Figure 4.2: Off-Table mouse.

Figure 4.3: Track-pad (Cirque)

For more information on the pointing devices used in the experiments see -

Figure 4.2 - http://www.trust.com/products/product.aspx?artnr=12772
Figure 4.3 - http://www.cirque.com/products/easy.html

4.2.2 Design

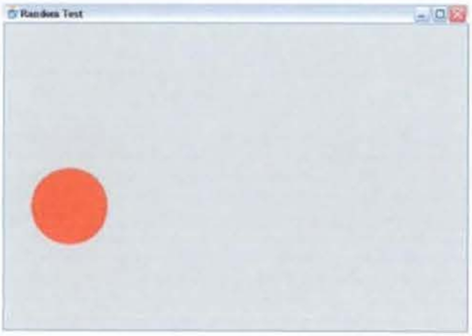
The experiment consisted of a 2 x 3 mixed factor design: a between-subjects design on the input device level and within subjects on the movement level. This design was chosen to avoid carry-over effects between the two movement conditions. The two levels of the movement condition *IV* were standing and walking. The three levels of target selection *IV* types were a Xybernaut stylus, a Cirque track-pad and an off-table mouse. There were three dependent variables measured. These were the time to select the target and the amount of target misses (errors) and the subjective workload of the participant, taken in each movement condition.

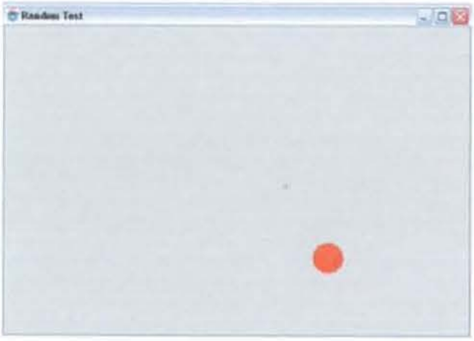
4.2.3 Procedure

Each participant was asked to put on the wearable computer in the vest-based system. The vest was then corrected to fit the participant. Depending on the counterbalancing, the participant either started the experiment standing or walking. They were asked to select each target that appeared on the screen. The target selection was started off by selecting the first target and then the participant would then use the said device to select the targets. Upon the last selection the software would stop producing targets. Once the practical experiment had finished the participant completed a NASA TLX assessment. This was carried out in each movement level. During the walking section of the experiment, the participant would select targets while walking around six road cones placed at one metre intervals.

A software package was used by IBM called IDtest was used to run the target selection experiment. 500 targets with a diameter of 10, 20, 30, 40 and 50 pixels (2.5mm, 5mm, 7.5mm, 10mm and 12.5mm) were used in the experiment, each participant selecting 1000 targets in all (500 standing, 500 walking). The target sizes were chosen to combine experiments previously done in the literature by (Smith *Et al.* 2000; Curry *Et al.* 1996) and to make the experiment more difficult for the participant to complete. It was theorised that having different target sizes would test the performance of the input device in a more thorough manner. It was found that this was an ideal number for the participants to select, as previous test runs had shown participants unwilling to select more targets. A screen size of 600 by 400 was selected to fit onto the portable Xybernaut screen. The position of the targets was randomly generated on the screen by the software.

The software allowed the user to select different randomisation patterns (where the targets appeared on the screen): these patterns were selected by entering a numerical value. To minimise carry over effects caused by using the same pattern in each movement condition, an online number randomisation tool (http://www.randomizer.org/) was used to create two numerical values. These were the numbers 10 and 13. At the end of each experiment a text file was created that gave the amount of errors during the experiment.





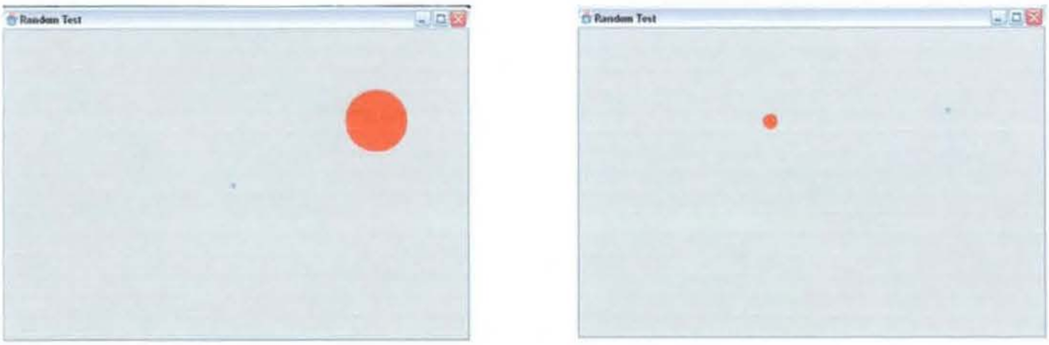


Figure 4.4: Screen shots from the target acquisition experiment.

The participant selected targets as they came on the screen; each participant selected 500 targets while standing/stationary and 500 while walking in a figure of eight through six traffic cones, spaced at one metre intervals. Overall 20,000 target selections were timed.

The experimental data collected was used to provide the mean movement time for each participant, mean errors made and workload rating. This data was analysed using a mixed factor ANOVA based statistical method (Field 2000). The participant's cognitive workload was measured in each condition using NASA TLX to provide us with the participant's subjective workload. This was important as it gave a reliable set of diagnostic data that could be used to examine and compare the participants' workloads in relation to the use of one of the input devices. The experiments were timed using a stopwatch. Errors made on the task were target misses. These were recorded by the IDtest software.

4.2.4 Training Task

In order that each participant understood the principles behind the task they were to complete, so they were asked to first take part in a pre-evaluation task. They selected five targets using the stylus, track-pad or thumb mouse. This proved beneficial as it allowed the participant to get a fuller understanding of the task and also allowed the user to computer was comfortable.

4.3 Results

Three sets of data were collected for each participant in the walking and standing sections of the experiment. Three sets of data were collected in both movement conditions. This was done while the participant was walking and while they were standing; this allowed the effect of movement upon the user to be examined. The three sets of data related to the time taken to complete the task in seconds, the amount of errors made during the completion of the task and the weighted workload of the participant calculated by NASA TLX. If there was a significant effect of the input device upon the participants' weighted workload, the results were further explored at the subscale level. The data relating to this experiment can be found in Appendix 2.

4.3.1 Results-Time

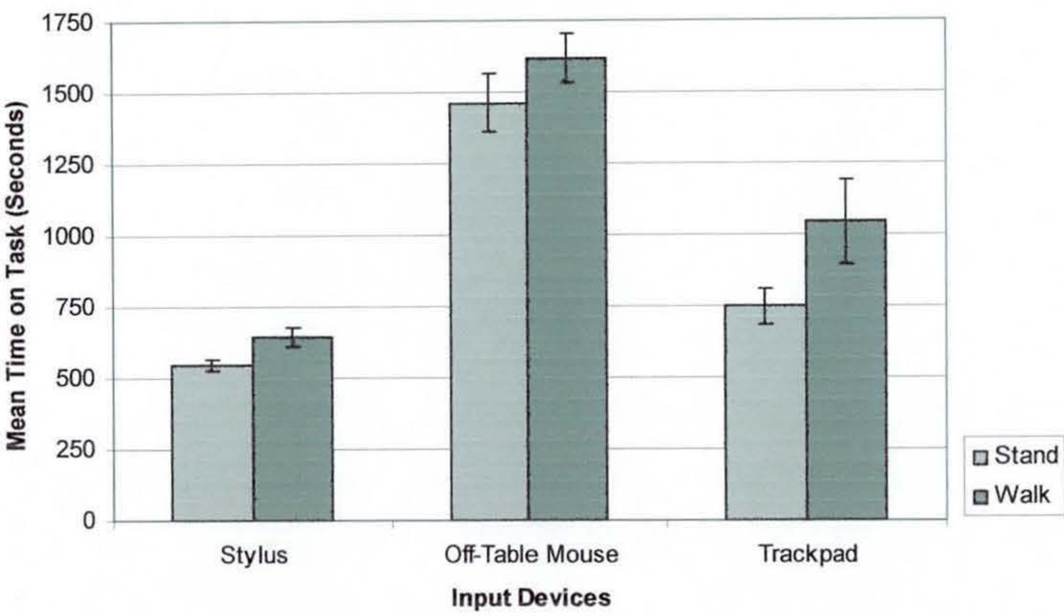
Three groups of 10 participants were timed, each using a different input device: group 1 used a stylus, group 2 used the off-table mouse and group 3 used a track-pad. This was done in order to find out which device was the quickest when used to select targets on the screen. Table 4.1 shows the mean times it took for each input device group to finish the task, along with their associated standard deviations.

Descriptive Statistics - Time			
Movement	Device Type	Mean	Standard Dev
Standing	Stylus	545.246	18.72427
	Off-Table	1463.552	102.4597
	Track-pad	747.209	64.00642
	Total	918.669	406.4954
Walking	Stylus	646.698	33.36
	Off-Table	1619.732	84.254
	Track-pad	1043.744	150.0511
	Total	1103.391	417.8717

Table 4.1: Wearable target selection. (Time - seconds)

Mean input times with associated standard deviations for the participants while standing and walking.

There was a significant main effect of the device type upon the time taken to complete the target acquisition task ($F(2,27) = 442.373$, $p<0.01$). Graph 4.1 further illustrates these differences. Similarly a significant main effect of movement was found ($F(1,27) = 109.507$, $p<0.01$). A highly significant interaction was reported between the input device and movement ($F(2,27) = 10.832$, $p<0.01$).



Graph 4.1: Wearable target selection. (Time - seconds).
This shows the mean input times for the 3 devices, while both walking and standing with standard deviation bars.

Tukey HSD *post-hoc* tests displayed a highly significant difference between the three different input devices. The largest difference was found between the stylus and off-table mouse with $p<0.001$, while the smallest difference was found between the stylus and track-pad as seen in Table 4.1, $p<0.001$. There was also a highly significant difference between the trackpad and off-table mouse $p<0.001$.

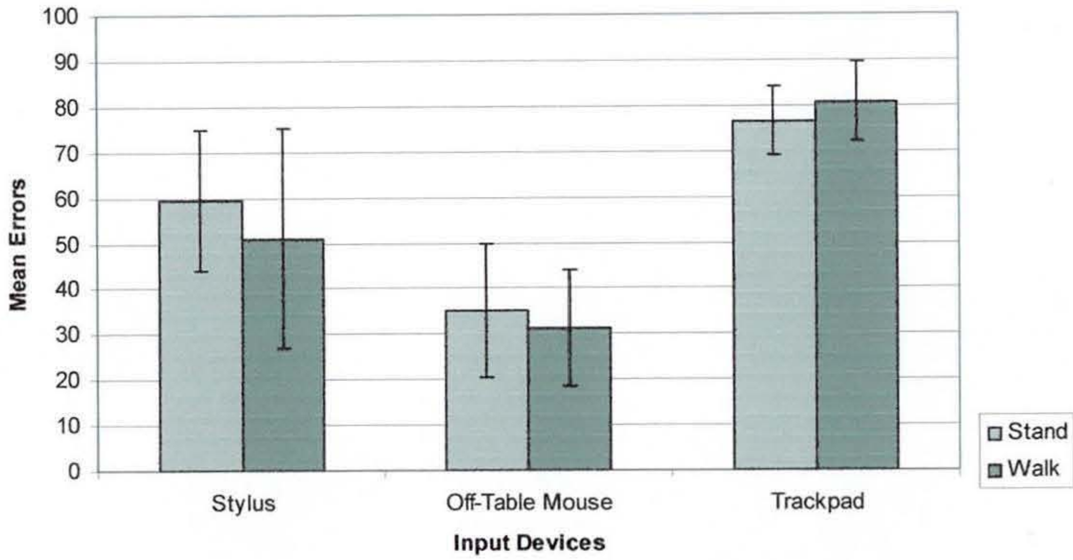
4.3.2 Results - Error

Table 4.2 presents the means and standard deviations for each input device group while standing and walking. These are based on the errors made while completing the target acquisition task.

Descriptive Statistics - Errors			
Movement	Device Type	Mean	Standard Dev
Standing	Stylus	59.5	15.58668
	Off-Table	35.3	14.69732
	Track-pad	76.6	7.560129
	Total	57.13	21.38116
Walking	Stylus	51	24.11546
	Off-Table	31.3	12.84134
	Track-pad	80.8	8.638415
	Total	54.367	26.13689

Table 4.2: Wearable target selection. (Errors made on task)
Mean errors made and standard deviations for participants while standing and walking.

There was a significant main effect of the device type upon the errors was found ($F(2,27) = 37.03$, $p<0.01$). No significant main effect of movement was shown ($F(1,27) = 0.686$, $p<0.5$). Similarly, no significant interaction between the movement and input device type was displayed ($F(2,27) = 1.239$, $p<0.5$).



Graph 4.2: Wearable target selection. (Errors made on task).
This shows the mean errors made for the 3 devices, both while walking and standing with standard deviation bars.

Following a Tukey HSD *post-hoc* test the pairwise comparison found highly significant differences between the stylus and off-table mouse at $p<0.01$, between the stylus and track-pad with $p<0.01$ and between the off-table mouse and track-pad $p<0.01$. The largest difference was between the trackpad and off-table mouse, while the smallest difference was between the off-table mouse and stylus. These differences are illustrated in Graph 4.2

4.3.3 Results - Workload

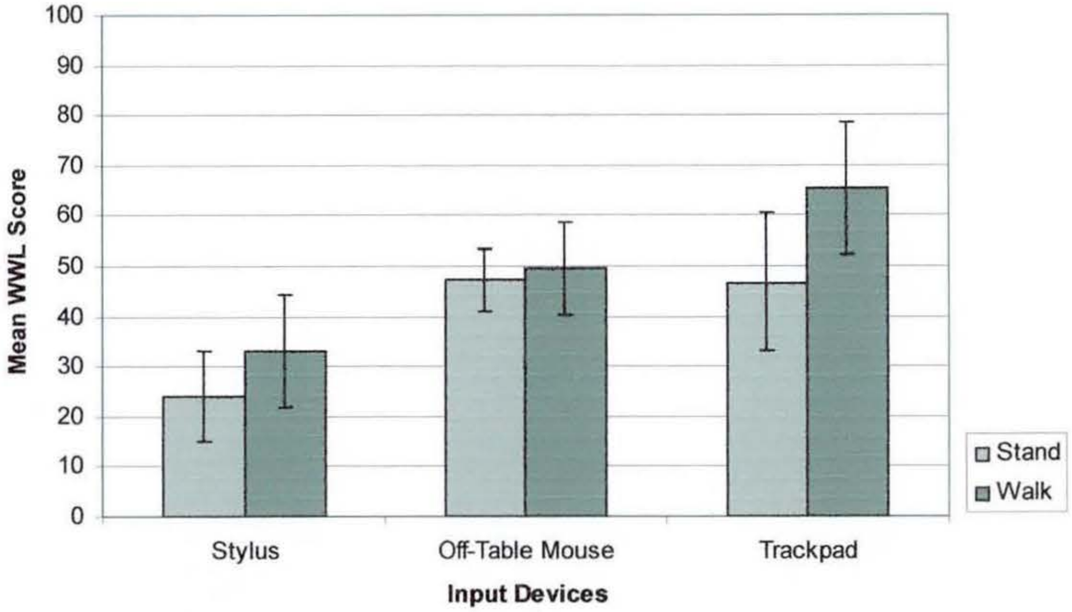
The table below (Table 4.3) shows the mean workload (WWL) scores from the NASA TLX assessment that the participants carried out after each of the two stages of the experiment. As we can see from the results displayed within the table, the stylus had the lowest workload while used standing and walking. The off-table mouse had the highest workload while standing, while the track-pad had the highest workload when walking.

Descriptive Statistics - Workload			
Movement	Device Type	Mean	Standard Dev
Standing	Stylus	24.1	8.9
	Off-Table	47.2	6.2
	Track-pad	46.8	13.9
	Total	39.3	14.7
Walking	Stylus	33.1	11.3
	Off-Table	49.5	9.2
	Track-pad	65.4	13.2
	Total	49.3	17.3

Table 4.3: Wearable target selection. (Mean – Workload)

Average workload scores and standard deviations.

Overall there is a significant main effect of device on the weighted workload (WWL) with stylus having the lowest WWL ($F(2,27) = 32.667$, $p<0.01$). A significant main effect of movement was also found ($F(1,27) = 13.338$, $p<0.05$). These differences are further represented in Graph 4.3. There was no significant interaction between the device type and movement ($F(2,27) = 3.004$, $p<0.1$).



Graph 4.3: Wearable target selection. (Mean - Workload).
This shows the mean workload scores for the 3 devices, while both walking and standing. They are shown here with their associated standard deviation error bars.

The Tukey HSD *post-hoc* tests resulted in the following: a highly significant difference between the stylus and off-table mouse $p<0.01$; similarly a highly significant difference between the stylus and track-pad was shown $p<0.01$; but no significant difference between the off-table mouse and track-pad $p<0.5$.

The results for this analysis have been tabulated into two paired table sets to more clearly represent the six sets of results from the NASA TLX assessment that was carried out by each participant (seen below). The tables show the significant main effects of device and movement; no interaction was found.

Table 4.4: Wearable target selection. (Workload Subscales 1.)

This shows the significance of device and movement for the NASA TLX subscales contained in the table. No significant interaction was found

NASA-TLX / Main Effects	Mental Demands	Physical Demands	Temporal Demands
Device	No Significance	F(2,27)=29.303,p<0.01	F(2,27)=24.916,p<0.01
Movement	F(1,27)=6.189, p<0.01	No significance	No significance

Table 4.4: Wearable target selection. (Workload Subscales 1.1)

This table is paired with table 6.3 and shows the Tukey HSD *post-hoc* results from the pairwise comparisons.

NASA-TLX / *Post-hoc*	Mental Demands	Physical Demands	Temporal Demands
Stylus – Mouse	No Significance	p<0.01 Mean Diff 34.75	p<0.05 Mean Diff 15.75
Stylus –Track-pad	No Significance	p<0.01 Mean Diff 34.25	p<0.01 Mean Diff 35.75
Track-pad – Mouse	No Significance	No Significance	p<0.01 Mean Diff 20.00

They also show the results from the Tukey HSD *post-hoc* comparisons and the mean differences that occurred between the pairwise comparisons. It was found that the device with the lowest mean average in each subscale was the stylus.

Table 4.5: Wearable target selection. (Workload Subscales 2.)

This shows the significance of device and movement for the NASA TLX subscales contained in the table. No significant interaction was found

NASA-TLX Main Effects	Effort	Performance	Frustration
Device	$F(2,27)=26.756, p<0.01$	$F(2,27)=4.596, p<0.01$	$F(2,27)=17.116, p<0.01$
Movement	$F(1,27)=12.322, p<0.01$	No significance	$F(1,27)=7.226, p<0.01$

Table 4.6: Wearable target selection. (Workload Subscales 2.1)

This table is paired with table 6.5 and shows the Tukey HSD *post-hoc* results from the pairwise comparisons.

NASA-TLX *Post-hoc*	Effort	Performance	Frustration
Stylus – Mouse	$p<0.01$ Mean Diff 22.50	No Significance	$p<0.05$ Mean Diff 27.25
Stylus –Track-pad	$p<0.01$ Mean Diff 37.00	$p<0.01$ Mean Diff 15.00	$p<0.01$ Mean Diff 33.75
Track-pad – Mouse	$p<0.05$ Mean Diff 14.50	No Significance	No Significance

4.3.4 Discussion

From Graph 4.1, it can be seen that the stylus was the fastest input device in both movement conditions. The slowest input device was the off-table mouse, followed by the track-pad. The reason for the off-table mouse being the slowest device may be due to the fact that the participants using this system had two mechanisms to control. The first is the trackball to control the movement of 'on-screen' items and the second is the trigger to select the items. Interestingly, participants using the off-table mouse made the least amount of errors. This could also be due the design of this device making the act of selection a two-part process; the participant first has to position the 'cursor' over the target with the trackball and then use the target to select. We can hypothesise that this two-part process leads to fewer errors. The stylus used a 'pecking' action, and the track-pad users dragged their finger across the pad and then tapped on the pad to select. It may also be the case that this made the task more difficult to perform, because the participants were doing three tasks: walking, controlling a cursor and using a trigger to select. Yet, the trackpad had a higher mean workload than the off-table mouse in the walking level of this experiment. It has been noted by Svennson *Et al.* (1997) that objective measures of performance cannot be correlated to perceived workload. Therefore, it is possible that a user who performs well on their task may see it as having a high workload. This is a case where further subjective questioning of the participants, could have given further insights into the properties of the three input devices.

In the literature review it was found that trade-offs can occur when a user is attempting to do two tasks simultaneously (referred to as dual-tasking). It was noted that during the experiment that the participants slowed down to select the targets and manoeuvre around the objects, so they could select the targets at a similar rate to the standing condition. It could be hypothesised that in relation to other studies (Pirhonen, Brewster and Holguin 2002; Brewster, Lumsden, Bell, Hall, and Tasker 2003; Lumsden and Brewster 2003), this slowing down was a performance loss in relation to the participant's walking speed, while there was an increase in accuracy-based performance. The participants using the off-table mouse and stylus, both made fewer errors while walking.

We can safely reject the first null hypothesis (***h1***), because we were able to report a significant time difference between the three different devices. This suggests that for tasks that involve a target acquisition task to be performed quickly, the stylus is the quickest device in this evaluation. It may therefore be suggested that it is the most appropriate device to use for fast target acquisition-based tasks.

There was a significant interaction between the movement of the participant and the input device, the track-pad having the largest time difference. It took longer to use walking, while the stylus took more time to use in the same condition and although the off-table mouse was the slowest device to use it showed the second-lowest time difference between the movement conditions, followed by the off-table mouse and stylus. This interaction shows that there is not a uniform effect shared by all of the input devices in relation to the way that movement affects the input time.

These tests were done under laboratory conditions where the participants could see and predict the obstacles that would be in their way. In a real world environment, where professionals, such as the armed forces, explorers and telecommunications engineers, may be interacting with a wearable system while on the move, they can often be placed in situations where the terrain/obstacles is/are not as predictable as a lab based environment. There may also be other factors that could impact upon the test in the real world, such as stress levels, differing light levels and the weather. It may be assumed from the results that for target selection the stylus is significantly faster to use than the off-table mouse used in these trials.

As previously mentioned, the off-table mouse was the slowest device to use, it produces the fewer errors; 35 while standing and 31 while walking. The device that had the largest amount of errors was the track-pad, these were almost double the amount of errors made by the users of the off-table mouse. The errors made with the track-pad were 67 while standing and 80 while walking, followed by the stylus at 59 and 51 errors made on the task. Surprisingly both the off-table mouse and stylus made more errors in the standing section of the experiment and less in the walking part. This may have been in part to people trying to be extra careful in the walking parts of the experiment or it.

The second null hypothesis (***h2***) can also be safely rejected as each input device's error rate significantly differed. The third null hypothesis (***h3***) can only be partially rejected, because although there was a significant effect of movement upon the input time of the device, there was no significant effect upon the errors made.

In examining the differences between the three input devices, it can clearly be seen from Graph 4.3 that the input device with the lowest significant workload in both movement conditions was the stylus. Therefore, the fourth null hypothesis (***h4***) can be rejected as the there was a significant effect of input device upon the participants' workload. Similarly there was also a significant main effect of movement on the participants. This means that we are also able to reject the fifth null hypothesis (***h5***), as the participants' workload while standing was significantly different (higher) while walking. This suggests that the act of being mobile can raise a wearable user's workload. Interestingly the input device with the highest workload while mobile was the track-pad; this may have been because it was worn on the participant's body, as opposed to being held by the participant.

On a further analysis of the six NASA TLX subscales only one scale, Mental Demands, showed no significant effect of device. The subscales Mental Demands, Physical Demands and Performance showed no significant effect of movement, suggesting that there was no significant difference between the participants' score while stationary and while mobile.

This evaluation was able to successfully compare the three input devices and gain a valid insight into the effect of mobility on the wearable user, while using the input device for target selection. There are other 'avenues' of exploration that this experimental procedure could also be used to investigate, such as an input device's appropriateness for selecting targets of different sizes while mobile. This could be used in the design of interfaces for mobile devices. In this experiment the mobile environment was in the same location as the stationary environment. This was done in order to not bias the experiment, but it would be enlightening to carry out the experiments in 'real world' scenarios.

4.4 Conclusion

This experiment has found that, of the three input devices evaluated (off-table mouse, track-pad and stylus) the stylus was the most effective for target selection while both stationary and mobile. This experiment used three evaluation techniques based on: the time taken to complete the task, the errors made during that task and the subjective workload of the participants involved in the task. The stylus performed best in relation to all of these.

These are interesting results, because it would be expected that the user would perform worse (in terms of errors) while mobile, as they are placed within a dual-task scenario. It may have been the case that participants slowed sufficiently to overcome the dual focus of attention that can cause an impaired performance on one task. It could also be the case that the task of target selection is so 'easy' that the act of being mobile has little effect upon the user. This would call for further investigation in relation to the user's movement as we have previously seen in the discussion

Evaluating wearable computers while the user is mobile is a complex and multi-facetted task. Although this experiment has found the stylus to be the most effective device for target selection, it may not be appropriate when used for other pointer-based interfacing techniques, for example dragging and dropping and trajectory based interaction. It is therefore necessary to further explore the pointing devices used in this experiment in further experiments that test their appropriateness in regard to dragging and dropping and trajectory based tasks.

Chapter 5: An Experiment to Evaluate Three Wearable Computing Input Devices for Dragging and Dropping

5.1 Introduction

At the desktop users primarily use a mouse to drag and drop. In a mobile environment it is difficult for the user to use a mouse because they are designed to be used on a smooth flat surface such as the desktop. This is not always an option that is available to the wearable user as they could be running though a jungle, in a desert or even working at the top of a telegraph pole 10 metres off the ground. Wearable users need input devices that allow them to use their interface while on the move. Drag and drop interfaces have been designed for wearable computer users (Kaefer and Weiss 2003; Butz *Et al.* 1999), but these systems were not evaluated while the user was mobile. This section of the thesis evaluates the three input devices used in the previous experiment and examines their performance for dragging and dropping. It may be the case that the use of the drag and drop widget is not appropriate for mobile interaction and this will be discussed in regard to the findings of the experiment, in the discussion section of this chapter.

For further explanations of the drag and drop process see Brewster (1998) in relation to sonically-enhanced dragging and dropping, and Inkpen (2001) for a comparison of drag and drop, and point and click interaction styles of children.

