University of Glasgow

Using Pressure Input and Thermal Feedback to Broaden Haptic Interaction with Mobile Devices

Graham Alasdair Wilson

Submitted for the degree of Doctor of Philosophy

School of Computing Science, University of Glasgow

June 2013

Abstract

Pressure input and thermal feedback are two under-researched aspects of touch in mobile human-computer interfaces. Pressure input could provide a wide, expressive range of continuous input for mobile devices. Thermal stimulation could provide an alternative means of conveying information non-visually. This thesis research investigated 1) how accurate pressure-based input on mobile devices could be when the user was walking and provided with only audio feedback and 2) what forms of thermal stimulation are both salient and comfortable and so could be used to design structured thermal feedback for conveying multi-dimensional information.

The first experiment tested control of pressure on a mobile device when sitting and using audio feedback. Targeting accuracy was >= 85% when maintaining 4-6 levels of pressure across 3.5 Newtons, using only audio feedback and a Dwell selection technique. Two further experiments tested control of pressure-based input when walking and found accuracy was very high (>= 97%) even when walking and using only audio feedback, when using a rate-based input method.

A fourth experiment tested how well each digit of one hand could apply pressure to a mobile phone individually and in combination with others. Each digit could apply pressure highly accurately, but not equally so, while some performed better in combination than alone. 2- or 3-digit combinations were more precise than 4- or 5-digit combinations. Experiment 5 compared one-handed, multi-digit pressure input using all 5 digits to traditional two-handed multitouch gestures for a combined zooming and rotating map task. Results showed comparable performance, with multitouch being ~1% more accurate but pressure input being ~0.5sec faster, overall.

Two experiments, one when sitting indoors and one when walking indoors tested how salient and subjectively comfortable/intense various forms of thermal stimulation were. Faster or larger changes were more salient, faster to detect and less comfortable and cold changes were more salient and faster to detect than warm changes. The two final studies designed two-dimensional structured ‘thermal icons’ that could convey two pieces of information. When indoors, icons were correctly identified with 83% accuracy. When outdoors, accuracy dropped to 69% when sitting and 61% when walking.

This thesis provides the first detailed study of how precisely pressure can be applied to mobile devices when walking and provided with audio feedback and the first systematic study of how to design thermal feedback for interaction with mobile devices in mobile environments.

Table of Contents

List of Tables

List of Figures

Acknowledgements

First of all I would like to thank my supervisor, Professor Stephen Brewster, for his invaluable help and support throughout this PhD research. I would also like to thank my second supervisor, Professor Rod Murray-Smith, for providing alternative insights and ideas. Thanks also to Dr Martin Halvey, who provided help and many useful discussions over the years.

I would like to thank the many past and present members of the Multimodal Interaction Group at Glasgow, who have given their advice, time and/or resources to helping me develop my research, in particular Dr Marilyn McGee-Lennon, Dr David McGookin and Dr Andrew Crossan.

Many thanks go to Nokia Research Centre, Helsinki for offering me a fantastic internship: to Dari Trendafilov, for initiating introductions; to Viljakaisa Aaltonen for agreeing to and organising the internship; and, most importantly, to Johan Kildal, for supporting and encouraging me throughout my time in Helsinki.

Finally, I would like to thank those closest to me. Most importantly, my thanks to Sarah, for the years of caring, supporting and understanding. Thanks also to my parents, Douglas and Marion, and to my brother Stephen, for all your support and encouragement.

This research was funded in part by the industrial partners of MobileVCE Virtual Centre for Excellence.

Declaration

The research presented in this thesis is entirely the author's own work. This thesis exploits only the parts of these papers that are directly attributable to the author:

The research in Chapter 3 has been published in MobileHCI 2010 [251]:
Wilson, G., Stewart, C. and Brewster, S. Pressure-Based Menu Selection for Mobile Devices. *Proceedings of MobileHCI 2010*, pp 181-190.

Experiment 2 in Chapter 4 has been published as a Works-in-Progress in CHI 2011 [244]:
Wilson, G., Brewster, S. and Halvey, M. The Effects of Walking and Control Method on Pressure-Based Interaction. *Proceedings of CHI 2011 Extended Abstracts*, pp 2275-2280.

Experiments 2 and 3 from Chapter 4 have been published in MobileHCI 2011 [248]:
Wilson, G., Brewster, S.A., Halvey, M., Crossan, A. and Stewart, C. The Effects of Walking, Feedback and Control Method on Pressure-Based Interaction. *Proceedings of MobileHCI 2011*, pp 147-156.

Experiment 4 in Chapter 5 has been published as a Works-in-Progress in CHI 2012 [250]:
Wilson, G., Hannah, D., Brewster, S. and Halvey, M. Investigating One-Handed Multi-Digit Pressure Input for Mobile Devices. *Proceedings of CHI 2012 Extended Abstracts*, pp 1727-1732.

Experiment 5 in Chapter 5 has been published as a Works-in-Progress in CHI 2013 [245]:
Wilson, G., Brewster, S. and Halvey, M. Towards Utilising One-Handed Multi-Digit Pressure Input. *Proceedings of CHI 2013 Extended Abstracts,* pp 1317-1322.

The research in Chapter 6 has been published in CHI 2011 [249]:
Wilson, G., Halvey, M., Brewster, S. and Hughes, S. Some Like it Hot? Thermal Feedback for Mobile Devices. *Proceedings of CHI 2011*, pp 2555-2564.

Experiment 7 in Chapter 7 has been published in HAID 2013 [246]:
Wilson, G., Brewster, S., Halvey, M. and Hughes, S. Thermal Feedback Identification in a Mobile Environment. *Proceedings of HAID 2013*, Article 2.

Experiment 8 in Chapter 7 has been published in MobileHCI 2012 [247]:
Wilson, G., Brewster, S., Halvey, M. and Hughes, S. Thermal Icons: Evaluating Structured Thermal Feedback for Mobile Interaction. *Proceedings of MobileHCI 2012*, pp 309-312.

1 Introduction

1.1 Motivation

Pressure input and thermal feedback are two under-researched aspects of touch in mobile human-computer interfaces and this thesis presents a study into how they could expand interaction options on mobile devices. Input and output options on mobile devices, such as phones, portable media players and tablets, are somewhat limited, at least in comparison to desktop machines. Touch-based interaction, focused on multitouch touchscreens, is becoming the primary means of input, often supported by a small number of physical buttons for basic navigation or specific tasks such as a camera shutter. Simple physical gestures, such as tilting and shaking the device, are also used. In current commercial devices, capacitive touchscreens can detect only a 2-dimensional contact point, along with any 2-dimensional movement across the surface. Multitouch gestures provide more complex interactions, but input options remains relatively limited.

Output from mobile devices is primarily visual, with large, high definition screens becoming increasingly popular, making consumption of more complex visual content easier or more enjoyable. In contrast, the non-visual output capabilities of mobile devices are underutilised. While devices are capable of outputting high-quality audio, conveying information through audio is generally limited to simple ringtones or discrete notifications. The recent addition of synthetic speech output, such as Apple's "Siri" service[1], provides a wider range of information, however. Other than audio, vibration is currently the only other means of conveying information non-visually in commercial devices. Vibration motors in mobile devices are simple and can produce a limited range of stimuli. Research has shown that novel auditory [47, 105, 160, 194, 236] and vibrotactile [22, 100, 156] feedback designs could improve interaction with mobile devices, but these channels have situational limitations. Different mobile environments call for the use of different feedback channels, such as vibrotactile for very quiet or very noisy environments. However, in very bumpy and/or loud environments neither channel may be suitable or desirable for the individual.

[1] http://www.apple.com/uk/ios/siri/

The hand is a multifaceted investigative and manipulative tool. Certain aspects of manual touch and haptic interaction have been well researched within HCI, including spatial gestures (for example, tilting or drawing shapes) [42, 76, 160, 178, 193, 195], force-feedback [28, 110, 163], vibrations and textures [1, 21, 41, 139, 152], finger orientation [199, 200] and tangible devices [34, 54, 86, 106, 107, 172, 186, 252]. The *application of pressure* and *thermal perception* are two other inherent aspects of manual touch, and therefore human-object interaction, that may have an enormous potential contribution to make to interaction with mobile devices. However, their relative merits in designing mobile interfaces have not been fully explored, despite being highly accurate and specialised systems. With the proliferation of mobile devices that focus on multitouch and gestural input, pressure input and thermal feedback sit as logical extensions of this touch-based interaction paradigm.

Every act of tactition or grasping necessarily includes a degree of contact pressure, in the form of, for example, touching, pushing, enclosing, grasping/lifting, squeezing and hitting. The amount of pressure applied by the individual depends on the purpose or intention of the action, but no act of touch is ever without a degree of applied pressure, and so the extent of pressure applied has a purpose and meaning. Applied pressure from the fingers could provide a wide, expressive range of input, one that can be controlled dynamically and continuously. To judge how wide a range of input is possible, and how precisely that range can be controlled, it is necessary to understand how well individuals can control the amount of pressure they apply to a mobile device when in both static (i.e., sitting stationary) and mobile scenarios. As part of the investigation into mobile interaction, it is important to consider the use of non-visual feedback. When walking, visual attention must be paid to the environment, to monitor the walking route and identify potential obstacles and hazards. Also, focussing on text or small icons on a mobile device may be difficult due to bodily motion, and resulting motion of the device. If visual attention cannot be focused on the screen of the device, then information must be conveyed through non-visual means. If pressure-based input is to be a feasible means of interacting with mobile devices, it is also necessary to understand how the use of non-visual feedback can facilitate mobile pressure-based input.

Thermal sensation is a vital facet of human touch, which continually provides information about our environments and each object that we touch. Every surface and object in our environment has a current temperature and a rate of thermal conductivity; no object can ever be without either and both of these give us information about the nature of the object, such as its material, its threat to the skin (i.e., very hot or cold) or whether it is alive. Beyond this base semantic information, there may also be an inherent hedonic, or emotional, element to

thermal feedback, something that is not necessarily present in other forms of feedback [223]. Therefore, thermal feedback is a natural way of conveying information in everyday scenarios. Because audio and vibrotactile feedback are not always usable, or desirable [101], thermal feedback may provide a salient alternative means of conveying information. It has been used in Virtual Reality to convey material property information [5, 94, 128] but its merits for conveying information in HCI are less established, especially in mobile interaction contexts, such as when walking or being in outdoor locations.

Both control of pressure and thermal perception have been extensively researched in psychophysical, physiological and medical literature. The deep understanding of the relevant neurological, perceptual and motor systems that this research allows provides a solid foundation on which to design user interfaces that are based on human ability. Also, pressure input and thermal feedback in HCI are not entirely new. Pressure input has been researched and used successfully in a number of static (seated), desktop applications [30, 191, 192, 209, 210], enjoying high accuracy and providing continuous input not available from other devices. However, research has not adequately investigated control of pressure applied to mobile devices, or when the user is walking. Also, the feedback provided during pressure interaction has been almost exclusively visual. Research needs to be done to understand if non-visual feedback can be used to facilitate eyes-free pressure input, so that the user can focus visual attention on their mobile environment. Thermal feedback has been used primarily in virtual reality to convey properties of virtual objects [5, 94, 122, 212, 258]. Other implementations of thermal feedback in HCI have merely been initial prototypes or proofs-of-concept, conveying basic information [52, 112, 143, 174, 239]. Little HCI research has systematically tested how well thermal feedback can be detected and identified in realistic scenarios, such as when walking and/or outdoors, and so how best to design and utilise thermal feedback.

Therefore, the research presented in this thesis aimed to test the feasibility and usability of pressure-input and thermal feedback for use in mobile HCI. Specifically, it focused on how well individuals can control the amount of pressure they apply with one or more digits of the hand to mobile devices, when they are sitting or walking and provided with audio feedback. This was to establish the fundamental input capabilities for mobile pressure-based input. For thermal feedback, the research aimed to identify what forms of thermal stimulation are reliably and comfortably detectable when the individual is sitting and walking, indoors and outdoors. Having identified these reliable forms, the research aimed to develop ‘thermal icons’: structured, multi-dimensional thermal stimuli capable of conveying multiple pieces of information thermally in a variety of interaction environments.

1.2 Thesis Statement

The hand is a multifaceted investigative and manipulative tool. The application of pressure and the reception of thermal feedback are inherent aspects of manual touch and provide new opportunities to broaden the input and output capabilities for mobile device interaction. Pressure input on mobile devices is highly accurate when walking, when provided with either visual or audio feedback and when applying pressure from multiple digits, both individually and in combination. Individuals can detect a range of thermal stimuli produced from limited hardware designed for mobile interaction when both sitting and walking indoors. Using these stimuli, structured thermal icons can be created to convey two pieces of information to users when in both indoor and outdoor environments.

1.3 Research Questions

This thesis aims to answer the following questions:

RQ1: How accurate is pressure-based input on a mobile device when using only audio feedback?

RQ2: How accurate is pressure-based input through the fingers when the individual is walking?

RQ3: How accurate is pressure-based input when multiple digits apply pressure to a mobile device?

RQ4: What parameters of thermal stimulation are most detectable and comfortable for use in mobile interaction?

RQ5: Can thermal stimulation be manipulated to convey multi-dimensional information?

1.4 Thesis Walkthrough

Chapter 2, *Literature Review*, reviews the literature on the application of pressure and thermal perception, from both a perceptual/psychophysical perspective, looking at the limits and precision of human ability, and an HCI perspective, looking at how applied pressure and thermal feedback have been used in various interfaces. The influence of feedback on the application of pressure is also discussed, as are the existing means of interacting with mobile devices and conveying information non-visually in mobile interfaces.

From the research questions the primary aims of the research into pressure input on mobile devices were to 1) develop an audio feedback design that allows for accurate pressure input when sitting 2) test pressure input accuracy when walking with both visual and audio feedback and finally 3) to test how accurately each digit individually, and in combination, can apply pressure to a mobile device. The research in the experimental chapters followed this progression path.

Chapter 3, *Non-visual Pressure-based Input When Sitting*, reports on Experiment 1, which tested the precision of pressure input applied by a single digit (thumb) to a pressure sensor during a linear targeting task, when participants were sitting in a chair and provided with visual feedback or audio feedback. The purpose was to develop a useful audio feedback design to facilitate eyes-free mobile pressure input, prior to testing control when mobile. Two different targeting selection techniques were compared, to judge which technique provides best performance, particularly when using audio feedback. This chapter answers RQ 1.

Experiment 1 Factors Tested: Feedback Modality (Visual, Audio), Selection Technique (Dwell, Quick Release) and Target Size.

Having tested pressure input using audio feedback when sitting, Chapter 4, *Mobile Non-visual Pressure-based Input*, includes Experiments 2 and 3, which extend Experiment 1 from Chapter 3 and test precision of pressure input using a single digit during linear targeting when walking a route indoors and provided with either visual or audio feedback. Two alternative input control methods, Positional and Rate-based, were also tested to determine which provides the better performance when walking. This chapter partly answers RQ 1 and also answers RQ 2.

Experiment 2 Factors Tested: Mobility (Sitting, Walking), Control Method (Positional,

Rate-based) and Target Size.
Experiment 3 Factors Tested: Feedback Modality (Visual, Audio) and Target Size.

Experiments 1 to 3 used only a single digit (thumb) for input, but mobile devices can be held by, and interacted with, using many digits at once. Chapter 5, *Multi-digit Pressure Input on a Mobile Device*, describes Experiments 4 and 5 using pressure input from all five digits of one hand, both individually and in various combinations. Input was provided to the sides, back and top of a mobile phone. Experiment 4 tested how precisely pressure could be applied by each digit or combination of digits and compared precision using two different ranges of pressure (called Pressure Spaces): a Fixed range that was used regardless of how many digits were used, and an Incremental range that increased by a set amount with the addition of each digit that was applying pressure. As multiple digits could provide multiple inputs simultaneously, more complex interactions may be controllable one-handed. Experiment 5 used the best-performing digits/grips and compared one-handed, multi-digit pressure input with common two-handed multitouch input during zooming and rotating (and a combination of both) in a map task, to determine if these tasks, which typically require two hands, can be carried out one-handed. This chapter answers RQ 3.

Experiment 4 Factors Tested: Grip (14 grips, including each digit individually) and Pressure Space (Fixed, Incremental).
Experiment 5 Factors Tested: Input Method (Pressure, Multitouch) and Task (Zooming, Rotation, Combined).

The aims of the research into thermal feedback for mobile devices were to 1) identify reliably salient stimuli for designing thermal feedback and 2) design and test identification of structured thermal feedback to convey multidimensional information.

Chapter 6, *Identifying Detectable and Comfortable Thermal Feedback Parameters*, describes Experiments 6 and 7, which tested detection of various thermal stimuli when sitting and walking indoors, to establish the influence of sitting and walking in a realistic interaction environment on thermal perception. The purpose was to identify which stimuli would be suitable for use in designing thermal feedback for HCI. The stimuli varied along three parameters (direction of change, rate of change and extent of change), which are known to result in varying sensations in the individual. The salience and comfort of each stimulus was measured and a set of guidelines was produced outlining which stimuli would be suitable for use in structured thermal feedback. This chapter answers RQ 4.

Experiments 6 and 7 Factors Tested: Direction of Change (Warming, Cooling), Rate of Change (1°C/sec, 3°C/sec), Extent of Change (1°C, 3°C and 6°C) and Body Location (fingertip, palm, forearm, upper arm).

Having identified which parameters of thermal stimulation are suitable for designing thermal feedback in Chapter 6, Chapter 7, *Conveying Multi-dimensional Information Thermally*, includes Experiments 8 and 9, which tested absolute identification of unique, structured thermal stimuli called 'thermal icons' when the individual was a) sitting indoors (Experiment 8), b) sitting outdoors (Experiment 9) and c) walking outdoors (Experiment 9). In Experiment 8, thermal icons were compared to "intramodal icons", where thermal and vibrotactile feedback parameters were combined, to test if two feedback channels from the same (tactile) modality could be interpreted together, and so whether thermal feedback could augment existing structured vibrotactile feedback. In Experiment 9, thermal icons were compared to purely vibrotactile Tactons, the most established means of conveying multidimensional information in mobile interaction scenarios, to provide context for the results using thermal icons. Identification of four different icons was tested, with each representing a different type of message being received: Standard Personal, Important Personal, Standard Work and Important Work. This chapter answers RQ 5.

Experiment 8 Factors Tested: Modality (Thermal, Intramodal) and Icon Type (Standard Personal, Important Personal, Standard Work, Important Work).
Experiment 9 Factors Tested: Mobility (Sitting, Walking), Modality (Thermal, Vibrotactile) and Icon Type (Standard Personal, Important Personal, Standard Work, Important Work).

Chapter 8, *Discussion and Conclusions*, reviews and summarises the research in this thesis, including its novel contributions and how it answered the research questions. Limitations of the research are discussed and possibilities for future research are proposed.

Topic	Experiment (Chapter)	Purpose	Factors Tested
Pressure Input	Experiment 1 (3)	• Test suitable audio feedback design • Identify optimal selection technique	• Feedback modality • Selection Technique
	Experiment 2 (4)	• Test control when walking • Identify optimal control method	• Mobility • Control Method
	Experiment 3 (4)	• Test non-visual control when walking	• Feedback Modality
	Experiment 4 (5)	• Test control using multiple digits • Test influence of pressure space	• Grip (no. of digits) • Pressure Space
	Experiment 5 (5)	• Compare pressure to multitouch	• Input Method • Task
Thermal Feedback	Experiment 6 (6)	• Identify suitable feedback parameters	• Direction of change • Rate of change • Extent of change • Body Location
	Experiment 7 (6)	• How walking influences perception	*Same as Experiment 6*
	Experiment 8 (7)	• Test identification of structured thermal feedback/"icons"	• Modality • Icon type
	Experiment 9 (7)	• Test icon identification outdoors	• Mobility • Modality • Icon type

Table 1-1: Summary of all experiments carried out, including the purpose of each and the experimental factors tested.

2 Literature Review

The aims of the research in this thesis are to understand and test 1) the control of applied pressure and 2) the identification of thermal stimuli as means of interaction with mobile devices. This chapter reviews the existing research literature related to the two topics. This includes research from within the fields of perception, psychophysics and haptic human-computer interaction. The review begins with a brief overview of the physiology and sensory networks of the hand, followed by research on the precision of prehensile action and the production of pressure from the fingers. These sections explain the limits of human ability: how well we can apply pressure in highly controlled laboratory studies. They provide the ideal baseline against which pressure input on mobile devices can be compared to judge the negative effects of control over pressure when walking, using only audio feedback and applying pressure one-handed to a mobile device. They also describe the influence of feedback on precision of applied pressure, particularly the reduction or removal of 'external' feedback such as visual and audio feedback, as compared to 'internal' kinaesthetic and cutaneous feedback. The research on feedback characteristics informs about how important feedback is to the accurate application of pressure, and so how the use of different forms of external feedback, including audio feedback, could impact the usability of eyes-free pressure input on mobile devices.