5.1.1 Rationale

The rationale behind this experiment is to further understand the three pointing devices used in chapter six and to further understand their appropriateness for use with a wearable computing system while the user is mobile. Many mobile systems have employed the drag and drop interaction metaphor (Rekimoto 1997; Rekimoto and Saitoh 1999; Iannizotto *Et al.* 2001; Kohno and Rekimoto 2002) yet little empirical research exists that has been conducted while the users were mobile.

5.1.2 Hypothesis

The null hypotheses are as follows:

h1. The input device type does not affect the time taken to complete the dragging and dropping task.

h2. The input device type does not have an effect on the error rate of the dragging and dropping task.

h3. The movement condition of the participant standing and walking has no effect on the time taken to complete the task or error rate.

h4. The input device type does not affect the subjective workload of the participant.

h5. The movement condition does not affect the subjective workload of the participant.

5.2 Method

5.2.3 Design

A 2 x 3 mixed factor design was used for the design of this experiment. This combined a between and within groups design as used in the previous experiment. The three pointing devices and wearable computer were the same as Chapter 4. The dependent variables recorded were: the time to complete the task, the errors made by the participant and the participant's subjective workload. Errors occurred if the participant dropped the selected item anywhere other than on the 'drop' target. These were recorded by the software. The participants were put into groups of ten. Each group was assigned one of three input devices: the stylus, off table mouse or track-pad. Overall thirty participants took part in the experimental trial. To stop carry-over effects the task used two different randomisation patterns. Each participant did one of the two different patterns in each movement condition. The three groups of ten participants were chosen at random through 'blind' selection.

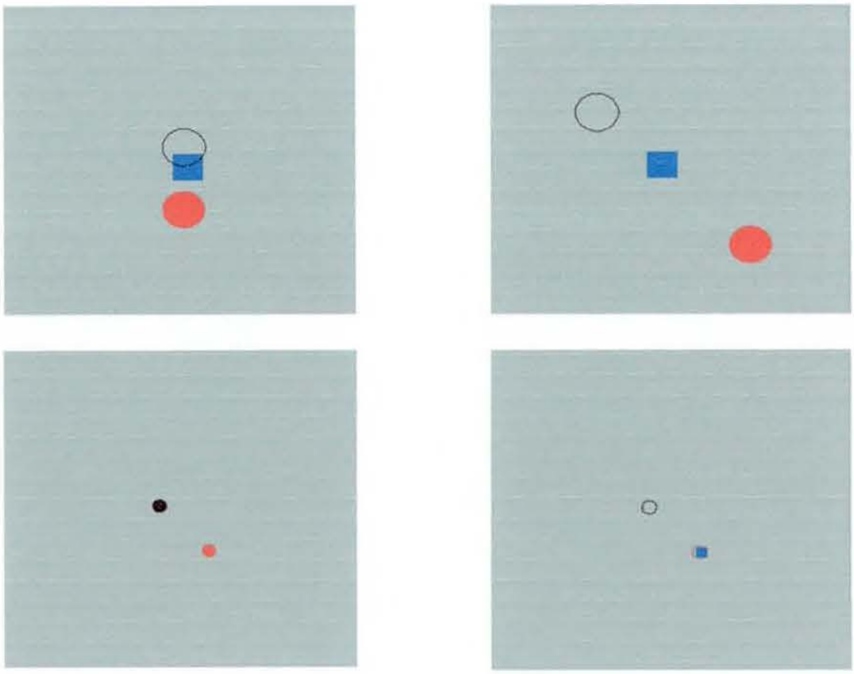


Figure 5.1: Screen shots from the dragging and dropping experiment.

Each participant dragged and dropped 540 items in both movement conditions. The order in which the participants took part in the experiments was balanced using a Latin square method. There were six different target diameter sizes of 10, 20, 30, 40, 50 and 60 pixels (2.5mm, 5mm, 7.5mm, 10mm, 12.5 mm and 15mm). There were three different dragging distances which were: 100, 200 and 300 pixels (25mm, 50mm and 75mm). Targets were dragged at three different angles; at 0, 45 and 90 degrees. There were 54 different combinations that the participant completed. The targets appeared on the screen in a random fashion. The randomisation patterns could be selected in relation to a numerical value, as explained in the previous chapter. To avoid carry over effects, the two numbers generated by the randomisation software used in Chapter 4 were used again. Screen shots from the experiment can be seen in Figure 5.1.

5.2.4 Procedure

The participant was first asked to put on the vest containing the computer. This was adjusted until the vest fitted the participant. The participant was either assigned an off-table mouse, track-pad or stylus depending on the input group that they had been assigned to. The participant was then shown how to use their device and then carried out of one of the dragging and dropping tasks on a monitor. Each participant did the experiment while standing and walking. The order of this was done in relation to the counterbalancing. The participant dragged and dropped 540 targets in each condition. To complete the task the participant had to drag each target, and drop it onto the other appearing on the screen; screen shots from this task can be seen in Figure 5.1. Two targets appeared on the screen, a black and a red target. The aim was to drag the black target onto the red target. This process had three steps that can be seen in Figure 5.2.

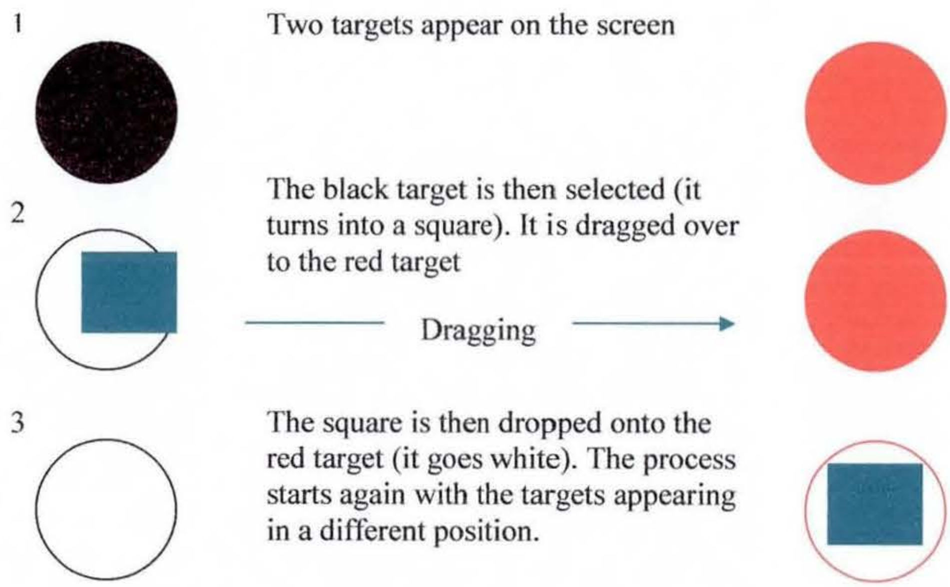


Figure 5.2: The Drag and Drop Process.

In the movement condition, the participant walked around six road cones in a figure of eight. The participant was timed and the error rate was recorded by the software used to create the task. After completing each part of the experiment the participant was asked to complete a NASA TLX assessment exercise.

5.3 Results

Three sets of data were collected for each participant in both the walking and standing sections of the experiment. The three sets of data were collected in both movement conditions. This was done while the participant was walking and while they were standing; this allowed the effect of movement upon the user to be examined. The three sets of data related to the time taken to complete the task in seconds, the amount of errors made during the completion of the task and the weighted workload of the participant calculated by NASA TLX. If there was a significant effect of the input device upon the participants' weighted workload, the results were further explored at the subscale level. The data relating to this experiment can be found in Appendix 3.

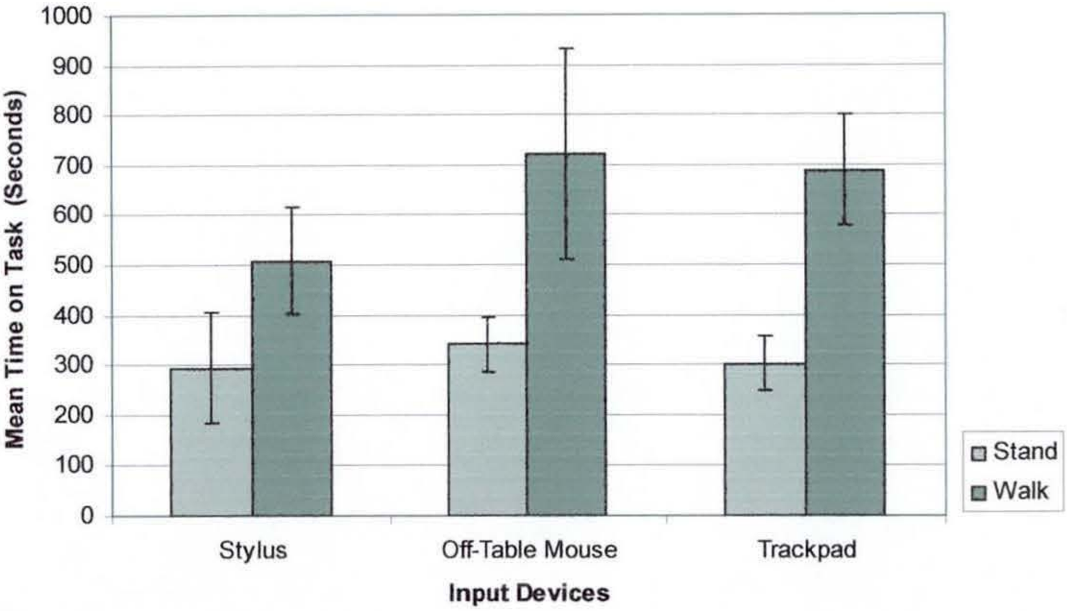
5.3.1 Results - Time

The mean time and their standard deviations for this study are presented in Table 5.1. Graph 5.1 illustrates these results in a graphical manner.

Descriptive Statistics - Time			
Movement	Device Type	Mean	Standard Dev
Standing	Stylus	295.056	110.912
	Off-Table	340.959	53.748
	Track-pad	301.577	54.280
	Total	312.530	77.807
Walking	Stylus	508.844	108.068
	Off-Table	720.920	209.568
	Track-pad	688.639	111.441
	Total	639.467	173.533

Table 5.1: Wearable dragging and dropping. (Time - seconds)
Mean input times with their associated standard deviations, for participant while standing and walking.

There was a significant main effect of the device type upon the time taken to complete the task ($F(2,27) = 74.23$, $p<0.05$). There was a highly significant main effect of movement upon the device type used to complete the task ($F(1,27) = 148.57$, $p<0.01$). There was significant interaction between device type and movement ($F(2,27) = 4.45$, $p< 0.05$). This further expanded upon within the discussion.



Graph 5.1: Wearable dragging and dropping. (Time - seconds).
This shows the mean input times for the 3 devices, both while walking and standing with standard deviation bars.

Tukey HSD *post-hoc* tests provided the following results, derived from the pairwise comparisons. There was a significant difference between the stylus and off-table mouse, p = 0.015, but there was no significance between the stylus and track-pad with p = 0.110, or between the off-table mouse and track-pad, p = 1.00.

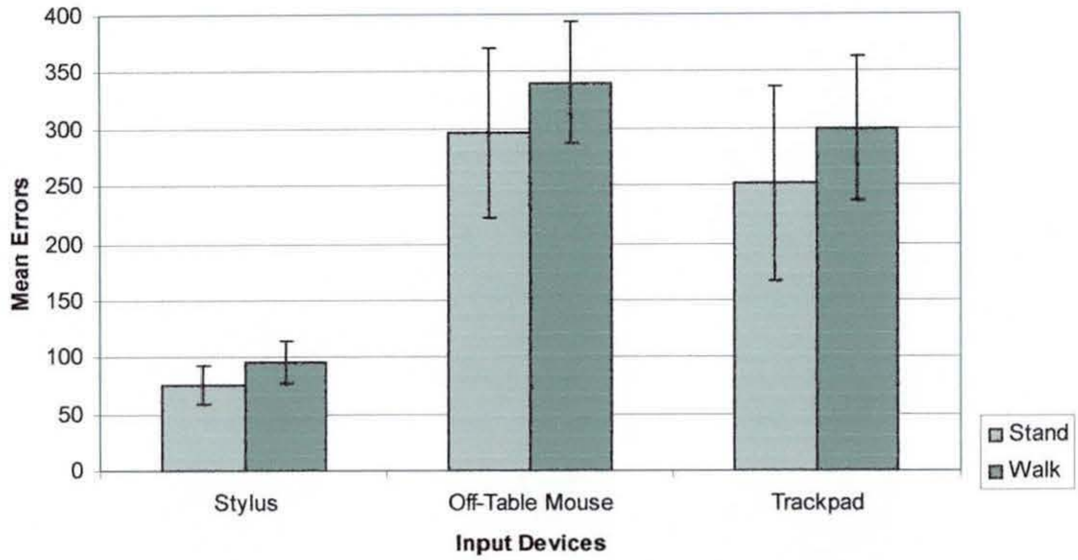
5.3.2 Results - Error

The means and standard deviations for this study are shown in Table 5.2, Graph 5.2 shows these results.

Descriptive Statistics - Errors			
Movement	Device Type	Mean	Standard Dev
Standing	Stylus	76.2	17.041
	Off-Table	295.8	74.118
	Track-pad	251.9	84.856
	Total	207.97	115.509
Walking	Stylus	96.3	18.625
	Off-Table	340.4	53.610
	Track-pad	298.9	63.320
	Total	245.2	118.358

Table 5.2: Wearable dragging and dropping. (Errors made on task). Mean number of errors made, with associated standard deviations

There was a highly significant main effect of the device type upon the errors made while completing the task ($F(2,27) = 751.33$, $p<0.01$). There was a significant main effect of movement upon the device type used to complete the task ($F(1,27) = 725.79$, $p<0.01$). There was no significant interaction between device type and movement ($F(2,27) = 1.37$, $p<0.5$).



Graph 5.2: Wearable dragging and dropping. (Errors made on task).
The graph shows the mean errors made for the 3 devices, both while walking and standing with standard deviation bars.

Tukey HSD *post-hoc* pairwise comparisons provided the following results: there was a highly significant difference between the stylus and off-table mouse, $p<0.01$, and between the stylus and track-pad, $p<0.01$, but not between the track-pad and off-table mouse, $p<0.5$.

5.3.3 Results Workload

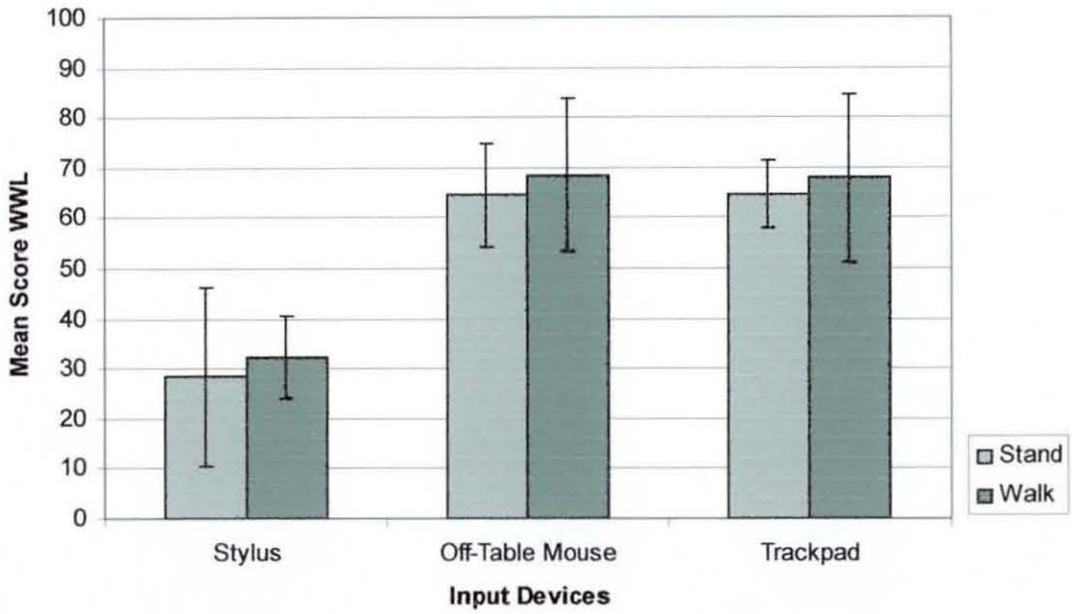
The table below (Table 5.3) shows the mean average workload (WWL) scores from the NASA TLX assessment. Each participant carried out the assessment after each of the two stages of the experiment. The results displayed in Table 5.3 demonstrated that the stylus had the lowest workload while used standing and walking. The track-pad marginally had the highest workload while standing, followed by the track-pad. In mobile movement condition the off-table mouse had the highest workload. By examining Graph 5.3 it can be seen how close the results for the off-table mouse and track-pad were.

.

Descriptive Statistics - Workload			
Movement	Device Type	Mean	Standard Dev
Standing	Stylus	28.4	18.00
	Off-Table	64.5	10.39
	Track-pad	64.7	6.91
	Total	52.5	21.21
Walking	Stylus	32.4	8.19
	Off-Table	68.6	15.24
	Track-pad	67.9	16.67
	Total	56.3	21.79

Table 5.3: Wearable dragging and dropping. (Mean – Workload)
Average workload scores and standard deviations

A mixed factorial ANOVA showed that the only significance that occurred was a highly significant main effect of device ($F(2,27) = 46.878$, $p<0.01$. No significant results could be reported for movement or for any interaction that may have occurred between device and movement.



Graph 5.3: Wearable Dragging and dropping. (Mean - Workload).
The graph shows the mean workload scores for the 3 devices, while both walking and standing. They are shown here with their associated standard deviation error bars.

Tukey HSD *post-hoc* found that the significant effect of device occurred between the stylus and off-table mouse, $p<0.01$ (mean difference 36.150), and also between the stylus and track-pad, $p<0.01$ (mean difference 35.900).

After a highly significant effect of device was found when comparing the overall weighted workload, further analyses of the six NASA TLX subscales was carried out. The paired tables below show the overall main effects, the Tukey HSD *post-hoc* results and the mean differences between each input device. These were put into a tabular format to enable an easier understanding of the results

Table 5.4: Wearable dragging and dropping. (Workload Subscales 1.)
The table shows the significance of device and movement for the NASA TLX subscales contained in the table. No significant interaction was found between device and movement.

NASA-TLX / Main Effects	Mental Demands	Physical Demands	Temporal Demands
Device	F(2,27)=9.472,p<0.01	F(2,27)=25.432,p<0.01	F(2,27)=33.809,p<0.01
Movement	No Significance	No significance	No significance

Table 5.5: Wearable dragging and dropping. (Workload Subscales 1.1)

This table is paired with table 5.4 and shows the Tukey HSD *post-hoc* results from the pairwise comparisons.

NASA-TLX / *Post-hoc*	Mental Demands	Physical Demands	Temporal Demands
Stylus – Mouse	p<0.01 Mean Diff 27.5	p<0.01 Mean Diff 41.50	p<0.01 Mean Diff 36.50
Stylus –Track-pad	p<0.01 Mean Diff 23.5	p<0.01 Mean Diff 31.50	p<0.01 Mean Diff 39.50
Track-pad – Mouse	No Significance	No Significance	No Significance

The tables also show the results from the Tukey HSD *post-hoc* comparisons and the mean differences that occurred between the pairwise comparisons. It was found that the device with the lowest mean average in each subscale was the stylus. Every subscale analysis showed a main effect of device. Only the subscales of frustration and effort displayed a main effect of movement. No significant interaction between device and movement can be reported.

Table 5.6: Wearable dragging and dropping. (Workload Subscales 2.)

This table shows the significance of device and movement for the NASA TLX subscales contained in the table. No significant interaction was found between device and movement.

NASA-TLX / Main Effects	Effort	Performance	Frustration
Device	F(2,27)=37.245,p<0.01	F(2,27)=18.303, p<0.01	F(2,27)=31.738,p<0.01
Movement	F(1,27)=0.41,p<0.05	No significance	F(1,27)=5.360,p<0.05

Table 5.7: Wearable dragging and dropping. (Workload Subscales 2.1)

This table is paired with table 5.6 and shows the Tukey HSD *post-hoc* results from the pairwise comparisons.

NASA-TLX / *Post-hoc*	Effort	Performance	Frustration
Stylus – Mouse	p<0.01 Mean Diff 38.75	p<0.01 Mean Diff 34.250	p<0.01 Mean Diff 29.450
Stylus –Track-pad	p<0.01 Mean Diff 35.00	p<0.01 Mean Diff 34.500	p<0.01 Mean Diff 32.0
Track-pad – Mouse	No Significance	No Significance	No Significance

The *post-hoc* analyses found only significant differences between the stylus and other devices. Overall the largest difference occurred between the off-table mouse and stylus. No significant results could be reported for the off-table – track-pad comparison.

5.3.4 Discussion

In examining the results it is clear that there was a significant difference between the devices relating to the input time, and therefore we can reject the first null hypothesis (***h1***). The fastest input device while the participants were standing was the stylus at 295 seconds, followed by the track-pad at 301 seconds, with the off-table mouse being the slowest input device, resulting in an input time of 340 seconds. From this we can say that the stylus was the fastest device to use for a dragging and dropping type task while stationary, although the stylus (while standing) had the largest standard deviation at 110 seconds (this was due to one participant taking 594 seconds to complete the task, nearly twice the mean time on task), as compared to the off-table mouse at 53 seconds and the track-pad at 54 seconds. This suggested that there was a larger fluctuation of input times for the stylus while standing and the off-table mouse while walking, when compared to the other devices. When examining the input times while the participant is walking we can see from Graph 5.1 that the stylus remained the fastest input device, with an input speed of 508 seconds, followed by track-pad at 688 seconds and the off-table mouse 720 seconds. These initial findings showed that the stylus was the fastest in both movement conditions, with only a 213 second difference, as opposed to the track pad at 387 and the off-table mouse at 380 seconds. So, although the off-table mouse was the slowest device overall, it had the second-smallest difference between the movement conditions.

The stylus was the fastest device to use. This may have been because people found this device easy to use and did not continually have to focus on the task, whereas the off-table mouse and track-pad have a level of novelty (newness); for most users it may have been the first time they had used such devices. Many of the participants will understand the concepts behind using the stylus almost immediately, as it relates to other systems in the world, such as a pen/pencil and paper. Participants using the other two systems may have found the input systems harder to use as they were not used to them. In examining the workload, it is apparent that both the off-table mouse and trackpad had higher workloads than the stylus.

A significant interaction between the input device and movement was reported. This suggests that the stylus users were able to complete the task much faster than the other device users. In examining Graph 5.1, we can clearly see that the stylus had the lowest input time difference between standing and walking, followed by the track pad and off-table mouse. This suggests some level of device-dependent effect on the time to complete the task in response to the movement conditions. If the results had been similar there would have been no significant interaction.

The device which had the least amount of errors in both conditions was the stylus, followed by the track-pad and off-table mouse, as seen in Table 5.2. The stylus had the smallest difference in errors between the two movement conditions; a difference of only 20 errors. The off-table mouse had an error difference of 45, while the track-pad had the highest difference with 47 errors. This suggests that while either stationary or mobile the stylus is the best input device to use, both in terms of time, as discussed earlier, and errors made, as discussed here. The results reported a significant difference between the devices, and therefore we can reject the second null hypothesis (***h2***). On further analysis of the *post-hoc* results it can be seen that a significant result was only found between the stylus and track-pad, and the stylus and off-table mouse, with the largest difference overall occurring between the stylus and off-table mouse. Interestingly, in this study there were more errors made while the participant was walking. In the previous experiment the participants made fewer errors while walking. This suggests that the pointing devices may be harder to use for dragging and dropping while mobile. It would normally be expected that the task would be harder to accomplish, as the participant would be dual-tasking (as described in section 2.10.2**)**. There would normally be an 'impairment' of performance; this could be in terms of: time, errors and/or workload.

The off-table mouse had the highest error rate and was also the slowest device to use. This may be explained by the following quote relating to the use of a trackball from another dragging and dropping based experiment (MacKenzie *Et al.* 1991), "The interaction between muscle and limb groups was considerable. This was not the case with the mouse or tablet". This is interesting, because, although called an off-table mouse this input device is effectively a small trackball which can be used while away from the desktop, which may be why it is so slow to use and makes the most errors.

The third null hypothesis (***h3***) can also be rejected as it was found that there was a significant effect on both input time and the errors made while doing the task. The overall differences were a mean rise of 38 errors in the walking condition, and a rise of 326.93 seconds. The timed results when related to the error-based results seem to suggest a proportional link between the time taken to complete the task and the errors made while completing the task. This may be explained by the fact that each time an error is made it has to be corrected by the users, and therefore it took extra time to complete the task.

The device with the lowest workload in both conditions was the stylus. It had a significantly lower workload than both the off-table mouse and the track-pad. The null hypothesis (***h4***) can therefore be rejected. No significant main effect of movement was reported so we are not able to reject the fifth null hypothesis (***h5***). This suggests that for a generic target selection task in this case there is very little difference on the participants' workload in relation to them standing still or walking while doing the task. If we examine the overall workload, the off-table mouse and track-pad barely differ: 0.2 while standing and 0.7 while walking. From these results we can surmise that the participants did not find the track-pad any more difficult to use than the mouse, and vice versa. The difference occurs between the types of device used and is not dependent on the way they were moving when they completed the task. It must however be noted that some participants did seem to walk slower than 'average' when completing the mobile part of the experiment. The experiment could have used a treadmill instead, but this was considered dangerous; if a participant slowed too much they could easily fall. Using a treadmill could have allowed the recording of the distance the participant walked and how far they walked, while using their input device. This could have then been correlated against their mean walking speed (when not using the input device) to see which device most affected the user's performance, in terms of their walking pace. There additionally could have been a way of measuring the participants' walking speed to examine how much slower they went when inputting. It might have been more revealing if the heart rate had been monitored to investigate stress levels. Like the other experiments this could be extended to focus on the most appropriate input device for certain target sizes to be dropped. It could also be further extended to accommodate real world settings and eventually task based scenarios.

5.4 Conclusion

This experiment found that the most appropriate device for dragging and dropping from the three evaluated was the stylus. This was in terms of the assessment criteria; these were the time on task, errors made while completing the task and the weighted workload of the participant. The stylus was significantly faster while stationary and moving, had significantly fewer errors while stationary and moving, and also had the lowest workload. One revealing finding of this study was the lack of a significant main effect of movement relating to the participants' workloads. This suggests that the act of being mobile for this task under these experimental conditions did not affect the users' workload. The worst device of the three was the off-table mouse, closely followed by the track-pad.

This study found that the performance of the participants was worse in the walking levels of the experiment. This contradicts the findings of the previous experiment, where the participants using the off-table mouse and stylus made fewer errors on the walking level. This may suggest that dragging and dropping as an interaction style is harder than target selecting, when the participant is mobile (walking). This itself brings into doubt whether the dragging and dropping is an appropriate style of interaction for mobile users. It can be hypothesised that the reduction in performance while walking was caused by the participants having to dual-task (Pirhonen, Brewster and Holguin 2002; Brewster, Lumsden, Bell, Hall, and Tasker 2003; Lumsden and Brewster 2003). There may also be a relation between the participants having to move objects on the screen as well as having to move themselves, because they are doing tasks that require similar types of attention simultaneously (this could be akin to patting one's head and rubbing one's tummy). This could relate to work carried out that found dual-task scenarios requiring the user to use the same modality for both tasks (such as speaking and listening simultaneously) are very difficult for the user to divide their attention between (*see* Allport, Antonis and Reynolds 1972). This often causes a loss in performance.

In this experiment participants walked around cones placed in their path, and as the results suggest, their performance while mobile was significantly worse than while stationary. One area that would be interesting to evaluate is the speed at which the participants move, walking and running. It could be hypothesised from the results of this experiment that a user running and trying to drag and drop may perform even worse than one who is walking.

This experiment could be further expanded in several ways. Firstly, it would be interesting to assess the users in different environments to see if that had an impact upon the users' workload. Although this study has evaluated the devices in terms of a generic pointer based task as specified by Mackenzie *Et al.* (1992), a next step forward would be to implement the knowledge gained into an interface for mobile users. This interface experiment could be an adaptive interface for wearable users that adapted elements such as: buttons, radio buttons and menu structures to relate to the movement conditions and environment. This experiment could easily be adapted to evaluate the effect of sonically enhanced drag and drop on mobile wearable users. After completing the experiment it was realised that important information about the walking level of the experiment (in regard to dual-tasking and the participant's performance) could have been gained if the participants walking speed and distance had been recorded as previously discussed.

Chapter 6: An Experiment to Evaluate Three Wearable Computing Input Devices for Trajectory-Based Interaction

6.1 Introduction

In this experiment, 27 different combinations of 'tunnels' were used for the participant to steer through as we found that this made the task more complicated. The original experiment used two tunnel amplitudes (A= 250, and 1000 pixels) and three path widths (W= 35, 45, and 70 pixels) see Figure 6.1. Another factor was introduced into the experiment: the angle at which the tunnels were presented to the participants. Tunnels were presented to the participant at three angles, these were: slanting at 45 degrees, horizontally at 0 degrees and vertically at 90 degrees. This was done in an attempt to mimic the way that a user carries out trajectory-based tasks such as navigating through nested menu structures. The 45 degree tunnels were used to represent the task of steering through pie menus. This was done to relate to existing wearable computer interfaces (Brewster *Et al.* 2003; Butz *Et al.* 1999; Lumsden *Et al.* 2003). The tunnel amplitudes (lengths) and widths (heights) also had to be altered. This was done to accommodate the use of a smaller screen, and is further discussed in the methodology section below.

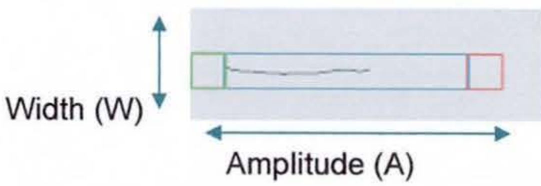


Figure 6.1, Width and amplitude of the tunnels.

6.1.1 Rationale

In Chapter 5 we found that the stylus was the most effective pointing device for target selection and in Chapter 4 it was found to be the best device for wearable users to use for dragging and dropping tasks. These experiments proved that the stylus was a better pointing device for selecting on-screen targets and for dragging and dropping tasks than the off-table mouse and track-pad. Pointing devices are not only used for target acquisition, but can also be used for dragging and dropping and trajectory-based interactions. As Acott (1997) States "Trajectory-based interactions, such as navigating through nested-menus, drawing curves, and moving in 3D worlds, are becoming common tasks in modern computer interfaces." This is further expanded upon in Section 2.7.2. As many wearable interfaces require the user to interact with trajectory based interfaces (Schmidt *Et al.* 2000, Blasko and Fiener 2002), it is important to evaluate the pointing devices used so far in regard to this.

6.1.2 Hypotheses

The null hypotheses are as follows:

h1. The input device type does not affect the time taken to complete the steering task.

h2. The input device type does not have an effect on the error rate of the steering task.

h3. The movement condition of the participant standing and walking has no effect on the time taken to complete the task or error rate.

h4. The input device type does not affect the subjective workload of the participant.

h5. The movement condition does not affect the subjective workload of the participant.

6.2 Method

6.2.1 Design

This experiment used the same design and apparatus as used in the previous two chapters. A 2 x 3 mixed factor design was used for the design of this experiment. This combined a between and within groups design. To stop carry-over effects the steering tasks used two different randomisation patterns (as used in Chapters four and five).

Each of the thirty participants steered through 540 tunnels in both movement conditions. The order in which the participants took part in the experiments was balanced using a Latin square method. The length, amplitude and angle of the tunnels were as follows: the lengths were 64, 128 and 256 (16mm 32mm and 64mm), the amplitudes were 12, 24 and 36 (3mm, 6mm and 9mm). The angles at which the tunnels were displayed were 0, 45 and 90 degrees. There were 27 different tunnel combinations that the participant completed. An example of what each participant saw is shown below in Figure 6.1.

6.2.2 Procedure

The participant was first asked to put on the vest containing the computer. This was adjusted until the participant felt comfortable. The participant was either assigned an off-table mouse or stylus, depending on the input group that they had been assigned to. The participant was then shown how to use their device, on an example of one of the steering tasks on a monitor. This was done a minimal amount of times to reduce carry-over effects. Each participant did the experiment while standing and walking. The order of this was done in relation to the counterbalancing. The participant steered though 540 tunnels in each condition. To complete each tunnel, the participant positioned the cursor in the green start box and had to steer through the tunnel until the cursor appeared in the red finish box; an example of the tunnels can be seen in Figure 6.1. In the movement condition, the participant walked around six road cones in a figure of eight. The participant was timed using a stopwatch and the error rate was recorded by the software used to create the steering test. Errors were steering outside of the 'tunnel' or stopping the task (lifting the stylus off the screen, the participant lifting their finger off the trackpad or releasing the trigger of the mouse). After completing each part of the experiment the participant was asked to complete a NASA TLX assessment exercise.

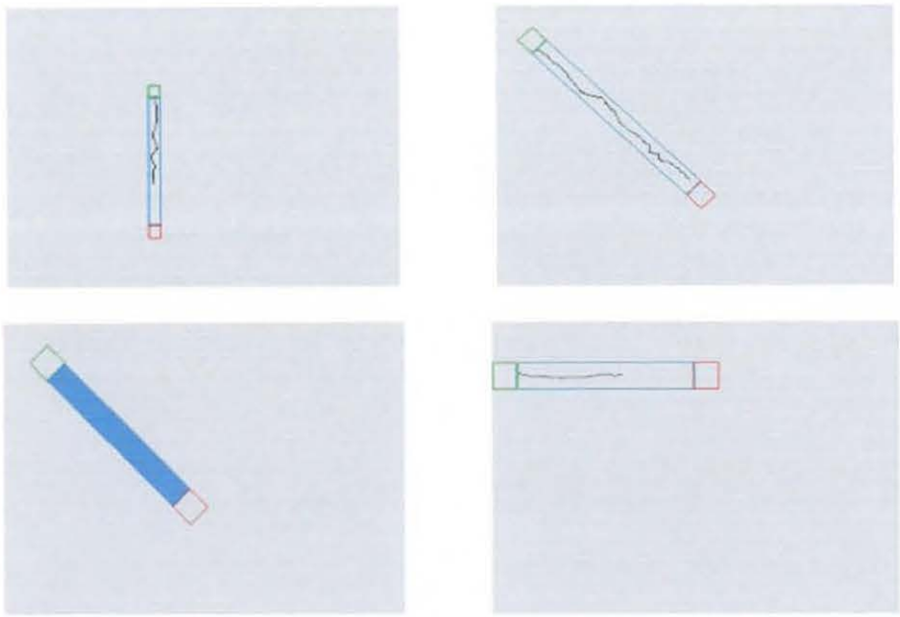


Figure 6.1: Wearable trajectory-based input (Experiment Screenshots).

Screen shots from the experiment, participants started at the green end of the tunnel and had to navigate their way through to the red end of the tunnel. The tunnels are at 0, 45, 90 degrees.

6.3 Results

Three sets of data were collected for each participant in the walking and standing sections of the experiment. Three sets of data were collected in both movement conditions. This was done while the participant was walking and while they were standing; this allowed the effect of movement upon the user to be examined. The three sets of data related to the time taken to complete the task in seconds, the amount of errors made during the completion of the task and the weighted workload of the participant calculated by NASA TLX. If there was a significant effect of the input device upon the participants' weighted workload, the results were further explored at the subscale level. A mixed factorial ANOVA was carried out upon each data set. Full results can be found in Appendix 4.

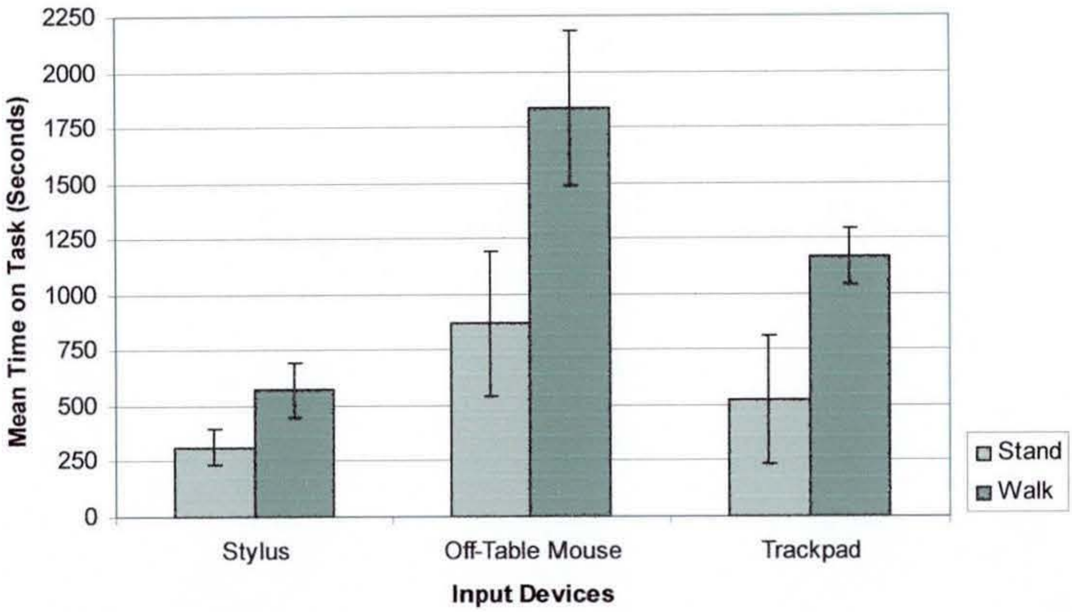
6.3.1 Results - Time

The means and standard deviations for this study are displayed in Table 8.1, Graph 8.1. illustrates the results.

Descriptive Statistics - Time			
Movement	Device Type	Mean	Standard Dev
Standing	Stylus	316.033	81.56315
	Off-Table	867.336	328.5203
	Track-pad	523.344	284.5992
	Total	568.9043	337.8933
Walking	Stylus	571.541	119.9539
	Off-Table	1836.563	349.6757
	Track-pad	1165.986	125.9976
	Total	1191.363	568.8434

Table 6.1: Wearable trajectory-based input. (Time - seconds).
Mean input times (seconds) and standard deviations for the trajectory-based task, while standing and walking.

There was a significant main effect of the device type upon the time taken to complete the task ($F(2,27) = 86.839$, $p<0.05$). Similarly, a significant main effect of movement upon the device type used to complete the task was found ($F(1,27) = 85.19$, $p<0.05$). There was significant interaction between device type and movement ($F(2,27) = 9.356$, $p<0.05$). This is further expanded upon within the discussion.



Graph 6.1: Wearable trajectory-based input. (Time - seconds).
The graph shows the mean input times for the 3 devices, both while walking and standing with standard deviation bars.

A Tukey HSD *post-hoc* test was carried out on the data. The pairwise comparisons found highly significant differences $p<0.01$ between all of the devices. If we examine the overall means, it can be seen that the greatest difference occurs between the stylus and off-table mouse at 1816.325 seconds, while the smallest difference occurred between the stylus and track-pad at 801.765 seconds, and the difference between the track-pad and off-table mouse was 1014.569. This is graphically illustrated in Graph 6.1.

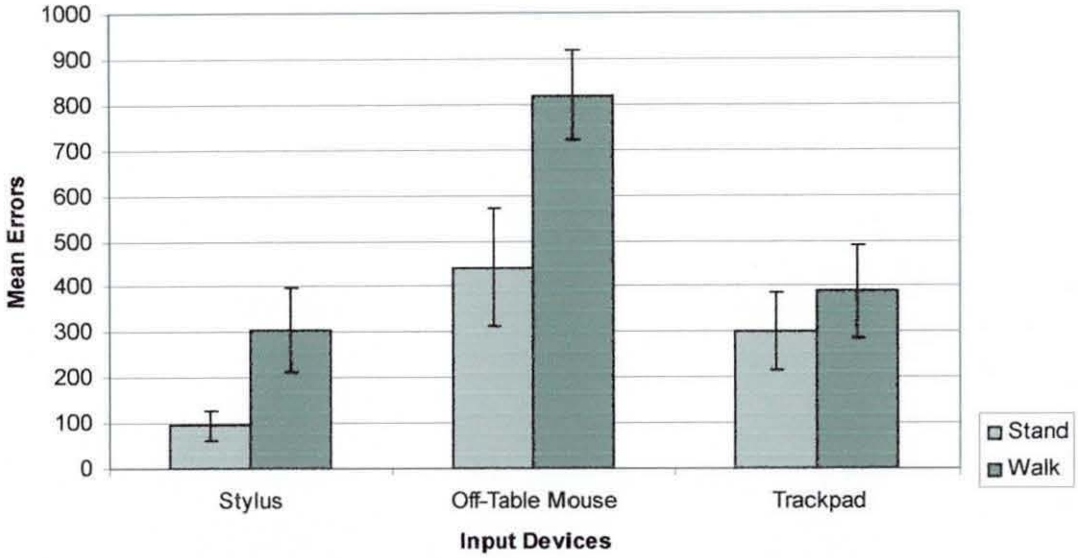
6.3.2 Results - Error

The means and standard deviations for this study relating to error are shown in Table 6.2. Graph 6.2 illustrates these results in a more intuitive manner.

Descriptive Statistics - Errors			
Movement	Device Type	Mean	Standard Dev
Standing	Stylus	94.90	34.35582
	Off-Table	441.00	128.8091
	Track-pad	302.00	85.24084
	Total	279.3	169.38
Walking	Stylus	303.90	93.17063
	Off-Table	819.40	97.66064
	Track-pad	388.20	103.8501
	Total	503.83	248.464

Table 6.2: Wearable trajectory-based input. (Mean – Errors made on task)
Mean errors and standard deviations for the trajectory-based task, while standing and walking.

A mixed factor ANOVA found a significant main effect of the device type ($F(2,27) = 74.23$, $p<0.05$). A significant main effect of movement upon the device type used to complete the task was revealed ($F(1,27) = 148.95$, $p<0.05$). There was significant interaction between device type and movement ($F(2,27) = 21.20$, $p<0.05$). The interaction is further discussed within the discussion section of the study.



Graph 6.2: Wearable trajectory-based input. (Mean - Errors made on task).
This shows the mean errors made for the 3 devices, both walking and standing, with standard deviation bars.

After a Tukey HSD *post-hoc* test was carried out, the pairwise comparison found highly significant differences all at $p<0.01$ between the devices. The combined overall difference between stylus and off-table mouse was 861.6 errors, between the off-table mouse and track-pad it was 570.2 errors, and between the stylus and track-pad there was a difference of 291.4 errors.

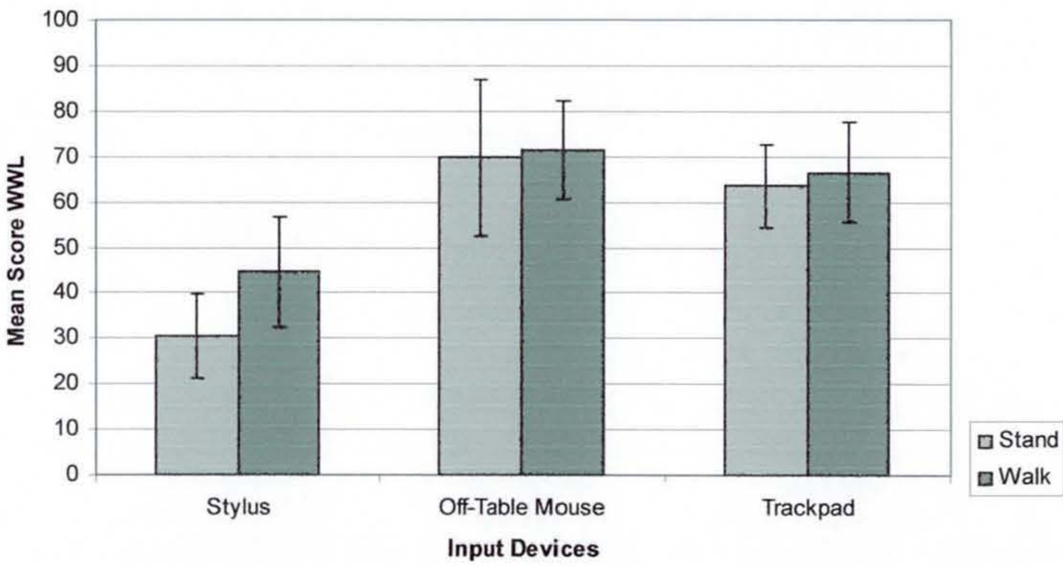
6.3.3 Results - Workload

After each part of the experiment the participant completed a NASA TLX (taskload assessment) exercise. This was done in order to find the participant's workload in both movement conditions while using one of the three input devices. The results for this are presented in Table 6.3 which shows the mean average scores for each input device grouping and the standard deviation. Graph 6.3 graphically represents this data. The results reported are further expanded upon in the discussion section of the report.

Descriptive Statistics - Workload			
Movement	Device Type	Mean	Standard Dev
Standing	Stylus	30.5	9.395
	Off-Table	69.7	17.204
	Track-pad	63.6	9.155
	Total	54.6	21.262
Walking	Stylus	44.6	12.020
	Off-Table	71.4	10.782
	Track-pad	66.6	11.177
	Total	60.8	16.141

Table 6.3: Wearable trajectory-based input. (Mean – Workload).
Mean workload scores and standard deviations for the trajectory-based task, while standing and walking.

A mixed factorial ANOVA revealed a highly significant effect of input device ($F(2,27) = 32.899$, $p<0.01$). The workload was rated highest by the off-table mouse users, while standing and walking. The lowest scores were given by the stylus users. There was a significant effect of movement, ($F(1,27) = 6.243$, $p<0.05$). The largest difference between the movement conditions was with the stylus; this was a 14 point difference. There was no significant interaction between input device and movement ($F(2,27) = 4.461$, $p< 0.5$).



Graph 6.3: Wearable trajectory-based input. (Mean - Workload).
The graph shows the mean workload scores for the 3 devices, while both walking and standing. They are shown here with their associated standard deviation error bars.

Tukey HSD *post-hoc* comparisons found the significant differences to lie between the stylus and off-table; $p<0.01$, and between the stylus and track-pad, $p<0.01$. No significance could be reported between the off-table mouse and track-pad, $p=0.667$. These results can be further interpreted from Graph 6.3.

To further understand the workload result, a series of ANOVAs were carried out on each of the six subscales that are used to make up the final weighted workload.

Table 6.4: Wearable trajectory-based interaction. (Workload Subscales 1.)
The table shows the significance of device and movement for the NASA TLX subscales contained in the table. No significant interaction was found between device and movement.

NASA-TLX / Main Effects	Mental Demands	Physical Demands	Temporal Demands
Device	$F(2,27)=6.763, p<0.01$	$F(2,27)=18.342, p<0.001$	$F(2,27)=19.045, p<0.001$
Movement	No Significance	No significance	No significance

Table 6.5: Wearable trajectory-based interaction. (Workload Subscales 1.1)

This table is paired with table 6.4 and shows the Tukey HSD *post-hoc* results from the pairwise comparisons.

NASA-TLX / *Post-hoc*	Mental Demands	Physical Demands	Temporal Demands
Stylus – Mouse	$p<0.001$ Mean Diff 24.00	$p<0.001$ Mean Diff 31.50	$p<0.001$ Mean Diff 36.75
Stylus –Track-pad	$p<0.05$ Mean Diff 21.00	$p<0.001$ Mean Diff 30.75	$p<0.001$ Mean Diff 33.75
Track-pad – Mouse	No Significance	No Significance	No Significance

The tables also show the results from the Tukey HSD *post-hoc* comparisons and the mean differences that occurred between the pairwise comparisons. It was found that the device with the lowest mean average in each subscale was the stylus. Every subscale analysis showed a main effect of device. Only the subscales of performance displayed a main effect of movement. No significant interaction between device and movement can be reported.

Table 6.6: Wearable trajectory-based interaction. (Workload Subscales 2.)

This table shows the significance of device and movement for the NASA TLX subscales contained in the table. No significant interaction was found between device and movement.

NASA-TLX / Main Effects	Effort	Performance	Frustration
Device	F(2,27)=19.464,p<0.001	F(2,27)=24.724, p<0.001	F(2,27)=12.681,p<0.001
Movement	No Significance	F(1,27)=4.537,p<0.05	No Significance

Table 6.7: Wearable trajectory-based interaction. (Workload Subscales 2.1)

This table is paired with Table 6.6 and shows the Tukey HSD *post-hoc* results from the pairwise comparisons.

NASA-TLX / *Post-hoc*	Effort	Performance	Frustration
Stylus – Mouse	p<0.01 Mean Diff 33.00	p<0.01 Mean Diff 34.00	p<0.001 Mean Diff 35.250
Stylus –Track-pad	p<0.01 Mean Diff 30.75	p<0.01 Mean Diff 25.111	p<0.01 Mean Diff 22.250
Track-pad – Mouse	No Significance	No Significance	No Significance

The *post-hoc* analyses found only significant differences between the stylus and other devices. Overall the largest difference occurred between the off-table mouse and stylus. No significant results could be reported for the off-table – track-pad comparison.

6.3.4 Discussion

Overall, participants using the stylus performed better than those using the off-table mouse and trackpad. As reported in the results section, the time taken to complete the task and the errors made almost doubled when the participants completed the task while walking. We can hypothesise that this was due to the participants having to dual-task. This hypothesis may be supported by the fact that during the mobile part of the experiment the participants had to be reminded to keep moving, as they tended to slow down while doing the task. It could have been possible to put the participants upon a walking/running machine, but if the participant slowed down too much they could fall off. This work could be extended by timing the participants' walking speed and correlating the results with that speed to see if there is a significant relationship between device usability and walking speeds (Pirhonen, Brewster and Holguin 2002; Brewster, Lumsden, Bell, Hall, and Tasker 2003; Lumsden and Brewster 2003).

One thing that could be questioned is the appropriateness of trajectory-based interfaces for mobile-wearable computer users. The results show that the participants' performance was worse while walking; participants had to be reminded to keep a constant pace. We can theorise that trajectory-based interaction of this type would be extremely difficult to use in situations where the user has to keep up a constant pace, or faster than average pace e.g. running. In a lab based situation the participant knew how the obstacles were laid out, but running in an unpredictable environment, such as a busy street or on in locations with uneven surfaces, the participant would not be able to focus all of their attention onto the screen without slowing; this may be a case where audio-centric interfaces may play an important role in mobile interaction (Brewster and Murray 2000). This is because the user does not have to focus their visual attention in more than one place (Brewster, Lumsden, Bell, Hall, and Tasker 2003)

This study found that there was significant difference between the three input devices, in terms of time and the errors made while completing the task. It was discovered that the device that was the quickest to use and had the lowest error rate was the stylus; the slowest device was the off-table mouse, which also had the highest error rate as described in the *post-hoc* results section. These findings meant that the null hypotheses *h1* and *h2* could be rejected.

A further analysis of the results found that the device with the smallest time difference between doing the task stationary and while mobile was the stylus. This suggests that for mobile tasks that consist of trajectory based interaction, and where interfaces such as those outlined in the introduction are used, the stylus is the most appropriate input device. The worst device was the off-table mouse. The interaction suggests that the type of input device and the movement condition interact in their effect on the time taken to complete the task. By looking at Graph 6.1, we can surmise that it takes more time to complete the task when the participant is walking than when they are stationary, but as we have seen, this difference was not uniform. The largest difference occurred while using the off-table mouse, a difference of 969.22 seconds; the second slowest device was the track-pad at 642.64, while the stylus had the smallest difference at 255.5 seconds. Interestingly we can see that the input times for the stylus while that participant was walking, in Table 6.1, were very similar to the input time for the track-pad while the participant is stationary.

The third hypothesis (***h3***) can now be rejected, as there was a significant effect of movement upon both the time taken to complete the task and the errors made while completing the task. In effect movement slows down the user and it may be suggested that this makes the user's task more difficult, causing more errors. It was found that the stylus and track-pad had the smallest difference both in terms of errors made and time; this may because they use a very similar form of physical manipulation. When using the track-pad, one uses one's finger to point and move the cursor and when using the stylus and touch screen the user uses a stylus, which becomes an extension of the hand to point. The findings in this experiment also found that like the previous experiment participants made more errors while walking, this was in contrast to the target selection experiment. The findings for this experiment may relate to the difficulty of the task.