Following the psychophysical literature review is a survey of practical and scientific HCI research carried out on the use of pressure as an input channel for interfaces. While psychophysical science illustrates the limits of ability in controlled lab studies, practical uses in realistic interaction scenarios can result in very different performance. The HCI literature review shows how well pressure input has been measured and mapped to control various interface elements, such a cursor, shape angle or zoom level. Research conducted within desktop interaction settings illustrates how accurately pressure can be applied in more physically stable locations, to compare with pressure input when walking. The limited research conducted on mobile devices is presented to show the existing state of the art against which the research in this thesis is compared, and includes how pressure can be applied in different ways (fewer/different digits and one-handed) to different form factors (smaller devices), compared to desktop interfaces.

Note that, in this thesis, the word "pressure" has the following meaning: *the exertion of force*

upon a surface by an object…in contact with it[2]. In general, psychophysical research uses the word "force" (as in "application of force") when referring to the application of pressure in prehensile actions, through, for example, pushing, pinching and squeezing. Within HCI research, however, the word "pressure" is more commonly used for the same actions. "Force" can also refer to "force-feedback" in HCI, to refer to mechanical resistance, or vibrotactile feedback, as an output from the system. Therefore, for clarity and consistency, the word "pressure" is used throughout this thesis. Exceptions include references to specific types of force, such as "normal force" (pressure applied directly onto a surface), "tangential force" (pressure applied across a surface), and "load force" (the vertical pull of an object on the skin due to gravity), as these are proper, and accepted, terms. In these cases the words are used interchangeably and mean the same.

The literature review continues with a discussion of the sensory and perceptual characteristics of thermal stimulation, including its uses in human-computer interaction. The wealth of psychophysical literature on thermal perception informs about the many factors which influence subjective experience and appreciation of thermal stimuli and so a) what factors need to be controlled or mitigated against to provide suitable feedback and b) what factors can be manipulated to produce a variety of sensations in the user for feedback purposes. HCI research is presented to show the limited ways in which thermal stimulation has been leveraged in computer interfaces thus far and so how the research in this thesis developing new forms of thermal feedback can expand these possibilities.

A brief overview of the influences of walking on interaction with mobile devices is given, to illustrate the importance of testing control of pressure input and thermal feedback when physically in motion and not relying on stationary laboratory studies for valid results. The literature review ends with a summary of how non-visual feedback has been used to overcome issues concerning interaction with mobile devices, including the negative effects of mobility, to highlight the benefits of utilising audio feedback for pressure input on mobile devices for when the user is walking. The chapter ends with a summary of the most important aspects of the research, which have shaped the research questions and so provide context for the contributions in this thesis.

Research Question 4 asks:

"What parameters of thermal stimulation are most detectable and comfortable when using equipment designed for mobile interaction?"

[2] Adapted rom www.dictionary.com (http://dictionary.reference.com/browse/pressure)

This question is partly answered in Section 2.3, which describes thermal sensitivity in detail, including which parameters of thermal stimulation influence perception when in highly controlled laboratory conditions, and so may be suitable for use in thermal feedback designs.

2.1 The Perception and Application of Pressure Through the Hand

This thesis research focuses on the application of pressure through the fingers, and so the physiology discussed will be limited to the muscles, nerves and perceptual processes governing flexion (pulling in towards the palm) of the five digits of the hand. This concerns the acts of prehension (applying pressure from opposing directions, such as grasping and squeezing) and unidirectional application of normal force (such as pressing or pushing). While pressure can be applied from the upper arm muscles (such as biceps, triceps and deltoid) and pectoral muscles through the hand and fingers (for pressing and pushing), this type of action is not discussed, as the research is primarily interested in manual/digital pressure applied to a device held in the hand. Extension, straightening of the digits away from palm, is not covered here.

The first two sections 2.1.1 and 2.1.2, describe the physiological basis of pressure application and sensation through the fingers. The following section 2.1.3, describes the precision in applying pressure and how the feedback provided from those actions influences behaviour.

2.1.1 Skeletal and Muscular Physiology

The bones and parts of the hand are shown in Figure 2-1. The muscles that control flexion of the fingers are predominantly outside of the hand, along the forearm (called “extrinsic” hand muscles), while a small number of muscles are located inside the hand (“intrinsic” hand muscles). There are no muscles in the fingers as all extrinsic hand muscles connect to tendons just above the wrist. The tendons then pass through the carpal tunnel (a passage through the wrist bounded by the carpal bones and a ligament called the *flexor retinaculum*) before connecting to the individual bones of the fingers (called phalanges, singular phalanx). Different muscles connect to different phalanges and so flex different interphalangeal joints.

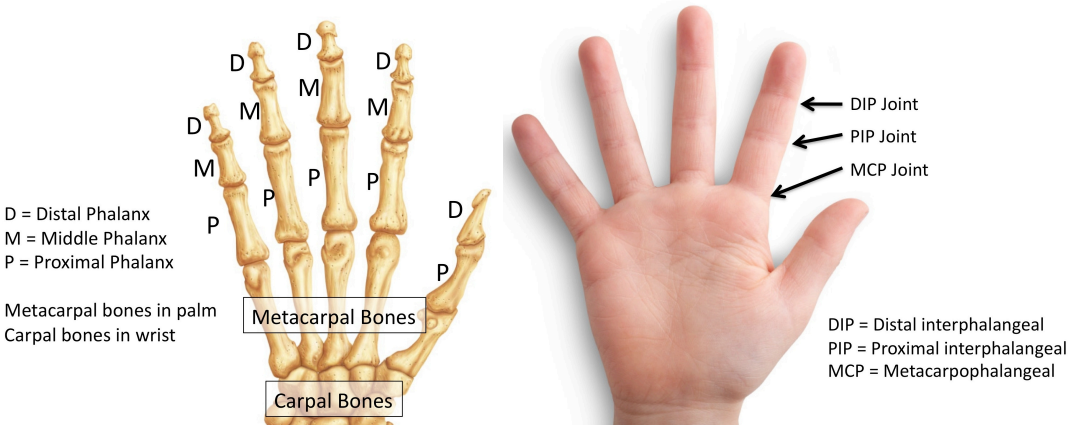

Figure 2-1: The bones of the right hand (left) and the names of each joint in the fingers (right)[3].

The main extrinsic flexor muscles (shown in Figure 2-2), and their connections to the digits, are:

- *Flexor pollicis longus* – Flexes the distal interphalangeal (DIP, see Figure 2-1) joint in the thumb (pulls the distal phalanx inwards). Connects to one tendon, which inserts (connects) into the base of the distal phalanx of the thumb. This muscle is unique to humans. An example use of this muscle is pressing the button at the end of a ballpoint pen, to protrude the tip.
- *Flexor digitorum profundus* – Flexes the DIP joint of each finger. Connects to four tendons, each of which inserts into the distal phalanx of one finger. An example use of this muscle is pulling the trigger of a gun.
- *Flexor digitorum superficialis* – Flexes the proximal interphalangeal (PIP) joint of each finger. Connects to four tendons, which insert into the base of the middle phalanx of one finger. Assists Flexor digitorum profundus in pulling fingers in towards palm as well as flexing the wrist. An example use may be to assist in gripping the butt of the gun.

[3] Left image © 2009 Pearson Education, Inc. Right image from http://www.ucl.ac.uk/news/news-articles/1006/10061602

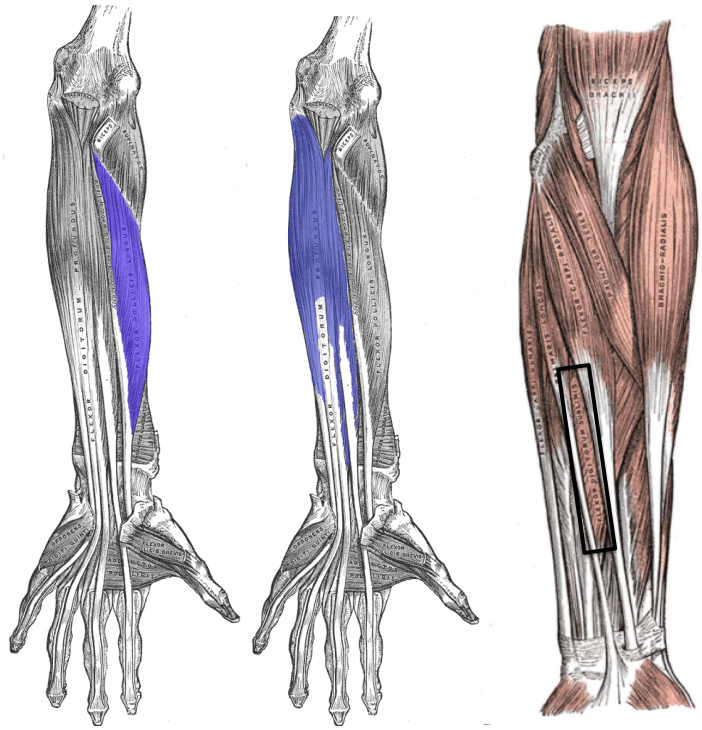

Figure 2-2: Extrinsic digit flexor muscles: Flexor pollicis longus (left; flexes thumb), flexor digitorum profundis (centre; flexes distal phalanx of each finger) and flexor digitorum superficialis (right; flexes proximal phalanx of each finger)[4].

The intrinsic hand muscles (shown in Figure 2-3) are responsible for abduction/adduction (lateral movement towards and way from each other) as well as being partly responsible for flexion and extension. They are also responsible for the extra degree of freedom enjoyed by the thumb: pronation (opposing the fingers with the thumb). The intrinsic hand muscles are:

- *Three thenar muscles* – Located at the thenar eminence, the bulbous area of the palm adjoining the thumb. These muscles pronate the thumb: *abductor pollicis brevis, opponens pollicis* and *flexor pollicis brevis*.
- *Three hypothenar muscles* – Located between the base of the little finger and the wrist. These muscles are responsible for flexion and abduction (away from the other fingers) of the little finger: *abductor digiti minimi, flexor digiti minimi* and *opponens digiti minimi*.
- *Four lumbrical muscles* – Located between each finger and flex the metacarpophalangeal (MCP, see Figure 2-1) joint of each finger, as well as extending the interphalangeal joints.
- *Four dorsal* and *three palmer interosseus muscles* – Located on either side of the metacarpal bones and abduct/adduct the fingers.

[4] Images from freely licensed Wikimedia (http://commons.wikimedia.org/) files, based on public domain lithograph plates from Gray's Anatomy 20th Edition (1918).

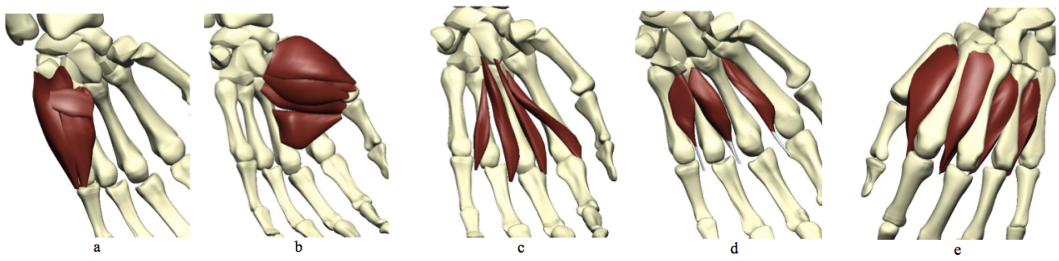

Figure 2-3: Intrinsic hand muscles of the left hand, adapted from Reynolds *et al.* [198]. From left to right: hypothenar, thenar, lumbrical, palmer interosseus and dorsal interosseus muscles.

All muscle contractions, including those involved in manual dexterity, are divided into two types: isotonic and isometric. Isotonic contractions involve changes to the length of a muscle, but not its tension, and so describe spatial movements where resistance remains the same. In the case of the hand/fingers, contraction (shortening) of the extrinsic or intrinsic muscles pulls on the connected tendon(s), which in turn pulls on the relative phalanx of the digit. In the face of no resistance, this results in movement of the digit 'inwards' towards the palm. An example of an isotonic device in HCI is the laptop trackpad, where the finger moves, unhindered, across the surface. The PC mouse and physical gestures are other examples. In contrast, isometric contractions involve changes to the tension of a muscle, while the length remains constant, resulting from increasing resistance against a static or rigid object. Isometric contractions are those that vary the amount of pressure applied to an object. When the extrinsic/intrinsic hand muscles cannot contract, due to the fingers contacting a rigid or static object, increasing muscular contraction is replaced by increases in muscular tension, which increases the pressure applied through the fingers. Examples of isometric devices in HCI include the IBM TrackPoint, force-sensitive resistors (FSR) and styli for computer-aided design, such as those from Wacom[5]. All of the extrinsic and intrinsic muscles listed above can be involved in the application of pressure; however, prehension (grasping, squeezing) is achieved primarily with the extrinsic muscles in combination with the pronating thenar muscles to oppose the thumb to the fingers. Objects are grasped by placing them between opposing fingers and thumb (or between the fingers and the palm).

2.1.2 Sensory Physiology

[5] http://www.wacom.com/

Feedback concerning pressure application comes from sensory receptors in the contracting muscles as well as from the skin contacting the object [126]. The muscle receptors give information about muscle length and tension, while cutaneous mechanoreceptors give information about skin compression and stretching.

2.1.2.1 Muscle Receptors

There are three types of muscle receptors: two stretch receptors (spindle receptors) and Golgi tendon organs [126]. The Golgi tendon organs are located at the junction between a small group of muscle fibres and the muscle tendon, and respond to being compressed by increasing tension within the local group of muscle fibres only. These receptors provide information about the level of muscle tension and so the amount of pressure being generated [114]. Stretch receptors provide information concerning muscle length and limb position and so are less important in understanding muscle tension/applied pressure, however, accuracy in judging pressure output was higher when the wrist and arm were free to move, compared to when restricted, suggesting arm movement/position may provide some feedback [72].

2.1.2.2 Cutaneous Mechanoreceptors

The hairless skin of the palm is referred to as *glabrous* and this type of skin has four types of mechanoreceptor, described based on the rate ("Slow" or "Fast") at which they stop responding to a sustained stimulus (called "adaptation" to the stimulus) and the size of the receptors receptive field (numerically designated 1 for small, and 2 for large). There are therefore, slow-adapting, small field (SA1), slow-adapting, large field (SA2), fast-adapting, small field (FA1) and fast-adapting, large field (FA2) receptors. The receptors most relevant to detecting the amount of applied pressure are the SA1, SA2 and FA1.

SA1 receptors are at high density at the fingertips, with some in the phalanges. While also sensitive to fine spatial details, such as points, edges and curvature [126] they are most sensitive to normal force: pressure applied directly upon the skin [9, 59] (opposed to at an angle or across the skin). This pressure is generated by having an object rest upon the skin parallel to gravity (for example resting a ball in an up-turned palm) or by pressing the fingers/hand against an object. SA2 receptors are at low density in the hand, with slightly more in the palm than the phalanges or fingertips. They are $1/6^{th}$ as sensitive as SA1 receptors but have receptive fields five times larger [126]. They contribute to perception of skin stretch [9] and the direction and pressure of object motion, when that object stretches the skin [180], such as during grasping, when a heavy object, under the influence of gravity, stretches the skin. These receptors also play a role in perceiving hand configuration and

finger position [45]. Finally, FA1 receptors, with twice the density at the fingertips as SA1 receptors [126], respond most to tangential force (across the skin surface) and so are critical during precision grip [116], as they detect object weight and slippage. As pressure is applied to an object through the fingertips, both the contact area of skin [238] and the displacement of skin at the fingertip [208] increase rapidly, reaching a peak and plateau within 3-4 Newtons. This rapid change in cutaneous feedback can provide rich information concerning the amount of pressure applied [121, 127] as well as indicate the compliance of the object [6, 217].

Having outlined the physiological basis of pressure application and sensation, the remainder of this section describes the uses and precision of applied pressure and how the feedback provided influences these actions. Lederman and Klatzky [148] make the distinction between the *sensory* subsystem of touch, through cutaneous, thermal and kinaesthetic information[6] and the *motor* subsystem of grasping and manipulation. As is discussed below, the application of pressure is a part of both subsystems, but used for different purposes: it is used as an exploratory behaviour to investigate object properties, and it is used to act upon objects, such as holding, squeezing, pushing or bending. Both aspects are discussed in this section, starting with investigative touch.

2.1.3 Psychophysics of Applied Pressure Control

This section describes research from psychophysical science on the limits of human ability to apply pressure, including the smallest amounts detectable, the largest amounts producible and the precision of pressure application. This research is indispensible in the design of appropriate and effective pressure-based human-computer interfaces, as an understanding of how humans can apply pressure leads to the creation of interfaces based on human ability, rather than the affordances or capabilities of facilitating hardware, for example. The SI unit for measuring applied pressure (force) is the *Newton* (N). One Newton is equivalent to the force required to accelerate a mass of one kilogram by 1 m/s^2 (meter per second, per second). In terms of normal force, it is equivalent to the force of gravity on an object of approximately 100 grams. Research in psychophysics examining human application of pressure refers to any given pressure level in terms of Newtons, thereby also describing the limits of human ability (for example the maximum that can be applied) in terms of Newtons.

[6] The umbrella term when considering cutaneous and kinaesthetic feedback together is also known as "haptic" feedback.

There is a considerable amount of research on how precisely humans can apply pressure through the fingers. This includes measuring control during the abstract (and explicit) application of pressure to immovable pressure-sensing apparatus, as well as more concrete applications of reactive (and implicit) pressure during active gripping and lifting of objects. These two areas both provide useful insights into the way that humans apply pressure and how they perceive the pressure they are applying. A number of studies testing control of applied pressure (both abstract and reactive) compare control when the individual is presented with different forms of feedback, to judge the usefulness or efficacy of each different feedback source. These sources include not only the inherent haptic feedback, but also external visual, audio or vibrotactile feedback. These studies are focused on particularly because external feedback is an important part of interaction with electronic devices, and it is also a focus of this thesis, and so it important to show how different forms of feedback influence control of pressure-based interfaces.

2.1.3.1 The Limits of Pressure Application/Perception

The maximum amount of pressure that an individual can apply in a particular manner or manipulation form (for example, pressing with a finger or squeezing a whole-hand grip) is referred to as the *maximum voluntary contraction*, or MVC. Different individuals have different MVCs for the same manipulation form, due to differences in muscle strength and flexibility [214]. Also the same individual will have different MVCs for different manipulation forms or when using the dominant vs. the non-dominant hand [87, 119]. The same individual may also produce different MVC values for the same manipulation over time [33, 109]. As the amount of pressure that an individual is asked to produce increases, the lower the total time that they can maintain that magnitude [125]. The MVC is an indicator of individual differences, and what influence these differences might have on the design of pressure-based interfaces (such as the need to adapt to the user's capabilities). Target pressures in psychophysical experiments are often described as a percentage of the individual's MVC.

The *just noticeable difference* (JND), also referred to as the *differential threshold* or *difference limen*, refers to the smallest difference between two magnitudes of applied pressure that the individual can identify as perceptually distinct. Pang *et al.* [183] and Tan *et al.* [231] tested JND of applied pressure via finger-thumb pinch, where a reference amount of pressure (between 2.5N and 10N) was produced followed by different comparative pressure. They both found a JND of 5-10% of the reference pressure, regardless of the reference pressure. This meant that a given comparative pressure level had to be 5-10% larger or smaller than a reference level of pressure to feel like a different magnitude.

2.1.3.2 Precision of Applied Pressure

The most common method of measuring control of applied pressure is through "magnitude production", where the individual attempts to produce a magnitude of applied pressure equal to a specified target magnitude. The target magnitude can be set by quantitative measures, such as by a visual representation of current and target pressure on an oscilloscope or other digital display, or by more qualitative measures, such as producing a set % of the individual's maximum effort or matching the pressure exerted by one limb with another limb (with no external feedback). The absolute error (difference between applied and target pressure) and variance (often the standard deviation of input) of applied pressure are generally used as measures of precision, in that a higher error or variance indicates poorer precision. A common method of judging an individual's perception of how much pressure they are applying is through "force-matching", where a reference pressure (force) is produced by the individual and is then matched in a subsequent attempt. This can be done unilaterally (reference and matching pressure are produced using the same arm) or bilaterally (reference and matching pressure are produced using opposite arms). An example: the experimenter asks the individual to increase the pressure produced by the right forefinger on a load cell to a set magnitude (seen by the experimenter as a line on a computer screen) and maintain it briefly. The individual is then asked to relax, and re-produce the same pressure level using the left forefinger. By measuring the error, force-matching shows how accurately the individual can produce and reproduce pressure levels.

There is a general trend that the precision with which we can apply pressure is related to the relative magnitude of the pressure compared to the individual's MVC. The relationship may be approximately U-shaped, as precision in applying low levels of pressure (relative to MVC) and high levels of pressure is worse than applying moderate pressure [119, 124, 213]. However, other research has simply found that error in maintaining pressure increases as the target pressure increases [218]. While research has found that it is more difficult to apply more pressure accurately, due to fatigue [123, 125], it also appears that applying very low levels of pressure (relative to the individual's ability) is also difficult. This suggests that user interfaces should potentially avoid both high and low levels of pressure, as well as tailor the interface to the individual.