In this discussion it is important to suggest how the experiment might be modified so that this work may be extended and perhaps modified, thereby leading into or expanding other studies. This study was able to perfectly examine the hypothesis, but there are areas that could be investigated further. One area that could be further explored is the long term use of wearable input-devices. This could be done by a participant who could do a whole series of trials over weeks to see if there is any improvement, and then we could also examine the learning rates of the participant. These results could be compared with real world users to explore the differences between everyday users and the participants in the experiment. Another factor that needs to be further understood is the interaction between the user and their environment, although if any data from these experiments were to be statistically analysed, each user might have to have had a set of uniform experiences, so as not to bias the experiment in any way. One interesting off-shoot of this analysis would be to examine each device's ability in relation to certain tunnel sizes, thereby judging the device in terms of its ability to be used for very accurate trajectory based tasks. Although this study did use different tunnel sizes and lengths, the errors made and time taken to complete a certain tunnel size were not recorded.

From the results and Graph 6.3, it is clearly illustrated that the stylus had the lowest workload (WWL). Therefore, we are able to reject the forth hypothesis (***h4***), because there was a significant main effect of input device upon the participants' subjective workloads. We are also able to reject the fifth hypothesis (***h5***), because a significant main effect of movement was reported, which means that the act of being mobile has an effect on the participants' subjective workload: it raises it. This suggests that mentally the participant finds the device more subjectively difficult to use while walking. As a significant difference was found between the input devices in relation to their workload, a further analysis of the subscales that make up the workload was carried out. This in depth analysis found that there was a significant effect of device within each of the scales examined. The analysis also found that the stylus gave the lowest result in each of the scales. The *post-hoc* results confirmed this by only finding significance between the stylus and two other input devices in every subscale analysis. Only the performance scale reported a significant effect of movement.

When applying these findings to wearable systems and users, it is important to take into consideration the output modality/technologies that the user is wearing. Out of the 3 devices that we evaluated only 2 could be used with an augmented reality system, as the stylus relies on a touch screen based system. This may be able to be replaced by some form of body-worn graphics tablet. However, what we can derive from these results is that overall for both standing and walking the errors made and time it took to complete the task were the lowest for the stylus, when used in conjunction with a fold-down touch screen. When a wearable user is mobile, physical movements are more erratic. A user may be walking, running or climbing; so controlling an input device such a stylus will prove more difficult while mobile than being at a desktop. There are also other environmental factors that can affect the type of device used; trackball devices (mice) are not good in dirty or dusty environments as their mechanisms can quickly get clogged and not work properly. Gloves in colder environments may also impede users' ability to use the device. This may certainly be so in the case of track-pads as they have to be used bare handed and a stylus would be difficult to hold.

There are further ways that future evaluators could expand this experiment. One would be to examine the devices with long term users of wearable technology. The devices could also be evaluated in different environmental settings and with different user groups such as the elderly. Future evaluations would aim to use more participants and to further explore the user's subjective opinions of input device use in a range of environments and for each user to use a range of input devices, instead of a single input device.

6.4 Conclusion

The experiment detailed in this chapter found that there were significant main effects of device in relation to: the time on task; the errors made while completing the task; and the participants' subjective workloads. Out of the three devices tested, the stylus had the lowest error rate, fastest input time and had the lowest subjective workload, followed by the track-pad and off-table mouse. This experiment has importantly shown that when a wearable user is mobile, it has a negative effect on their ability to use input devices to perform trajectory-based tasks, such as navigating through menus. These findings can also help software developers, wearable computer users and wearable hardware designers to develop strategies to cope with the difficulties that mobility can case the wearable user.

The findings in this chapter also lead to the questioning of trajectory-based interaction as an appropriate interaction style for mobile wearable users. This was because participants completing the task while mobile performed worse in terms of the time taken to complete the task and the errors the participants made, than in the standing condition.

For generic trajectory based tasks the stylus should be the preferred choice of input device. In situations where it is not possible to use the stylus, the track-pad is an ideal secondary input device. The stylus is an ideal input device as its weight is minimal, it is economically viable, it is cable-free and its usage mimics other 'real world' devices that the user should be familiar with, such as a pen or pencil. The problems associated with using are that it may need to be attached to the body to prevent it being lost, it is difficult to use with head-mounted displays and it may be problematic for users with little manual dexterity, such as older people.

Chapter 7: Experimental Review

7.1 Introduction

This chapter presents and discusses the work carried out thus far within this thesis. It presents an overall review of the findings from the experiments. From the scientific evidence it identifies the most appropriate text input and pointing device for wearable computer users. It discusses the empirical and subjective evidence, and also examines the evaluation methods that were used to collect the data. The outline of the findings in this chapter finishes the research in an appropriate place to form the final conclusions and discussions put forth within the next chapter.

7.2 Wearable Text Input

There is strong evidence to suggest that the act of being mobile has a negative effect upon the wearable computer user's time to complete a text input task, the amount of errors they make and their cognitive workload. This means that the participant is mentally working harder to input text, it takes longer to input text and the user makes more errors when doing so. Secondly, the experimental data in Chapter three proved that there was a significant difference in terms of time, error and workload between the four text input systems in this trial. The results for this experiment can be seen in sections 3.4.1 to 3.4.3 and pertain only to the use of Roman alphabet script. This experiment required the participants to walk around road cones in a figure of eight motion while they inputted text, but it may be the case that different tasks and environments may impact differently upon the user's input time, error rate and cognitive workload. However, at this stage, this is hypothetical.

The speech recognition system had the lowest input time (33 seconds) while the user was stationary and the fastest while the user was in motion (35 seconds). The speech recognition system also had the lowest difference between the stationary and mobile conditions: 33 seconds standing and 35 while the user was mobile. There was no significant difference between the handwriting recognition system, wrist keyboard and the on-screen keyboard system. The significance lay between the speech recognition and other devices only. The worst input device in both movement conditions was the on-screen keyboard, followed by the wrist keyboard.

Another measurement of the usability of the wearable text input systems was the amount of errors that the user made while completing the experimental tasks. The system with the lowest error rate was the handwriting recognition system as we can see from the results in section 6.3 of the thesis. The errors produced by users inputting text with this system did not differ while the user was stationary and mobile. Previously we saw that the on-screen keyboard had the lowest time difference between the two movement conditions; but this was not mirrored in the errors that the users made while using this system. While stationary, the average amount of errors that the users made was 3.5, this almost doubled to 6.50 while the users were mobile. It can be theorised that, because the participants were dual-tasking (walking and inputting) it was more difficult to input the text and navigate around the obstacles simultaneously. Users made the highest amount of errors in both movement conditions using the speech recognition system; while users using the wrist keyboard made 0.4 more errors than users using the on-screen keyboard while stationary and 0.8 fewer errors while mobile. This suggests that where mobile text input-based tasks need a high level of accuracy, a handwriting recognition system would be the most appropriate system from the four evaluated.

We should also consider the issue of hands-free use. If the task requires a user to use both hands and input text, speech recognition is one of the few viable input methods available; but there are drawbacks to using speech recognition, as we have previously seen (Chapter 3). In an environment with a high level of background noise the recognition rate of the device can be affected; so specialised microphones may need to be used that cancel background noise or are placed on the user's throat, called a throat microphone, as seen below in Figure 9. In considering hands-free interaction it is important to note that gestural input may also be used as a hands-free input method (see section 2.9.1). In many systems the users gesture with their hands (Tsukada and Yasumura 2002; Bowden *Et al* 2004), or involve arm and upper body movements (Kang, Lee and Jung 2004) which can be physically difficult to accomplish if the user requires their hands to be free. Wearable gesture-based systems such as Brewster's (*Et al* 2003) have taken this into account and instead provide the user with a method that allows the user to interact with the computer by using a head-based gesturing system. This allows the users to interact, and to have their hands free to accomplish other tasks. One branch of research that could stem from this initial research is the comparison of hands-free input methods for users in mobile environments.

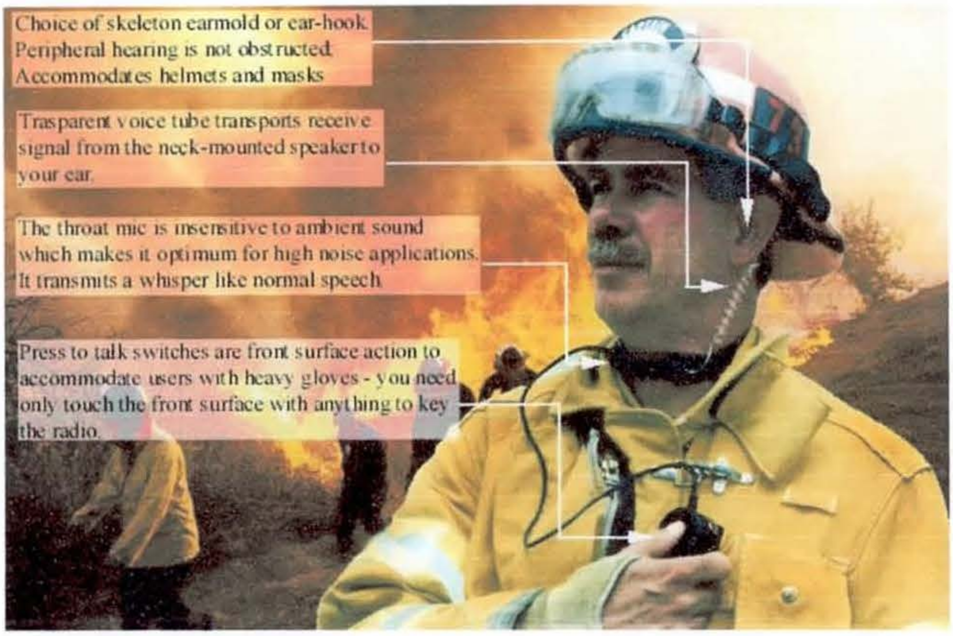


Figure 7: Fire-fighter wearing a throat microphone.
See - http://www.swatheadsets.com/ICONS/lash2.jpg

In tasks that require accurate text input an on-screen keyboard or a handwriting recognition system should be used, but for conversational information to another user it may be more appropriate to use a speech based system, if the user is in an environment where they can hear sufficiently. An additional property of using a stylus is the ability to notate diagrams and text using systems such as Microsoft Journal; this allows text and images to be notated, and handwriting to be 'translated' into text. Although, 'it is very hard to handwrite while on the move.' (Brewster, Lumsden, Bell, Hall and Tasker 2003), it must be noted that the participants using the handwriting recognition system were using a system that heavily mimics the use of a pen/pencil and paper; tools that many users use on a daily basis. Normally this is while the user is stationary, but it is important to recognise that it may have been harder to use the wrist-keyboard and on-screen keyboard while mobile. In examining the workload, users had a heavier workload when using the wrist-keyboard, speech recognition and on-screen keyboard; this may have also been because of the stylus mimicking a pen and paper.

The text input system with the highest cognitive workload was the wrist-worn keyboard. This may have been because the participants using this system had to input text on their wrist-worn keyboard and look in a separate place (the vest-mounted screen) for feedback, therefore making their task more complicated. Also, as earlier considered in the discussion section of Chapter three, it could be due to the small key sizes on the wrist-worn keyboard and the user being expected to type with one hand. As noted earlier a participant actually commented on this after finishing the text input task.

When carrying out tasks that have a high cognitive workload, it is important not to cognitively over burden the user so they can carry out their task in an appropriate manner. Handwriting recognition had the lowest cognitive workload in both movement conditions. This may be due to the participants already being familiar with the system as it mimics handwriting on paper with a pen or pencil (Feng 2003). It might even be said that speech recognition, on-screen keyboards and wrist-worn keyboards are not yet culturally 'accepted' methods of text input for a majority of users. This could have been the case for the participants in this study. The on-screen keyboard had a lower cognitive workload than speech recognition while the participant was stationary, but higher than speech recognition while the user was mobile. The different cognitive workload scores can be seen in Graph 3.3. There was a significant main effect of movement in regard to the errors made, the time taken to complete the task and the participants' cognitive workload. The devices were faster to use, the participants made fewer errors and had a lower cognitive workload score in the stationary movement condition.

The scale table below illustrates the virtues of each of the four text input systems that were used in the experiment. The scale at the top of the tables shows a plus and a minus sign. The nearer to the plus sign the text input system is, the better it performed in the experiment; the nearer it is to the minus sign (the right hand side of the screen), the worse it performed.

Table 7.1: Text Input – Scale Table. (Chapter 3)

Scale	**+**	⟵⟶		**-**
Time				
Standing	Speech Rec	Handwriting	Wrist Keyboard	On-Screen
Walking	Speech Rec	Handwriting	Wrist Keyboard	On-Screen
Errors				
Standing	Handwriting	On-Screen	Wrist Keyboard	Speech Rec
Walking	Handwriting	Wrist Keyboard	On-Screen	Speech Rec
Workload				
Standing	Handwriting	On-Screen	Speech Rec	Wrist Keyboard
Walking	Handwriting	Speech Rec	On-Screen	Wrist Keyboard

Key:

+ = Faster input time
Lower error rate
Lower workload score

- = Slower input time
Higher error rate
Higher workload score

By using the scale table we can make choices that relate to the user's and task's requirements. If we need a text input system that is fast to use, needs to be accurate and has a low cognitive workload, we would chose handwriting recognition. If we wanted a system that required the user to input text quickly while mobile, we may opt for the on-screen keyboard. It becomes clear that after completing this set of experiments, the scientific evidence gathered can be used to support the requirement of both the user and the task. These impact upon the type of input device they choose, and ultimately how well they perform their task.

7.3 Wearable Pointing Devices

Three different experiments were carried out to evaluate the appropriateness of the use of pointing devices as part of a wearable computing system. The three pointing devices were assessed in terms of their appropriateness for selecting targets (as explained in chapters section 2.9), for steering-based tasks and for dragging and dropping. The three devices were: a stylus, an off-table mouse and a track-pad. These were assessed while the participants used the input devices while stationary and while mobile.

7.4 Target Selection

Chapter four was the first investigation into the use of pointing as part of a wearable computing system; it directly followed on from the initial investigation into pointing devices for target selection in Appendix 5. This initial investigation was a pilot study and was done in order to find any problems and issues that might negatively influence the experiment. By conducting this initial experiment in chapter five, we could ascertain that the use of speech recognition would be problematic, because the recognition rate was low and as a consequence some of the participants seemed to get frustrated. In Chapter four we extended the evaluation and replaced the speech condition with the track-pad as it had been used in previous wearable computing systems. For this experiment it was found that there was significant main effect of movement on the time taken to complete the task and on the participants' cognitive workload, with the speed and cognitive workload score being lower in the stationary movement condition. Movement had no significant effect upon the errors made by the participants using the pointing devices.

As discussed in Chapter three, this may have been due to: carry-over effects, the task being too simple or the participants walking slowly and therefore taking more time to complete the task, in order to make fewer errors (Lemmens *Et al.* 2000). Interestingly, the stylus and off-table mouse both produced fewer errors while the participant was walking. This may have been due to the participants stopping or walking slowly as stated earlier in the discussion section of Chapter four. Further investigation could be done in this area to evaluate the participant's walking speeds and distances, and also which target size was easiest to select while mobile.

This experiment evaluated the three pointing devices for selecting targets in terms of accuracy (errors made), speed (time to complete the task) and the participants' cognitive workload. Target selection is important as it relates to the selection of buttons, icons, folders and other graphical items that can be selected by a pointing device on wearable interfaces (WetPC 2001; Carr *Et al.* 2002; Rensing *Et al.* 2002). In Table 7.2 (below) it can be seen that the best performing device was the stylus in terms of time and workload, but the off-table mouse in terms of the errors made (a full discussion of this can be found in the discussion section of Chapter four. Sections 4.3.1 to 4.3.3).

Table 7.2: Target Selection – Scale Table. (Chapter 4)

Scale	**+**	←——————→	**-**
Time			
Standing	Stylus	Track-Pad	Off-Table Mouse
Walking	Stylus	Track-Pad	Off-Table Mouse
Errors			
Standing	Off-Table Mouse	Stylus	Track-Pad
Walking	Off-Table Mouse	Stylus	Track-Pad
Workload			
Standing	Stylus	Track-Pad	Off-Table Mouse
Walking	Stylus	Off-Table Mouse	Track-Pad

Key:

+ = Faster input time	**-** = Slower input time
Lower error rate	Higher error rate
Lower workload score	Higher workload score

In cases where a stylus cannot be used, the selection of an input device for target selection is dependent on the task requirements. If the task requires a high level of accuracy then the off-table mouse should be used, but if speed is a requirement then the track-pad should be selected. For tasks that require the user to have a low workload while standing, the track-pad is marginally better than the off-table mouse, yet the off-table mouse is more ideally suited to mobile usage, as it produced a lower workload rating than the track-pad (see Table 4.3 for further data relating to these workload results).

7.5 Dragging and Dropping

Chapter five was the second investigation in the series of three experiments that assessed the performance of the three pointing devices used as part of a wearable computing system. Dragging and dropping is used on many WIMP based interfaces, but has also been used in mobile computing interfaces (Rekimoto 1997; Rekimoto and Saitoh 1999). Again, these devices were evaluated in terms of their speed, their accuracy and the participants' cognitive workload. From the scale Table 7.3, it can be seen that the device that was the most accurate, fastest to use and had the lowest cognitive workload was the stylus. The device with the second best performance in terms of time and errors was the track-pad, followed by the off-table mouse. However, the results for the participants' cognitive workload scores show that the off-table mouse performs better than the track-pad, while the participant is stationary, whereas the track-pad out performs the off-table mouse while the user is mobile.

In Chapters three and four it was noted that the participants slowed dramatically when completing their task. In these chapters it was hypothesised that this was mainly due to the participants having to dual-task. The participants were doing two tasks simultaneously, which led to trade-offs in terms of the users' performance; the participants walking slower than their normal pace was one of these trade-offs. In this experiment it was also noted that participants walked very slowly when completing their task. It is therefore important that in any other experiments evaluating mobile drag and drop techniques some level of evaluation is implemented that takes into consideration the participants' walking rate and distance (this is further expanded upon in the discussion section of Chapter five).

It is now starting to appear that the stylus is the most appropriate input device for both text input and as a pointer. This may be because, as earlier theorised, the participants were more familiar with the use of stylus-based devices. In Chapter five the appropriateness of drag and drop interaction while mobile was discussed. It was hypothesised that this interaction style would be difficult for users to use while mobile, especially on small screen devices when users have to walk constantly faster than their normal rate.

The findings from Chapter five also showed that there was a significant effect of movement upon the user's time to complete the task and the errors made while completing the task. No significant effect of movement was found in regard to the participants' cognitive workload. (The results relating to this experiment can be found in sections 5.3.1 to 5.3.3).

Table 7.3: Dragging and Dropping – Scale Table. (Chapter 5)

Scale	+	⟵⟶	-

Time			
Standing	Stylus	Track-Pad	Off-Table Mouse
Walking	Stylus	Track-Pad	Off-Table Mouse
Errors			
Standing	Stylus	Track-Pad	Off-Table Mouse
Walking	Stylus	Track-Pad	Off-Table Mouse
Workload			
Standing	Stylus	Off-Table Mouse	Track-Pad
Walking	Stylus	Track-Pad	Off-Table Mouse

Key:

+ = Faster input time
Lower error rate
Lower workload score

- = Slower input time
Higher error rate
Higher workload score

7.6 Trajectory-Based Input

Menu based interaction is found on many wearable computer based interfaces (Carr *Et al.* 2002; Rensing *Et al.* 2002; Basko *Et al.* 2002; Schmidt *Et al.* 2000) as well as Windows based interfaces (which come as standard with Xybernaut systems). Chapter six was an experiment that further evaluated the pointing devices used in Chapters four and five, to assess their appropriateness as input devices for trajectory based tasks.

The stylus had the overall best performance for the trajectory-based interaction task, but is trajectory-based interaction an ideal interaction style for mobile users? Performance in terms of input time almost doubled and when the participants came to do the walking sections of the experiment they moved very slowly and some, as in the other experiments, had to be reminded to keep walking. This is very important, as it showed that users found it difficult to dual-task while using this interaction style. Further evaluations using evaluation methodologies that examine walking speed and distance could have given further insights into the type of dual-task interaction that was happening in these experiments. Further discussion can be found in Chapter six.

Unlike the findings from Chapters four and five, there is a clear order in terms of the stylus being the best performing overall device for trajectory-based tasks while both stationary and mobile; the track pad has the second best performance in each condition and the off-table mouse has the worst performance. There was a significant effect of movement on the time on task, the errors made and the participants' cognitive workload; these were lower in the stationary movement condition. These findings are represented in the scale table below, Table 7.4. For task and user requirements that need, for example, the user to select menus quickly and accurately and the user to have a low workload, the stylus should be selected. (The results for this experiment can be found in sections 6.3.1 to 6.3.3).

Table 7.4: Trajectory-Based Input – Scale Table. (Chapter 6)

Scale	**+**	← →	**-**
Time			
Standing	Stylus	Track-Pad	Off-Table Mouse
Walking	Stylus	Track-Pad	Off-Table Mouse
Errors			
Standing	Stylus	Track-Pad	Off-Table Mouse
Walking	Stylus	Track-Pad	Off-Table Mouse
Workload			
Standing	Stylus	Track-Pad	Off-Table Mouse
Walking	Stylus	Track-Pad	Off-Table Mouse

Key:

+ = Faster input time Lower error rate Lower workload score	**-** = Slower input time Higher error rate Higher workload score

7.7 Evidence to Support the Choice of Input Device

The evidence that was used in these trials was based around the time on task, errors made and the cognitive workload of the participant. Each experiment was designed to avoid carry-over effects and different participants were used for each evaluation. Future evaluations aim to gather questionnaire and ethnographically based data, to further understand the interactions that occur between the user, their task, the input devices used and their movement condition.

Initially this thesis aimed to contribute to the 'field' of research in four ways (see section 1.1.2). This will be discussed in the final concluding chapter. Although input speeds and error rates were collected while the user is stationary and mobile, these had their limitations. After carrying out the experiments, it was found that they could have been expanded in regard to collecting data on the distance and speed that the participants moved at in the mobile conditions of the experiments. If this data had been recorded, it could have led to a more thorough understanding of the participants' interaction while using input as part of a wearable computing system in a dual-task paradigm (walking and inputting). For future evaluations this is highly important as it can show how useable the device is in a mobile environment. It may also be the case that two input systems have a similar error rate and input time; by comparing walking speeds and distances we are able to further analyse and evaluate the systems.

7.9 Conclusion

The experiments and literature review carried out in this thesis have identified and demonstrated the following:

- Differing input speeds across identical tasks using different input devices.
- Differing levels of accuracy across identical tasks using different input devices.
- Differing cognitive workload scores across identical tasks using different input devices.

- Differing input speeds across different movement conditions using identical input devices.
- Differing levels of accuracy across different movement conditions using identical input devices.
- Differing cognitive workload scores across different movement conditions using identical input devices.

The results above are supported by the experimental outcomes. This chapter has reviewed the results from the experiments and has discovered that overall the stylus had the best performance, as part of a handwriting recognition system. The findings of the three experiments that were used to judge the performance of the three pointing devices were that the most appropriate pointing device for use with a wearable computing system, while stationary and mobile was the stylus.

These experiments provided a valuable insight into the behaviour of wearable users when wearing a computer and inputting text and using pointing devices, while stationary and mobile. This discussion is carried into the final conclusion of the thesis and further expanded.

Chapter 8: Thesis Conclusion

8.1 Introduction

This chapter completes the thesis. It presents an outline of the contribution of knowledge that this thesis has made to the field of wearable computing. It examines ways in which the research presented in the thesis could be expanded upon and also provides an insight into future research that could provide a better understanding of wearable computers, mobile use and input/output modalities for mobile/wearable computers. This section ends by focusing upon the expansion of the evaluation methodologies used in the thesis.

8.2 Original Contribution to Knowledge

In section 1.1.2 the aims of the thesis were listed. These have been accomplished, but there is still further work that can be done, based on the experiments that have already been completed. Techniques for the evaluation of input systems for use with wearable computers were examined and used throughout the thesis and as discussed in the previous chapter, empirical research based on input speeds and error rates have shown that overall the stylus was the best text input and pointing device. There were limitations to the studies in regard to the recording of the distance that the participants walked and the speed at which they walked when carrying out the experimental tasks. This is further discussed in the following section.

The literature in Chapter two suggested that there were some research areas that had yet to be addressed in regard to wearable computer use. This thesis identified a lack of empirical evaluations relating to the use of input devices for text input and pointing while the user was wearing a computer and mobile. Previous studies had offered an initial insight into this area, but had failed to use wearable computer systems and focused on desktop evaluations (Curry, Hobbs and Toube 1996) or, had importantly, not assessed the user while mobile (Thomas *Et al.* 1997).

The research within this thesis addressed these research issues to a PhD standard and by doing so offered new insights and perspectives on the use of input devices for wearable computers. Off-the-shelf devices were used in the evaluations as it was theorised that these were the most accessible to a wide range of users. After discussing the findings from the studies in the previous chapter we are able to see where this thesis has made original contributions to scientific knowledge in the following key areas.

1. It provided experimental evidence of significant differences between the use of input devices, as part of a wearable computing system, when used in a stationary and mobile manner.
2. It provides evidence to suggest that the stylus was the most appropriate input device for inputting text, when used with a handwriting recognition system while stationary, and when used with a virtual keyboard while mobile.
3. It provides evidence to suggest that the stylus was overall the most appropriate pointing device in terms of target selection, steering and dragging and dropping while used in both a stationary and mobile manner
4. It provided a scientific foundation in terms of results to further explore the use of input devices with wearable computers.

The main contribution of this thesis has been a combination of the points listed above. It has given an insight into the way that movement impacts upon a wearable computer user's ability to use text input devices and pointing devices. This thesis investigated and reported both empirically and subjectively on the way in which different input devices could be used with wearable computer systems.

8.3 Further Work

This research has focused on generic tasks that apply to many users, using GUI based interfaces and input devices with wearable computing systems, but there are other non-GUI-based systems such as audio-centric wearable devices and wearable computing systems that use the user's gestures for input. The evaluation methodologies applied in the studies carried out in this thesis could be applied to a variety of other input and output systems. The studies carried out in this thesis used a fold-down touch-screen. Other wearable computer-based systems used augmented reality displays, but we found these systems to cause binocular rivalry (Laramee and Ware 2002), which led to the participants experiencing bouts of dizziness and feeling sick. *An initial comparative experiment was carried out and further extended the trials in this thesis to evaluate the impact of different visual output systems on wearable users. These could not be further carried out because of the symptoms of the fore mentioned participants.* Some research has been carried out into the use of wearable computers and three-dimensional augmented-reality (Piekarski and Thomas 2002). There needs to be exploration into the use of such systems and empirical studies that evolve from this work could be used in such a manner.

The six studies carried out in this thesis related to generic task activities, while the participants were stationary and mobile. Further research aims to focus on specific user and task domains, such as the use of wearable computers in domestic pervasive environments.

After examining the experiments they could have been expanded in several ways to gather more information. One further route of exploration would have been through the use of a subjective questionnaire. This would have provided open-ended responses and user opinions to the devices evaluated. Since the completion of these trials a questionnaire has been designed for future evaluations.

In these trials we used only one output system, which was a small vest-worn touch screen. Other wearable systems have used various augmented reality systems and have started to examine the impact of these various output devices upon mobile users. Will the way that each output modality is presented, whether sound based or visual, affect our use of input modalities? Although some work was initially started in this area it was a desktop-based evaluation and as such is difficult to apply to mobile/wearable users.

As wearable computing is still an evolving area, there is still much research to be done to examine the most appropriate input and output appropriate for different users and environments. It would also be feasible to develop and assess adaptable systems that may alternate between different input and output modalities depending on the users' preferences, task, other input/output modalities, their placement on the body and the ever-changing context of some wearable users.

One direction this research could advance in would be the understanding of a walking-based assessment such as PPWS (Brewster 2003). In all of the experimental discussions it was noted that participants walked very slowly. It was hypothesised that this was due to their doing two tasks simultaneously (dual-tasking) and therefore there were 'trade-offs' in terms of the participants' walking speed and the distance they travelled. This was one of the limitations of these studies. In future evaluations it is therefore important to record the distance and speed that the participants walk. Although studies have used this to assess interfaces, there still needs to be further investigation into this assessment methodology that originally stemmed from assessing the stress levels of visually impaired users. Initially, heart rate monitoring was to be used, but it was found that visually impaired user's heart rates were too high when they were asked to navigate around unknown environments. It could easily be established if this method was accurate by correlating a user's heart rate, PPWS, error rates, time on task and the participant's subjective workload. The diagram in Figure 8, below, shows the methodologies used in the studies explored within this thesis and the evaluation methods yet to be further explored and evaluated.

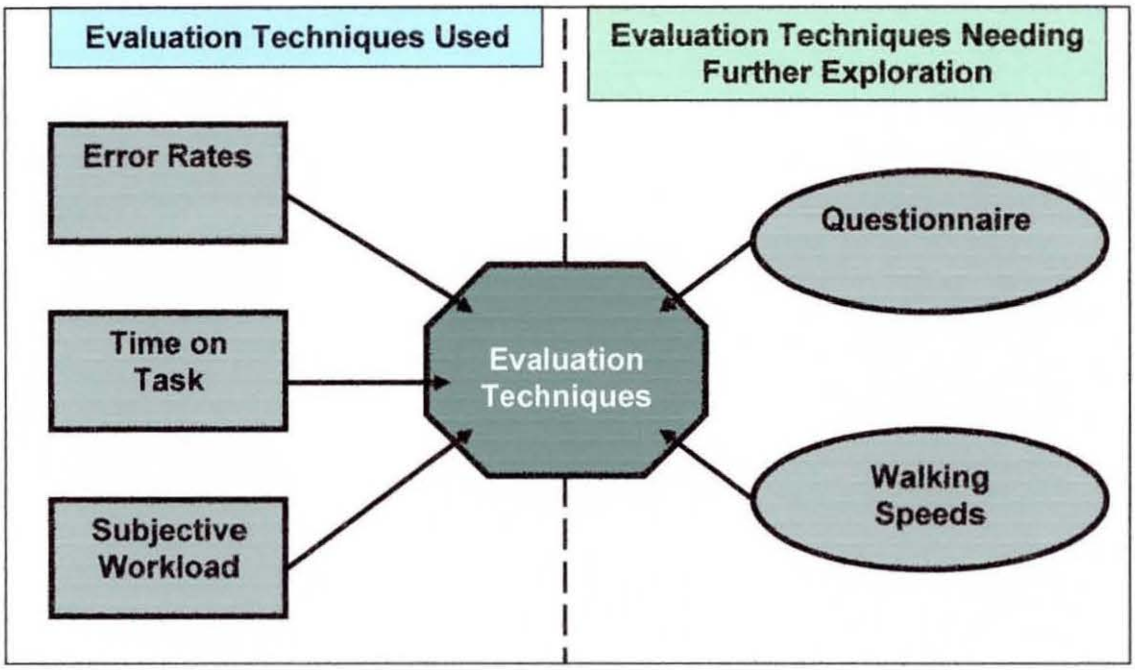


Figure 8: Evaluations and explorations.

Using the methods given within this thesis combined with the questionnaire and a walking-based evaluation will lead to a fuller evaluation methodology. This multi-faceted approach could provide further valuable insights into the way that wearable users interact with input devices and the interplay between the user, the input device they are using and the way they are moving. In this study, lab based conditions were used, but it is hoped that further experimentation will be carried out in 'real world' settings.

8.4 Conclusion

This thesis initially started by identifying the core characteristics of wearable computers as identified by (Mann 1998; Bass 1997 and Rhodes 1997). One of the common characteristics (seen in Chapter two) was: *The user should be able to enter data and control the system.* Based on this a review of the literature was carried out, and identified that little empirical research had been undertaken in relation to the input systems/devices used to enter data and manipulate the interface. This led into an evaluation of text input systems and pointing systems/devices for system control and interface manipulation (although, there are other means of input available, such as sensor based input). One of the specific aims of this research was to report and understand the effect that movement (walking and standing) has on usability for a wearable computer user, when inputting text and doing generic pointing tasks such as dragging and dropping, selecting targets and steering-based interaction.

Drawing on this initial research, two sets of experimental trials were conducted. These were based around text input and pointer based input. Initially Chapter three explored text entry systems for wearable computers and compared four different systems, while the user was standing stationary and walking. The systems evaluated were: handwriting recognition, speech recognition, a wrist-worn keyboard and a virtual keyboard. It was found that the handwriting recognition was the most appropriate method of input. Full details of this experiment can be found in Chapter three.

The second strand of experimentation focused upon the use of pointing devices with wearable computing systems. The devices used were a track-pad, off-table mouse and stylus, used in conjunction with a touch screen. The literature found that there were three main interaction styles associated with pointing: target selection, dragging and dropping and steering. Importantly this thesis took "off the shelf" devices that are commonly available to many wearable users and evaluated their appropriateness for inputting text and also for use as pointing devices. This is important as it gives both the systems designer and user a range of 'safe' options to use when developing interfaces for wearable computing-based systems. These experiments found that overall, out of the devices used, the stylus was the most appropriate device to use while standing and walking. For a more detailed account of the experiments turn to Chapters four, five and six.

The experimental evidence shows that this thesis has found that there are significant differences between the input devices used in the experiments. This evidence found that the movement conditions that wearable computer users encounter, standing stationary and walking, have an impact upon the user's ability to input text and use pointing devices. These differences occurred in terms of the time taken to complete the tasks, the errors made while completing the task and the participant's subjective workload score.

Interestingly, there have been developments in the 'off the shelf' wearable computer market that are supported by the findings in this thesis. The miniaturisation of computing technology has made it is possible to develop a tablet computer that is small enough to be worn in a vest Figure 8.1; and has therefore eliminated the cabling that connected the CPU (central processing unit) to the screen and to the power-pack. This has led to a convergence of the tablet-computer and the wearable computer and has resulted in products such as the Xybernaut Atigo.

Figure 8.1. A Vest Worn Xybernaut Wearable Computer (MA Range).
See - http://www.usc.edu/dept/architecture/mbs/thesis/anish/2_mobile.jpg

The research in this thesis both mirrors and supports these new developments in the wearable computing market. It also leads us to conclude that the next generation of wearable computers will be pen-based computers that are reliant on the stylus for text input and pointing. Based on the findings of this thesis, the stylus is one of the most appropriate means of physically interacting with a wearable computer. Worn in a vest based configuration, this allows the user to fold down the screen, interact with the system and be cable-free. It also allows the user to have one hand free to do other tasks.

Figure 8.2: Xybernaut Attigo Computer.
See - http://www.xybernaut.com/images/Atigo_L_product_small.jpg

Currently, Xybernaut computers use the Windows platform, and the WIMP interaction style still prevails in these systems. After completing the evaluations within this thesis, it is evident that there is still much work yet to be done into understanding the effect of movement upon a wearable computer user's ability to input, and the type of input devices that are appropriate for wearable users.

References

Accot, J. and Zhai, S. (1997). Beyond Fitts' Law: Models for Trajectory-Based HCI Tasks. In the Proceedings of the SIGCHI conference on Human factors in computing systems - CHI'97, Atlanta, Georgia, USA: ACM Press, pp. 295-302.

Accot, J. and Zhai, S. (1999). Performance Evaluation of Input Devices in Trajectory - based Tasks: An Application of Steering Law. In the Proceedings of the SIGCHI conference on Human factors in computing systems: the CHI is the limit - CHI'99, Pittsburgh, Pennsylvania, USA: ACM Press, pp. 466-472.

Adams, N., Gold, R., Schilit, B.N, Tso. M. and Want, R. (1993). An Infrared Network for Mobile Computers. In the Proceedings of the Winter Symposium on Mobile and Location-independent Computing – USENIX'93, Cambridge, Massachusetts, USA: ACM Press, pp. 41-52.

Albers, M.C. (1994). The Varèse System, Hybrid Auditory Interfaces, and Satellite-Ground Control: Using Auditory Icons and Sonification in a Complex, Supervisory Control System. In the Proceedings of the Second International Conference on Auditory Display – ICAD'94. Santa Fe, New Mexico, USA: Addison-Wesley, pp. 3 – 14.

Allport, D. A., Antonis, B. and Reynolds, P. (1972). On the division of attention: a disproof of the single-channel hypothesis. Quarterly Journal of Experimental Psychology, 24, pp. 225-235.

Avancha, S., Chakraborty, D., Chen, H., Kagal, K., Perich, F., Finin, T. and Joshi, A. (2002). Issues in Data Management for Pervasive Environments. In the Proceedings of the NSF Workshop on Context Aware Mobile Database Management – CAMM'02, Providence, Rhode Island, USA: Brown University Computer Science, pp.24-25.

Baber, C. and Noyes, J.M. (1996). Automatic speech recognition in adverse environments. Human Factors, 38, pp. 142-155.

Braarud, Ø. (2001). Subjective Task Complexity and Subjective Workload: Criterion Validity for Complex Team Tasks. International Journal of Cognitive Ergonomics, 5(3), pp. 261-273.

Bass. L., Siewiorek. D., Smailagic. A., Stivoric. J. (1995). On Site Wearable Computer System. In the Proceedings of the SIGCHI conference on Human factors in computing systems - CHI'95, Denver, Colorado, USA: ACM Press, pp. 83-84.

Bass, L., Kasabach, C., Martin, R., Siewiorek, D., Smailagic, A. and Stivoric, J. (1997). The Design of a Wearable Computer. In the Proceedings of the SIGCHI conference on Human factors in computing systems - CHI'97, Atlanta, Georgia, USA: ACM Press, pp. 139-146.

Bates, R. (1999). Multimodal Eye-based Interaction for Zoomed Target Selection on a Standard Graphical User Interface. In the Proceedings of the Seventh IFIP Conference on Human Computer Interaction - Interact'99, Riccarton, Edinburgh, Scotland: British Computer Society Press, Volume II, pp.7-8.

Batista, E. (2001). Sore Thumb? Talk More, Type Less. Wired News. http://www.wired.com/news/wireless/0,1382,43875,00.html

Beck, E., Christiansen, M., Kjeldskov, J., Kolbe, N. and Stage, J. (2003). Experimental Evaluation of Techniques for Usability Testing of Mobile Systems in a Laboratory Setting. In the Proceedings of the OZCHI'03 Conference, Queensland, Brisbane, Australia: CHISIG, pp. 106-115.

Bernard, M., Mills, M., Peterson, M. and Kelsey-Storrer, A. (2002). Comparison of Popular Online Fonts: Which is Best and When? Online - Usability News, 4.1.

Billinghurst, M. (2002). Wearable Appliances: The Future of Wearable Computing. Appliance Design, 2, pp. 5-6.

Billinghurst, M., Bowskill, J., Jessop, M., and Morphett, J. (1998). An Evaluation of Wearable Information Spaces. In the Proceedings of the Virtual Reality Annual International Symposium – VRAIS'98, Atlanta, Georgia, USA: IEEE Press, pp. 20-27.

Blaskó, G. and Feiner, S. (2002). A Menu Interface for Wearable Computing. In the Proceedings of the Sixth International Symposium on Wearable Computers - ISWC 2002, Seattle, Washington, USA: IEEE press, pp.164-165.

Blattner, M.M., Sumikawa. D.A., Greenberg, R.M. (1989). Earcons and Icons: Their Structure and Common Design Principles. Human-Computer Interaction, 4, pp.11-44.

Blattner, M. and Dannenburg, R.B., Eds. (1992). Multimedia Interface Design. ACM Press, Addison-Wesley, New York.

Blickenstorfer, C. H. (1995). Graffiti: Wow!!!! Pen Computing Magazine, January. Pen Computing Publishing USA, pp. 30-31.

Bohan, M., Phipps, C.A., Chaparro, A. and Halcomb, C. (1999). A psychophysical comparison of two stylus-driven soft keyboards. In the Proceedings of the Graphics Interface Conference - GI '99, Kingston, Ontario, Canada: ACM Press, pp. 92-97.

Bourke, P.A. (1997) Measuring attentional demand incontinuous dual-task performance. The Quarterly Journal of Experimental Psychology: Section A. 50(4), pp. 821-40.

Bowden, R., Windridge, D., Kadir, T., Zisserman, A. and Brady, M. (2004). A Linguistic Feature Vector for the Visual Interpretation of Sign Language. In the Proceedings of the

Eighth European Conference on Computer Vision – ECCV'04, Prague, Czech Republic: LNCS3022, Springer-Verlag, pp. 391- 401.

Bregler, C. and Omohundro, S.M. (1995). Nonlinear manifold learning for visual speech recognition. In the Proceedings of the Fifth International Conference on Computer Vision – ICCV'95, Cambridge, Massachusetts, USA: IEEE Press, pp. 494-499.

Brewster, S.A., Wright, P.C. and Edwards, A.D.N. (1992). A detailed investigation into the effectiveness of earcons. In G. Kramer (Ed.), Auditory display, sonification, audification and auditory interfaces. In the Proceedings of the First International Conference on Auditory Display - ICAD, Santa Fe, New Mexico: pp. 471-498.

Brewster, S.A., Wright, P.C. and Edwards, A.D.N. (1993). An evaluation of earcons for use in auditory human-computer interfaces. In the Proceedings of InterCHI'93 Conference, Amsterdam, Holland: ACM Press, pp. 222-227.

Brewster, S.A., Wright, P.C. and Edwards, A.D.N. (1994). The design and evaluation of an auditory-enhanced scrollbar. In the Proceedings of the SIGCHI conference on Human factors in computing systems: celebrating interdependence - CHI'94, Boston, Massachusetts, USA: ACM Press, pp. 173-179.

Brewster, S.A. (1994). Providing a structured method for integrating non-speech audio into human-computer interfaces. PhD Thesis, University of York, UK.

Brewster, S.A. (1998). Sonically-enhanced drag and drop. In the Proceedings of the Fifth International Conference on Auditory Displays - ICAD'98, Glasgow, UK: British Computer Society, ISBN 1 902505 05 0.

Brewster, S.A. and Cryer, P.G (1999). Maximising Screen-Space on Mobile Computing Devices. In the summary Proceedings of the SIGCHI conference on Human factors in computing systems: the CHI is the limit - CHI'99, Pittsburgh, Pennsylvania, USA: ACM Press, Addison-Wesley, pp. 224-225.

Brewster, S.A. and Walker, V.A. (2000). Non-Visual Interfaces for Wearable Computers. IEE Workshop on wearable Computing (00/145), London, UK: IEE Press.

Brewster, S.A. and Murray, R. (2000). Presenting dynamic information on mobile computers. Personal Technologies, 4(2), pp. 209-212.

Brewster, S.A. (2002). Overcoming the Lack of Screen Space on Mobile Computers. Personal and Ubiquitous Computing, 6(3), pp. 188-205.

Brewster, S.A., Lumsden, J., Bell, M., Hall, M. and Tasker, S. (2003). Multimodal 'Eyes-Free' Interaction Techniques for Wearable Devices. In the Proceedings of the conference on Human factors in computing systems - CHI '03, Fort Lauderdale, Floida USA: ACM Press, Addison-Wesley, pp. 463-480.

Buck, L. (1980). Motor performance in relation to control-display gain and target width. Ergonomics, 23, pp. 579-589.

Buergy, C., Garrett, J. H., Klausner, M., Anlauf, J., Nobis, G. (2001). Speech-Controlled Wearable Computers for Automotive Shop Workers. The Society of Automotive Engineers – SAE'01 World Congress, Detroit, Michigan, USA: SAE Technical Paper Series No. 2001-01-0606.

Bürgy, C and Garrett, J. (2002). Wearable Computers: An Interface between Humans and Smart Infrastructure Systems. Bauen mit Computer, Bonn, Germany. VDI Verlag GmbH, Duesseldorf, Germany.
See - http://www.ce.cmu.edu/~wearables/docs/Buergy_BauenMitComputern.pdf

Bussemakers, M. P. and de Haan. A. (2000). When it Sounds like a Duck and it Looks like a Dog .Auditory icons vs. Earcons in Multimedia Environments. In the Proceedings of the Sixth International Conference on Auditory - ICAD'00, Atlanta, Georgia, USA: ICAD, pp. 184 – 189.

Butts, L. and Cockburn, A. (2002). An Evaluation of Mobile Phone Text Input Methods. In the Proceedings of the Third Australasian User Interface Conference – AUIC'02, Melbourne, Australia: Australian Computer Society, pp. 55-59.

Butz, A., Höllerer, T., Feiner, S, MacIntyre, B. and Beshers, C. (1999). Enveloping Users and Computers in a Collaborative 3D Augmented Reality, In the Proceedings of the International Workshop on Augmented Reality - IWAR '99. San Francisco, California, USA: IEEE Press, pp.52-62.

Card, S. K., Moran, T. P, and Newell, A. (1983). The psychology of human-computer interaction. Hillsdale, NJ: Erlbaum.

Card, S., English, W.K., and Burr, B.J. (1978). Evaluation of Mouse, Rate-Controlled Isometric Joystick, Step Keys, and Text Keys for Text Selection on a CRT. Ergonomics, 21(8), pp. 601-613.

Carr, C.E., Schwartz, S.J. and Rosenberg, I. (2002). A Wearable Computer for Support of Astronaut Extravehicular Activity. In the Proceedings of the Sixth International Symposium on Wearable Computers – ISWC'02, Seattle, Washington, USA: IEEE Press, pp. 23 - 30.

Chi, S. and Oh, Y. (1996). Lombard effect compensation and noise suppression for noisy Lombard speech recognition. In the Proceedings of the Fourth International Conference on Spoken Language Processing – ICSLP'96. Philadelphia, PA, USA: CD Rom. Vol 4 - SuA1L1.

Clark-Carter, DD., *Et al.* (1986). The efficiency and walking speed of visually impaired people. Ergonomics; 29, pp. 779-789.

Cohen, J. (1994). Monitoring background activities. Auditory Display: Sonification, Audification, and Auditory, Interfaces. Reading MA: Addison-Wesley.

Comdex News (2004) Press Releases.
www.comdex.com/news/press_releases/ index.php?s=default&pr_id=8204

Connell, S.D. (2000). Online Handwriting Recognition Using Multiple Pattern Class Models.PhD Thesis, Michigan State University (Department of Computer Science and Engineering).

Cross, J., Woolley, S. and Baber, C. (2000). Wearing Computers - An Archaeological Case Study. IEE Workshop on wearable Computing (00/145), London, UK: IEE Press. http://www.eee.bham.ac.uk/pcg/docs/wearcomp.pdf

Cross, J. Woolley, S. Baber, C. Gaffney, V. (2001). Wearable Computing for field Archaeology. In the Proceedings of the Fifth International Symposium on Wearable Computers – ISWC'01, Zurich, Switzerland: IEEE Press, pp. 169 – 171.

Curry, M., Hobbs, J. and Toub, R. (1996). Will Using a Head Mounted Display Effect Human Computer Interaction on Wearable Computers? In the Proceedings of the Second International Conference on Mobile Computing and Networking - MobiCom 96, White Plains, New York, USA: ACM Press, -
http://www.cs.uoregon.edu/~kortuem.htbin/download.cgi?/cs/www/home/research/wearables/Papers/AdvUIpaper.html

Deshmukh, O., Espy-Wilson, C. and Juneja A. (2002). Acoustic-phonetic speech parameters for speaker-independent speech recognition. In the Proceedings of the International Conference on Acoustics, Speech, and Signal Processing – ICASSP'02, Orlando, Florida, USA: IEEE Press, pp. 593-596.

Ditlea, S. (2000). The PC Goes Ready to Wear, IEEE Spectrum, October 2000. 37 (10), IEEE Press, pp. 34-39.

Dix, A., Finlay, J., Abowd G. and Beale R., (2004). Human-Computer Interaction, Prentice Hall, Third Edition.

Dunlop, M.D. (2004). Watch-Top Text-Entry: Can Phone-Style Predictive Text-Entry Work With Only 5 Buttons?. In the Proceedings of Mobile HCI 04. Glasgow, Sotland: Lecture Notes in Computer Science, Vol. 3160, pp. 342 – 346.

Epps, B. W. (1986). Comparison of Six Cursor Control Devices Based on Fitts' law Models. In the Proceedings of the Annual Meeting of the Human Factors Society – HFES'86, Dayton, Ohio, USA: HFES, pp. 327-331.

Evans, K.B., Tanner, P.P. and Wein, M. (1981). Tablet-Based Valuators that Provide One, Two or Three Degrees of Freedom. Computer Graphics, 15(3), pp. 91-97.

Eves, D., Green, J., van Heerden, C., Mama, J.,Marzano, S. (2000). New Nomads: An exploration of wearable electronics, by Philips. 010 Publications: Rotterdam.

Faruk Özer, O., Özün. O., Tüzel, O., Atalay, V. and Çetin E.A. (2001). Vision-Based Single-Stroke Character Recognition for Wearable Computing , IEEE Intelligent Systems and Their Applications, 16(3), pp. 33-37.

Feiner, S., MacIntyre, B., Höllerer, T., and Webster, T. (1997). A touring machine: Prototyping 3D mobile augmented reality systems for exploring the urban environment. Personal Technologies, 1(4), pp. 208-217.

Feng, W. (2003). An Integrated Multimedia Environment for Speech Recognition Using Handwriting and Written Gestures. In the Proceedings of the 36th Hawaii International Conference on System Sciences, IEEE Press, Hawaii USA, pp.128b-c.

Field, A. (1998). A bluffer's guide to ... sphericity. The British Psychological Society: Mathematical, Statistical and Computing Section Newsletter, 6, pp. 13-22.

Field, A. and Hole G.J. (2002). How to Design and Report Experiments, Sage.

Fitts, P.M. (1954). The information capacity of the human motor system in controlling amplitude of movement, Journal of Experimental Psychology, 47(6), pp.381-391.

Freed, L. (1997, March 25) 'Future user interfaces' PC Magazine 16, pp. 206-207.

Fukumoto, M. *Et al.* (1994). "Finger Ring": A Full-time Wearable Interface. In the Proceedings of the SIGCHI conference on Human factors in computing systems: celebrating interdependence, Boston, Massachusetts, United States, ACM Press, pp. 81-82.

Fukumoto, M. and Tonomura. Y. (1997). "Body Coupled Finger Ring": Wireless Wearable Keyboard. In the Proceedings of the SIGCHI conference on Human factors in computing systems - CHI'97, Atlanta, Georgia, USA: ACM Press, pp. 47-154.

Gales, M. J. F and Young, S.J. (1995). Robust speech recognition in additive and convolutional noise using parallel model combination. Computer Speech and Language 9, pp. 289–307.

Gaver, W. (1986). Auditory Icons: Using sound in computer interfaces. Human - Computer Interaction. 2, pp. 167-177.

Gauvain, J., Gangolf, J., and Lamel, L. (1996). Speech recognition for an information kiosk. In the Proceedings of the Fourth International Conference on Spoken Language Processing - ICSLP '96, Philadelphia PA, USA: pp. 849–852.

Gemperle, F., Kasabach, C., Stivoric, J., Bauer, M. and Martin, R. (1999). Design for Wearability. In the Proceedings of the International Symposium on Wearable Computing - ISWC'99, San Francisco, CA, USA: IEEE Press, pp. 116–122.

Gentner, D. R. (1983). Keystroke timing in transcription typing. In W. E. Cooper (Ed.), Cognitive aspects of skilled typewriting, New York: Springer-Verlag: pp. 95-120.

Gentner, D. R., Grudin, J. T., Larochelle, S., Norman, D. A., and Rumelhart, D. E. (1983). A glossary of terms including a classification of typing errors. In W. E. Cooper (Ed.), Cognitive aspects of skilled typing. Springer-Verlag, pp. 39-44.

'Gili' Wearable Interface (1989).
See - http://wearables.essex.ac.uk/sulawesi/io/gili.html

Goldberg, D. and Richardson, D. (1993). Touch-typing with a stylus, In the Proceedings of the InterCHI'93 conference, Amsterdam, Holland: ACM Press, pp. 80-87.

Goldstein, M., Book, R., Alsio, G. and Tessa, S. (1998). Ubiquitous Input for Wearable Computing: Qwerty Keyboard without A Board. In the Proceedings of the First Workshop on Human Computer Interaction with Mobile Devices, Glasgow, Scotland.
http://www.dcs.gla.ac.uk/~johnson/papers/mobile/HCIMD1.html

Goldstein, M. and Chincholle, D. (1999). The Finger-Joint Gesture Wearable Keypad, In the Proceedings of the Second Workshop on Human Computer Interaction with Mobile Devices, Edinburgh, Scotland. - http://www.dcs.gla.ac.uk/mobile99/

Goldstein, M., Chincholle, D. and Backström, M. (2000). Assessing two new wearable input paradigms: The finger-joint-gesture palm-keypad and the invisible phone clock. Personal Technologies, 4(2), pp.123 -133.

Goodman, J., Brewster, S.A. and Gray, P.D. (2004). Using Field Experiments to Evaluate Mobile *Guides*. In Proceedings of HCI in Mobile Guides 2004 workshop, Mobile HCI2004, Glasgow, Scotland.
http://www.dcs.gla.ac.uk/~stephen/papers/MobileGuides04_Goodman.pdf

Gopher, D., and Raij, D. (1988). Typing with a two-handed chord keyboard: Will the QWERTY become obsolete? IEEE Transactions on Systems, Man, and Cybernetics, 18, pp. 601-609.