There is also a trend that the inaccuracy at low levels of pressure results from inadvertently over-exerting and applying too much pressure (by under-estimating the extent of pressure being applied), and the inaccuracy at high levels results from under-exerting (over-estimating) [119, 124]. These results came from studies that gave no external feedback

concerning how much pressure was being applied, so participants had to rely on only the inherent haptic feedback from the skin and muscles, coupled with the efferent (signals sent from the brain to the muscles) motor signals. The influence of external feedback on pressure application is discussed in more detail below in Section 2.1.3.3, however, an additional influence, related to precision at varying levels of pressure, is the number of digits used to apply the pressure. Newell and McDonald [177] found that accuracy depends on the number of digits used, relative to the magnitude of target pressure, so that a finger + thumb pinch could accurately apply low pressure (10% of MVC) but not higher pressure (50%) whereas whole-hand grip (thumb + all four fingers) could apply higher levels accurately but not lower levels. They conclude that more digits introduce a greater number of *degrees of freedom* in the behaviour. Using more digits for higher pressure improves accuracy but using more digits for lower pressure introduces *redundancy*, leading to poorer accuracy.

There is evidence for a short-term motor memory effect during repetition of pressure output, specifically when pulling on a lever. Fowler and Notterman [51] found that if participants attempted to pull to the same target pressure magnitude within 10 seconds of a previous attempt, they were 30% more accurate than if the gap between attempts was greater than 10 seconds, or if different target magnitudes were attempted on successive trials.

2.1.3.3 The Influence of Feedback on Precision

This section will show how the feedback available to the individual can strongly influence precision of applied pressure. Feedback sources include internal haptic (cutaneous and kinaesthetic) sensations from mechanoreceptors and muscle receptors, as well as external visual, audio or vibrotactile feedback. As external feedback is a major component of mobile device interaction, this section will show how the input range of an interface, and how accurately we can interact with it, could be improved by the addition of external feedback.

Section 2.1.2 briefly described the muscle receptors and mechanoreceptors in the hand and extrinsic hand muscles, as well as the feedback they provide on how much pressure is being applied. The following research has looked at the accuracy of applied pressure, both when the individual has access only to this internal information, as well as when they have access to external visual, audio or vibrotactile feedback.

There has been a considerable amount of research on how cutaneous feedback is necessary, and sufficient by itself, for accurate control when gripping objects [72, 115, 116]. This precise control is in fact partly subconscious, as reactions to changes in “load force” are faster than cognitive reaction time [115, 116]. However, it appears that haptic feedback alone

may be insufficient for very precise *deliberate* application of pressure, as is the case when participants are asked to apply target levels of pressure to a load cell or dynamometer.

Jones [121] found that maintenance of low levels of pressure (2-6 N) pressing with the index finger was accurate to within 1 N (50-83% accuracy), but that the addition of visual feedback dropped this error value to 0.22 N (78-96% accuracy). Participants overshot lower target pressure and undershot higher target pressure when no visual feedback was available.

Henningsen *et al.* [87] found that pressing on a conical pad resulted in lower applied pressure than was required, compared to pressing on a flat pad, during concurrent bilateral force-matching using the fingertip. They explain this result through the sensitivity of the mechanoreceptors in the fingertip responsible for detecting pressure, as they respond more to edges and pointed surfaces, and the intensity of the associated tactile sensation is more closely tied to the extent of indentation itself rather than the level of pressure *per se.* This may lead to a reduction in the pressure applied to a conical pad as a greater magnitude of tactile sensation arises, relative to the pressure applied. This study suggests that tactile feedback is more important than kinaesthetic feedback in the application of low pressure at the fingertip, as kinaesthetic information is still available while pressing on the conical pad.

Jones and Piateski [127] tested bilateral force-matching using the index finger, index + middle + thumb together and the elbow when the user had full haptic (tactile and kinaesthetic) feedback available and when tactile feedback was removed by the use of a rigid splint placed between the fingers and the apparatus. The magnitude of the matching pressure was lower than the reference pressure when tactile feedback was removed, while accuracy was much better when tactile feedback was available. This suggests that tactile feedback from skin deformation/stretch and contact area is important in accurately judging the extent of applied pressure. Removing tactile feedback through anaesthetizing the fingertips results in a matching pressure higher than reference pressure [137] (the opposite of the results from Jones and Piateski), so the present, but impoverished, tactile feedback from pressing on the splint may still have provided some information.

External feedback

These studies described thus far show that haptic feedback from the skin and muscles helps us to judge how much pressure we are applying. However, as is shown in the research in this section, for very precise application of specific levels of pressure, external feedback is needed, and so external feedback is likely to be needed in mobile pressure-based HCI.

Hong *et al.* [108] measured the accuracy of pressure applied normally (parallel to gravity) by the index finger and varied both the gain (vertical resolution) and the frequency (regularity) of visual feedback available to the individual. The gain varied from 2 pixels per Newton (p/N) to 512 p/N, while the frequency varied from 0.4 Hz to 25.6 Hz. Both factors influenced the accuracy of applied pressure but in different ways. Reducing the gain to 2 p/N, forcing the individuals to rely more on tactile sensation, led participants to significantly undershoot (i.e., press less than) the target pressure, a similar finding to Jones and Piateski [127]. The standard deviation (or variation) of input was also significantly worse when gain was 2 p/N or 32 p/N, compared to 512 p/N. Lowering the frequency at which the level of pressure was shown on screen significantly increased the variation of input, a similar result to that found in the authors' previous study [214]. These results suggest that more continuous visual feedback is needed for highly accurate control; however, increasing the gain past 128 p/N or increasing the frequency past 0.4 Hz (>=3.2 Hz) had no effect on performance, so there appears to be a limit to the benefits of the spatial and temporal resolution of visual feedback. Hong *et al.* [108] also found that the natural individual differences in accuracy were more pronounced as visual feedback was reduced.

Mai *et al.* [157] also varied the visual feedback available while gripping a device between forefinger and thumb at low levels of pressure (less than 2.5 N), and maintaining the pressure for 20 seconds. Average error (distance from target pressure) was highest when no visual feedback was provided. Error was slightly lower when discrete visual feedback (which only indicated whether too much, too little or the correct amount of pressure was being applied) was available. Best performance occurred when continuous visual feedback was provided.

As well as showing that the number of digits used to apply pressure influences the accuracy of the pressure produced, Newell and McDonald [177] found that the gain (vertical resolution) of visual feedback influenced different grip formations differently. Accuracy when squeezing with a thumb-finger pinch did not improve with increased visual resolution, but accuracy during whole-hand grip did improve when more visual information was available. The authors argue that the whole-hand grip, with redundant 'degrees of freedom', was better able to make use of extra visual feedback.

When both visual and haptic feedback is available, but the information they provide is deliberately contradictory, there is evidence that visual feedback is given precedence. Srinivasan *et al*. [215] asked participants to push on springs of varying stiffness while showing a visual representation of the springs' compression and judge which spring was

stiffer. This visual representation showed either the correct and corresponding compression relative to the extent the user pressed on the spring, or gave a false impression of stiffness. The change in visual-real discrepancy was denoted by λ, with 0λ meaning no discrepancy, 0.5λ meant equal visual displacement for left and right, regardless of underlying displacement and 1λ being a complete reversal/swapping, with the visual displacement equal to the other spring's displacement, given the same pressure. Srinivasan *et al*. found that, as the discrepancy between visual and tactile feedback increased, correct discrimination (based on stiffness) decreased. Therefore, discrimination of stiffness is not solely based on haptics when visual information is available, as with no visual info, stiffness discrimination was 98% accurate. However, at 0.5λ the visual feedback showed the spring reacting in the same way for the given pressure, when physical displacement was different. In this case, accuracy was 67%, higher than chance, suggesting that tactile sensation still influences discrimination.

From the research summarised here it is clear that the feedback available to the individual is of utmost importance when attempting to accurately apply pressure. To produce suitable pressure-based interfaces for mobile devices, it will be necessary to create feedback designs that allow for quick and accurate application of pressure, and so optimize performance during a pressure-based interaction. In particular, it is clear that some form of external feedback, particularly visual feedback, will be necessary for accurate control of the interface. User input is likely to benefit if the external feedback is continuous and makes use of as high a spatial resolution as is available. When the individual is provided with external feedback concerning how much pressure is being applied, it is evident that increasing the amount of feedback (or the information contained within it) improves the accuracy of applied pressure, up to a point.

2.2 The Application of Pressure as Input in HCI

Section 2.1 described psychophysical research relevant to the research questions on pressure-based input: how accurately humans can control the amount of pressure they apply when presented with varying feedback methods, and through various combinations of fingers when sitting in a lab. However, to understand how these results relate to more realistic HCI scenarios, and so better frame the research questions, this next section discusses how our ability to apply pressure has been successfully leveraged, to an extent, as a means of providing input to a computer interface. Section 2.2.1 describes the applications of pressure as an input to a computer system, including the research on accuracy of pressure-based input.

Section 2.2.2 then discusses the research examining the influence of varying external feedback on pressure-based input. These summaries will go some way to framing each research question in context, by showing how accurate pressure-based input can be when sitting and being provided with visual feedback.

2.2.1 Applications and Measuring Precision of Input

This section describes the ways in which digital pressure input has been used as a part of interfaces in HCI, including those from static desktop scenarios as well as mobile devices. While many examples are given, the focus is on pressure-based linear targeting, a variation on magnitude production tasks in psychophysics, which helps to show how accurate pressure-based input in HCI can be. The term 'pressure space' is used here to refer to the total amount of pressure that is used in an interaction, be that the maximum amount the given sensor can detect or a limit enforced by the experimenters.

2.2.1.1 Control Methods

Before discussing the uses of pressure input, this section will briefly outline the differences between the two primary means of using pressure to control interface elements; they are referred to as 1) *positional* control and 2) *velocity* or *rate-based* control. The relative merits of each are briefly discussed in Section 2.2.1.2. For illustration, control of a pointing cursor is used as an example, using an isometric joystick: a pressure-sensitive omni-directional joystick with no physical travel.

Positional Control

It is called positional because the *position* of the interface element, i.e., the cursor, within the interaction space (such as the viewable dimensions of a computer screen) is controlled by the amount of pressure applied. The more pressure that is applied (and maintained) the further the cursor moves from its starting point. Releasing the amount of pressure applied returns the cursor to its starting point, as if elasticated. In the joystick example, the pointing cursor would start in the middle of a screen/display and pressing on the joystick would move it in the relevant direction away form the centre point. Pressing (and holding) the joystick lightly to the left would move (and maintain) the cursor at a position slightly to the left-of-centre. Pressing (and holding) hard to the right would move (and maintain) the cursor at a position at the far right. This method is a 1:1 relationship between applied pressure and cursor movement/position and is akin to pressing on a spring. It is the method used in the traditional

PC mouse.

Rate-based (Velocity) Control

In this control method, the *speed* of the interface element's (i.e., cursor) movement within the interaction space (screen) is controlled by the amount of pressure applied. Speed increases as the amount of pressure applied increases and the cursor comes to a halt when no pressure is applied. In the joystick example, pressing (and holding) the joystick lightly to the left would make the cursor move continuously at a low speed to the left, until the edge of the screen is met or the pressure is removed from the joystick. Pressing (and holding) hard to the right would move the cursor continuously at a high speed to the right. This method is similar to pushing an object along a smooth surface: how hard the object is pushed dictates how fast it moves, and stopping pushing stops the object's movement.

2.2.1.2 Static and Desktop Applications

One of the earliest common uses of pressure in HCI was two-dimensional pointing using an isometric joystick such as the IBM TrackPoint and velocity-based control. Many studies have measured pointing performance using the TrackPoint (or variations of it) including comparisons to other pointing devices. Its performance is generally slower and less accurate than regular mouse movement through isotonic (and positional) control [113, 167, 260]. Campbell *et al*. [27] found that adding tactile feedback to a TrackPoint-style device improved control over cursor movement through a narrow tunnel. When looking at the microstructure of pointing movements using both an isometric joystick and a mouse, Mithal and Douglas [167] found that the pressure-based input of the joystick led to faster and more precise initial movements but that finer homing movements were slower and less accurate, whereas the mouse had a slower and less accurate first movement but more accurate homing behaviour. This suggests that isometric devices are not as well suited to 2D pointing as isotonic devices.

However, when using isometric devices such as the TrackPoint or force-sensing resistors (FSR), velocity (or rate-based) input is the most suitable control method, compared to positional input [260]. In contrast, positional control is better suited to isotonic devices [260]. Pressure has since been used for other methods of movement or traversal, such as zooming [23, 191], where the pressure applied to a Wacom stylus controlled the magnification level of zoom in a document, and object rotation [210], where two force-sensitive resistors attached to a computer mouse provided bi-directional rotation of shapes, in conjunction with x-y mouse movement. In this latter study, Shi *et al.* [210] compared

positional and rate-based control for object rotation and found that rate-based control resulted in better performance.

A common use of pressure input, especially using pressure-sensitive stylus pens, is the creation of buttons, switches or gestures for changing the visual features or function of an interface element. The main benefit of this is to provide quicker and easier access to common actions, rather than pressing keyboard buttons or on-screen menu options. Li *et al.* [153] used a high level of pressure applied to a stylus to activate mode-switching with moderate success, however, they mentioned that tailoring the interaction to individual differences in ability would improve performance. Forlines *et al.* [50] used a light stylus press to provide a preview state for GUI actions, such as zoom level, colour selection and window management. A heavy press would confirm the action while releasing the light press returned to the original state. Ramos *et al.* [190] also used a stylus to generate movement + pressure-based gestures called "Pressure Marks", which could be used as shortcuts to common GUI commands. Marks ranged from simple straight-line movements accompanied by low, high or low-to-high/high-to-low pressure changes to "lasso" and "tail" movements, which circled onscreen objects. A major limitation of this type of implementation is that it makes little use of the range of available pressure: they are effectively swapping physical button clicks for threshold pressure values.

Similar research that has used a much wider range of pressure input includes the mapping of pressure to cursor size in object selection, where both Raisamo [188, 189] and Ren *et al.* [197] used the extent of pressure applied to a pressure-sensitive kiosk screen or stylus, respectively, to control the radius of an on-screen cursor. Raisamo found the pressure-based method to be the least preferred method of area selection, compared to time-based or direct manipulation alternatives, but Ren *et al.* found their pressure-based "Adaptive Hybrid Cursor" to have the fastest object-selection time, lowest error rate and highest preference. The vast difference in findings could be down to the hardware and manipulation method. Raisamo used a single, unsupported finger to press on a screen which detected pressure based on area of finger contact (rather than the extent of pressure *per se*). Ren *et al.* used a pressure-sensitive stylus, which is held and supported by multiple fingers and the holding hand and arm are rested on a flat surface. The sensor within the stylus also directly measures pressure. Therefore, input in Ren *et al.* is likely to have been more stable, both in terms of physical support and reliability of input. Kildal [139, 140] also used the range of stylus pressure, however, not for functional input to an interface but for tactile feedback from it. He mimicked surface friction and compliance of virtual objects through a vibrotactile transducer attached to the stylus, an effect he showed to be convincing.

Finally, using squeezing for input and vibrotactile feedback for output has been investigated as a means of communicating emotions or actions between remote individuals. Rantala *et al.* [193] compared user preferences towards three gesture types for creating haptic messages with a hand-held device: moving (in 3D space), squeezing (the sides of the device) and stroking (the surface of the device). Participants generated and performed gestures to convey or represent: excitement, agreement, alerting (request for action) and love/longing. Vibrotactile feedback was generated relative to the intensity of input. Squeezing and stroking were reported as the most preferred methods for generating messages, with pressure input being rated as "easy" and "pleasant", but not very "expressive", to use. Users considered squeezing as an applicable means of conveying excitement, agreement and drawing attention, but less so for love/longing. The study only investigated the generation of messages from a sender and so Heikkinen *et al.* [83] tested interpretation of feedback by a receiver, using the feedback designs created by participants in Rantala *et al.* [193]. Participants were able to interpret the intensity of a squeeze-based tactile message from the stimulus alone, but other details about the messages meaning or purpose could only be interpreted when the message had context, in terms of where the receiver was and what they were doing.

Suhonen *et al.* [230] extended this research by utilising squeezing as both an input and an output method, where participants wore a constricting wristband that contracted relative to the extent of squeezing on the input device. As is discussed in Section 2.4.2, this study also utilised thermal feedback as a means of communication, relative to squeezing input. Squeezing (as well as warmth and cold) was used to convey emotions and actions during discussions of positive ("happy") or negative ("sad" or "angry") events that the participants had experienced, as well as a neutral, hypothetical event ("restaurant"). In the study squeezing was often used to convey physical touch or for emphasis of something said. Squeezing was also used the most, compared to warmth and cold, and was rated as "easy", "natural", "less confusing" and "pleasant". While this research did not measure the precision with which participants could apply pressure, it gives useful insight into how participants think about pressure/squeezing, both as a natural action with inherent properties/purposes and as an input to a communication system. In this way it shows that applied pressure in a natural means of manipulating, interacting and using, and so it is well suited for use in HCI.

Linear Targeting

The most important application of pressure-based input in relation to this thesis research is that of linear targeting, as it is a variation of magnitude production tasks used in

psychophysics, and so is a useful means of measuring the precision of pressure-based input. Pressure-based linear targeting divides the pressure space into a set number of levels, or bins, of a given width (in Newtons or sensor values) and each level/bin is a potential 'target'. For example, if a pressure sensor can detect approximately 4N of pressure, the 4N could be divided into 10 levels/bins, each 0.4N wide. Level/target 1 covers 0-0.4N, target 2 covers 0.4-0.8N, target 3 covers 0.8-1.2N and so on. The task requires the individual to apply sufficient pressure to be within a target level: to acquire target 3 the individual must press between 0.8 and 1.2N. When the individual has applied enough pressure to be within the target level/bin, he/she activates a selection mechanism (discussed in this section) to confirm selection of the target level. Pressure-based linear targeting is illustrated in Figure 2-4. In the figure, the pressure space (detectable pressure) is laid out vertically from top-to-bottom, starting at the top with 0-pressure. Figure 2-4, left, shows the GUI from Ramos *et al.* [192], where the pressure space is divided into 4 levels and the amount of pressure applied (using positional input) is shown by the position of a blue cursor (seen at the top). The individual would apply pressure to move the cursor into a target level (or Distance, labelled D1-D4 in the left-hand image) and select that level to proceed. As shown in Figure 2-4, right, Ramos *et al.* [192] divided the same space into 4, 6, 8 and 10 levels, thereby decreasing the size of the targets.

In HCI, the task is used not only to measure the control of pressure-based input, but it also serves as the evaluation of targeting-based pressure interactions, such as a menu or other single-axis selection task. Making the target levels thinner, such as by increasing the number of levels/targets within the same space (e.g., 16 levels across 4N rather than 10) gives an indication of how precisely users can apply pressure as well as how many levels can feasibly be used in an interaction such as menus, zoom levels or paint brush thickness. Selecting targets at different positions along the pressure space axis also gives an indication of control at different magnitudes of pressure. Because linear targeting is based around levels/bins of pressure, conclusions drawn about how accurately users can apply pressure in HCI is framed in terms of the number of levels that can be selected with an acceptable level of accuracy (although what is considered "acceptable" varies). A larger number of levels (within the pressure space) would result in thinner levels suggesting higher accuracy of pressure input. This is different than the average error (distance from target pressure, measured in Newtons) used to indicate precision in psychophysics. The number of recommended levels is, in fact, less important than the width of those levels; however, in the research, the number of levels acts as an implicit indication of level width: a recommendation of fewer, larger levels suggests poor precision of input, as more, thinner levels cannot be selected accurately.

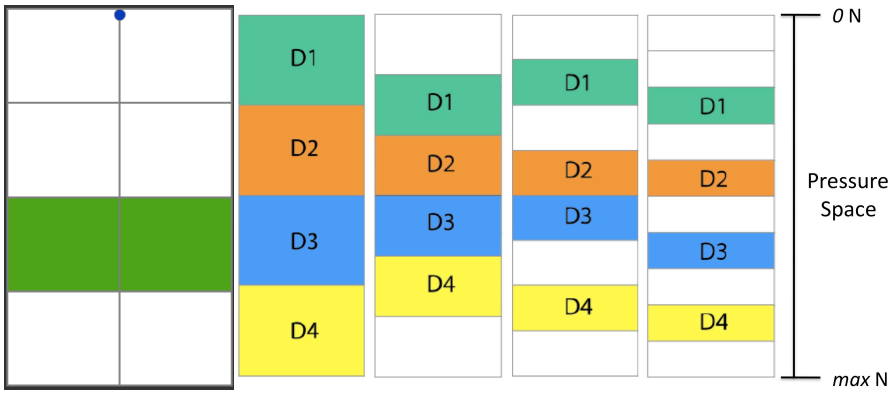

Figure 2-4: Pressure-based linear targeting illustration, adapted from Ramos *et al.* [192]. The pressure space ranges from the top (no pressure applied) to the bottom (max pressure applied). Applied pressure is indicated by position of a cursor (left, as a blue dot) through the pressure space. Applying more pressure moves the cursor further down the pressure space and the task is to move cursor into a given target (D1-4, right image) and select it

Ramos *et al.* [192] used a Wacom Intuos pressure-sensitive stylus to investigate the feasibility of using pressure for general GUI interactions. They used a Fitts' law-based [49] linear targeting task to establish how many levels of pressure users can accurately discriminate between. While they only describe the size of the pressure space (the range of pressure used) in terms of the 1024 sensor values that the stylus outputs, Wacom stylus pens are reported to have a range of 4 N (400g)[7]. They divided the pressure space into 4, 6, 8, 10 and 12 levels of equal width (see Figure 2-4). They also compared performance when presented with continuous visual feedback, where both the target level and a cursor indicating the extent of applied pressure was always visible, and partial visual feedback where the cursor was only visible when the trial started and only the target level was shown on screen. The latter partial feedback condition was to simulate 'expert behaviour'. They compared four different selection mechanisms: Click, Dwell, Quick Release and Stroke.

- *Click: Pressing the barrel button on the stylus*
- *Dwell: Remaining within the same pressure level for 1 second*
- *Quick Release: quickly lifting the stylus, removing all pressure input*
- *Stroke: Making a spatial movement to the right*

Performance was measured in terms of *error* (number of selections outside of the given target), *number of crossings* (the number of times the cursor crosses the boundary of the

[7] http://www.wacom.eu/index2.asp?pid=8025&lang=en

target level) and *movement time* (time from first non-0 pressure reading to level selection). They concluded that performance degraded markedly when more than 6 levels of pressure were used or when only partial visual feedback was available. This latter finding mirrors those from psychophysical science on the influence of impoverished feedback [108, 121, 214]. The error rate for 4 or 6 levels was between 1-8%, increasing to approximately 25% for 12 levels. They also found that control of pressure was worse at low magnitudes, resulting in higher errors and crossings for low-pressure target levels. The Dwell selection technique produced the least error-prone control, but it was also the slowest method, whereas Quick Release was the opposite: higher errors but fast. Both Click and Stroke performed poorly, as the inherent movement required by the mechanisms resulted in unintended changes in applied pressure, leading to higher errors or target crossings. For these reasons they recommended separation of movement and selection mechanism.

This task, or variations of it, has been used several times in HCI research to test input using different devices. Cechanowicz *et al.* [30] looked at several factors in pressure input using force-sensing resistors (FSRs) attached to a computer mouse, including the placement of sensors around the mouse, the use of two different sensors for concurrent input and the manner in which the pressure space is divided into levels (discretization). The FSRs used could detect a maximum of 1.5 N, outputting 1024 sensor values across this range. Cechanowicz *et al.* found that placing a sensor on the side of the mouse, to be activated by the thumb, produced best targeting accuracy across 6 levels of pressure, the same number as Ramos *et al.* [192], although error rates were slightly higher, at 14% for 6 levels. The best discretization function was a quadratic function centred at the lower range, which made the lower levels wider and increasingly thin further along the pressure space. This avoided the problem of poor control at low levels found by Ramos *et al.* [192]. Finally, they attached two pressure sensors to the mouse, one controlled by the thumb (on the left of the mouse) and one by the middle finger (on the front of the mouse) to traverse through up to 64 theoretical levels. One sensor activated coarse-grained traversal (through levels 1, 7, 13, 19 and so on) and the other activated fine-grained traversal (levels 1-6, or 7-12 etc.). Using discrete "taps" to traverse the coarse-grain levels (rather than continuous pressure input) task time and numbers of target crossings were similar when selecting the 64 theoretical levels, compared to 16 levels.

Shi *et al*. [209] extended this research on using FSRs attached to the mouse by exploring different visualisations of movement through the target levels/pressure space. Using the same 1.5 N sensor, they found that a fish-eye visualisation makes a higher number of levels more accurately selectable, improving control, however, error rates remained higher than

those of Ramos *et al.* [192] for the same number of levels (10-20% for 6 levels). In an adaptation of the linear targeting task, Shi *et al.* [210] used two sensors attached to the mouse to rotate shapes on screen to three set angles using alternative mappings of pressure to movement: "naïve" (linear mapping), rate-based (more pressure rotates faster), hierarchical (rate-based, coarse and fine-grained rotation using two sensors) and hybrid (linear coarse-grained rotation, rate-based fine-grained rotation). They found the rate-based method to be the fastest, most accurate and least mentally demanding interaction method.

One significant issue with the body of linear targeting research is that the size of the pressure space, and the number and size of levels within it, may vary between studies. Some research does not report the size of the pressure space in Newtons, only the range of values that the chosen sensor outputs for example, "a pressure space from 0 to 255" [23] and there is no mention of how sensor values relate to pressure. Some research does quantify the pressure space in Newtons, but these values also vary, from 1.5 N [29, 162, 209] to 4 N [170]. Research from psychophysics has shown that the amount of pressure to apply influences the precision with which it can be applied [119, 177, 213]. Therefore, recommendations for limiting pressure-based input to 6 [170, 192] or 12 [209] levels may be meaningless without the context of how much pressure those levels are spread across. It is, therefore, difficult to know how the results from one study may relate to others.

Another important issue to discuss is the manner in which pressure was applied. A stylus is held in a grip using at least two fingers, but more likely three, with the index finger and thumb providing downward pressure and the middle finger providing support. The hand holding the stylus, as well as the arm, may also be supported, lying on the desk. Similarly, a mouse with pressure sensors attached is stabilised sat on a desk and cupped in the palm and fingers, providing strong support against pressure applied to it. Raisamo [188, 189] used only a single, unsupported digit to provide pressure input and found poor results. Research from psychophysics shows that the number of digits used to apply pressure, in relation to the target level of pressure, influence accuracy [177].

This thesis focuses on pressure input for mobile devices, but summarising this research on desktops is necessary, as it provides context for the development of mobile pressure interfaces. It also shows how accurate pressure input can be when the individual is sitting in a stable and controlled environment, which can then be used for comparison when evaluating pressure-based input when walking.

2.2.1.3 Mobile Applications

In comparison to the physical stability of pressing on a stylus or FSR attached to a computer mouse while sitting at a desk, applying pressure to mobile devices is quite different. The same device is used for input and output and the device must be held and supported in the hand while pressure input is provided. This can lead to more complex manipulation methods, as a moving hand, not a rigid desk, provides resistance to pressure. Many proposed uses of pressure as an input to mobile devices have been similar to those proposed in desktop scenarios, however, some implementations have been specifically suited to mobile interaction.

More common uses include spatial traversal tasks such as zooming or scrolling using continuous input from a stylus [23, 187] or using FSRs located underneath traditional mobile phone keypads [35] or placed between a mobile touchscreen device and a case [168]. In one of the first examples of pressure input on mobile devices, Harrison *et al.* [78] suggested various uses for pressure sensors attached around a mobile device, such as navigating digital documents and turning pages (mimicking the real action of stroking a page across) and detecting how the device was being held (if at all). Gummi [205, 206] used a bendable surface and connected display to provide map zooming via bending the device up and down, where the amount of bend controlled the level of zoom. Subsequent research has investigated ways of augmenting touchscreen mobile devices with pressure-sensitivity. Heo and Lee [90] placed multiple FSRs around the back and sides of a touchscreen device to detect normal and tangential force applied to the device, against a surrounding case. They generated a set of "Force Gestures" for two-dimensional navigation of Web pages and e-books combining touchscreen x-y movement, discrete normal force and continuous tangential force. Harrison and Hudson [79] created a custom device consisting of an LCD display with attached touch-sensitive screen. Connecting these two devices were two elastic (self-centring) analog joysticks, which detected lateral motion from shear/tangential force, which could be coupled with the x-y coordinates of touch input from the connected screen. The research described thus far in this section has made use of dedicated pressure sensors, such as a pressure-sensitive stylus or FSRs. Goel *et al.* [57] looked at leveraging non pressure-related sensors and actuators in existing commercial devices to retrofit them with pressure sensitivity. They used the orientation of the device (through a gyroscope), the contact area of fingers on the touchscreen, and the dampening of vibration from squeezing the device to infer the amount of pressure being applied and the way that they device was being held (posture).

While these examples show a range of interesting implementations, unfortunately none were tested empirically to determine how well users could control them, either while sitting or walking. Baglioni *et al.* [3] used finger pad contact area on a touchscreen device as an

indication of applied pressure for the purpose of "braking" or slowing flick-based document scrolling. While not a direct measure of pressure input, applying more pressure increases the spread or the finger tip pulp, increasing the contact area [208]. They found that an automatic scrolling method, with pseudo-pressure braking, performed faster than traditional scrolling, which generally requires multiple flicks to move a large distance, due to virtual inertia. The pressure method was also more preferred.

Uses more specific to interaction with mobile devices include text entry. McCallum *et al.* [162] placed FSRs under a traditional MultiTap 12-button numeric keypad, where each key is used to input multiple letters based on the number of presses (for example, the number 2 key inputs a, b, or c when clicked once, twice or three times, respectively). McCallum *et al.* substituted multiple presses for three pressure levels where a single soft press inputs "a", a moderate press "b" and a hard press "c". They compared the pressure-based version to the traditional MultiTap method and pressure resulted in more words-per-minute typed. Brewster & Hughes [14] also used pressure for text entry, employing the resistive screen of a Nokia N800 (which converts contact area to a pseudo-pressure value). In this study, a low amount of pressure entered a lower case letter and a high amount of pressure entered a capital letter, removing the need for frequent movements to the shift key. They compared a Dwell and Quick Release selection mechanism for pressure input and compared both to the traditional text input method using the shift key. In line with Ramos *et al.*'s [192] findings, Dwell was more accurate than both the traditional method and Quick Release, but Quick Release was faster than both traditional and Dwell. Brewster & Hughes [14] also looked at the effect of walking on pressure-based text input performance, finding that walking significantly decreased typing accuracy but had no effect on words-per-minute. A final example of pressure used in hypothetical mobile scenarios is for biometric authentication [84, 85] where individuals tap rhythms in unique patterns of tap-pressure and inter-tap timing.

Linear Targeting

A number of studies have also used linear targeting to test how well users can apply pressure to mobile devices. The motivation for this research tends to be one of three: testing control of an interface relevant to mobile interaction (for example text entry); testing pressure control against existing means of interacting with mobile devices, such as tilt or buttons; or testing control of pressure applied to compact form factors to judge the influence of manipulation method. Despite the focus on "mobile interaction" only one study has tested control of pressure when the user is in motion.

Brewster & Hughes [14] is the only study to test the precision of deliberate pressure input while walking. This implementation only used two target levels of pressure, yet they found that walking significantly degraded accuracy. This study is limited in that it used so few levels and an impoverished sensor for input. Other research has tested targeting of larger numbers of thinner target levels of pressure, but has only done so when the participants were sitting down. As mentioned above, McCallum *et al.* [162] used three levels of pressure to input letters on an augmented MultiTap keypad, and found an error rate of 8.7%. This rate is as high as Ramos *et al.* [192] observed for selection of 6 levels, twice as many levels as McCallum *et al.* [162]. Ramos *et al.* did not report the size of pressure space in Newtons, but, as mentioned, Wacom stylus pens may have a range of 4 N. McCallum *et al.* [162] used 1.5 N of space, and so it is possible that the smaller range made the targets slightly thinner (~0.5 N wide compared to ~0.67 N wide in Ramos *et al.*) and so more difficult to select. In subsequent research, Stewart *et al.* [227] measured the extent of inadvertent changes in how much pressure was applied to a mobile phone during normal use (holding the device and talking on a phone call) when walking and sitting. Walking led to larger amounts of pressure being applied to hold the device, as well as more variant pressure input. From the findings they suggest mobile pressure input is likely to be more variant and less controlled than when sitting down.

Scott *et al.* [207] augmented an Ultra Mobile PC with FSRs in order to detect isometric bending and twisting of the device. While they envisioned its use for such things as page turning in e-books or application switching, they tested user accuracy in bending or twisting by target amounts through a linear targeting task. As both bending (up and down) and twisting (clockwise and anti-clockwise) are bi-directional, they measured targeting in both directions. They found that the time to acquire a target increased as the number of targets increased, with near targets (in either direction) requiring the most time to select when both bending and twisting (far targets also required more time when twisting). They did not measure selection error. The result that the nearest targets were difficult to select supports both Ramos *et al*. [192], where low levels of target pressure were difficult to select, and psychophysical research where low levels are also less precisely applied [119, 124, 213].

Stewart *et al.* [228] compared targeting performance when pressure was applied to a mobile device in different ways. These were: a single finger pressing on the device when it sits on a desk, pressing from both front and back sides (gripping) with thumb and forefinger when held in two hands, and pressing from the front and back individually, when held in the hands. Pressing only from the front (while held in the hands) was significantly slower than

the other methods, which were not different. Gripping (squeezing from both sides) had the lowest selection time. Lee *et al.* [149] carried out a systematic analysis of how precisely users could apply 2-dimensional tangential force (up-down and left-right) to a mobile touchscreen using either one hand (holding the device in one hand and using the thumb to apply pressure) or two hands (where one hand holds the device and the thumb of the other hand applies pressure). They also used a linear targeting task and found that handedness (one vs. two) did not influence input precision (targeting error, input variability or selection time), suggesting one-handed input can be as accurate as two-handed input. The direction of tangential force significantly influenced selection time, with up-down presses being faster than left-right. The overall error rate was low, at around 3-7% across targets 70-140 pixels wide.

Research examining various uses of pressure input on mobile devices is increasing in volume, and yet only one study has tested pressure input when the individuals were walking. Also, most research has used a low number of target pressure levels (up to three) while desktop counterparts have managed 6 or 12 levels. It is, therefore, fundamentally important to investigate a) how mobility influences pressure input and b) whether more, and consequently thinner, pressure levels can be accurately selected when walking. This led to Research Question (RQ) 2:

> *RQ2: "How accurate is pressure-based input through the fingers when the individual is walking?"*

Another important factor is the way in which pressure is applied to the device. The results from Stewart *et al.* [228] support those of Newell & McDonald [177], that applying pressure with different numbers of digits results in varying precision, as does the manner in which it is applied, such as where pressure is applied around the device and how it is held. Research has shown that more than one pressure input can be used successfully in desktop interaction [30, 209, 210], and multiple inputs on a mobile device could greatly expand the interaction possibilities. Most mobile studies have only used a single digit, or single sensor, for input, even though all five digits from the hand are in contact with the device when it is being held. This aspect of pressure input needs more research in the context of mobile interaction, which led to RQ 3:

> *RQ3: "How accurate is pressure-based input when multiple digits apply pressure to a mobile device"?*

Several other studies have measured precision of pressure applied to mobile devices, but they have focused on the effects of varying the feedback available to users. The influence of feedback on pressure input is of central importance to this thesis research, as the use of non-visual feedback is an important consideration when designing mobile interfaces, to allow visual attention to be paid to the environment. Therefore, the research investigating the effects of feedback on pressure input is discussed separately.

2.2.2 The Influence of Feedback on Pressure Input

There is considerable evidence from psychophysics that we need external feedback to apply very precise magnitudes of pressure, as the haptic feedback received from the skin and muscles is insufficient [108, 121, 127]. Therefore, this is also likely to be true during pressure-based interaction. Visual feedback is the most common means of conveying information in mobile devices; however, the use of non-visual feedback has been shown to improve mobile interaction [11, 47, 158]. This section discusses the HCI research on how removing or varying the amount of visual feedback influences accuracy of pressure-based input, and how non-visual audio and vibrotactile feedback has been used to substitute for visual feedback.

2.2.2.1 Removal or Variation of Visual Feedback

As mentioned in Section 2.2.1.2, Ramos *et al.* [192] compared accuracy of input when provided with continuous visual feedback and partial visual feedback (in the latter the cursor was not visible while applying pressure). They found that providing only partial visual feedback resulted in significantly more errors, more erratic input and longer selection times. Every participant took part in the partial feedback condition after having completed a full visual condition, to give him or her training in how much pressure is needed for each level. However, benefits of muscle memory in pressure application are said to last only 10 seconds [51]. Also, when both visual and inherent haptic feedback are presented together, there is a suggestion that the individual becomes dependent on visual feedback [215]. Therefore, the participants are unlikely to have been as "expert" in the behaviour as the authors had initially intended.

Srinivasan & Chen [216] tested the accuracy of maintaining constant pressure, controlled increases in pressure and matching sinusoidal variations in pressure (controlled increase/decrease) with the index finger. They compared performance during the presentation and removal of visual feedback and when the finger was anaesthetized,

removing tactile sensation. When visual feedback was removed, the average error (distance to target pressure) for maintaining constant pressure was significantly worse, and got progressively worse at higher pressure levels. Removing tactile feedback also increased error, but error remained constant across target levels. They also suggest that haptic interfaces require a pressure resolution of at least 0.01 N in order "to make full use of human haptic capabilities" (p. 125).

Mizobuchi *et al.* [170] spread 10 levels of pressure across 4 N and tasked users with selecting targets using a stylus pressed against a mobile device screen. Pressure was sensed by FSRs located underneath the screen and performance was measured under three feedback conditions:

- Continuous: a gauge was shown indicating the level of applied pressure throughout the entire pressure space.
- Discrete: only the number corresponding the pressure level currently being applied was shown on screen.
- No visual feedback: no indicator was shown at all.

Accuracy was worst when no feedback was provided, ranging from approximately 60% for the lowest level to as little as 10% for the highest level. Continuous feedback provided significantly better accuracy than discrete feedback. They also found that the 4 N pressure space was too large, leading to fatigue and that, when presented with no feedback, participants consistently over-estimated the amount of pressure they applied, leading to under-shooting all but the lowest two levels. While this study had users interacting with a hand-held mobile device, pressure was applied by multiple digits gripping a stylus, which may provide more support than a single digit. It also required two hands: one to hold the device and one to apply pressure.

This HCI-related research is consistent with psychophysical research on the negative effects of reducing or removing external feedback (particularly continuous visual feedback) [108, 121, 127]. HCI interfaces, therefore, appear to require continuous feedback of some sort, however, to better facilitate mobile interaction, non-visual audio or vibrotactile feedback may be necessary.

2.2.2.2 Non-Visual Feedback

Some research has looked at the effect of substituting visual feedback for forms of non-visual feedback such as audio or vibrotactile, while others have augmented visual

feedback with non-visual feedback to determine if the addition of more feedback affects performance.

As discussed in Section 2.2.1.2, Rantala *et al.* [193] and Heikkinen *et al.* [82, 83] used pressure input (squeezing) and vibrotactile output to send and convey emotional messages during communication. In the research, participants squeezed on a hand-held device with varying patterns and intensity and these characteristics were mapped to vibrotactile feedback. In Rantala *et al.* [193], the same participants provided pressure input and felt the corresponding feedback, to judge whether the intended feedback matched their desired intention. While participants used the form of the feedback to evaluate the manner in which they squeezed on the device, the authors did not specifically measure the precision of pressure input, nor how the form of vibrotactile feedback aided participants in the appropriate generation of desired pressure/squeezing.

Hoggan *et al.* [104] used the same hand-held device as Rantala *et al.* [193] and Heikkinen *et al.* [83] and compared whole-hand squeezing to tilting in a targeting-style menu interaction task. Participants had to squeeze or tilt the device by a target extent and maintain that extent for five seconds. They compared performance (time taken and precision of maintained squeeze/tilt) when provided with continuous visual feedback and when provided with both visual feedback and vibrotactile feedback. A brief 250Hz vibrotactile pulse was provided when the target extent/level had been reached and a 170Hz pulse was played when transitioning from one level to the next, with the amplitude increasing as the number of levels passed increased. They found that squeezing was significantly faster than tilting, and that being provided with tactile feedback (in addition to visual) made pressure input significantly faster. They also found that applying more than 4 N resulted in significantly higher variation of pressure input (as also found by Mizobuchi *et al*. [170]), especially when maintained for 3 or more seconds. Although visual feedback was still used in this study, it does suggest that additional non-visual feedback, which indicates the transition between pressure levels, is beneficial.

Stewart *et al.* [228] also used a 'level-transition' design of non-visual feedback, and compared selection of three levels of pressure when the individual was provided with different forms of feedback. Continuous visual feedback (cursor showing input was always visible but target to select disappeared upon first movement), discrete audio feedback (short tones of increasing pitch played as the level of pressure transitioned), discrete vibrotactile feedback (unique vibration patterns played at level transition) and combined audio + vibrotactile feedback were used. Visual feedback was significantly more precise (99%

accuracy) than all other forms, with audio (69% accuracy), vibration (82%) and audio + vibrotactile (71.3%) being similarly accurate. Modality did not influence the time to make a selection. This study used only 3 levels of pressure, but still found low accuracy for entirely non-visual feedback. While other research has extolled the virtues of continuous feedback [170, 192], Stewart *et al.* deliberately chose discrete feedback, as they report that participants can become less sensitive to continuous vibrotactile feedback [235] and pilot tests suggested continuous audio feedback to be annoying. While non-visual pressure interaction seems possible, interaction design may need to devise a continuous form of feedback that remains useful and does not put users off.