Gould, S. J. (1987). The panda's thumb of technology. Natural History, January edition pp. 14-23.

Gravetter, F.J. and Wallnau, L.B. (1985). Statistics for the behavioural sciences. St Paul, MN: West Publishing Company.

Grudin, J. T. (1983). Error patterns in novice and skilled transcription typing. In W. E. Cooper (Ed.), Cognitive aspects of skilled typewriting. New York: Springer-Verlag. pp. 121-143.

Hare, C.B. (2002). Redefining user input on handheld devices. In the Proceedings of the 3G International Conference on Mobile Communication Technologies, Cambridge, MA, USA: IEEE Press, pp. 388 – 393.

Hart, S.G and Wickens, C.D. (1990). Workload Assessment and Prediction. In: Booher, H.R. (Ed), MANPRINT: An emerging technology, advanced concepts for integrating people, machine, and organizations. New York: Van Nostrand Reinhold.

Hazeltine, E., Teague, D., and Ivry, R. B. (2002). Simultaneous dual task performance reveals parallel response selection after practice. Journal of Experimental Psychology: Human Perception and Performace, 28, pp. 527-545.

Hommel, B. (1998). Automatic stimulus-response translation in dual-task performance. Journal of Experimental Psychology: Human Perception and Performance, 24, pp. 1368-1384.

Howard, B. Howard, S. (2001). Lightglove: wrist-worn virtual typing and pointing. In the Proceedings of the Fifth International Symposium on Wearable Computers – ISWC'01, Zurich, Switzerland: IEEE Press. pp. 172-173.

Howell, D. C. (2002). Statistical methods for psychology. (5th. edition.). Belmont, CA: Duxberry Press.

Hunter, M., Zhai, S., and Smith, B. A. (2000). Physics-based graphical keyboard design. In the Extended Abstracts of the Conference on Human Factors in Computing Systems – CHI'00, The Hague, The Netherlands: ACM Press. pp.157-158.

IBM Embedded ViaVoice, Multiplatform Edition (2004). http://www-306.ibm.com/software/pervasive/embedded_viavoice_multiplatform/

Iannizzotto, G., Villari, M. and Vita, L. (2001). Hand tracking for human-computer interaction with graylevel visualglove: Turning back to the simple way. In the Proceedings of the Workshop on Perceptive User Interfaces – PUI'01, Lake Buena Vista, Florida, USA: ACM, pp. 1-7.

Ingmarsson, M., Dinka, D., Zhai, S. (2004). TNT—A numeric keypad based text input method, In the Proceedings of the 2004 conference on Human factors in computing systems – CHI'04, Vienna, Austria: ACM Press, pp.639-646.

Inkpen, K.M. (2001). Drag-and-Drop versus Point-and-Click: Mouse Interaction Styles for Children. In ACM Transactions on Computer Human Interaction, 8(1), pp.1-33.

Jacko, J. A. and Vitense, H.S. (2001). A review and reappraisal of information technologies within a conceptual framework for individuals with disabilities Universal Access in the Information Society 1, pp 56-76.

Johnson, P. (1998). Usability and Mobility: Interactions on the move. In the Proceedings of the First Workshop on Human Computer Interaction with Mobile Devices (BSC/EPSRC), Glasgow, UK, Department of Computing Science. http://www.dcs.gla.ac.uk/~johnson/papers/mobile/HCIMD1.html

Kaefer, G. and Weiss.R. (2003) Application of Wearable Guards to Increase Motor Racing Safety. Telematik, 1, pp. 10-13.

Kang, H., Lee, C.W. and Jung, K. (2004). Recognition-based gesture spotting in video games, Pattern Recognition Letters 15(25), pp. 1701-1714.

Karat, C.M., Halverson, C., Horn, D., & Karat, J. (1999). Patterns of entry and correction in large vocabulary continuous speech recognition systems. In the Proceedings of the Conference on Human Factors in Computing Systems - CHI'99, Pittsburgh, Pennsylvania, USA: ACM Press, pp. 568-574.

Karvonen, K. and Parkkinen J. (2001). A Short Note on the Social Consequences of Wearables. In the Proceedings of the Fifth World Multi-Conference on Systemics, Cybernetics and Informatics - SCI'01, Orlando, Florida, USA: IIIS, - www.cs.hut.fi/~pmrg/Publications.html

Kawachiya, K. and Ishikawa, H. (1998). NaviPoint: An input device for mobile information browsing. In the Proceedings of the Conference on Human Factors in Computing Systems - CHI '98, LA, California, USA: ACM, pp.1-8.

Kjeldskov J. and Skov M. B. (2003). Evaluating the Usability of Mobile Systems: Exploring Different Laboratory Approaches. In the Proceedings of the 10th International Conference on Human-Computer Interaction, Crete, Greece: Lawrence Erlbaum Associates, pp. 123-127.

Knight, J.F., Baber, C., Schwirtz, A. and Bristow, H.W. (2002). The comfort assessment of wearable computers. In the Proceedings of the Sixth International Symposium of Wearable Computers, October 7-10, Seattle, Washington USA: IEEE Press, pp. 65-72.

Kohno, M. and Rekimoto, J. (2002). New Generation of IP-Phone Enabled Mobile Devices. In the Proceedings of the Fourth International Symposium Mobile HCI 2002, Pisa Italy: LNCS 2411, pp. 319-323.

Kortuem, G, Bauer, M., Heiber, T. and Segall, Z. (1999). Netman: The Design of a

Collaborative Wearable Computer System. Journal on Mobile Networks and Applications (MONET), 4(1), pp. 167-175.

Kristensson, P.O. (2004). Pattern Matching on Stylus Keyboards: A Powerful Approach to Faster and Easier Pen-based Text Entry. In the Conference Companion of the Symposium on User Interface Software and Technology – UIST'04, Santa Fe, New Mexico, USA: ACM, pp. 59–62.

Krum, D. M., Melby, R., Ribarsky, W. and Hodges, L. (2003). Isometric Pointer Interfaces for Wearable 3D Visualization. In the Extended Abstracts of the Conference on Human Factors in Computing Systems – CHI'03, Ft. Lauderdale, FL,USA: ACM Press, pp. 774 – 775.

Kumar, K.G., Cheok, A.D. and Zhi, Q. (2001). Hardware and software tracking for smart pen interface in wearable computing and mixed reality. In the Proceedings of the International Conference on Multimedia and Expo – ICME'01, Waseda, Tokyo, Japan: IEEE Press, pp.853-856.

Kurata, T., Kato, T., Kourogi, M., Keechul, J. and Endo, K. (2002). A functionally-distributed hand tracking method for wearable visual interfaces and its applications. In the IAPR Workshop Proceedings on Machine Vision Applications, Nara, Japan: IAPR, pp. 84–89.

Laramee, R. and Ware C. (2002). Rivalry and Interference with a Head-Mounted Dsplay. ACM Transactions on Computer-Human Interaction 9(3), pp.238-251.

Lehikoinen, J. and Roykee, M. (2001). N-fingers: A finger-based interaction technique for wearable computers. Interacting with Computers, 3, pp. 601-625.

Lemmens, P.M.C., Bussemakers, M.P. and De Haan, A. (2000). The effect of earcons on reaction times and error-rates in a dual-task vs. a single-task experiment. In the Proceedings of International Conference of Auditory Displays - ICAD'00, Atlanta, Georgia, USA: ICAD pp. 177-183.

Leplâtre, G. and Brewster, S. (2000). Designing Non-Speech Sounds to Support Navigation in Mobile Phone Menus. In the Proceedings of International Conference of Auditory Displays - ICAD'00, Atlanta, Georgia, USA: ICAD, pp. 190-199.

Litterick, I. (1981). QWERTYUIOP - dinosaur in a computer age. New Scientist, Jan 8th, pp. 66-68.

Lukowicz, P., Anliker, U., Ward, J., Tröster, G., Hirt, E. and Neufelt, C. (2002). AMON: A Wearable Medical Computer for High Risk Patients. In the Proceedings of the Sixth Intenational Symposium on Wearable Computers, Seattle, USA: IEEE Press, pp.133-134.

Lumsden, J. and Brewster, S.A. (2003). A Paradigm Shift: Alternative Interaction Techniques for Use with Mobile and Wearable Devices. In the Proceedings of the 13th Annual IBM Centers for Advanced Studies Conference - CASCON'03, Toronto, Canada: CASCON, pp. 97 – 110.

Lyons, K., Gandy, M. and Starner, T. (2000). Guided By Voices: An Audio Augmented Reality System. In the Proceedings of the International Conference of Auditory Displays - ICAD'00, Atlanta, Georgia, USA: ICAD, pp. 57 – 62.

MacKenzie, I.S., Sellen, A., and Buxton, W. (1991). A comparison of input devices in elemental pointing and dragging tasks. In the Proceedings of the SIGCHI conference on Human factors in computing systems: celebrating interdependence - CHI '91, Boston, Massachusetts, USA: ACM Press, pp. 161-166.

MacKenzie, S., Nonnecke, B., Riddersma, S., McQueen, C., and Meltz, M. (1994). Alphanumeric entry on pen-based computers. International Journal of Human-Computer Studies, 41, pp. 775-792.

MacKenzie, S., Nonnecke, B., McQueen, C., Riddersma, S., and Meltz, M. (1994). A comparison of three methods of character entry on pen-based computers. In the Proceedings of the Human Factors and Ergonomics Society 38th Annual Meeting – HEFS'94, Santa Monica, CA, USA: Human Factors Society, pp. 330-334.

MacKenzie, I.S. and Zhang, S. (1997). "The immediate usability of Graffiti", In the Proceedings of the Graphics Interface conference - GI '97, Kelowna, British Columbia, Canada: ACM pp.129-137.

MacKenzie, I.S. and Zhang, S. Z. (1999). The design and evaluation of a high performance soft keyboard. In the Proceedings of the SIGCHI conference on Human factors in computing systems: the CHI is the limit - CHI '99, Pittsburgh, Pennsylvania, USA: ACM Press, pp. 25-31.

MacKenzie, I. S., Zhang, S. X., and Soukoreff, R. W. (1999). Text entry using soft keyboards. Behaviour and Information Technology, 18, pp. 235-244.

MacKenzie, I. S., Kober, H., Smith, D., Jones, T., Skepner, E. (2001). LetterWise: Prefix-based disambiguation for mobile text input. In the Proceedings of the Symposium on User Interface Software and Technology - UIST 2001, Orlando, Florida, USA: ACM Press, pp. 111-120.

MacKenzie, I. S., Kauppinen, T., and Silfverberg, M. (2001). Accuracy measures for evaluating computer pointing devices. In the Proceedings of the Conference on Human Factors in Computing Systems – CHI'01, Seattle, Washington, USA: ACM Press, pp. 9-16.

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

MacKenzie, I.S., and Soukoreff, R. W. (2003). Phrase sets for evaluating text entry techniques. In the Extended Abstracts of the Conference on Human Factors in Computing Systems - CHI'03, Ft. Lauderdale, Florida, USA: ACM, pp. 754-755.

Mankoff. J and. Abowd, G.D. (1999). Error Correction Techniques for Handwriting, Speech, and other ambiguous or error prone systems. Georgia Tech GVU Center Technical Report, GIT-GVU-99-18.

Mankoff. J and. Abowd, G.D. (1989). Cirrin: A word-level unistroke keyboard for pen input. In the Proceedings of the Symposium on User Interface Software and Technology - UIST '98. Technical note, San Francisco, California, USA: ACM Press, pp.213-214.

Mann, S. (1998). Wearable Computing as Means for Personal Empowerment, Keynote Address for The First International Conference on Wearable Computing – ISWC'98, Cambridge, MA, USA: IEEE. See - http://wearcomp.org/wearcompdef.html

Masui, T. (1998). An efficient text input method for pen-based computers. In the Proceedings of the conference on Human factors in computing systems – CHI'98, LA, California, USA: ACM Press, pp. 328 - 335.

Matias, E., MacKenzie, I. S. and Buxton, W. (1994). Half-QWERTY: Typing with one hand using your two-handed skills. In the Conference Companion of Human Factors in Computing Systems, Boston, Massachusetts, USA: ACM Press, pp. 51-52.

Matias, E., MacKenzie, I. S. and Buxton, W. (1996). A wearable computer for use in microgravity space and other non-desktop environments. In the Conference Companion of Human Factors in Computing Systems – CHI'96, Vancouver, British Columbia, Canada: ACM Press, pp. 69-70.

Matias, E., MacKenzie, I. S. and Buxton, W. (1993). Half-QWERTY: A one-handed keyboard facilitating skill transfer from QWERTY. In the Proceedings of the INTERCHI '93 Conference on Human Factors in Computing Systems, Amsterdam, The Netherlands:ACM, pp. 88-94.

Mayzner, M. S., and Tresselt, M. E. (1965) Table of single-letter and digram frequency counts for various word-length and letter-position combinations, Psychonomic Monograph Supplements 1, pp. 13-32.

McClintock M. and Hoiem D. (1993). Minimal target size in a pen-based system. In the Abridged Proceedings of the 5th International Conference on Human-Computer Interaction, Amsterdam, The Netherlands: ACM, p.243.

McConnell, V.P. (Sept 2004), Computing Technology: Wearable Computing. Aviation Today.
www.aviationtoday.com/cgi/am/show_mag.cgi?pub=am&mon=0601&file=0601via.htm

McQueen, C., MacKenzie, S., Nonnecke, B. and Riddersma, S. (1994). A comparison of four methods of numeric entry on pen-based computers. In the Proceedings of the Graphics Interface Conference - GI '94, Toronto: Canadian: ACM, pp. 75-82.

McQueen, C., MacKenzie, I. S. and Zhang, S. X. (1995). An extended study of numeric entry on pen-based computers, In the Proceedings of Graphics Interface Conference – GI'95. Toronto: Canada: ACM, pp. 215-222

Michie, D. and Johnson, R. (1985). The Creative Computer; Machine Intelligence and Human Knowledge, Pelican Books UK, 1985.

Mills, T. and Beliveau, V. (1998). Virtual Site Visit Using Wearable Computers. The International Conference on Wearable Computing, Fairfax, Virginia, USA:

Murray, P. (2000). Wearable Computers in Battle: Recent Advances in the Land Warrior System. In the Proceedings of the Fourth International Symposium on Wearable Computers – ISWC'00, Atlanta, Georgia,USA: IEEE Press, pp. 169-170.

Najjar, L.J, Ockerman J.J and Thompson, J.C. (1998). User interface design guidelines for wearable computer speech recognition applications. In the Virtual Reality Annual International Symposium (Workshop on Interfaces for Wearable Computers) – VRAIS'98, Atlanta, Georgia, USA:
See - http://www.media.mit.edu/wearables/lizzy/vrais98.html

Nakamura, S., Tsukamoto, M. and Nishio, S. (2001). A Method of Key Input with Two Mice. In the Proceedings of the Fifth International Symposium on Wearable Computers - ISWC'01, Zurich, Switzerland: IEEE Press, pp.13-20.

NASA (2000). Body Wearable Computer Applications.
http://science.ksc.nasa.gov/payload/projects/borg/

Navon, D. (1984). Resources a theoretical soup stone? Psychological Review, 9(1), pp. 216-234.

Newsome, C. (1997). Speech therapy for slow typists, PC User 26, pp.16.

NIST (National Institute of Standards and Technology) (2001). Handwriting: An Easy Way to Enter Computer Data, from - In Performance of 50 Completed ATP Projects: Status Report - Number 2 , NIST Special Publication, pp. 950-952.

Noble, B.D., Satyanarayanan. M., Narayanan. D., Tilton, J. E., Flinn, J. and. Walker K. R. (1997). Agile application-aware adaptation for mobility. In the Proceedings of the Sixteenth ACM Symposium on Operating Systems Principles - SOSP'97, Saint Malo, France: ACM Press. pp.276 – 287.

Noyes, J. (1983). The QWERTY keyboard: A review. International Journal of Man-Machinc Studies, 18, pp. 265-281.

Nusser, S., Miller, L., Clarke, K. and Goodchild, M. (2001). Future views of field data collection in statistical surveys. Proceedings of the National Conference on Digital Government Research - DG'01, Los Angeles, CA, USA:
http://www.dgrc.org/conferences/2001_proceedings.jsp

Ockeman, J.J. and Pritchett, A.R. (1998). Preliminary Investigation of Wearable Computers for Task Guidance in Aircraft Inspection. In the Proceedings of the International Symposium on Wearable Computers - ISWC'98, Pittsburgh, PA,USA:IEEE Press, pp. 33-40.

Orth, M., Post, R. and Cooper, E. B. (1998). Fabric Computing Interfaces. In the Proceedings of Conference on Human Factors in Computing Systems - CHI '98, Los Angeles, USA: ACM Press, pp.331-332.

Panasonic (2002). Toughbook Range.
http://www.panasonic.com/computer/toughbook/home.asp

Peckham, J. (1994).Behavioral aspects of speech technology: industrial systems. In Syrdal, A., Bennett, R. and Greenspan, S. (Eds) Applied Speech Technology, CRC press, pp.469-486

Petrie, H., Furner, T. and Strothotte, T. (1998). Design Lifecycles and Wearable Computers for Users with Disabilities. In the Proceedings of the First Workshop on Human Computer Interaction with Mobile Devices Glasgow, University UK.
See - http://www.dcs.gla.ac.uk/~johnson/papers/mobile/HCIMD1.html

Piekarski, W. and Thomas, B. H. (2002). ARQuake: The Outdoor Augmented Reality Gaming System. ACM Communications, 45(1), pp.36-38.

Pirhonen, A., Brewster, S.A. and Holguin, C. (2002). Gestural and Audio Metaphors as a Means of Control for Mobile Devices. In the Proceedings of the conference on Human factors in computing systems: Changing our world, changing ourselves - CHI'02, Minneapolis, MN, USA: ACM Press, pp. 291-298.

Potosnak, K. M. (1988). Keys and keyboards. In Helander, M. (Ed.), Handbook of Human-Computer Interaction. Amsterdam: Elsevier, pp. 475-494.

Randell, C. and Muller, H. (2000). The Shopping Jacket: Wearable Computing for the Consumer. In Thomas, P. (Ed.) Personal Technologies 4(4), Springer Verlag, pp. 241-244.

Randall, C. (2000). Computerised Clothing will benefit Textile Manufacturers, Technical Textiles International, 10(7), pp. 3-27.

Randall, C. and Muller, H. (2002). The Well Mannered Wearable Computer, Personal and Ubiquitous Computing, 6(1), pp.31-36.

Rantanen, J., Alfthan, N., Impiö, J., Karinsalo, T., Malmivaara, M., Matala, R., Mäkinen, M., Reho, A., Talvenmaa, P., Tasanen, M. and Vanhala, J. (2000). Smart Clothing for the Arctic Environment. In the Proceedings of the Fourth International Symposium on Wearable Computers – ISWC'00, Atlanta, Georgia, USA: IEEE Press pp.15-23.

Rantanen, J., Impiö, J., Karinsalo, T., Malmivaara, M., Reho, A., Tasanen, M. and Vanhala, J. (2001). Smart Clothing Prototype for the Arctic Environment. Personal and Ubiquitous Computing. 6(1), pp. 3-16.

Rekimoto, J. (1997). Pick and drop: a direct manipulation technique for multiple computer environments. In the Proceedings of the Symposium on User Interface Software and Technology - UIST'97, Banff, Alberta, Canada: ACM Press, pp. 31 - 39.

Rekimoto, J. and Saitoh, S. (1999). Augmented Surfaces: A Spatially Continuous Workspace for Hybrid Computing Environments. In the Proceedings of the SIGCHI conference on Human factors in computing systems: the CHI is the limit - CHI'99, Pittsburgh, Pennsylvania, USA: ACM Press, pp.378-385.

Rekimoto, J. (2001). GestureWrist and GesturePad: Unobtrusive Wearable Interaction Devices. In the Proceedings of the Fifth International Symposium on Wearable Computing – ISWC'01, Zurich, Switzerland: IEEE Press, pp.21 – 31.

Reitmayr, G. and Schmalstieg, D. (2001). A Wearable 3D Augmented Reality Workspace. In the Proceedings of the Fifth International Symposium on Wearable Computers – ISWC'01, Zurich, Switzerland: IEEE Press, pp.165 – 167.

Rensing, N.M., Weststrate, E., Zavracky, P.M., Chandler, M., Nobel, K.R., Helfter, S., Kinsky, M., Gold, M. and Martin, B. (2002). Threat response: a compelling application for wearable computing. In the Proceedings of the Sixth International Symposium on Wearable Computing – ISWC'02, Seattle, USA: IEEE Press, pp. 152-153.

Rhodes, B.J. (1998). The Wearable Remembrance Agent: a system for augmented memory. In the Proceedings of the Second International Symposium on Wearable Computing – ISWC'98, Pittsburgh, PA, USA: IEEE press, pp. 123-128.

Rhodes, B. and Crabtree, B. (1998). 'Wearable computing and the remembrance agent' BT Technology Journal, 16,3: pp.118-124.

Rosenberg, R. (1998). Computing without Mice and Keyboards: Text and Graphic Input Devices for Mobile Computing. Doctoral dissertation. Department of Computer Science, University College, London.

Ross, A. and Blacsh, B, (2000). Wearable interfaces for orientation and wayfinding. In the Proceedings of the Fourth International Conference on Assistive Technologies, Arlington, Washington, USA: ACM Press. pp. 193 – 200.

Rosson, M.B., Carrol, J.M. and Bellamy, R. (1990). Smalltalk Scaffolding: A case study of minimalist instruction. In the Proceedings of the SIGCHI conference on Human factors in computing systems: Empowering people – CHI'90, Seattle, Washington, USA: ACM Press, pp. 423-429.

Rudnicky, A.I., Reed, S.D., Thayer, E.H. (1996). SpeechWear: A mobile speech system. In the Proceedings of the International Conference on Spoken Language Processing - ICSLP'96, Philadelphia, PA, USA: ICSLP, pp. 538-514 CD Rom.

Sawhney, N. and Schmandt, C. (2000). Nomadic Radio: Speech and Audio Interaction for Contextual Messaging in Nomadic Environments. ACM Transactions on Computer-Human Interaction with Mobile Systems, 7(3), pp. 353-383.

Sawhney, N. (1997). Message Swatch: Temporal Representation for Information Augmentation. Technical Report for Intelligence Augmentation (MAS 963), Pattie Maes, MIT Media Lab.

Sazawal, V., Want, R., and Borriello, G. (2002). The Unigesture Approach: One Handed Text Entry for Mobile Devices. In the Proceedings of Mobile HCI 2002, Pisa, Italy: LNCS, pp.256 -270.

Schmidt, A., Gellersen, H.W. and Merz, C. (2000). Enabling Implicit Human Computer Interaction - A Wearable RFID-Tag Reader. In the Proceedings of the International Symposium on Wearable Computers – ISWC'00, Atlanta, GA, USA: IEEE Press, pp.193-194.

Schmidt, A., Gellersen, H.W., Beigl, M. and Thate, O. (2000). Developing User Interfaces for Wearable Computers: Don't Stop to Point and Click. In the proceedings of the International Workshop on Interactive Applications of Mobile Computing – IMC'00, Warnemünde. Institut für Grafische Datenverarbeitung, Fraunhofer Gesellschaft. http://www.comp.lancs.ac.uk/~albrecht/pubs/pdf/schmidt_imc2000_uis-for-wearables.pdf

Schumacher, E. H., Seymour, T. L., Glass, J. M., Kieras, D. E., and Meyer, D. E. (2001). Virtually perfect time sharing in dual-task performance: Uncorking the central attentional bottleneck. Psychological Science, 12, pp. 101-108.

Schvaneveldt, R. W., Reid, G. B., Gomez, R. L., and Rice, S. (1998). Modelling mental workload. Cognitive Technology, 3, pp.19-31.

Sears, A. (1991). Improving Touchscreen Keyboards: Design issues and a comparison with other devices. Interacting with Computers, 3(3), pp. 252-269.

Sears, A. and Shneiderman, B. (1991). High precision touchscreens: Design strategies and comparisons with a mouse, International Journal of Man-Machine Studies, 34(4), pp.593-613.

Sears, A., Plaisant, C. and Shneiderman, B. (1992). A new era for high precision touchscreens. Advances in Human-Computer Interaction, Vol.30, pp.1-33.

Sears, A., Revis, J., Swatski, R., Crittenden, D. and Shneiderman. B.(1993). Investigating touchscreen typing: the effect of keyboard size on typing speed. Behaviour and Information Technology, 12(17), pp. 17-22.

Sears, A., Jacko, J.A., Chu, J. and Moro, F. (2001). The role of visual search in the design of effective soft keyboards. Behaviour and Information Technology, 20(3), pp. 159-166.

Shneiderman, B. (1993). Sparks of Innovation in Human-Computer Interaction, Ablex Publishing Corporation, Norwood, NJ.

Siegel, J. and Bauer, M. (1997). A field usability evaluation of a wearable system. In the Proceedings of the First International Symposium on Wearable Computers, Cambridge, Massachusetts, USA: IEEE Press C, pp.18-22.

Simpson, C.A., McCauley, M.E., Roland, E.F., Ruth, J.C. and Williges, B.H. (1985). System design for speech recognition and generation, Human Factors 27, pp.115-141.

Silfverberg, M., MacKenzie, I. S., and Kauppinen, T. (2001). An isometric joystick as a pointing device for handheld information terminals. In the Proceedings of the Graphics Interface Conference – GI'01, Toronto, Canada: Canadian Information Processing Society, pp. 119 -126.

Sitemaster UK, Wearable Asbestos Control (2004).
http://www.sitemasteruk.com/

Slater, M. (1997) User interfaces: beyond keyboards. June, Microprocessor Report pp.11-15.

Smailagic, A. and Siewiorek, D. (2002). Application Design for Wearable and Context-Aware Computers. IEEE Pervasive Computing 1(4), pp.20-29.

Smailagic, A. (1997). ISAAC: A Voice Activated Speech Response Computers. In the Proceedings of the First International Symposium on Wearable Computers - ISWC'97, Cambridge, Massachusetts, USA: IEEE Press, pp.183 – 184.

Smith, B., Bass, L. and Siegel, J (1995), On site maintenance using a wearable computer system. In the Proceedings of the SIGCHI conference on Human factors in computing system – CHI'95, Denver, Colorado, USA: ACM Press, pp. 119-120.

Smith, B.A., Ho, J., Ark, W. and Zhai, S. (2000). In the Proceedings of the ACM Symposium on Eye Tracking Research and Applications, Palm Beach Gardens, FL, USA: ACM, pp.117 – 122.

Soukoreff, R. W. and MacKenzie, I. S. (1995). Theoretical upper and lower bounds on typing speed using a stylus and soft keyboard. Behaviour and Information Technology, 14, pp.370-379.

Soukoreff, R. W. and MacKenzie. I.S. (2004). Recent Developments in Text-Entry Error Rate Measurement. In the Extended abstracts of the 2004 conference on Human factors and computing systems - CHI 2004, Vienna, Austria: ACM Press, pp.1425 - 1428.

Starner, T., Mann, S., Rhodes, B., Healey, J., Russell, K., Levine, J. and Pentland, A. (1995). 'Wearable Computing and Augmented Reality', Technical Report, Media Lab Vision and Modelling Group RT-355, MIT.

Starner, T. (2004). Mobile Interaction Interfaces (Keynote Speech).International Conference on Advances in Computer Entertainment Technology – ACE'04, Singapore.

Steadman, A. W., Moore, P .M., Kalawsky, R.S., Aung, M., Purcell, J., Reech, C. and York, T. (1999). It's not what you wear, it's how you wear it: Human factors of wearable computers. In the Proceedings of the Human Factors and Ergonomics Society 43rd Annual Meeting, Houston, Texas, USA: pp. 1050 – 1054.

Strayer, D. L. and Johnston, W. A. (2001). Driven to distraction: Dual-task studies of simulated driving and conversing on a cellular phone. Psychological Science, 12, pp. 462-466

Strayer, D. L., Drews, F. A., and Johnston, W. A. (2003). Cell phone induced failures of visual attention during simulated driving. Journal of Experimental Psychology, 9, pp. 23-52.

Sugimoto, M. and Takahashi, K. (1996). Single hand key card for mobile devices. In the Conference companion on Human factors in computing systems: common ground – CHI'96, Vancouver, British Columbia, Canada: ACM Press, pp. 7 – 8.

Svensson, E., Angelborg-Thanderz, M., Sjoberg, L., Olsson, S. (1997). Information complexity--mental workload and performance in combat aircraft. Ergonomics. 1997 Mar; 40(3), pp. 362-380.

Tang, A., Owen, C., Biocca, F., and Mou, W. (2003). Comparative effectiveness of augmented reality in object assembly. In the Proceedings of the conference on Human factors in computing systems - CHI'03, Ft. Lauderdale, Florida, USA: pp. 73 – 80.

Thomas, B.H., Tyerman, S. and Grimmer. K. (1997). An Evaluation of Three Input Mechanisms for Wearable Computers. In the Proceedings of the First International Symposium on Wearable Computers - ISWC'97, Cambridge, Massachusetts, USA: IEEE Press, pp. 2-9.

Thomas, B.H., Grimmer, K., Zucco, J. and Milanese, S. (2002). Where Does the Mouse Go? An Investigation into the Placement of a Body-Attached TouchPad Mouse for Wearable Computers. Personal and Ubiquitous Computing 6(2), pp. 97-112.

Toney, A., Mulley, B., Thomas, B. H. and Piekarski, W. (2002). Minimal Social Weight User Interactions for Wearable Computers in Business Suits. In the Proceedings of the Sixth International Symposium on Wearable Computers, Seattle, Washington, USA: IEEE Press, pp. 57 – 64.

Toub, R. (1996). The Importance of Usability Studies and Experimental Design Evaluations for Wearable Computers, Position paper for the Workshop on Wearable Computer Systems, Seattle, USA: IEEE Press pp.19-21.

Tsukada, K. and Yasumura, M. (2002). Ubi-Finger: Gesture Input Device for Mobile Use.Proceedings of APCHI 2002, In the Proceedings of the Fifth Asia Pacific Conference on Computer Human Interaction, Beijing, China: IFIP, pp.388-400.

Veniola, D.D. and Neiberg, F. (1994). "T-cube: A fast, self-disclosing pen-based alphabet. In the Proceedings of the SIGCHI conference on Human factors in computing systems: celebrating interdependence, Boston, Massachusetts,USA: ACM Press, pp. 265-270.

Wearable Computing Lab, (2004).
http://www.wearable.ethz.ch/qbic.0.html

WetPc, (2002).
http://wetpc.com.au/html/technology/interface.htm

Wigdor, D., Balakrishnan, R. (2003). TiltText: Using tilt for text input to mobile phones. In the Proceedings of the Symposium on User Interface Software and Technology – UIST'03, Vancouver, BC, Canada: ACM Press, pp. 81-90

Wolf, C., Glasser, G. and Fujisaki. A. R. (1991). An evaluation of recognition accuracy for discrete and run-on writing. In the Proceedings of the Human Factors Society 35th Annual Meeting, Santa Monica, CA: Human Factors Society, pp. 359-363.

Woolley, S. I., Cross, J. W., Ro, S., Foster, R., Reynolds, G., Baber, C., Bristow, H. and Schwirtz, A. (2003). Forms of Wearable Computer, In IEE Euro Wearable, Birmingham, UK, BCS. http://conferences.iee.org/eurowearable/

Xybernaut Commercial Product page. (2004).
http://www.xybernaut.com/Solutions/product/ma_v_tech.htm

Xybernaut – White paper on Speech (2004*).
www.xybernaut.de/downloads/daten/ anwendungen/white_paper_on_speech.pdf

Yankelovich, N., Levow, G and Marx, M. (1995). Designing Speech Acts: issues in speech user interfaces. In the Proceedings of the SIGCHI conference on Human factors in computing systems - CHI'95, Denver, Colorado, USA: pp. 369 - 376.

Zhai, S., Hunter, M. and Smith. B.A. (2000). The Metropolis Keyboard - an exploration of quantitative techniques for virtual keyboard design. In the Proceedings of the 13th Symposium on User Interface Software and Technology – UIST'00, San Diego, California, USA: ACM Press, pp. 119-218.

Zhai, S., Kristensson, P-O, Smith, B.A. (2004). In search of effective text interfaces for off the desktop computing, Interacting with Computers, in press.

Appendices

Appendix 1.

The table below shows the Time and Errors for Chapter 3, the text input experiment. The table shows the time and errors for the participants while both standing and walking.

Participant	Time	Errors
Speech Stand		
1	25.89	4
2	42.03	2
3	27.02	12
4	32.54	4
5	36.00	9
6	41.04	5
7	26.70	15
8	39.53	4
9	33.34	7
10	26.45	12
Speech Walk		
1	34.01	13
2	43.04	5
3	28.33	23
4	33.21	4
5	37.23	8
6	42.98	9
7	28.04	17
8	41.54	6
9	34.01	8
10	28.00	13
On-Screen Stand		
1	138.08	3
2	242.68	0
3	163.69	2
4	263.59	3
5	163.45	4
6	175.88	2
7	349.72	5
8	289.52	6
9	262.23	5
10	259.63	5
On-Screen Walk		
1	147.04	6
2	277.83	4
3	204.98	6
4	263.12	3

5	188.90	4
6	177.07	7
7	464.21	8
8	346.66	7
9	203.14	8
10	231.15	12
Handwriting Stand		
1	144.41	1
2	134.89	1
3	143.76	2
4	189.96	5
5	132.94	2
6	228.75	5
7	154.88	3
8	267.94	2
9	126.27	4
10	159.07	3
Handwriting Walk		
1	167.13	2
2	193.02	3
3	225.21	2
4	230.12	2
5	164.09	2
6	230.00	5
7	173.01	3
8	289.50	2
9	181.01	3
10	177.24	4
Wrist Key Stand		
1	105.09	2
2	172.70	5
3	184.86	2
4	212.90	3
5	126.80	5
6	232.02	4
7	294.07	7
8	128.56	3
9	273.19	2
10	193.22	6
Wrist Key Walk		
1	181.39	5
2	198.70	6
3	207.62	8
4	233.51	3
5	132.95	3
6	249.07	5

7	330.83	7
8	257.41	8
9	240.00	5
10	209.28	7

The table below shows the workload score and six sub-scale scores for Chapter 3 the text input experiment. The table shows the time and errors for the participants while both standing and walking.

Key:
md = mental demands, pd = physical demands, td = temporal demands, ef = effort, pe = performance and rf = frustration.

Participant	md	pd	td	ef	pe	rf	WWL
Speech Stand							
1	20	50	20	20	75	50	41
2	5	10	25	70	75	50	39
3	10	15	10	10	60	35	25
4	20	10	15	15	60	60	38
5	5	40	10	15	20	15	17
6	5	5	20	15	45	5	21
7	15	25	35	45	50	45	41
8	15	10	20	45	40	30	29
9	10	10	5	40	20	20	16
10	5	5	5	50	80	25	33
Speech Walk							
1	15	20	20	25	45	40	32
2	15	15	10	25	45	70	43
3	20	15	40	30	25	30	29
4	25	55	35	50	35	40	38
5	35	70	20	30	65	30	40
6	100	20	35	55	65	35	52
7	25	40	40	55	45	50	42
8	20	25	15	40	55	70	37
9	25	30	65	50	35	60	44
10	25	35	30	20	30	20	31
OnScreen Stand							
1	15	30	20	10	5	40	14
2	75	15	55	75	10	30	33
3	10	5	25	10	5	5	9
4	5	5	15	15	10	15	11
5	25	45	40	35	35	40	37
6	30	15	50	45	30	30	33
7	10	20	25	25	40	30	27
8	25	30	25	40	40	40	33
9	35	50	50	55	50	50	50
10	20	20	25	30	30	35	27
OnScreen Walk							
1	20	25	40	15	10	5	19

2	80	55	35	90	20	25	57
3	75	5	30	75	80	25	47
4	70	55	35	35	45	30	44
5	40	40	30	50	15	35	35
6	35	60	60	65	30	65	53
7	65	40	50	40	40	40	46
8	55	40	55	55	25	25	42
9	80	80	50	70	50	65	66
10	60	35	65	30	80	15	48
Handwriting Stand							
1	10	45	40	10	10	10	27
2	5	10	10	5	5	5	7
3	20	25	20	20	40	50	35
4	5	55	50	75	20	90	58
5	5	5	20	10	5	20	10
6	10	20	25	15	5	20	17
7	35	30	30	45	50	50	44
8	20	5	45	25	25	25	27
9	5	5	25	25	5	15	8
10	5	5	25	30	5	20	12
Handwrting Walk							
1	10	20	45	35	30	25	30
2	25	25	35	35	40	25	33
3	35	40	15	40	15	15	25
4	20	40	50	20	15	40	32
5	20	45	35	55	40	40	41
6	35	45	40	15	25	10	28
7	45	40	55	45	50	50	50
8	60	50	35	30	25	25	40
9	20	20	30	20	20	25	21
10	50	40	40	40	25	15	32
Wrist Key Stand							
1	70	100	20	75	35	65	75
2	20	40	55	45	60	25	43
3	20	30	60	45	35	25	35
4	60	20	70	35	30	60	48
5	40	20	60	45	35	25	35
6	60	40	55	35	30	60	43
7	20	90	50	55	45	60	52
8	20	85	40	40	50	55	54
9	15	35	75	75	85	30	63
10	20	70	45	45	50	50	52
Wrist Key Walk							
1	30	65	70	60	35	30	50
2	15	70	55	60	35	40	46
3	90	75	70	45	15	20	52

4	50	40	35	40	25	75	44
5	75	30	75	75	35	75	61
6	60	70	65	70	80	90	73
7	55	75	30	20	25	30	36
8	100	90	90	70	30	80	69
9	45	35	50	40	40	40	41
10	30	30	25	25	50	30	35

Appendix 2.

The table below shows the Time and Errors for Chapter 4, the target selection experiment. The table shows the time and errors for the participants while both standing and walking.

Participant	Time	Errors
Stylus Stand		
1	572.21	86
2	523.29	49
3	566.17	37
4	562.11	38
5	551.32	61
6	544.31	70
7	536.39	52
8	541.09	73
9	512.28	63
10	543.29	66
Stylus Walk		
1	679.10	78
2	675.19	24
3	693.16	87
4	612.21	20
5	615.09	33
6	665.19	34
7	644.03	54
8	660.85	62
9	592.10	41
10	630.06	77
Mouse Stand		
1	1562.17	62
2	1487.22	28
3	1495.25	26
4	1413.41	29
5	1401.50	21
6	1509.21	40
7	1519.22	24
8	1524.24	21
9	1515.26	51
10	1208.04	51
Mouse Walk		
1	1793.40	22
2	1537.51	55

3	1669.40	25
4	1602.24	36
5	1651.26	30
6	1694.25	19
7	1587.97	18
8	1598.27	21
9	1549.02	48
10	1514.00	39
Track Stand		
1	653.15	77
2	811.94	94
3	713.90	67
4	806.45	72
5	641.17	70
6	789.49	79
7	765.94	80
8	737.45	71
9	731.33	77
10	821.27	79
Track walk		
1	843.96	79
2	938.67	89
3	1098.72	65
4	1032.37	70
5	935.15	78
6	1142.65	86
7	1356.34	78
8	894.29	87
9	1074.26	93
10	1121.03	83

The table below shows the workload score and six sub-scale scores for Chapter 4, the target selection experiment. The table shows the scores for the participants while both standing and walking.

Key:
md = mental demands, pd = physical demands, td = temporal demands, ef = effort, pe = performance and rf = frustration.

participant	md	pd	td	ef	pe	rf	wwl
stylus stand							
1	15	20	25	10	40	5	19
2	15	5	30	10	35	40	22
3	5	5	5	5	30	5	9
4	30	25	5	45	25	30	27
5	35	40	25	10	25	5	23
6	20	40	40	40	25	25	32
7	20	15	5	15	15	10	13
8	30	30	35	35	35	35	33
9	30	20	50	45	40	45	38
10	15	25	35	30	25	20	25
stylus walk							
1	40	55	40	30	40	10	36
2	65	45	55	40	40	25	45
3	35	40	15	5	10	10	19
4	40	55	50	40	50	50	47
5	5	30	40	35	35	15	27
6	30	25	20	20	20	15	22
7	15	65	60	60	30	30	43
8	20	15	10	10	25	25	17
9	25	20	30	40	45	35	32
10	50	55	45	35	45	30	43
mouse stand							
1	40	65	60	55	20	65	51
2	25	90	45	50	55	50	52
3	30	80	45	55	50	65	54
4	35	30	20	35	90	40	42
5	55	65	35	30	30	35	42
6	60	40	25	20	65	25	39
7	10	60	60	60	30	60	47
8	25	45	70	80	50	70	57
9	25	100	65	10	30	10	40
10	50	85	65	25	20	45	48
mouse walk							
1	65	75	50	70	50	90	67

2	30	70	65	50	60	65	57
3	30	50	50	70	50	90	58
4	20	65	55	55	25	55	46
5	30	70	50	50	25	55	47
6	50	80	10	75	45	55	52
7	35	80	25	65	35	55	49
8	10	45	45	60	45	20	37
9	35	60	40	45	35	40	42
10	25	70	55	50	20	20	40
track stand							
1	40	45	55	55	40	70	53
2	25	65	40	50	45	35	43
3	25	40	45	35	25	20	31
4	15	100	90	85	60	75	66
5	45	70	75	90	75	65	66
6	35	70	65	65	55	55	60
7	60	60	75	45	30	25	43
8	35	30	40	35	25	30	31
9	20	95	90	30	40	10	45
10	15	40	65	30	10	15	30
track walk							
1	70	80	75	70	45	90	69
2	45	60	60	70	55	85	62
3	15	90	80	80	45	80	53
4	40	50	65	90	40	35	49
5	55	70	35	45	45	25	45
6	45	80	45	70	75	80	70
7	55	65	85	80	40	90	75
8	20	45	95	90	50	80	66
9	85	85	80	95	75	85	85
10	40	75	75	90	60	90	80

Appendix 3.

The table below shows the Time and Errors for Chapter 5, the dragging and dropping experiment. The table shows the time and errors for the participants while both standing and walking.

Participant	Time	Errors
Stylus Stand		
1	248.46	81
2	253.87	57
3	275.51	90
4	232.23	69
5	210.61	48
6	318.63	99
7	243.87	72
8	253.80	63
9	318.60	90
10	594.98	93
Stylus Walk		
1	386.64	66
2	480.63	105
3	410.42	75
4	534.05	93
5	543.53	96
6	540.43	99
7	648.33	108
8	469.80	111
9	372.60	81
10	702.01	129
Mouse Stand		
1	263.03	201
2	354.54	223
3	321.66	352
4	273.58	253
5	401.57	377
6	357.69	345
7	355.78	216
8	288.56	260
9	425.93	409
10	367.25	322
Mouse Walk		
1	614.52	263
2	522.17	354
3	1026.11	321

4	594.19	273
5	507.66	401
6	702.67	357
7	594.73	355
8	918.45	288
9	1080.39	425
10	648.31	367
Track Stand		
1	339.65	113
2	244.67	167
3	281.85	325
4	257.51	242
5	400.00	364
6	310.73	212
7	307.81	275
8	255.43	236
9	248.04	207
10	370.08	378
Track walk		
1	602.88	179
2	551.63	231
3	627.53	341
4	594.53	271
5	719.41	387
6	829.11	325
7	765.19	332
8	733.59	297
9	588.32	265
10	874.20	361

The table below shows the workload score and six sub-scale scores for Chapter 5, the dragging and dropping experiment. The table shows the scores for the participants while both standing and walking.

Key:
md = mental demands, pd = physical demands, td = temporal demands, ef = effort, pe = performance and rf = frustration.

Participant	md	Pd	td	ef	pe	rf	WWL
Stylus Stand							
1	100	55	50	45	90	50	73
2	5	20	55	35	25	20	30
3	10	5	10	5	10	5	7
4	5	30	25	30	25	40	29
5	25	15	10	10	10	25	15
6	45	30	40	25	25	25	29
7	15	30	35	30	10	25	23
8	30	25	45	55	30	25	37
9	10	20	15	5	5	40	15
10	30	10	15	35	40	45	26
Stylus Walk							
1	40	45	20	40	25	30	32
2	30	45	35	45	50	45	42
3	5	5	30	40	50	45	26
4	5	10	5	30	55	55	21
5	35	35	15	15	5	25	25
6	20	15	25	25	30	45	26
7	35	40	40	20	25	35	32
8	55	50	50	40	35	55	46
9	35	25	45	60	25	60	40
10	5	40	40	45	30	40	34
Mouse Stand							
1	45	60	70	95	75	45	68
2	55	95	45	60	70	55	70
3	55	50	50	50	45	25	48
4	35	40	75	70	50	74	55
5	55	40	60	70	85	90	67
6	60	65	85	75	40	70	68
7	85	85	60	100	95	45	82
8	35	55	95	35	75	40	60
9	45	65	65	55	45	55	53
10	65	65	55	75	85	65	74
Mouse Walk							
1	20	75	85	55	90	90	77

2	75	85	95	95	95	80	90
3	70	65	50	45	30	100	54
4	30	100	20	55	30	90	49
5	60	60	65	65	60	65	62
6	85	95	80	80	70	35	72
7	20	50	65	60	65	60	53
8	70	100	90	100	75	100	93
9	80	70	55	75	45	45	62
10	45	60	70	95	60	95	74
Track Stand							
1	30	50	45	70	55	65	59
2	35	55	70	80	80	70	73
3	25	40	65	80	70	80	69
4	65	70	85	60	45	35	63
5	40	45	50	70	60	75	59
6	70	60	65	80	75	70	69
7	55	60	60	55	55	60	58
8	35	65	70	90	60	90	70
9	50	65	75	85	70	80	73
10	55	45	65	75	60	40	54
Track walk							
1	65	25	85	25	60	60	50
2	15	35	65	75	65	90	62
3	65	70	90	85	90	50	80
4	70	85	60	65	55	90	77
5	45	65	65	45	55	70	58
6	55	10	40	35	60	50	44
7	30	100	100	80	80	95	95
8	70	85	90	65	85	80	80
9	55	80	80	35	25	40	53
10	80	70	70	80	85	85	80

Appendix 4.

The table below shows the Time and Errors for Chapter 6, the trajectory-based interaction experiment. The table shows the time and errors for the participants while both standing and walking.

.

Participant	Time	Errors
Stylus Stand		
1	216.09	52
2	275.23	81
3	367.21	155
4	270.18	76
5	324.56	89
6	485.92	150
7	264.53	74
8	291.61	90
9	259.11	68
10	405.89	114
Stylus Walk		
1	649.80	227
2	523.19	316
3	529.32	197
4	593.37	201
5	459.13	435
6	729.28	453
7	547.73	279
8	459.99	331
9	431.51	232
10	792.09	368
Mouse Stand		
1	513.67	482
2	853.02	165
3	685.17	441
4	1101.42	476
5	518.47	451
6	534.57	392
7	1355.34	626
8	703.04	323
9	1075.19	478
10	1333.47	576
Mouse Walk		
1	2099.02	785
2	1863.46	751
3	1911.95	892

4	1758.15	676
5	2173.56	959
6	1582.67	710
7	1263.13	807
8	1501.39	792
9	2462.85	961
10	1749.45	861
Track Stand		
1	495.32	228
2	487.31	314
3	885.65	196
4	599.38	445
5	453.45	410
6	73.58	242
7	672.02	331
8	922.43	206
9	567.44	354
10	76.86	294
Track walk		
1	1209.91	266
2	1155.77	510
3	1252.39	205
4	1382.41	453
5	993.85	494
6	1188.17	389
7	1090.97	416
8	966.33	290
9	1269.39	477
10	1150.67	382

The table below shows the workload score and six sub-scale scores for Chapter 6, the trajectory-based interaction experiment. The table shows the scores for the participants while both standing and walking.

Key:
md = mental demands, pd = physical demands, td = temporal demands, ef = effort, pe = performance and rf = frustration.

	md	pd	td	ef	op	fr	wwl
stylus stand							
1	25	25	35	10	15	15	21
2	35	35	20	40	35	35	31
3	20	60	20	40	35	25	38
4	5	40	25	15	25	45	27
5	15	50	50	55	30	35	42
6	25	35	30	60	45	60	42
7	35	25	40	35	35	40	35
8	30	25	25	15	10	10	17
9	35	30	60	10	20	20	34
10	10	35	20	20	25	15	18
stylus walk							
1	15	50	35	25	5	35	29
2	30	80	45	65	20	90	56
3	10	20	25	35	45	55	36
4	55	55	45	50	50	45	50
5	50	65	70	65	55	55	61
6	5	35	30	60	35	70	40
7	30	45	30	75	60	25	43
8	20	25	15	15	25	45	26
9	15	40	45	75	65	85	58
10	60	65	35	25	35	40	47
mouse stand							
1	85	85	80	100	80	100	87
2	40	90	75	75	85	65	79
3	95	75	100	100	100	100	95
4	20	65	65	75	75	50	66
5	75	90	70	65	35	100	70
6	10	30	40	35	70	60	39
7	45	65	40	65	65	80	60
8	50	55	65	60	65	60	61
9	70	95	95	80	60	75	86
10	60	35	40	40	80	70	54
2	30	75	95	85	70	90	69
mouse walk	50	70	90	75	65	95	80
4	50	[illegible]	60	[illegible]	85	[illegible]	[illegible]

5	60	95	85	70	80	90	84
6	70	80	35	75	50	25	49
7	45	60	70	65	65	75	64
8	70	95	70	75	60	80	78
9	10	95	100	95	90	85	71
10	30	95	80	90	100	95	83
track stand							
1	55	35	50	65	20	60	48
2	50	75	80	85	25	85	70
3	15	65	40	40	60	35	46
4	100	80	65	80	50	60	67
5	25	100	75	80	45	55	65
6	20	60	60	65	55	70	71
7	65	70	75	70	60	70	62
8	30	85	75	70	55	65	69
9	10	95	90	75	65	70	68
10	65	55	75	85	60	80	70
track walk							
1	60	70	90	90	75	80	76
2	5	100	95	20	60	25	50
3	45	25	70	70	75	30	54
4	65	75	95	90	95	95	88
5	90	90	85	70	55	55	72
6	20	80	60	70	45	75	63
7	70	60	45	65	60	65	59
8	60	65	90	80	70	70	73
9	25	80	30	80	65	90	66
10	70	90	75	55	55	55	65

Appendix 5.

An Initial Experiment to Evaluate Three Wearable Computing Input Devices for Target Selection

An Investigation into three pointing devices for use with a wearable computing system examining usability, speed of use and cognitive workload.

5.1 Introduction

This experiment compares three input methods that can be used for target selection with a wearable computing system. The input methods were voice recognition, stylus and an off-table mouse. The experiment contained two movement conditions and used one target size. This is the first experiment in this chapter and is a preliminary investigation into pointing devices for wearable computers. This first experiment uses time as a dependent variable and a cognitive workload scoring system (NASA TLX). This is an initial study was designed with three goals in mind:

1. To introduce the user to the style and experimental research methodology in this chapter.
2. To compare the three input methods in terms of speed and workload.
3. To detect if any of the input methods or any experimental techniques used could cause future problems in the next generation of experiments.

5.2 Rationale

This experiment is the first experiment to evaluate wearable pointing devices. In the second chapter of the thesis it was identified that there was a lack of empirical research into the use of pointing devices, when used as part of a wearable computing system. This experiment acts as a pilot study to examine the pointing devices' performance and also gives an insight into any problems that may occur during the experiment, so that these may be corrected in future experiments.

5.3 Null Hypothesis

h1. The pointer type does not affect the time taken to complete the text input tasks.

h2. The movement condition of the participant standing and walking has no effect on the time taken to complete the task.
h3. The pointer type does not affect the subjective task load of the participant.

5.4 METHOD

5.4.2 Apparatus

The apparatus used in this trial consisted of a Xybernaut MAIV wearable computer worn in an Agora vest based system with a fold down touch screen (figure 5), the underlying operating system was Win 2000. Microsoft Explorer 5 was used as the base input software. The following were used in the trial to input text: a Xybernaut stylus, finger-worn mouse and Dragon Naturally Speaking voice recognition software, along with a Labtec single ear piece-microphone headset. Two HTML files contained the target objects to be selected.

Figure 5: Participant using the Xbernaut touch-screen.

5.4.3 Design

The experiment consisted of a 2 by 3 between-subjects design. The 2 independent variables (IV) were movement conditions and pointer type. The 2 levels of the movement condition IV were standing and walking. The levels of text input IV types were stylus, finger mouse and voice recognition. There were 2 dependent variables measured, the time to point and the cognitive workload rating (NASA TLX). 30 participants took part in the trials; 10 were assigned to each input condition at random. The movement condition (standing and walking) were counterbalanced using a 2x2 Latin-square design.