Tang *et al.* [232, 233] looked at concurrent pressure input with three digits (fore, middle and ring fingers) creating pressure 'chords'. Each digit applied one of three levels of pressure (low, medium or high) when presented with visual, audio or vibrotactile feedback or no feedback at all. In their first study [232], they found that task time (to create a chord) and error rate were significantly better when presented with feedback. Vibrotactile feedback produced significantly lower error rates than either visual or combined visual + vibrotactile feedback, and both the tactile and combined feedback was faster than visual. This was a surprising result, which the authors attribute to the abstract nature of the visual feedback. Rather than showing a continuous meter of applied pressure, colours were used to indicate the different pressure levels: green, blue and red for low, medium and high pressure, respectively. The vibrations increased in frequency as pressure increased, which was a potentially more logical mapping. Tang *et al.* [232] also suggest that having the same input and output channel (in this case haptic) can improve selection performance. A follow-up study [233] also included audio feedback, where the pitch/frequency increased as the pressure increased as well as a slightly different vibrotactile feedback design where the burst frequency for each pressure level was reduced. In this study, vibrotactile feedback led to slower chord formation than either audio or visual feedback, the opposite finding to the first study, but it remained the most accurate feedback method. Audio feedback produced more errors but was quicker.

Pressure and non-visual feedback have also been combined for exploratory or experiential purposes, not tied to specific interactions or interface elements. As mentioned before, Kildal [139-141] looked at mimicking surface compliance and friction through tactile feedback and a pressure-sensitive stylus. The study focused on how users described the sensations, to understand the relationship between pressure and perceived physical properties. Changing the frequency of vibration pulses, their regularity and their amplitude led to varying reports of compliance, elasticity, displacement and texture, all when applying pressure to an

isometric device. Lai *et al.* [145] followed on from this research by looking at the use of audio feedback and pressure input. Finally, Hoggan *et al.* [103] presented ForcePhone, which could be used to send vibrotactile messages (called "pressages") between phones during voice calls by squeezing (pressing) of the phone. Four levels of pressure were mapped to four different textures, and they were used to convey a variety of meanings between users, including greetings, playfulness or emotions.

This section has shown how accurate pressure input can be when visual feedback is altered or reduced or when non-visual feedback is used. The results from non-visual pressure input have been mixed, but it does suggest that accurate control using only audio or vibrotactile input is possible. Aside from accuracy, the main limitation of existing research is the small number of pressure levels used in the interaction (usually three), and so the more limited usefulness that non-visual pressure input may be suggested to have thus far. More research needs to be done looking at non-visual feedback designs that can facilitate interaction with a wider range of pressure input and so expand the bandwidth of non-visual input. It is also important to investigate whether or not this non-visual feedback design can facilitate interaction while walking, and so determine whether mobile eyes-free pressure input is feasible. Therefore, Research Question 1 asks:

> *RQ1: "How accurate is pressure-based input on a mobile device when using only audio feedback"*

2.3 The Human Thermal Sense

The human thermal sense encompasses two systems: the homeostatic system, monitoring internal core body temperature, and the cutaneous thermal sense, monitoring changes in external thermal stimulation. Thermal feedback in HCI acts to stimulate the skin in such a way as to produce detectable and interpretable sensations. To understand how thermal stimulation could be used to convey information in mobile HCI, it was necessary to understand the psychophysical characteristics of the cutaneous thermal sense: how changes in stimulation relate to changes in subjective sensation. Therefore, the literature summary here is limited to the cutaneous system.

Section 2.3.1 describes the physiological basis of thermal perception around the body, including sensory receptors and spatial sensitivity. Section 2.3.2 describes the limits of the

cutaneous thermal sense, including the smallest detectable changes. Section 2.3.3 describes various phenomena of thermal stimulation that can influence the internal subjective appraisal of a stimulus.

2.3.1 Physiology and Homeostasis

2.3.1.1 Thermal Receptors and the Basis of Thermal Sensation

The human skin rests in a relatively small 'neutral' homeostatic thermal state, on average ranging from around 28°C up to maximum of 39°C, when in moderate thermal environments [118]. The size of this neutral zone (between the minimum and maximum) is relatively constant across individuals at around 6-8°C, but due to individual differences in thermal sensitivity, the relative position of each individual's neutral zone varies (for example, 30-36°C or 28-34°C). Thermal perception is inextricably linked to skin temperature [223] and the behaviour of thermal receptors varies by how a thermal stimulus differs from current skin temperature [43, 60, 65, 118, 136]. Within the neutral zone there is no discernable thermal sensation of warmth or cold [128]; warm and cool receptors fire spontaneously at these temperatures, with no resulting thermal sensation [223]. Adaptation (where the sensation of thermal neutrality returns after heating or cooling the skin to a different temperature) only occurs within the neutral range [223]. Outside of this range a constant sensation of warmth (above) or cold (below) is perceived [134]. Kenshalo [132] suggests that cold perception has a more immediate onset whereas warm sensations grow slowly before 'blooming'. Note that throughout this review the terms "warm/warmth/warming" and "cool/cold/cooling" generally refer to increases or decreases in temperature, respectively. On occasion, if they are used to refer to subjective appraisals of specific temperatures, for example those that feel "warm" or "cold", they are used in terms such as "sensation of warmth/cool".

Scientific understanding of the characteristics and behaviour of the thermal sense is both wide and deep. However, the physiology underlying thermal perception is less well understood. While some nerve-endings and receptors in the skin have been identified as serving specific functions, others serve different (sometimes conflicting) functions and some perceptual processes are yet to have an identified neural/physiological source. There are specific warmth, cold and pain perception channels in the skin, but the three also interact [65]: there are purely warm-sensitive fibres, which activate and send signals to the nervous system when the skin is subjected to increases in temperature and result in sensations of non-painful warmth [65, 223]; there are also purely cold-sensitive fibres, which are active as a result of decreases in skin temperature, resulting in non-painful cooling sensations [60, 65, 223]. Finally there are purely nociceptive receptors, which respond to noxious (dangerous)

thermal stimuli, resulting in sensations of pain [223].

The afferent sensory fibres (or axons) identified as responding purely to warmth and cold are unmyelinated c-fibres and myelinated A-fibres, respectively [67, 88]. However, there are other fibres that respond to both heat and cold and are also involved in nociceptive sensations, thereby providing ambiguous information about skin temperature [65]. Some thermal perceptual phenomena may arise because of these mixed signals (see Section 2.3.3). Therefore, there can be non-painful warm and cold signals, painful non-noxious sensations (from receptors jointly sensitive to heat/cold and pain) and purely painful sensations (from areas only innervated by nociceptors). Purely warm and cool receptors/fibres increase in the extent and frequency of activation as the stimulating temperature increases/decreases, respectively. However, receptor activity can also be inhibited by activation of an opposing thermal channel: activation of cold fibres is inhibited by concurrent warm and cold stimulation [237, 259]. Warmth, cold and pain stimuli, in conjunction with information from central core temperature, are likely to be processed together in pre-cortical areas [67].

Sensory innervation in the skin is not even across the body. There are 30x as many cold-sensitive fibres as there are warm, and so we are more sensitive to cold than we are to warm (see Section 2.3.2), and the density of cold- and warm-sensitive 'spots' (receptive fields that are a few millimetres wide) varies from region to region [70, 223]. There are spots that are only sensitive to one sensation: warmth, cold or pain. Green and Cruz [68] found warmth-insensitive fields as large as $5cm^2$ on the arm of healthy, young people. On the forearm it is estimated that there are approximately 7 cold spots and 0.24 warm spots per $100mm^2$ [120]. The fibres identified as responsible for (at least part of all) warmth perception, C-fibres, have single, spot-like receptive fields [44, 73, 89], and in surveys of human cutaneous nerves these fibres have been difficult to find. The scarcity of C-warm fibres has been interpreted as evidence of low innervation density [73]. The uneven distribution of sensitive areas has an impact on pain perception (see Section 2.3.1.3) and may also contribute to certain sensory phenomena (Section 2.3.3).

Thermal perception is said to include an inherent hedonic element, where appraisal of pleasantness/unpleasantness is as important in the subjective judgement as the intensity of the stimulus [53, 171, 223, 229]. However, the processing of pleasantness/comfort and thermal sensation may occur in different parts of the brain, specifically the insula, cingulate gyrus, somatosensory areas SI and SII and thalamus for the thermal sensation and amygdala for comfort [131].

2.3.1.2 The Effects of Skin Temperature and the Environment on Perception

As mentioned above, thermal perception, and the resulting sensations, is tied to skin temperature, which is influenced by environmental temperatures, which can cool and warm the skin both within and beyond its natural homeostatic neutral zone. As skin temperature varies, our sensitivity to tactile stimulation, including thermal, texture and vibrotactile stimuli, varies with it. Internal thermal state, including homeostatic, hypothermic (low core body temperature) or hyperthermic (high core body temperature) does not influence intensity ratings of thermal stimuli, so set stimuli are rated as similarly "warm" or "cold" regardless of internal temperature [171]. Only comfort/pleasantness ratings are influenced by internal state, as cold temperatures are less pleasant when hypothermic and warm temperatures are less pleasant when hyperthermic [171]. When in a normal homeostatic state, any stimulus that is further away from neutral skin temperature (33°C in this case) was rated as less pleasant. Ambient humidity is an inherent aspect of environmental conditions, however, it may not have any substantial influence on thermal sensations [56].

Hirosawa *et al.* [92] found that the relationship between environmental temperature and skin temperature (tested on the fingertip) is sigmoidal. In a climate-controlled room that was set to ~11°C, skin temperature at the fingertip was 16°C. As the room warmed to 20°C, skin temperature rose quickly to 30-31°C. At room temperatures above 20°C, skin temperature only increased slightly, and rooms below 12°C causes skin to cool only slightly again. They concluded that the relationship between skin temperature (X) and warm (Wt) and cool (Ct) thresholds (the smallest detectable change in thermal stimulation) was near linear. Their equations for warm and cool thresholds were described by the equations:

$$\mathrm{Wt} = X0.69 + 11.56$$
$$\mathrm{Ct} = X0.1 + 2.64$$

Strigo *et al.* [229] found a similar relationship between room and skin temperatures. In their study, skin temperature was 30.1°C on average in a cooler room of 15°C, 33.4°C in a room or 25°C and 34.5°C in a warmer room of 35°C. They found that the perceived intensity of both cold (0-25°C) and hot (44-50°C) stimuli was reduced when in the cooler 15°C room, suggesting that thermal sensitivity in general drops as skin temperature drops.

There is research suggesting that skin temperature strongly influences tactile perception. Stevens & Hooper [219] found that skin temperature and object temperature interact to influence the perceived weight of the object. At a neutral skin temperature of 33°C, either warming or cooling the object led to the object feeling heavier than it was. When the skin

was warm (38°C), colder objects felt heavier, while warm objects felt no different, and when the skin was cooler (25°C) the perceived weight of all objects *fell*. Skin temperature can also influence vibrotactile sensitivity. Green [62] tested the effect of skin temperature on vibrotactile thresholds (the minimum peak-to-peak amplitude (in micrometres) required to generate a sensation) and found that cooling the skin (below 30°C) led to higher thresholds (i.e., it reduced sensitivity) compared to when the skin was warmed (above 35°C). However, skin temperature only affected higher frequency vibrations of 150Hz and 250Hz, as cooling the skin had no effect on frequencies of 80Hz or lower. Finally, cooling the skin also results in lower sensitivity to roughness [69]. The reduced sensitivity may be due to cold-induced vasoconstriction (contraction of the blood vessels) which may reduce the blood flow to receptors, reducing their activity [69]. Cooling may also increase skin stiffness, which can also result in poor detection of surface textures [147].

These influences have immense impact on the design of thermal feedback for mobile HCI. As sensitivity varies with skin temperature, and skin temperature varies widely, the feedback design needs to take skin temperature into account when choosing the starting temperatures from which to change and the extent by which it changes, to make stimuli reliably detectable. Is also means that there may be issues with combining thermal feedback with other forms of tactile feedback such as vibrotactile, as the individual's sensitivity to the vibrations may vary if the thermal feedback is particularly cold or warm (and the stimuli are presented to proximate locations on the skin).

2.3.1.3 Thermal Pain and Comfort

As mentioned above, there is an inherent hedonic quality or pleasantness associated with thermal stimulation [223]. Pleasurable responses to thermal stimuli may also be anticipatory, as brief cooling on a warm day (or *vice versa*) does not cool or warm the skin, and yet it provides a pleasant sensation [53]. Once thermal neutrality is reached (from a previous cold or warm state), the pleasurable sensation ceases. Conversely, unpleasant sensations will persist while an unpleasurable state, such as over-heating or cooling, continues.

Beyond sensations of discomfort are those of pain. There are receptors in the skin that respond purely to noxious stimuli and result in negative painful sensations [223], and others which respond to both noxious and non-noxious stimuli [65]. C-polymodal nociceptors (CPN [8]) are likely to contribute to both sensations of heat and heat pain, but activity in warm fibres contributes to the quality of sensation. Painful heat and heat pain depend on integration of activity including warm fibres and heat-sensitive nociceptors [65]. The point at which a thermal stimulus changes from one of simply heat to one of heat-pain (or from cold

to cold-pain) is called the pain threshold, specifically heat-pain threshold and cold-pain threshold. There are individual differences in the specific temperatures at which heat/cold-pain arise, but have been identified at around 45°C for warm pain and 11-15°C for cold pain [118, 223]. Because damage from cold takes longer than damage from heat, cold-pain onset is much longer than the almost immediate heat-pain onset (such as from burning).

There are several factors, other than individual differences, which can influence pain thresholds, including the location around the body, the skin type [81], body size [146] and gender [146, 166]. There is conflicting evidence on whether there are gender differences in thermal pain thresholds. Lautenbacher & Strian [146] found no difference in pain thresholds between males and females, but they did find an influence of body size, as the warm and cold pain thresholds increased (they moved further away from skin temperature) as body size increased. In contrast, Meh & Denislic [166] found that females had lower pain thresholds than males, although they suggest that the difference in results may come from the use of a larger thermal stimulator than Lautenbacher & Strian. This influence of areal extent of stimulation, called spatial summation, is discussed in Section 2.3.3.1. Regarding the influence of skin type, Harrison & Davis [81] compared pain thresholds on hairy skin (like that on the dorsal surface of the arm) against glabrous skin (hairless skin on the palm) and found that the hairy skin was more sensitive, requiring lower intensities to reach sensations of pain. Starting from 34°C, the smallest cold pain threshold found was 15°C for glabrous skin and 20°C for hairy skin, so hairy skin required less cooling to elicit pain.

The influence of body location on pain thresholds is more complex, as it not only relates to different sensitivity on different body parts but also within the same body part. Investigating warmth-sensitivity on the forearm, Green & Cruz [68] found areas of skin 5cm^2 in area that were generally insensitive to warm stimulation. At these positions, heating was not felt until temperatures exceeded 41°C and even heating to 44°C, near the pain threshold, only elicited sensations described as “barely detectable” or “weak”. When heat was eventually detected, the sensation was one of burning or stinging rather than warmth and heat pain thresholds were 2°C larger within these fields. Stimulating adjacent areas that possess normal warmth sensitivity produced much stronger warmth and heat sensations that increased as temperature increased. The results provide evidence for a (partially) separated sensitivity to pain and thermal sensation, but that warmth-sensitive fibres contribute to sensations of heat pain, not just nociceptors.

Thermal feedback that is painful to receive is unlikely to be accepted by users. An exception may come from specific usage scenarios where grabbing the attention of the user is

imperative and other feedback methods are not suitable or available. The pain will be more noticeable than more moderate thermal changes. In everyday usage scenarios, thermal feedback needs to avoid stimuli that would cause pain, and so it is important to:

- Use temperatures away from the pain thresholds, within the range of approximately 20-40°C.
- Reduce the extent of thermal change as the area of stimulation increases.

2.3.2 Thermal Sensitivity

One of the most important features of temperature perception is the rate of temperature change. Measures of thermal sensitivity generally focus on the size of the thermal "threshold". A threshold is the smallest change in temperature needed for an individual to notice a change in stimulation: it indicates how much warming or cooling is required before the individual feels the stimulation as warmer or cooler. It is the thermal equivalent of a 'just noticeable difference' and, the smaller the threshold, the more sensitive the skin is considered. Thresholds are measured from a set baseline temperature and are inextricably linked to both this starting temperature and the rate of change (ROC) of the stimulus.

At low rates of change (ROC) of up to 3°C/sec, the size of both warm and cold thresholds decrease as ROC increases, with the most dramatic decrease occurring from ~0.01°C/sec to 0.3°C/sec [133]. Cold thresholds are generally smaller and reduce faster as ROC increases, compared to warmth, as the sensory receptors identified as responsible for conveying cold perception are faster myelinated A-fibres, compared to unmyelinated warm C-fibres [67, 88]. Above 3°C/sec (up to ~7°C/sec) thresholds then begin to increase again, with this being attributed to the conduction velocities of thermal receptors [36, 185] as well as reaction and cognition time [81]. This suggests that stimuli become more salient as ROC increases; however, above a set speed, even if salience increases, the ROC 'overtakes' reaction time so that further increases in stimulation have occurred by the time the participant could react. As the skin adapts to the warm or cool extreme of the neutral zone, warm and cold thresholds, respectively, decrease and decrease more as the stimulus intensity approaches the heat/cold pain thresholds (~45°C and 11-15°C respectively [118, 223]). In other words, we become more sensitive to thermal changes if they move the skin closer to pain thresholds. Conversely, warm and cold thresholds increase as the skin is cooled and warmed respectively. From this, and other evidence, it is clear that the thermal sense is more sensitive to dynamic changes in temperature, rather than absolute temperature itself.

The effect of ROC on perception provides means for thermal feedback to produce a variety of temporally variant notifications. Slower changes, which take longer to detect, can be used for ambient notification, to gradually make the user aware of the feedback and not grab attention away from a current task. Increasing the ROC can then produce more immediately noticeable feedback for time-dependent events. Using faster ROC can also be used to increase the subjective intensity of the feedback (due to temporal summation; Section 2.3.3.2), so different thermal sensations can be produced by manipulating either the extent of thermal change or the ROC (or both together).

The research summarised thus far has looked at identifying thermal thresholds by using the Method of Limits. This method starts from a set base temperature and increases or decreases at a constant rate until the participant responds that a change in stimulation has been detected. There are other means of testing thresholds that provide further insight. Harding & Loescher [77] compared the Method of Limits with the Method of Levels, also known as a staircase method, where the stimulus is decreased and increased between levels of imperceptibility and perceptibility until the lowest level of perceptibility is identified. They found that thresholds were smaller when using the Method of Levels, as delays in reaction time did not artificially increase the size of thresholds. Darian-Smith & Johnson [43] looked at the just-noticeable difference (JND) between a reference and a test stimulus, to identify the smallest necessary difference between two successive changes for them to be felt as perceptually distinct. Starting from a baseline temperature of 34°C, they found different results for warming and cooling stimuli. The JND between reference and test cooling stimuli increased as the extent of the reference change increased from skin temperature: differences between cooling stimuli need to be larger as they increase in magnitude in order to feel different. For warming, however, the JND *decreases* as the reference temperature increases, up to a maximum of 6°C, above which JNDs become bigger.

The impact of these factors on the design of thermal feedback is significant. First, it is necessary to design feedback relative to skin temperature. To produce a subjective "warm" stimulation when the skin is 32°C requires the stimulation be set around 35-38°C, whereas a stimulation of 32°C would be considered warm with a skin temperature of 25°C. Secondly, as is discussed in the section on the hedonic aspects of thermal stimulation, what would be considered a "pleasant" stimulation changes depending on skin temperature. Although a warm temperature might feel homely and loving when the skin is cool, it may feel uncomfortable or oppressive when it is warm. In the latter circumstances a cooler stimulation would be more pleasant

2.3.2.1 Spatial sensitivity

While thermal HCI stimulators could be placed anywhere around the body, the immediately logical places are the hand and along the arm, as devices are already held in the hand. "Watch-phones", such as the Sony SmartWatch©[8], can also receive communications or act as proxies for mobile communications, and could feasibly be augmented with thermal stimulators to stimulate the wrist and forearm. The upper arm is a common place for placing mobile devices while exercising, held in place by elastic straps. These could also be similarly augmented with stimulators. There are two different aspects of spatial sensitivity covered here: how sensitivity to thermal changes varies around the body, and how localization of thermal changes varies. It is important to understand how sensitivity varies around the body, so that feedback can be properly designed to be salient and comfortable when produced at the desired location. Localization is also important if multiple stimulators are to be used: if they must be differentiated, they must be suitably far apart.

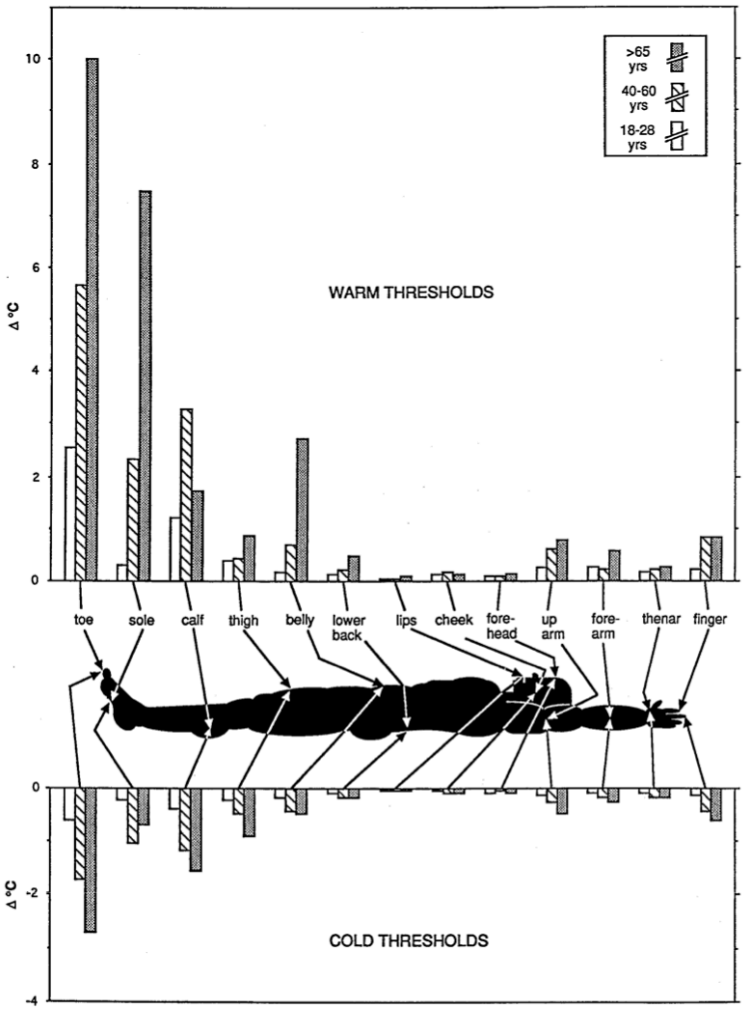

Figure 2-5: Warmth (top) and cold (bottom) sensitivity around the body (from [224]): smaller bars = higher sensitivity.

[8] http://www.sonymobile.com/gb/products/accessories/smartwatch/

Thermal sensitivity is not uniform across the body; it varies approximately 100-fold over the body surface, however all body regions are more sensitive to cold than to warm and as the intensity of warmth increases, the differences in sensitivity between body regions reduce [224]. There are no significant differences between the left and right side of the body [36, 166]. In general, thermal sensitivity is best on the head and trunk but worse towards the extremities [36, 71], but an approximate ranking of body locations in terms of thermal sensitivity (highest first [36, 224, 226], see also Figure 2-5) is: Lips, Forehead, Cheek, Palm, Shoulder, Lower back, Forearm, Upper arm, Fingers, Thigh, Belly, Calf, Sole of foot, Toe.

Thermal sensitivity drops with age, with the largest decline seen in the belly, thigh, sole of foot and the fingers. In contrast, sensitivity remains constant over time on the face, thenar eminence, fore- and upper arm and the lower back [224]. In relation to the hand and arm locations, potentially more relevant to mobile HCI, glabrous skin (hairless skin as found on the fingertips or palm) is generally less sensitive to changes in thermal stimulation than non-glabrous (hairy) skin, with thermal thresholds being generally larger and slower to occur on glabrous skin due to skin thickness [81, 185, 240]. The thenar eminence (the bulbous region of the palm adjoining the thumb) has higher sensitivity than the rest of the palm [71, 118], but is still not as sensitive as non-glabrous skin on the hand [81]. The fingertips are less sensitive than other hand and arm locations [118, 224, 240].

Along the forearm, localization of cold is better than warm localization and identification of two contact points was better when they were stimulating two different dermatomes (transdermatomal) than when stimulating the same dermatome (intradermatomal) [150], particularly for warm stimuli. Dermatomes are areas of skin that are innervated by different spinal nerves (that connect to the spine at different vertebrae) and run longitudinally along the arms. This means that, if multiple stimulators are to be used, placing them on opposite sides of the forearm will make them more differentiable. When stimulating within the same dermatome (such as in a linear pattern along the forearm), localization of stimulators improves if the gap between them is 8-15cm [46, 150] or more. In general, increasing or decreasing the temperature of a stimulator further from skin temperature improves localization [221], as does using a contact stimulator compared to the use of radiant stimulation [25, 234].

Given the vast differences in thermal sensitivity around the body, thermal feedback will need to be tailored for the location to which it is presented. More sensitive areas will be able to detect smaller and more fine-grained changes, allowing for more complex nuanced feedback

designs. Less sensitive areas will require larger changes and so may not be able to differentiate subtle differences between stimuli. Simpler feedback designs, perhaps employing gross changes (simply warming or cooling by a large amount) may be required. The locations most commonly associated with mobile device use, such as the hands, arms and face (for talking on the phone, Bluetooth earpieces) are highly sensitive to thermal changes, and so are optimal locations for detecting and differentiating thermal feedback.

2.3.3 Perceptual Influences and Phenomena

The human thermal sense is complex and influenced by several perceptual phenomena, which result in different internal subjective sensations than might be predicted by the veridical (true) form of stimulation. They involve summing the total extent of stimulation, both spatially and temporally and likely arose due to the influence of the total magnitude of stimulation on maintaining homeostasis. They are important for designing thermal feedback for two reasons: firstly, it is necessary to understand how perception of a piece of feedback will vary depending on the size of the stimulator and how quickly it changes temperature, as these can vary between devices. But, secondly, it means that feedback designs can leverage these phenomena to produce different sensations (or produce the same sensations but using different stimuli).

2.3.3.1 Spatial Summation

The body's focus on the overall magnitude of stimulation leads to increased importance of the areal extent of stimulation [135]. The larger the area of stimulation, the more effect it is likely to have on body temperature. Therefore, the body sums the veridical temperature over the area of stimulation to produce a subjective perception of greater magnitude than the veridical temperature provides [25, 135, 220, 223, 225, 226]. What this means is that stimulating a larger area of skin at a given temperature, for example warming to 38°C, produces a stronger sensation, in this case of warmth, than stimulating a smaller area at the same temperature. Not only this, but the area of stimulation and the extent of temperature change (distance of stimulatory temperature from skin temperature) trade off almost equally, so that the same subjective perception of stimulation is achieved by halving the area and doubling the intensity as doubling the area and halving the intensity [135, 223, 225]. Spatial summation of warmth reduces as the extent of warmth increases to the point where no summation occurs at the pain threshold [223, 225, 226]. In contrast, spatial summation of cold continues, regardless of temperature. The difference may be to improve localization of imminently damaging heat-sources, as damage from cold takes much longer [225].

This phenomenon is linked to poor spatial localization of thermal stimulation described above. Spatial summation limits thermal feedback as it means that accurate localization cannot be used as a parameter in feedback design, for example using arrays of stimulators placed close together and utilising spatial patterns. As mentioned in Section 2.3.2.1, stimulators must be placed at least 8cm apart on the forearm to be identified as separate [46, 150]. If stimulating the fingertips, localization is better between individual fingers than within the same finger [257]. However, spatial summation opens up opportunities for creating a greater magnitude of stimulation without increasing the temperature of the stimulators used, such as by using a larger stimulator, or by activating numerous stimulators placed in an array. The power requirements for a thermal display may potentially be quite large [239], which is particularly problematic for mobile devices, which have limited battery capacity.

2.3.3.2 Temporal Summation

Just as the skin sums stimulation over a given area, it also sums over time to produce a larger subjective stimulation. Up to a limit of approximately one second, extent of thermal change and duration of stimulation trade off almost proportionately (like area and intensity above) so that half the duration at twice the extent of change is equal in subjective magnitude as half the extent of change over twice the duration [223]. Beyond the limit of one second the duration of stimulation no longer affects perceptual magnitude.

The consequence of temporal summation for HCI is that faster changes feel stronger, so that larger magnitudes of sensation can be achieved with smaller changes in stimulator temperature, simply by changing temperature quickly. Therefore, more intense stimuli can be produced in shorter periods of time. Also, any feedback that includes short "pulses" of thermal change must increase the output temperature as the pulse duration shortens, in order to provide the same level of perceived stimulation.

2.3.3.3 Referral, Enhancement and Synthetic Heat

Synthetic heat is the most intriguing phenomenon of thermal perception and it is linked to spatial summation. Green [63] discovered that, when both warm and cold stimulators contact the skin close by each other, stimulation produces a sensation of heat that is both more intense than, and perceptually distinct from, the warmest veridical temperature. In the study he had participants place three fingers on three Peltier heat modules simultaneously. He then heated and cooled the Peltiers in different combinations and asked the participants to focus

on the sensation in the middle finger of the three, before recording their responses.

When the outer two were warm and the middle one neutral, warmth was reported in the middle finger. The same effect occurred when the outer Peltiers were cold and the middle neutral, only with the production of a phantom cold sensation. This is defined as referral. When all three were heated (or cooled) to the same temperature the sensation in the middle finger was reported as higher than when only the middle Peltier was touched. This is called enhancement. Both of these can be largely explained by spatial summation. In referral, the two surrounding temperatures sum together and transfer over to the middle finger and in enhancement the same happens, but, because the middle finger already had an equal level of stimulation, the subjective stimulation increased further.

Synthetic heat occurred when the outer two Peltiers were warmed and the central one cooled. It should be noted that it did not occur every time this combination of warming and cooling was presented, but it did occur on several occasions. In referral and enhancement we see what Green referred to as "domination" where the sensation at the middle finger was enhanced or altered by the quality of the sensation produced at the outer fingers. However, rather than simply enhancing the stimulation at the middle finger, the participants reported an inversion of stimulation, shifting from the veridical cold to feeling warmer than the outer stimulators. This did not occur when the outer Peltiers were cold and the middle warm, in this case the warm was simply described as colder than the veridical temperature describes.

It is not clear exactly why this phenomenon occurs, however it could be due to the role cold afferents play in our perception of heat and also heat pain. As mentioned above, purely cold afferents do not fire as temperature stimulation rises, only the warm afferents are active. However, there are cold receptors that are also involved in sensations of pain and when the temperature reaches a more intense heat, especially around the pain threshold, the cold afferents then begin firing again [223]. It is the combination of these signals in unison that gives the sensation of heat and heat pain. In the case of synthetic heat, the co-occurrence and amalgamation of the warm and cold stimulations, which would be considered part of the same stimulation due to spatial summation, may be perceived as heat, beyond the lesser veridical warmth from the outer Peltiers.

What synthetic heat means for feedback design is that multiple stimulators in close proximity can provide a wide array of stimulations, not just utilising spatial summation to provide more intense stimulations more easily, but by providing different stimulations than could be provided with individual stimulators.

2.4 Thermal Stimulation as Feedback in HCI

This section summarises existing HCI research into the use of thermal feedback to convey information. It will show what progress has been made, but also the various limitations of existing research, including the use of complex apparatus, overly simple feedback designs and the lack of real-world testing. Section 2.4.1 summarises research from Virtual Reality while Section 2.4.2 includes less developed uses of thermal stimulation in other, more traditional, static interaction scenarios. Section 2.4.3 summarises research into the use of thermal stimulation for conveying information in mobile HCI, before Section 2.4.4 gives summary conclusions about existing research on thermal feedback.

2.4.1 Material Properties and Virtual Reality

One of the most common uses for thermal feedback in HCI is in helping to convey material properties of touched virtual objects, particularly in Virtual Reality (VR). This generally involves mimicking the changes in thermal stimulation when the skin contacts an object. This includes simulating thermal properties such as the material's thermal conductivity (SI unit *watts per meter kelvin* or W·m−1·K−1), a rating of the material's ability to conduct heat (in this case away from the skin). Materials with low thermal conductivity, such as paper (0.05 W·m−1·K−1), take very little heat away slowly, resulting in very small changes in skin temperature during contact, whereas materials of high thermal conductivity, including many metals such as copper (~390 W·m−1·K−1) conduct heat away from the skin quickly, leading to a larger, faster drop in skin temperature. The material can then be mimicked by changing the temperature of thermal stimulators to change the skin temperature at the same rate as a given material would when contacted.

Research has generally focused on using Peltier-based apparatus, as the stimulators can be both warmed and cooled to a high-level of precision. Due to the focus on VR and material properties, stimulation most often occurs on the fingertips, conveying materials of objects touched in the virtual environment. Research has found that participants were able to identify materials based purely on thermal cues, when presented with two alternatives and/or given a fixed choice of material types [5, 26, 93, 94, 111]. However, accuracy in identifying materials varies widely, from just 16% identifying rubber [111] up to 100% accuracy for aluminium [111] or a 'burning kettle' [26]. The materials used for comparison vary between

studies, but conclusions have generally stated that materials have to have quite large differences in thermal properties to be identifiable through thermal cues alone. For example, Jones & Berris [122] suggest the thermal conductivities must differ by at least 80 times. Ho & Jones [94] concluded that the 'contact coefficient' plays an important role in identifying materials. The contact coefficient ($J/m^2s^{1/2}$ K) is the square root of the product of the thermal conductivity, object density and specific heat. They concluded that materials could only be distinguished or identified reliably, based on thermal cues alone, if the ratio of their contact coefficients exceeded three. Also, stimulating a larger area of skin, specifically the number of fingertips stimulated, increases material discriminability [257].

It can be seen that, while some research here suggests that thermal feedback can be used to convey a range of material properties, and so convey a range of information through thermal stimulation alone, it is also clear that there are several issues which make this means of conveying information troublesome. There are vast differences in identification accuracy between materials, and vast differences in material properties are needed to be able to tell them apart. Therefore, there is likely to be a very small subset of thermal cues that can be reliably differentiated. Also, the results emerged from fixed-choice and comparative studies, which may support higher accuracy rates. The final issue is one of the technical requirements for such displays. In a summary of the requirements for thermal interfaces in VR, Jones & Berris [128] laid out the "desired features" of a thermal display:

- A maximum operating range of 20°C (from 22-42°C)
- Heating resolution of 0.001°C
- Cooling resolution of 0.002°C
- 2-10 elements in the display
- Cooling rate of up to 20°C/sec
- Warming rate of up to 10°C/sec

These are the requirements to make full use of the thermal sense and to be able to convey materials thermally. However, when considering thermal feedback as a means of conveying information for interaction with mobile devices, these requirements may be infeasible and perhaps even unnecessary. In highly controlled psychophysical studies, often after many hours of training in identifying small changes in thermal stimulation, humans can detect very small thermal stimuli (as small as 0.02°C [117, 118]). The high thermal and transient resolutions are necessary to accurately mimic thermal conductivity, but in laboratory studies the accuracy in identifying material properties based on those can be very low. As is discussed in Sections 2.4.2 and 2.4.3, a small number of HCI studies have looked at

perception of thermal changes in more realistic interaction scenarios, and the results from these suggest that, when in indoor and outdoor interaction environments, it is unlikely that these very small thermal changes or small differences between stimuli will be sufficiently salient to reliably convey information in realistic mobile environments.

2.4.2 Thermal Feedback in Non-Mobile HCI

This section describes research into thermal feedback outside of Virtual Reality, in more traditional interaction scenarios, such as interpersonal communication or interacting with desktop PCs. Many are simple prototypes or proofs-of-concept and so have not been tested empirically and few have gone into detail on how effective the hardware or feedback designs are in terms of conveying information.

Lee & Lim [151] investigated inherent associations made between thermal stimulation and personal experiences and every day events, to determine what information could be conveyed through thermal feedback. They conclude that thermal feedback only has meaning when detected in context, for example when presented with relevant visual or tactile stimuli (for example light and the colour red would be associated with warmth). They gave pairs of participants (either a mother and a daughter or colleagues who share an office) a wristband containing a single Peltier inside and "thermal messages" could be conveyed to the other person in the pair by pressing a button on the wristband. No details are given concerning what temperatures were produced from the Peltier. The authors recorded usage of the device and participant perception of the feedback:

- Participants generally only liked warming feedback when they were cold.
- "Radical" (perhaps fast and/or large) changes were perceived as negative signs.
- Participants could interpret varying degrees of warmth and cold, not just simply warm or cold.
- People felt negative towards coldness and positive towards warmth.
- Interpretation of thermal signals was context-dependent.
- Thermal feedback was unobtrusive, to both the receiver and those around him or her.
- Warmth was associated with physical touch and emotional closeness.

Suhonen *et al.* [230] aimed to investigate the role of context by allowing participants to use thermal feedback as a means of conveying emotions and actions during discussions on positive ("happy") or negative ("sad" or "angry") events that they had experienced, as well

as a neutral, hypothetical event ("restaurant"). The results showed a clear dichotomy between warmth and cold, as warmth was used to represent or reinforce agreement and positivity, while cold was used to represent disagreement and negativity. These results largely echo those of Lee & Lim [151], where participants also felt positively towards warmth and negatively towards cold. These results may be limited by their cultural and sample homogeneity but they provide interesting insight into innate preconceptions about thermal feedback and the unique sensations and experiences that could be provided by using thermal feedback in HCI. Given the communicative possibilities, its potential contribution to mobile HCI is significant.

2.4.2.1 Abstract Uses and Prototypes

There are several examples of thermal interfaces for HCI that have been proposed but have only been subjected to functional testing or basic user-perception evaluation. Oron-Gilad *et al.* [181] tested a proprietary Peltier-based hardware design utilising three stimulators placed along the volar surface of the forearm close enough together to create sensations of synthetic heat (which they refer to as the Thermal Grill Illusion). In an initial test, the sensation was successfully aroused and the authors intend to use it in future interfaces.

Kushiyama and colleagues [2, 142-144] have developed a horizontal thermal display consisting of 80 Peltiers arranged in an 8 x 10 grid. The prototype has been used for a variety of implementations, including augmenting visual art with thermal feedback [144] and providing spatial patterns both thermally, through the sense of touch [142], and visually, through the use of thermo-reactive plastic sheet, which changes colour based on thermal stimulation [143]. Finally, they attached two Peltier modules to a video game controller (one on each side, to contact the palm of the hands holding the controller) to provide game-relevant thermal feedback. None of these examples was tested scientifically, and so little is known about the range of thermal stimuli that the devices can output, or what stimuli users could perceive when using them. Their potential contribution to thermal feedback in HCI is, therefore, also unknown. The multi-stimulator display is unique and could potentially provide sensations not possible with smaller apparatus, including spatial patterns, but it is large, at 120 x 150mm, and likely to require large amounts of power to run. It is therefore not suitable for use in mobile HCI.

Sato & Maeno [202] have subsequently designed a much smaller grid-array of Peltiers, measuring just 16.6mm^2 (each Peltier was 8.3mm^2) and used spatially-divided, alternating warm and cold temperatures to create an illusion of rapid temperature change. The motivation was to produce strong and rapidly detectable thermal stimuli that require less

power and overall extent of thermal change than traditional hardware, which warms or cools all elements in an array. This was achieved by manipulating spatial summation so that sensations from across the array are combined and treated as a singular stimulation/sensation. They did this by starting two Peltiers at a warm temperature (4°C from skin temperature) and the other two at 1.5°C below skin temperature. All 4 Peltiers were then cooled together, so those set 1.5°C below skin temperature reached a colder temperatures as the warm Peltiers reach neutrality. Therefore half the time is needed to change from warm to cool as using the traditional method. The same method was used for warming changes, only with the starting temperatures reversed. They found their spatially-divided method led to faster stimulus detection times than the traditional method for both warm and cold changes (cold changes were detected faster than warm, in line with psychophysical research [43, 133, 223, 224]). Warm changes were also reported as stronger than using the traditional method, but there were no differences in cold stimulus strength between methods. This study is important in showing how the characteristics of thermal perception, in this case spatial summation, can be used to generate salient stimuli using hardware that is limited in both size and power requirements.

2.4.2.2 Affective Computing

Because of innate associations between thermal sensations and emotion [223] or interpersonal warmth/closeness [243], one of the most common implementations of thermal feedback has been to convey emotional or social information.

Gooch [58] had pairs of participants communicate remotely over an instant messaging (IM) application during two tasks: a personal task describing a holiday and an impersonal task ranking items in terms of importance for being stuck on a desert island. Only one of the pair could receive thermal feedback (heatee), while the other could only "give" thermal feedback to the other (heater). Gooch measured whether receiving or not receiving thermal feedback would influence subjective reports of "social presence" (a feeling of being physically or emotionally connected with someone). The feedback was specifically designed to act as a thermal "hug" and so he placed 3 Peltier elements along the back and waist of the heatee, in positions similar to those of hugging arms. At any time during the IM task, the heater could cause the heatee's Peltiers to warm up by pressing a "hug" button, or setting a "love-o-meter" 10-point rotary knob to 7 or higher, while the heatee was asked every 2 minutes to rate their feelings of social presence.