Standing 5 Participants Objects 1	Walking 5 Participants Objects 2
Walking 5 Participants Objects 1	Standing 5 Participants Objects 2

Figure 5.1: Experimental balancing.

5.4.4 Procedure

10 participants were randomly assigned to each of the 3 pointer conditions. Each participant put on the vest based wearable computing system. Depending on the handedness of the participant, the touch screen could be moved to either the left or right had side of the participant. They were then asked to walk around the room to see if the vest was comfortable; if not, the vest could be adjusted to fit. Each of the 30 participants spent a 5 minute session, training/understanding the principles of their assigned input system. The participants were requested to use their pointing device whilst walking and whilst standing still; the order of this was done in relation to the counterbalancing. The pointing systems used were off-table mouse, voice and stylus. If voice was used, the participants did a small voice-training exercise which was to read a set amount of 'Alice in Wonderland' as specified by the speech recognition system.

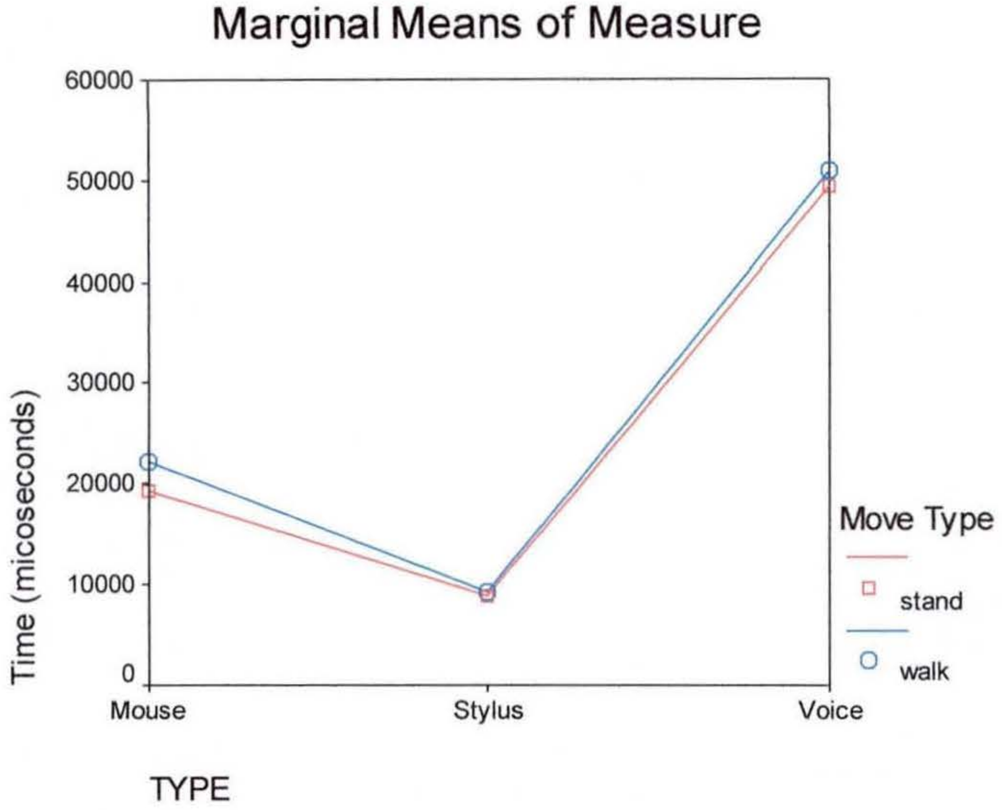
The objects they had to point to were displayed on the screen in MS Explorer. Each object was randomly positioned, the random screen positioning given by a software package called id Test. Each target was 25mm square. This size had been previously used by Toube (1996), and it could spatially accommodate a textual number from one to ten. The words representing the numbers from one to ten were selected as voice activators for the targets, as they are commonly used and are widely known. The numbers were selected 'blind' so the same order of words would not be repeated during the experiment. This was done for both the standing and walking conditions. The target links for voice selection were given a label relating to the chosen word so they would move on to the next target. Numbers were marked on the target. Each subject selected fifty targets and was timed via stopwatch.

The walking section of the experiment was designed so the participant had to walk around objects placed in their path in a figure of eight motion. The objects (six cones) were placed at one metre intervals in a straight line. There was no background noise during the experiments as this might have conflicted with the voice recognition software. After finishing the experiment each participant completed a NASA task load assessment and a post-test questionnaire.

5.5 Results

To analyse the results a mixed factor ANOVA was used and supported by Tukey HSD *post-hoc* tests (Field 2000), with the movement on the within-subjects level and device (input method) on the between-subjects level. The software used to analyse these results was SPSS 11.

Graph 5 gives us an idea of the performance of the three input methods. The blue line representing walking is very close to the red line representing standing. There is plainly very little difference between the times for standing and walking while using each device, suggesting there is no interaction between the type of device and the movement condition. On further visual examination of the graph we can see that it looks as though the voice recognition system used was almost five times slower than the stylus, while the off-table mouse was twice as slow.



Graph 5: Wearable target selection (Device types against time).

The descriptive statistics in Table 5 further illustrate the points made about the graph. We can see that the mean total speeds for standing and walking are very close, but we must look at the next table (Table 5) to see if there is significance. In terms of speed the stylus was the fastest, the off-table mouse was almost twice as slow and voice recognition was the slowest.

Descriptive Statistics time (ms)

	TYPE	Mean	Std. Deviation	N
Total stand time as h of secs	Mouse	19251.00	4301.301	10
	Stylus	8803.40	1026.328	10
	Voice	49471.60	9303.794	10
	Total	25842.00	18454.200	30
total time walk in h secs	Mouse	22241.80	5673.953	10
	Stylus	9198.30	1365.000	10
	Voice	50866.30	9024.643	10
	Total	27435.47	18685.449	30

Table 5: Wearable target selection. (Mean-Time)

The table shows the mean times for the three input devices while standing and walking with standard deviations.

Table 5.1 shows there is a significant main effect of 'moving' ($F(1,27)=12.960$, $p<0.05$). Standing was significantly faster than walking in relation to all the input methods, as illustrated in the descriptive results. There was no significant interaction displayed between 'moving' and 'type' with ($F(2,27= 2.917$, $p > 0.05$ (0.071).

Tests of Within-Subjects Contrasts

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
MOVE	38087040.267	1	38087040.267	12.960	.001
MOVE * TYPE	17143053.433	2	8571526.717	2.917	.071
Error(MOVE)	79345168.300	27	2938709.937		

Table 5.1: Wearable target selection. (Movement)

5.5.1 Input Method (TYPE)

In relation to Table 5.2, a highly significant main effect of 'device' was found ($F\ (1,27) = 126.80$, $p<0.01$) with the stylus being faster than the off-table mouse and voice recognition.

Tests of Between-Subjects Effects

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	42577326816.267	1	42577326816.267	600.244	.000
TYPE	17989712339.433	2	8994856169.717	126.807	.000
Error	1915201326.300	27	70933382.456		

Table 5.2: Wearable target selection. (Input Type)

5.5.2 Post-hoc Tests

The *post-hoc* test further illustrates the significant main effect of 'TYPE' (input method/device) and provided a set of pairwise comparisons. Each comparison was highly significant at $p<0.01$. So overall the three input methods we have evaluated are significantly different from each other in terms of speed of use.

Multiple Comparisons

Tukey HSD

		Mean Difference (I-J)	Sig.
(I) TYPE	(J) TYPE		p<0.01
Mouse	Stylus	11745.55	p<0.01
	Voice	-29422.55	p<0.01
Stylus	Mouse	-11745.55	p<0.01
	Voice	-41168.10	p<0.01
Voice	Mouse	29422.55	p<0.01
	Stylus	41168.10	p<0.01

Table 5.3: Wearable target selection. (Post-hoc)

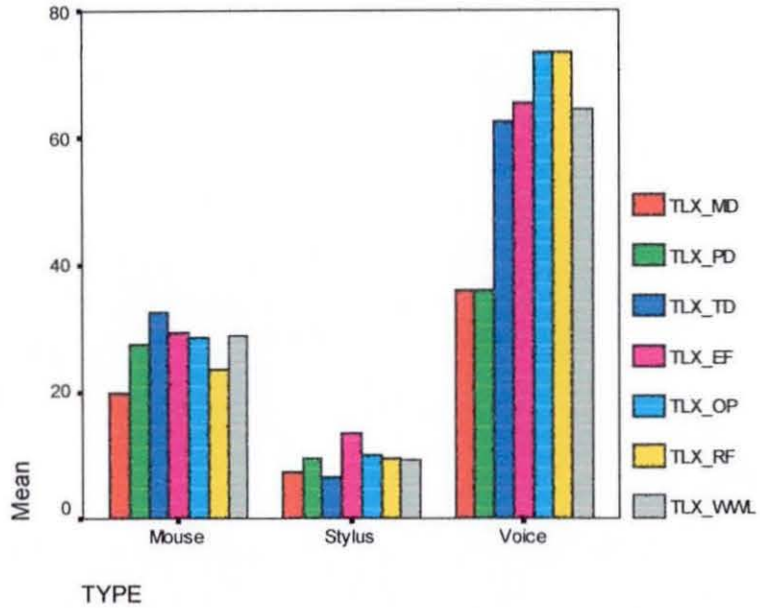
5.5.3 NASA TLX

After the task was completed the participant completed a NASA TLX assessment to give a cognitive workload rating known as a weighted workload or WWL. NASA TLX is a system for assessing the participant's subjective workload. It derives a weighted workload (WWL) from the way the participants rate six subscales: mental demands, physical demands, temporal demands, participant's performance, effort, and frustration. For the workload result it is the case that the lower the score the lower the workload.

5.5.4 Workload Results

The Graph 5.1 (below) gives us an intuitive insight into the results. The graph shows each of the six subscales along with the weighted workload. We can clearly see that the workload is much higher for voice, with stylus having the lowest workload and the mouse scores were in between. It is interesting to note that a lower time relates to a lower task load in this study.

Graph 5.1: Wearable target selection. (Workload) The graph below shows each of the six NASA TLX subscales as scored (mean) by each of the input device groups.



Key:

TLX_MD = Mental Demands	TLX_PD = Physical Demands	TLX_TD = Temporal Demands
TLX_EF = Effort	TLX_OP = Performance	TLX_RF = Frustration
TLX_WWL = Weighted Work load		

The Tukey HSD *post-hoc* analysis broke down the ANOVA table further and does a complete set of pairwise comparisons. From this table we can ascertain that there is highly significant difference between voice and stylus $p<0.01$, voice and mouse $p<0.01$ and mouse and stylus $p<0.05$.

Tukey HSD Pairwise Comparisons

TYPE	(J) TYPE	Mean Difference	Sig.
Mouse	Stylus	19.7000	.003
	Voice	-35.5000	.000
Stylus	Mouse	-19.7000	.003
	Voice	-55.2000	.000
Voice	Mouse	35.5000	.000
	Stylus	55.2000	.000

Based on observed means.

* The mean difference is significant at the .05 level.

Table 5.4: Wearable target selection. (Workload)

5.6 Discussion

After examining the results we can reject the first and third of the null hypotheses.
The pointer type does not affect the time taken to complete the text input tasks. The method of input does affect the input time significantly ($F(1,27) = 126.80$, $p<0.01$). It was clear to see from a comparison of the mean time that using voice recognition took significantly longer to select targets than both off-table mouse and stylus. Although errors were not dependent variables in this assessment, they did play a key part in the reason that it took so long to select targets using voice. It was noted that one person said the same word 22 times before they had to resort to touching the screen to move onto the next target. All of the users using voice for target selection had to touch the touch screen at least once, because the voice recognition was not recognising their command. From these actions, it may be suggested that voice recognition is not appropriate in a 'hands free' environment as this experiment clearly demonstrated that the recognition rate will have an impact upon the user's ability to perform the task that they want to accomplish.

This could have contributed to the high score that voice recognition also got on the NASA TLX assessment. When we look at the graph 5.1, at a glance we can see that the mental demands and physical demands are not substantially different when we compare the off-table mouse and voice recognition, but the performance and frustration are much higher and overall significantly higher. Using this system could prove extremely problematic in terms of time and recognition errors putting an undue amount of stress and pressure on the user. These trials were carried out with no background noise, so one can only imagine the effect of a high level of background noise on a system already having difficulty recognising the user's commands.

We could not reject the null hypothesis number 2 (2. The movement condition of the participant standing and walking has no effect on the time took to complete the task.) as there was no significant effect of movement. This in part was due to the fact that it was noted that some participants walked around the objects very slowly so as to avoid the obstacles in their path and still do their task.

We can reject the third null hypothesis number 3 (3. The pointer type does not affect the subjective task load of the participant.), as there is a significant difference between the three different input methods. With the difference between voice and the stylus and off table mouse being highly significant at $p<0.01$, while a little less significant between mouse and stylus at $p<0.03$.

5.7 Conclusion

In conclusion, the slow results recorded from voice recognition have led us to decide to eliminate it from the next set of pointer trials because of its poor performance and high cognitive workload rating. It may be one of the few viable ways of total hands-free interaction along with gaze control, but for further trials it would bias the results too heavily. It may be that after a sustained period of use that its error rate would go down, but these experiments used novice participants to complete the tasks. This suggests the need for more real world trials on voice recognition systems in the field, with an emphasis on its use while users are mobile.

We have found from this experiment that the stylus was the fastest method of selecting targets and had the lowest cognitive workload; the second fastest was the off table mouse in terms of speed and cognitive workload. Voice recognition was the worst of the three.

After the trials it became clear that there were three other areas that we could further explore and examine. The first area that we will include in the next level of experimentation is error rates, so we can see how many errors a user makes with a given pointing device and thereby examine if there is any relation between movement, device and error rate. The second area of development is using a NASA TLX assessment on both movement conditions of the next experiment instead of an overall assessment. This would give us the ability to use the data in a mixed factorial ANOVA to examine whether movement has an effect on the workload rating. The next experiment will also be made longer and have a higher level of difficulty, by giving the participants more targets to select and also different sizes of targets. The decision was also made to replace the speech recognition system with a track-pad (sometimes known as a touchpad) as these had been used in previous wearable systems (Blaskó and Feiner 2002; Thomas, Grimmer, Zucco and Milanese 2002).

Appendix 5.1

The table below shows the Time for the initial target selection experiment in Appendix 5. The table shows the time for the participants while both standing and walking.

Participant	Time	Time
Speech	Stand	Walk
1	54696	54943
2	41968	42434
3	51756	52074
4	36435	36998
5	54489	54054
6	39296	39147
7	46342	53809
8	65480	65523
9	44664	49895
10	59590	59786
Off-Table Mouse		
1	18250	22598
2	28616	31487
3	24000	27432
4	15806	18208
5	17286	20356
6	20749	21784
7	16965	14757
8	13728	14062
9	18016	27861
10	19094	23873
Stylus		
1	8800	8789
2	8676	10659
3	9586	9597
4	6497	6171
5	9851	10277
6	8567	8733
7	9112	9492
8	10086	10857
9	8853	9172
10	8800	8789

The table below shows the workload score and six sub-scale scores for Appendix 5, the dragging and dropping experiment.

Key:
md = mental demands, pd = physical demands, td = temporal demands, ef = effort, pe = performance and rf = frustration.

Participant	md	Pd	td	ef	pe	rf	WWL
Speech							
1	40	5	85	85	75	85	81
2	20	5	35	20	75	70	40
3	60	50	40	70	80	85	74
4	20	45	90	75	80	75	69
5	60	60	90	60	85	90	79
6	15	25	60	60	65	25	43
7	15	30	55	80	65	60	51
8	55	55	60	55	55	55	55
9	20	30	45	80	55	90	70
10	55	55	65	70	100	100	83
Off-table mouse							
1	10	15	15	15	50	20	18
2	10	15	15	15	35	35	22
3	40	25	10	30	15	10	16
4	20	25	25	35	40	45	35
5	10	25	30	20	25	5	23
6	15	25	30	35	25	20	26
7	10	25	50	30	20	20	31
8	40	15	55	30	15	20	36
9	25	65	60	60	45	45	54
10	20	40	35	25	15	15	29
Stylus							
1	5	5	5	5	5	5	5
2	5	5	5	5	10	5	5
3	10	20	5	25	5	5	13
4	5	10	5	5	15	5	6
5	5	5	5	5	5	5	5
6	5	5	5	5	5	5	5
7	5	5	5	5	5	5	5
8	25	30	10	45	15	30	24
9	5	5	15	30	30	20	19
10	5	5	5	5	5	10	6

Appendix 6.

The subject instructions given to the participants before they complete a workload assessment (as stipulated by the software instructions).

Subject Instructions: Ratings (Mouse Version)

We are interested not only in assessing your performance but also the experiences you had during the different task conditions. Right now we are going to describe the technique that will be used to examine your experiences. In the most general sense we are examining the "workload" you experienced. Workload is a difficult concept to define precisely, but a simple one to understand generally. The factors that influence your experience of workload may come from the task itself, your feelings about your own performance, how much effort you put in, or the stress and frustration you felt. The workload contributed by different task elements may change as you get more familiar with a task, perform easier or harder versions of it, or move from one task to another. Physical components of workload are relatively easy to conceptualize and evaluate. However, the mental components of workload may be more difficult to measure.

Since workload is something is experienced individually by each person, there are no effective "rulers" that can be used to estimate the workload of different activities. One way to find out about workload is to ask people to describe the feelings they experienced. Because workload may be caused by many different factors, we would like you to evaluate several of them individually rather than lumping them into a single global evaluation of overall workload. This set of six rating scales was developed for you to use in evaluating your experiences during different tasks. Please read the descriptions of the scales carefully. If you have a question about any of the scales in the table, please ask me about it. It is extremely important that they be clear to you. You may keep the descriptions with you for reference during the experiment.

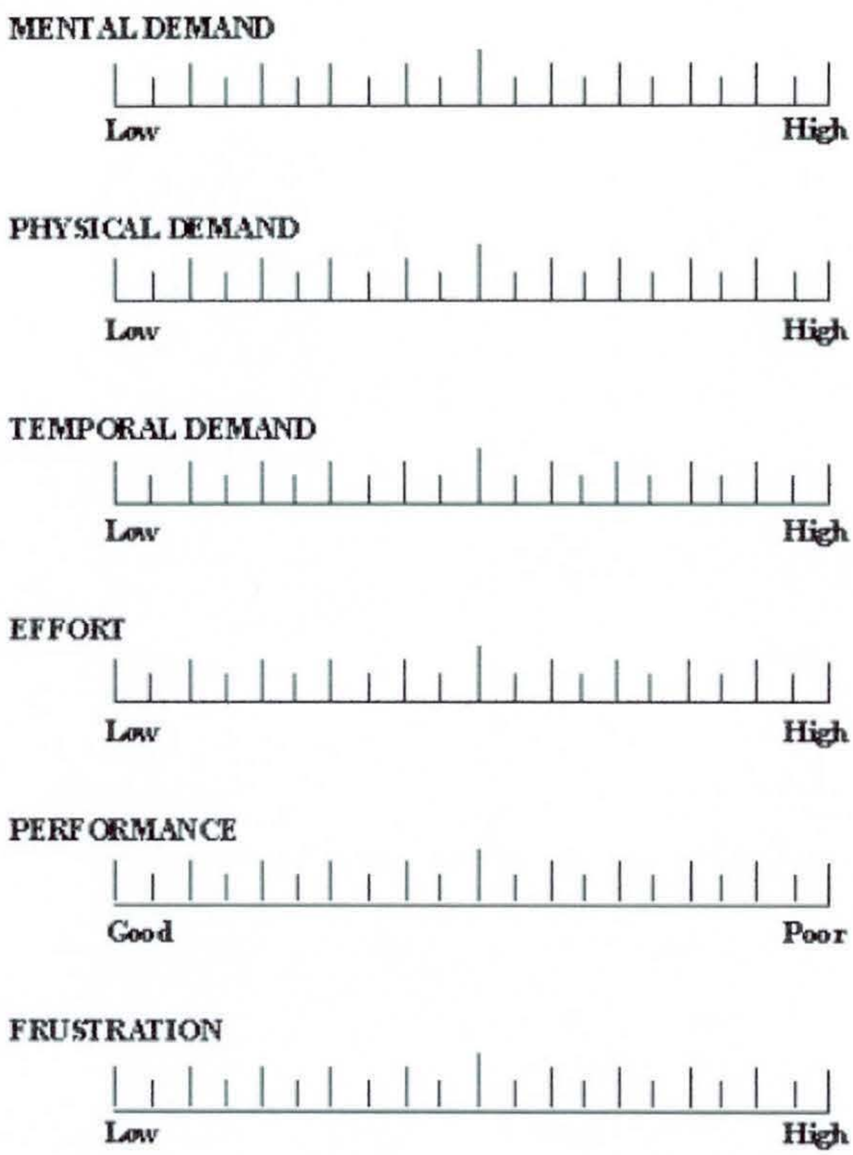
After performing the task, six rating scales will be displayed. You will evaluate the task by marking each scale at the point which matches your experience. Each line has two end point descriptors that describe the scale. Note that "own performance" goes from "good" on the left to "bad" on the right. This order has been confusing for some people. Move the arrow to the right or left with the mouse until it points at the desired location. When you are satisfied, press either button to enter your selection. Please consider your responses carefully in distinguishing among the task conditions. Consider each scale individually. Your ratings will play an important role in the evaluation being conducted, thus your active participation is essential to the success of this experiment, and is greatly appreciated.

Appendix 6.1
The Nasa TLX rating scale definitions that were given to the participant with the subject instructions

RATING SCALE DEFINITIONS		
Title	Endpoints	Descriptions
MENTAL DEMAND	Low/High	How much mental and perceptual activity was required (e.g., thinking, deciding, calculating, remembering, looking, searching, etc.)? Was the task easy or demanding, simple or complex, exacting or forgiving?
PHYSICAL DEMAND	Low/High	How much physical activity was required (e.g., pushing, pulling, turning, controlling, activating, etc.)? Was the task easy or demanding, slow or brisk, slack or strenuous, restful or laborious?
TEMPORAL DEMAND	Low/High	How much time pressure did you feel due to the rate or pace at which the tasks or task elements occurred? Was the pace slow and leisurely or rapid and frantic?
EFFORT	Low/High	How hard did you have to work (mentally and physically) to accomplish your level of performance?
PERFORMANCE	Good/Poor	How successful do you think you were in accomplishing the goals of the task set by the experimenter (or yourself)? How satisfied were you with your performance in accomplishing these goals?
FRUSTRATION LEVEL	Low/High	How insecure, discouraged, irritated, stressed and annoyed versus secure, gratified, content, relaxed and complacent did you feel during the task?

Appendix 6.2

An example of the NASA TLX assessment completed by all of the participants. This shows the six scales that the user uses to rate their workload (High or Low). This is provided with the instructions of the computerised version of NASA TLX. This is provided to illustrate what the participants saw on screen when using this system.



Appendix 7.

These are the participant instructions for Chapter 3.
Text input
1. Instructions to Participants

Dear participant

Thank you for your time and supporting our research.

After a training session we will start the experiment. You will be able to ask any questions prior to the start of the experiment to make sure that you fully understand the task that you are to accomplish.

There are two parts to this experiment. The experiment will be done while walking around road cones placed in your path (figure of eight style as demonstrated by the experimenter) and while standing stationary. You will be told the order in which you will do each of these prior to the experiment.

The aim of this experiment is to copy text. In order to complete your task you must copy the text displayed on the screen using the device you used during your training. You must not correct any mistakes that you may make while copying the text or stop if you are walking.

The experimenter will tell you when you can start the task, when you have finished the task simply say, 'stop'. You will be wearing a vest-based computer while completing the tasks.

After completing each part of the experiment you will be asked to complete a NASA TLX assessment.

Do you have any questions?

These are the participant instructions for Chapter 4.

Target Selection.
1. Instructions to Participants

Dear participant

Thank you for your time and supporting our research.

After a training session we will start the experiment. You will be able to ask any questions prior to the start of the experiment to make sure that you fully understand the task that you are to accomplish.

There are two parts to this experiment. The experiment will be done while walking around road cones placed in your path (figure of eight style as demonstrated by the experimenter) and while standing stationary. You will be told the order in which you will do each of these prior to the experiment.

The aim of this experiment is to use the input device, which you earlier used in the training session to select the targets that appear on the vest-mounted screen. Select all the targets that will appear in a random order upon the screen.

The experimenter will tell you when you can start the task, when you have finished the task simply say, 'stop'. You will be wearing a vest-based computer while completing the tasks.

After completing each part of the experiment you will be asked to complete a NASA TLX assessment.

Do you have any questions?

These are the participant instructions for Chapter 5.

Dragging and Dropping
1. Instructions to Participants

Dear participant

Thank you for your time and supporting our research.

After a training session we will start the experiment. You will be able to ask any questions prior to the start of the experiment to make sure that you fully understand the task that you are to accomplish.

There are two parts to this experiment. The experiment will be done while walking around road cones placed in your path (figure of eight style as demonstrated by the experimenter) and while standing stationary. You will be told the order in which you will do each of these prior to the experiment.

The aim of this experiment is for you to drag and drop the items that appear upon the vest-mounted screen as in the training session. You will use the same input device that you used in the training session.

The experimenter will tell you when you can start the task, when you have finished the task simply say, 'stop'. You will be wearing a vest-based computer while completing the tasks.

After completing each part of the experiment you will be asked to complete a NASA TLX assessment.

Do you have any questions?

These are the participant instructions for Chapter 6.

Trajectory-Based Interaction.
1. Instructions to Participants

Dear participant

Thank you for your time and supporting our research.

After a training session we will start the experiment. You will be able to ask any questions prior to the start of the experiment to make sure that you fully understand the task that you are to accomplish.

There are two parts to this experiment. The experiment will be done while walking around road cones placed in your path (figure of eight style as demonstrated by the experimenter) and while standing stationary. You will be told the order in which you will do each of these prior to the experiment.

The aim of this experiment is for you to steer through the tunnels that appear upon the vest-mounted screen as was done in the in the training session. You will use the same input device that you used in the training session.

The experimenter will tell you when you can start the task, when you have finished the task simply say, 'stop'. You will be wearing a vest-based computer while completing the tasks.

After completing each part of the experiment you will be asked to complete a NASA TLX assessment.

Do you have any questions?