"Hugs" were given for several reasons, including humour/playfulness and to indicate/accompany expressions of sympathy, apology or forgiveness. The results showed

that the heatee reported higher levels of social presence, as well as mutual awareness, however these were not significantly higher than the heater. A stronger influence appeared to come from a halo effect, as the reports of social presence were significantly higher during whichever of the two tasks participants completed first. Thermal feedback is a novel means of conveying information, and participants are unlikely to have experienced thermal sensations arriving without real-world context or cause (such as a cold breeze, cold water splashed on them or touching a hot pan). The halo effect here could be because of experimental design, as thermal feedback may have little true influence on social presence, but it also recommends being cautious in drawing firm conclusions about the effects of thermal feedback from small studies which involve only one task.

Nakashige *et al.* [174] put a Peltier element inside the body of a trackball device to contact the palm of the holding hand and had participants hover the on-screen trackball cursor over images containing hot and cold materials, such as fire or snow, or over images of food. In a basic study they accompanied the food images with either a relevant temperature (such as heat for soup) or the wrong temperature (cold for soup) and asked participants to rate the "deliciousness" of the food items. They found that foods were rated as more delicious when accompanied by the corresponding temperature. While this is a questionable experimental design which may have biased participants to rate foods as more delicious, the study did elicit unexpected emotional reports from some participants, including 'a strong impression of a loving home' from warm thermal feedback and an image of miso soup (the participant sample was Japanese).

Finally, Salminen *et al.* [201] measured both self-reported emotional responses as well as physiological (galvanic-skin) responses to thermal stimuli, specifically 4°C changes from skin temperature, which was measured prior to presentation. They looked at two different methods for presenting the thermal change, either "dynamic" or "pre-adjusted". During dynamic presentation, the participant placed his/her dominant palm on the Peltier before any change began and they felt the change towards the end-point temperature. During pre-adjusted presentation the Peltier was changed to the end-point before the participant placed the palm on top of it. They found that warm stimuli resulted in higher subjective reports of arousal than the neutral starting (skin) temperature, but that no thermal stimuli had any effect on galvanic-skin response.

While there are some faint influences of thermal feedback on emotional responses, both Gooch [58] and Salminen *et al.* [201] failed to gain strong responses, and those gained from Nakashige *et al.* [174] may be slightly unreliable. The participants that Lee and Lim [151]

surveyed said that thermal feedback only has meaning in context, which was missing from Salminen *et al.* [201], and Gooch [58] admits that thermal feedback may simply not influence social presence, so no positive results would be expected. If there are inherent ties between thermal stimulation and emotion, more research needs to be done to identify which associations there are and so how to best design feedback that elicits these sensations in relation to appropriate events or tasks.

2.4.3 Thermal Feedback in Mobile HCI

There have only been a small number of studies that have looked at thermal feedback on mobile devices, and most peer-reviewed papers are merely prototypes with no empirical user evaluations. Only one study has tested perception of thermal feedback while users are walking outdoors [239], but the study has very few details on the feedback provided and how well participants could actually perceive or interpret the feedback. Therefore, there are large gaps in the research that need to be addressed to be able to evaluate the feasibility of using thermal feedback to convey information in mobile HCI.

Most research into thermal feedback reported thus far has used Peltier elements to provide the thermal sensations, as they can output a wide range of both warm and cold temperatures to a high degree of precision. However, Peltiers require large amounts of power to change temperature quickly or substantially [202, 239]. Concerned that the power requirements were too large for realistic mobile interaction design, Wettach *et al.* [239] used a much simpler, low-power device, consisting of a 10-ohm power resistor with 0.5W power supply. While providing the benefit of lower power consumption, it was limited in only being able to produce varying degrees of warm stimuli, and was unable to produce cold stimuli. Wettach *et al.* created a research prototype, which included the heating element and 5 LEDs in a key fob to make it fully mobile. At first they found that participants could identify three extents of warmth at up to 75% accuracy: 32°C, 37°C and 42°C. During longer-term training, participants could identify "five different temperature levels" (p. 184) at 75% accuracy, after 10 days training. Unfortunately, the authors do no state what the five temperature levels are, although they later conclude that five temperature levels can be identified "within a…range of approximately 10°C" (p.184), which may be the 10°C between 32°C and 42°C. They used these five temperature levels in an outdoor navigation task where the heating element warmed as the participant faced in the correct direction and got less warm and they deviated. Very few details are given as to how well users could actually differentiate the levels, only that they completed the tasks, and no mention is given as to any effects of walking or being outdoors on perception of the thermal feedback.

This study suggests that participants can identify varying degrees of warmth, with 75% accuracy, supporting the responses from Lee and Lim [151] that thermal feedback is interpreted along a spectrum. Therefore, thermal feedback may be able to make use of varying degrees of intensity as a parameter for conveying information, even when walking and outdoors. The results were gained with very simple apparatus, which is likely to have had very low rates of thermal change. Utilising more capable hardware, in terms of available power, and the production of cold stimuli, may provide more salient stimuli, potentially increasing the identification rate above 75%.

Other research into thermal feedback in mobile HCI has been less thorough and generally only involved initial functional or exploratory testing. Narumi *et al.* [175, 176] placed Peltier elements into earmuffs and asked participants to walk around and explore an open, empty indoor space. Hot, warm, neutral, cool and cold stimuli (ranging from approximately 15-40°C) were presented to the ears (via the earmuffs) based on the participant's location in the space, with designated "warm" and "cool" spots. The study showed that participants spent more time in the warm areas, but the study was conducted in winter, which may have biased the results. However, it still suggests that thermal feedback can be used to facilitate or encourage people to behave in a certain way, in specific contexts. Unfortunately, the study did not test perception or identification of thermal changes.

Two other studies have suggested means of using thermal feedback to convey emotional information in mobile HCI but have only provided prototype descriptions. Fujita & Nishimoto [52] attached a Peltier element to a wearable device to convey the air temperature around a partner. Pressing on a touchpad could then warm a Peltier element on the partner's device. The Affect Phone [112] also conveyed heat from a Peltier but placed it on the back of a phone and the feedback temperature was based on the level of physiological arousal of another person. No testing was done, so nothing can be concluded about the ability to convey emotional information, or even perceivable thermal changes.

2.4.4 Conclusions

While a wealth of psychophysical research exists showing the various influences of stimulus characteristics on thermal perception, little research has been done into how well thermal changes can be detected or identified in more realistic interaction scenarios. Only one HCI study has looked at thermal feedback when walking outdoors and it contains very few useful details for determining how best to design thermal feedback to be salient and useful in

mobile interaction environments. Research needs to be done measuring perceptual fidelity in realistic indoor and outdoor environments, while the individual is both sitting and walking. Therefore, Research question 4 asks:

"What parameters of thermal stimulation are most detectable and comfortable when using equipment designed for mobile interaction?"

Further, to be able to use thermal feedback to convey information, thermal changes need to be uniquely identifiable. Simply detecting warmth or cold may be useful for very simple feedback designs, as shown in some affective research, but conveying more than one piece of information will require more complex feedback designs that can still be identified. To that end, Research Question 5 asks:

"Can thermal stimulation be manipulated to convey multi-dimensional information?"

2.5 Interaction with Small Devices when Mobile

The research in this thesis sought to expand the possibilities for interacting with mobile devices through applied pressure and thermal feedback. This section provides a general overview of issues affecting interaction with mobile devices in general. Section 2.5.1 reviews the literature on the negative effect walking has on interacting with mobile devices. Section 2.5.2 reviews novel and alternative means of providing input to mobiles to overcome some of these issues.

2.5.1 Walking-Induced Detriment in Performance

Interaction when 'mobile' includes a wide range of scenarios and includes not only physical motion (for example walking or being on moving vehicles) but also environmental and contextual factors, which can influence how we can interact with mobile devices. However, being physically in motion, and particularly walking (either indoors or outdoors), has received the most attention in mobile HCI research. There have been slightly conflicting results as to the influences of walking on our ability to effectively interact with mobile devices, but, in general, walking leads to poorer motor-control based performance (such as pointing/targeting [7, 39, 155, 169, 204]), cognitive task performance [4, 173, 204] as well as slower interaction times [4, 130, 155, 173] and higher cognitive or physical workload [4,

155, 173]. Interacting with a device also commonly leads to slower walking speeds compared to when walking with no device [4, 7, 155, 204].

Due to the proliferation of touchscreen devices, a lot of research has tested how accuracy in pointing to virtual buttons onscreen is influenced by walking. Recent research has shown that touchscreen targeting (touching onscreen buttons) is less accurate if the user is walking [7, 11, 39, 155, 204] and accuracy reduces as walking speed increases [7]. This drop in accuracy is due to additional motion in both the hand holding the device and the hand pointing at targets (or the same hand if targeting is being done one-handed), although there is evidence that individuals subconsciously time their tapping behaviour to coincide with the most stable moments in their walking gate, thus minimising errors [7, 39]. Reading comprehension and other cognitive tasks on mobile devices also suffer when the individual is walking [173, 204] and more so when the walking route is more complex, requiring visual attention to be paid to potential obstacles [4].

The issue of visual attention is critical in mobile interaction, to avoid obstacles or gain necessary information, such as an approaching station when on a train. Visual attention regularly switches from the device to the environment and back again during interaction, in small 'bursts' of only a few seconds [182] and these regular changes can lead to slower walking speeds [4, 7, 155, 204] and longer interaction times [4, 155, 204]. Researchers have sought ways to alleviate some of these issues by developing non-visual means of interacting with mobile devices, so that visual attention can be paid fully to the environment, and this research is reviewed in Section 2.6.

What is clear is that being in motion negatively impacts an individual's ability to interact with their mobile device. It is necessary to keep these influences in mind when developing interfaces for mobile devices, to be able to design them in such a way as to avoid or mitigate the negative influences of motion on user input and output perception. Placing pressure sensors around the edges of a mobile device would remove the need for precise pointing on a touchscreen, and utilising audio feedback, interaction could occur without looking at the device, even squeezing it in a pocket. Providing notifications non-visual through thermal feedback will also help alleviate visual attention demands, so they can be paid to the environment.

2.5.2 Alternative Input Methods

Mobile devices typically have comparatively limited means of receiving input, due to small form factors, which leave little room for physical buttons, and small touchscreens which require the user to cover the content shown on screen in order to interact with it, known as the "fat finger" problem [211]. Research has sought to expand input options for mobile devices, or replace existing ones with more effective alternatives. The research presented in this thesis suggests applied pressure as such an alternative means of providing input, as it can potentially provide continuous, multi-dimensional control without covering the screen. Therefore, alternative input methods are reviewed here, for comparison.

Physical gestures, where the user performs a recognisable motion with the device or other such sensor in space, have been a popular means of providing input. Tilting the device in three-dimensional space has been used for scrolling text [178], traversing menus and zooming maps [195], inputting text [241] and interpersonal communication [82, 83, 193, 230]. This input method provides the benefit of not requiring accurate pointing or pressing buttons, and it also provides continuous control over input. However, by tilting the device, the screen that is presenting the content being interacted with becomes less visible.

To avoid these issues Crossan and colleagues [38, 42] looked at attaching external accelerometer sensors to either the head [38] or the wrist [42] for input. Tilting the head, or rotating the wrist, to the left or the right could then control the one-dimensional, bi-directional movement of an onscreen cursor for linear targeting. Tilting the wrist provides the benefit of leaving the screen fully visible. Used in conjunction with non-visual feedback also provides eyes-free input to the device. However, using the wrist with visual feedback requires two hands for interaction: one holds the device while the other wrist rotates. Moving the head can be uncomfortable and awkward, with a limited range of movement.

Other examples of gesture-based input include the use of the feet to provide 'kick-based' gestures, where a phone camera can detect the motion of the users feet and the direction and velocity of motion provide input [76]. This method leaves the screen fully visible and can be used while the user is sitting or standing. However, physical space is needed to perform the gesture and it may be awkward to perform while retaining balance. Detecting and coding the length and number of foot-taps whilst seated has also been suggested as a means of non-visually interacting with menus [37].

Pressure-based input can be provided from around the body of a mobile device, by placing sensors in various positions where the fingers can both hold the device and interact with the sensors [83, 193]. This would also leave the screen fully visible. Other research has proposed

providing input from the sides and body of mobile devices, but in limited ways. In a relatively early example, Hinckley *et al.* [91] suggested placing touch sensors around the sides and back of mobile devices to detect that the device is being held. Other research has expanded upon this by classifying specific ways of holding a mobile device to automatically change the function based on hand position [31]. Butler *et al.* [24] provided contactless input to the sides of a mobile device through infrared light reflected off the fingers tips. Harrison & Hudson [80] suggested similar input by placing a magnet on the tip of the finger, where the position and relative movement of the magnet could be detected and used for contactless input around the sides of very small mobile devices, such as one that might be worn around the wrist.

2.6 Non-Visual Feedback

There is a wealth of research on the use of non-visual feedback as a means of conveying information in HCI. One of the potential benefits of thermal feedback proposed in this thesis is to provide an alternative means of conveying information to the user when he or she is mobile, in environments where audio feedback cannot be heard, vibrotactile feedback cannot be felt, or in situations where neither is desirable nor appropriate. Therefore, the research summarised here focuses specifically on creating non-visual feedback for interaction with mobile devices, specifically conveying multi-dimensional information non-visually so that visual attention can be paid to the mobile environment.

Earcons [10, 16-18] are structured, abstract non-speech sounds. Information is encoded in the sound's auditory parameters, such as the timbre, rhythm and pitch and, by using several different timbres, rhythms and pitches, a single Earcon can convey up to three pieces of information. Using Earcon design as a basis, Tactons [12, 20-22] are structured vibrotactile icons that can convey multiple pieces of information, mapped to unique vibration parameters including rhythm, roughness and spatial location. Tactons can convey up to three pieces of information at up to 80% accuracy, when identification was tested seated indoors [21]. Vibrotactile feedback has been used to improve interaction in mobile tasks [13, 100], and the unique identification of multidimensional structured vibrations has been tested when walking, both indoors [179] and outdoors [48, 154]. Oakley & Park [179] found that walking reduced identification accuracy of two-dimensional (body location and roughness) vibrotactile stimuli, but other research found high identification rates when navigating [154] or simply walking outdoors [48]. Therefore, Tactons could be considered an established and

effective means of conveying information when mobile. They represent an upper bound against which to compare thermal icons.

Because some environments are not suitable for audio or tactile feedback, and because user preference varies regarding which modality is desirable when, Hoggan and colleagues developed crossmodal audio and tactile icons which can be interchanged to suit the user's current environment or preference [98, 99, 101, 102]. In these examples, icon parameters are chosen based on their having the same perceptual properties in both audio and tactile domains, such as the same rhythms, timbre/textural quality or spatial location. Icons can be learned in one modality and recognized in the other [98]. While vibrotactile feedback is commonly used for private notifications, it is not entirely private, as the vibration is often audible to, or even felt by, others. Thermal icons, multi-dimensional structured thermal feedback, could provide an entirely silent means of conveying information, or provide a replacement parameter for less reliable (roughness) or less feasible (spatial location) Tacton parameters [21].

2.7 Conclusions

The application of pressure through the fingers and the use of the thermal sense to convey information are two inherent aspects of the haptic system that are under-studied in the domain of mobile human-computer interaction. Both are highly specialised and precise systems, and their merits for providing input to, or output from, computer systems have been demonstrated to a limited extent in static desktop scenarios, or highly specialised niche implementations (such as Virtual Reality). However, very little research has been conducted on how they could augment and improve interaction on small devices when the user is mobile. With the proliferation of mobile devices that focus on multitouch and gestural input, both pressure input and thermal feedback sit as logical extensions of this interaction paradigm. Every act of touch inherently includes the deliberate control of some degree of applied pressure, and every act of touch inherently includes the reception of thermal feedback from the object. The research in this thesis will establish the usefulness of both pressure input and thermal feedback in mobile HCI. Towards that end, this chapter has summarised existing research on the control of applied pressure through the fingers, the use of pressure in human-computer interaction, the characteristics of the thermal sense and the ways that thermal feedback has been developed to convey information in HCI. It also summarised research on alternative means of interacting with, and receiving information

from, mobile devices.

Research Questions (RQ) 1 and 2 ask:

RQ1: "How accurate is pressure-based input on a mobile device when using only audio feedback"
RQ2: "How accurate is pressure-based input through the fingers when the individual is walking?"

The research reviewed in Section 2.1 and Section 2.2 shows that pressure can be applied highly accurately when sitting down and provided with full, continuous visual feedback. The research in Section 2.1.3.3 and Section 2.2.2 shows that varying the amount of available feedback influences the precision of control, with less continuous feedback resulting in poorer control. The HCI research on non-visual feedback and pressure has been limited to a relatively small range of pressure input and only when the user is sitting down. Therefore, the research in Chapters 3 and 4 investigates how the use of audio feedback influences control of pressure when the individual is walking and controlling a wide range of pressure.

Research Questions 3 asks:

"How accurate is pressure-based input when multiple fingers apply pressure to a mobile device?"

The psychophysical research in Section 2.1.3 has tested how control of pressure varies when more than a single digit applies pressure, but HCI research has generally been limited to just a single digit (usually the thumb) or a two-digit thumb-finger pinch. As all five digits of the one hand can be in contact with a mobile device when holding it, there is the opportunity to provide input from several digits concurrently. The use of multiple digits can potentially provide multiple different inputs to the system at one time, greatly expanding interaction design options. Chapter 5 includes a study into multi-digit application of pressure to a mobile phone.

Research Question 4 asks:

"What parameters of thermal stimulation are most detectable and comfortable when using equipment designed for mobile interaction?"

The research in Section 2.3 summarizes the main influences on thermal perception and the ways that sensation changes based on changes in the characteristics of thermal stimuli. This research provides several candidate parameters for the design of thermal feedback for mobile interaction, and how to manipulate them to provide safe and salient stimuli. The parameters include rate of thermal change, extent of thermal change, area of stimulation and spatial/bodily location. While their influences are well understood when the individual receiving the stimuli is sitting, engaged in protracted learning in controlled laboratory environments and receiving stimuli from large and complex hardware, how well stimuli can be perceived when the individual is walking in less controlled environments and using more compact, simple hardware, is not known. The research in Chapter 6 tests detection of thermal stimuli that vary in their rate of change, extent of change and direction of change.

Research Question 5 asks:

"Can thermal stimulation be manipulated to convey multi-dimensional information?"

Existing thermal feedback designs are either highly complex and demanding (Section 2.4.1) or simple and low-bandwidth (Sections 2.4.2 and 2.4.3). The complex and demanding designs from Virtual Reality are likely to be unusable for realistic outdoor scenarios, and only very limited information is available from simple warming and cooling changes proposed thus far in more traditional HCI. The research in Chapter 7 evaluates the design of multi-dimensional thermal feedback that could convey two pieces of information thermally, to be reliably detected and identified when the user is sitting or walking indoors or outdoors.

3 Non-Visual Pressure-Based Input When Sitting

3.1 Introduction

Requiring constant visual attention to be paid to visual feedback on a device while the user is walking means that visual attention cannot be paid to the environment, which could put the user in the way of hazards and potentially lead to injury. Therefore it is necessary to offer alternative non-visual feedback for mobile interaction to avoid these dangers. Non-visual feedback has been shown to facilitate certain mobile interactions, by freeing visual attention from the device, so that it can be focused on the surrounding environment [11, 158, 164]. Therefore, it is an important consideration in the design of interfaces for mobile devices. Before testing control of applied pressure when walking, it was necessary to design an effective non-visual feedback design for pressure input. A few studies have tested control of pressure input using non-visual feedback [228, 232, 233] but none have done so when the user is walking, and only a small range of input (three levels of pressure) has been used. Control of applied pressure improves when continuous external feedback is provided [170, 192, 214] yet the non-visual pressure interfaces all used discrete feedback. Stewart *et al.* [228] suggested that continuous audio feedback was reported as annoying during pilot testing, and so it was deliberately replaced by discrete feedback. While the discrete design may have been less annoying (user appreciation was not reported), this decision led to poor performance. Non-visual feedback design needs to improve to be able to provide the same support for pressure input as visual feedback. Therefore, RQ 1 asks:

"How accurate is pressure-based input on a mobile device when using only audio feedback?"

This chapter describes an initial preliminary experiment followed by Experiment 1, investigating control of a pressure-based interface when the user was seated indoors and provided with only audio feedback. For comparison, performance was also measured when provided with visual feedback. Section 3.2 describes the experimental task used in the studies, while Section 3.3 describes the visual and audio feedback designs. Section 3.4 includes the initial preliminary experiment and Section 3.5 describes Experiment 1, both of

which tested visual and non-visual control of pressure input when the users were sitting down. The apparatus used for the preliminary study and Experiment 1 were different, and so they are described separately in the relevant sections (3.4.2 and 3.5.3.1). Section 3.6 discusses the limitations of the research in this chapter before Section 3.7 gives overall discussion and conclusions.

3.2 Task

Target selection along a single axis has been a common and effective way of demonstrating control of pressure in many other studies [30, 170, 192] and so is used again here as well, to provide better comparison to existing findings from static desktop research. The task is described in detail in Section 2.2.1.2 in Chapter 2. Previous work [30, 170, 228] suggests that user accuracy can remain high at up to 10 distinct levels of pressure, and so this was chosen as the maximum number of divisions in the preliminary and Experiment 1. A pressure space of approximately 3.5 N was divided into 4, 6, 8 or 10 equal-sized bins/levels visualized on-screen as a vertical menu of as many menu items, running from top-to-bottom (see Figure 3-1). Therefore, target levels had a width of approximately 0.87 N, 0.58 N, 0.44 N and 0.35 N for the 4, 6, 8 and 10-level menus, respectively.

Applying more pressure moved a cursor further through the pressure space, and so further down the menu shown on screen. The task consisted of multiple trials and each trial involved selecting a single target level (i.e., menu item). Once the participant had applied enough pressure to place the cursor within the target level, he or she had to use a selection mechanism (Section 3.2.1) to confirm acquisition of the target. By using thinner targets (a higher number of levels, or menu items) the task tests how precisely the participants can control the amount of pressure applied.

3.2.1 Selection Techniques

Ramos *et al.* [192] and Cechanowicz *et al.* [30] tested different selection techniques in desktop settings and found each to have its own merits. Quick Release involves lifting the finger/thumb from the pressure sensor when the cursor is in the target level, and the amount of pressure applied immediately before 'lift-off' is used as the selection point. Dwell requires the user to remain in the given target level for a set length of time to confirm selection. In general, the Quick Release technique is more error prone than Dwell but is usually much

faster [14, 30, 192]. The preliminary and Experiment 1 both compared the Quick Release and Dwell target-selection techniques. In each study in this chapter a Dwell duration of 1 second was chosen. 500ms was used in initial preliminary testing to increase the speed of interaction and this was also the length of time used successfully by Brewster & Hughes [14]. However, after a high number of erroneous selections, the time was increased to 1 second. This length has been found to be a suitable length of time in a similar interaction [30].

3.3 Feedback Design

3.3.1 Visual Feedback

To give the task more relevance to real-life mobile use, the interaction was designed to resemble traversing a flat linear menu and selecting menu options, with each pressure level being given a unique label that one might find in a typical application (see Figure 3-1**)**. The labels chosen are common menu items found in various applications:

File, Edit, View, Format, Bookmarks, Text, Tools, Window, Help, Exit.

The order of the items never changed, only the number that were placed on screen, starting with File, so pressure menus with 4 items ended at "Format", 6 items went up to "Text" and so on. The visual feedback displayed the pressure levels as equal-sized grey rectangles aligned vertically in the middle of the screen (see Figure 3-1), measuring 200 x 400 pixels, giving a sufficiently high visual gain (in p/N) for good performance [108]. A small cursor (10 x 10 pixels) moved vertically just outside the menu, indicating the amount of pressure being applied in a continuous form. During Experiments 2 and 3 in Chapter 4, the target level (menu item) that the cursor was currently within was highlighted by making its boundaries green (see Figure 3-1). This indicator was added after Experiment 1 and the preliminary, and so was absent from these tasks. The active target for any given trial was displayed briefly in bright green at the start of the trial. A continuously moving cursor was chosen over discrete feedback (for example simply highlighting levels relative to applied pressure), as continuous feedback is necessary for successful pressure-based target acquisition [170, 192, 214]. Additionally, the pressure levels were given common labels to aid familiarisation with the interaction for when audio feedback was used.

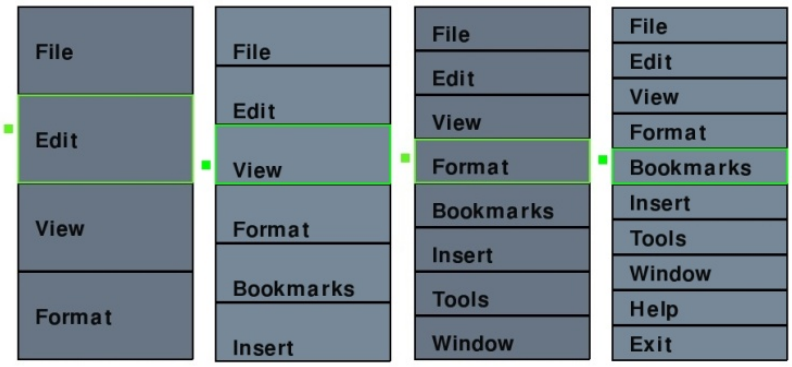

Figure 3-1: Visual feedback showing menu layouts for 4, 6, 8 and 10-item menu sizes, with relative target widths.

3.3.2 Audio Feedback

Audio feedback was chosen over vibrotactile feedback, as it was judged capable of conveying more information, even though audio feedback designs had resulted in less accurate pressure input in other research, compared to vibrotactile [228, 232, 233]. Two audio feedback designs were used in the research. An initial design was used during the preliminary study, which informed the design of a more useful implementation for the subsequent experiments. In each case the screen was always left blank: the movement of the cursor in relation to pressure input, as well as the positions and layout of the menu items, were all the same as in the visual feedback, only they were not visible to participants.

3.3.2.1 Preliminary Study Audio Design

To inform users of which menu item they were in, the item's label was spoken in synthetic speech once as the cursor entered the item from either side: entering it by increasing pressure (moving down the menu) or by decreasing pressure (moving back up the menu). If the cursor moved so fast as to enter another item before the synthetic speech had finished playing, the initial audio was stopped and the newly entered item's label was spoken. To help users identify when they were on the verge of crossing over into the next menu item, a warning tone (chord of 2 sine wave notes: F4 (349.23Hz) and A#4 (466.16Hz)) was played when the cursor entered the last 25% of any menu item. This was added to help avoid accidentally moving into a menu item unintentionally, or to help participants know how much further to go to deliberately move into the next item. All audio was played monaurally through a set of headphones. The current target menu level to select was indicated at the beginning of each trial by the phrase "Get {label} " spoken in synthetic speech, where {label} is one of the ten menu items.

3.3.2.2 Main Experimental Audio Design

The audio feedback design for Experiments 1 to 3 (Sections 3.5, 4.2 and 4.3) was changed significantly from the preliminary study. The main problem from the preliminary came from a lack of positioning: participants complained of being "lost" in the menu, not knowing where they were or where other items were in relation to their position (see Section 3.4.5). Given the spatial nature of visual feedback it is easy to see where display elements are relative to others. The audio was simply presented monaurally in the preliminary study, and so provided no such spatial information. This was rectified by using egocentric panned audio around the head for the main evaluation.

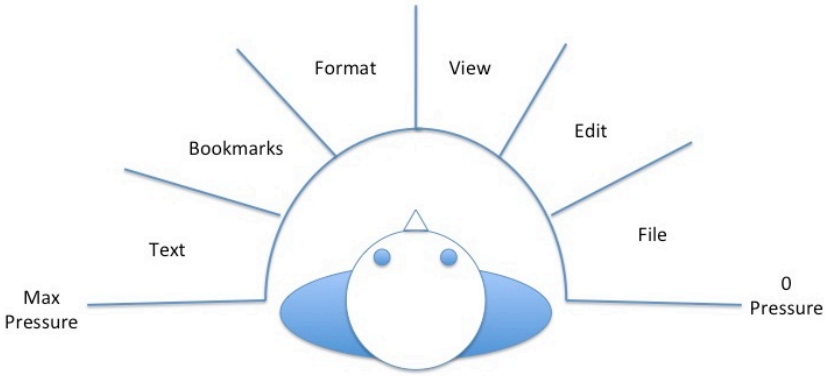

Figure 3-2. Panned audio design for Experiment 1.

The audio menu was now laid out across ~180° of horizontal space in front of the user, so that the first menu item was always on their far right and the last item always on their far left (see Figure 3-2). Spatialisation was achieved by simply altering the stereo volume output (0 to 100) to the left and right ears so that, for example, a volume of 0 (left) and 100 (right) indicated positioning at the far right and 70 (left) and 30 (right) indicated position left of centre. Initially, audio interaction was envisaged as taking place "in pocket", where the individual could apply pressure to a mobile device without having to take it out. This was initially intended to be in the right trouser or jacket pocket, and so right-to-left panning was chosen (rather than more common left-to-right), to fit a metaphor of pressing on the device from the right-hand side of the body and "pushing" the cursor away, towards the left-hand side. This kind of interaction was never implemented and so the direction was changed to left-to-right for Experiments 2 and 3. Although the horizontal dimension is a different orientation compared to the vertical visual menu, several studies have found spatialised audio around the head to be suitable for mobile interactions [15, 159, 203].

The label of each menu item was still spoken by synthetic speech whenever the cursor entered that item from either side, identical to the preliminary. Each item was also given a unique musical note[9] that played for the duration of the time that the cursor was in that item. While this feedback is discrete (it only changes when the cursor moves to a neighbouring level) it plays continuously. Both the label and the note were played in panned audio in the position of where the item lay in the menu: for example, "File" was always heard on the far right. Mizobuchi *et al.* [170] reported that users instinctively aimed for the centre of the targets in their study and some participants in our preliminary study reported gaining no benefit from the warning tone at an item's edge. Therefore, rather than a warning tone, a second note, one octave above the given item's unique note, was played when the cursor was in the central third of the item.

The cursor moved along the menu based on pressure in exactly the same way as when presented with visual feedback, only now moving invisibly from right-to-left (or left-to-right in Experiments 2 and 3). When the cursor entered the boundaries of an item, from either direction, that item's audio feedback was played: the label and unique note. As the cursor moves up or down the menu, the names and notes of each menu item play in the 1-dimensional egocentric horizontal location around the head relative to that item's location in the menu, from right-to-left (or left-to-right on the way down). From hearing the location of the label and note in their position, relative to left and right, the user got a spatial clue as to the cursor's location in the whole menu. For example, hearing "Bookmarks" slightly to the left of centre tells the user it is quite far up the menu (in Experiment 1).

One final addition changed the way the user was informed of which target to select. In the visual feedback, the user can see what the active target is (it is briefly coloured bright green) and, automatically, can see how far down the menu it is. In the preliminary study it was simply spoken to the user in the form of, for example, "Get Bookmarks". But unless the user is familiar with the layout and ordering of the menu items, this does not indicate where *Bookmarks* is in the menu, unlike the visual feedback. Therefore, for the main experiments, during the phrase "Get {label}", the name and note of the target item was played in its relative panned position before each trial, indicating where in the menu that item was.

3.3.3 Experimental Measures

[9] 10-note scale, from A4 (440Hz) to C6 (1046.50Hz): A, B, C, D, E, F, G, A, B, C.

The same three performance measures were used for both the preliminary and Experiment 1: Errors, Movement Time and Number of Crossings.

- *Errors* (ER) – When a selection is made outside of the target level. This occurred when the 'lift-off' point was outside of the target when using the Quick Release selection technique, or by remaining in any non-target menu item for 1 second when using the Dwell technique. Errors are reported as the percentage of incorrect selections.
- *Movement Time* (MT) – Measured the time from the first non-zero reading from the pressure sensor up until selection, be that an error or a correct selection.
- *Number of Crossings* (NC) – If the cursor entered a target level and subsequently exited it again, this was counted as a crossing, and was used as a measure of control, where a lower number of crossings was equated with a higher degree of control (a lower level of input variation).

3.4 Preliminary Study

The initial preliminary study, followed by Experiment 1, tested control of pressure-based linear targeting (in the guise of a menu interaction task) on a Nokia N810 mobile device (Figure 3-3) when the participants were seated at a desk. The task used is that described in Section 3.2 above, and uses up to 10 pressure levels, a far higher number than have been tested non-visually thus far. Control was tested when presented with visual feedback and when presented with the audio feedback design in Section 3.3.2, to test control over a wider range of pressure input than previous research. The studies also compared the Dwell and Quick Release selection techniques to judge their relative merits when applying pressure to mobile devices. The preliminary study made use of the N810's pseudo-pressure-sensitive resistive touchscreen for input, while the main experiment used a force-sensitive resistor (FSR) connected to the N810 for input. The N810 screen is not a proper pressure sensor as the values it outputs are based on how much of the screen is pressed, rather than the extent of pressure applied when pressing. However, it is a commercially available device and made development and testing of real-world usage of pressure input quicker.

3.4.1 Participants

Fourteen participants (7 Male, 7 Female) aged between 20 and 32 (mean = 22.2 years) took part in the evaluation, all of whom were studying or working in the University of Glasgow. All were right-handed and paid £10 for participation, which lasted no more than 90 minutes.

3.4.2 Experimental Design and Procedure

The N810's resistive touchscreen sensor outputs a value between 0 and 1, relative to the size of contact area being pressed (contact area increases as more pressure is applied by the fingertip [208]) and this was divided into 1024 pressure levels to allow comparison with previous work [30, 192, 209]. Unfortunately, the sensitivity of the sensor was not uniform around the screen, resulting in uneven behaviour depending on where the screen was pressed. To minimise this effect, a specific location was chosen as the contact point for all participants. A black square outline was placed on screen to indicate where participants were to press. Due to the uneven behaviour of the sensor it was not possible to accurately calibrate the sensitivity of the sensor (and so the size of the pressure space) in Newtons.

Figure 3-3: Nokia N810 Internet tablet used in preliminary study and Experiment 1. A mock-up of the experimental interface is shown on-screen (right).

Participants held the device in both hands in the landscape orientation, using their right thumb to press on the screen. This pressing action closely resembled a pinch between thumb and first or second finger, due to the way the device was held. The thumb pressed against the device, which was then resisted by the fingers behind the device, mimicking a thumb-finger squeeze. This action, or similar, has been used in other HCI pressure research to some success [14, 162, 228]. Audio feedback was delivered through headphones from the audio jack of the N810.

3.4.2.1 Variables

Ramos *et al.* [192] designed their experimental task so as to be able to measure conformity

of pressure-based target selection to Fitts' Law [49]. Because dividing the pressure space into differing numbers of levels produced targets of different width, the targets were different distances away from the start point. The authors therefore chose four targets from each number of divisions which have within them a common distance. This meant they could compare acquisition of smaller targets at similar distances. These same four distances were used in the preliminary study and equate to 205, 410, 615 and 820 sensor values, out of the 1024 value range output by the stylus used by Ramos *et al*. For the experimental task used here, these values relate to the menu items shown in Figure 3-4.

If an error occurred, an error tone was played (short, 2-note melody). There was no extra feedback for correct selections and the next trial started after a pause of 2 seconds, which was accompanied by a blank screen in the visual condition. It was decided to make it impossible for an individual to "overshoot" the last item in a menu, as it was assumed that this would be the case in a real-life implementation of this type of task in an application.

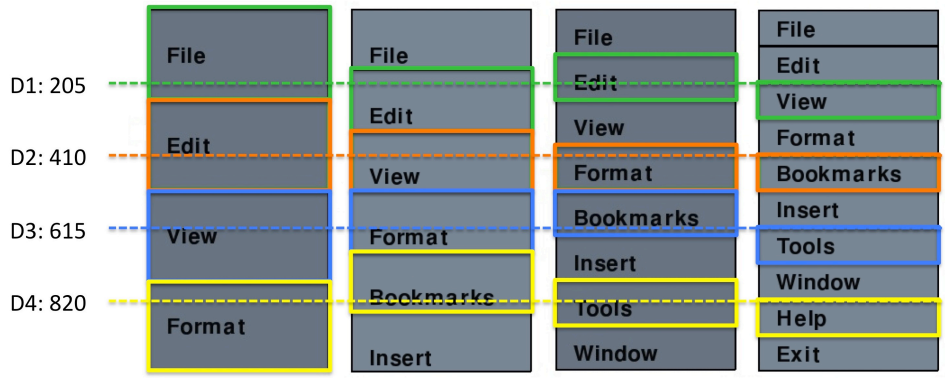

Figure 3-4: Common target distances D1-D4, used to compare performance across different menu sizes. Adapted from Ramos *et al.* [192].

The experiment used a within-subjects repeated-measures design with 4 Independent Variables: *Number of menu items* (4, 6, 8 and 10), *Target distance* (205, 410, 615 and 820), *Selection method* (Quick Release and Dwell) and *Feedback modality* (Visual and Audio). The dependent variables were: *Errors* (ER, % of missed targets), *Movement time* (MT) and *Number of target crossings* (NC).

3.4.2.2 Procedure

The whole task was split into 2 halves: one using only the Quick Release selection technique and one using only the Dwell technique. Within these conditions were one visual-only and one audio-only feedback condition, giving 4 conditions referred to here as Quick-Visual,

Quick-Audio, Dwell-Visual and Dwell-Audio. In order to remove potential ordering issues, half of the participants took part in the Quick Release conditions first and the other half took part in the Dwell conditions first. All participants took part in all conditions, with the ordering of conditions counterbalanced except for the first 2 feedback conditions. To facilitate familiarisation with the interaction as a whole, all users first engaged in a visual feedback condition under their first selection technique, followed by an audio condition. This was a similar tactic used by Ramos *et al.* [192], as they imagined experts would be better able to use the interface with more impoverished visual feedback. The use of audio feedback could also be considered an expert choice, as experts are familiar enough with the interface to be able to control it non-visually. The order of conditions under the second selection technique was fully counterbalanced to reduce (but not eliminate) possible bias towards audio feedback. Experimental instructions can be found in Appendix A.

Under each selection technique x feedback pairing (for example, Quick-Visual) there were 3 blocks of trials. Each block presented each of the 4 menu sizes (4, 6, 8 and 10) once, and within each menu size, all 4 target distances (shown in Figure 3-4) were selected twice in a random order. This gave a total trial count of: 14 participants x 2 selection techniques x 2 feedback techniques x 3 blocks x 4 target sizes x 4 target distances x 2 repetitions = 5376 trials. All analyses involved 2 x 2 x 4 repeated measures ANOVA. Raw data for all measures can be found in Appendix A.

3.4.3 Hypotheses

H1: Quick Release will be a more error-prone selection technique than Dwell
H2: Quick Release will be a faster selection technique than Dwell
H3: Performance using audio feedback will be worse than using visual feedback

3.4.4 Results

3.4.4.1 Error Rate

Analysis revealed a significant main effect of selection technique ($F_{(1,153)} = 72.463$, $p < 0.001$): Quick Release had significantly fewer errors (32%) than Dwell (50%), leading to a rejection of hypothesis H1. There was also a significant effect of feedback modality ($F_{(1,153)} = 313.672$, $p < 0.001$): Audio feedback produced significantly more errors (56%) than Visual Feedback (26%). Error rate also increased as the number of menu items increased (target width decreased) with mean error rates of 31%, 36%, 45% and 50% for 4, 6, 8 and 10 items

respectively.

3.4.4.2 Movement Time

Both selection technique ($F_{(1,1234)} = 22.752$, $p < 0.001$) and feedback modality ($F_{(1,1234)} = 59.115$, $p < 0.001$) had a significant effect on movement time: Dwell produced a significantly higher average movement time (2.7s) compared to Quick Release (2.3s). This result leads to acceptance of H2. Audio feedback had a significantly higher average MT (2.8s) than Visual Feedback (2.2s) and so, combined with the result that Audio feedback was also more error-prone than visual feedback, H3 can be accepted.

3.4.4.3 Number of Crossings

Similarly, both selection technique ($F_{(1,1234)} = 74.289$, $p < 0.001$) and feedback type ($F_{(1,1234)} = 44.434$, $p < 0.001$) had a significant effect on control. Dwell technique led to a significantly higher average number of crossings (2.3) per target compared to Quick Release (1.4). Visual feedback produced a significantly higher average number of crossings (2.2) per target compared to Audio feedback (1.5).

3.4.5 Discussion

The findings of the preliminary study were somewhat disappointing with high error rates for the Dwell technique and for Audio feedback. Also the error rates found for all numbers of menu items were well above those that other research achieved with higher numbers of levels (12 or even 16) [30, 192]. From the data and from subjective reports by users, two primary contributing factors were identified for the poor results: the pressure-sensitive screen used and the audio feedback.

3.4.5.1 Sensor Deficiencies

Although different sensors use different analogue-to-digital converters, there is a common problem in that they are often disproportionately more sensitive to light touches compared to moderate or high pressure. This was found to be the case by other research [30, 192, 209, 228] and was the case with the N810 screen, as users complained that the low levels were much less controllable and error prone than farther levels. This lack of a uniform, or linear, relationship between pressure and cursor behaviour confused users and made holding the cursor at a given level (particularly low levels) much more difficult and more frustrating. Being less able to accurately hold the cursor at a desired level had a greater negative effect on the Dwell condition, which required precise control over time. For the Quick Release

condition users stated that they simply lifted their thumb “as soon as they entered” the target level, requiring little ‘fine-tuning’ of cursor position. Given the common problems across digital pressure sensors this suggests a fundamental problem with their use in HCI. Because the screen was such a poor sensor and led to participant reports of frustration and annoyance, the decision was made to abandon the use of the N810 screen as an input in favour of a more stable sensor, to provide more controllable input.

3.4.5.2 Audio Feedback

The poor results for audio feedback suggest that the design choices were not as useful for orienting around the menu as initially hoped. Participants were encouraged to familiarise themselves with the order and layout of the menu items during their first visual condition, which was intended to aid them in navigating the audio feedback, as it was hoped that they would know where each menu item was in relation to the others. This proved highly troublesome, however, as users were apparently unable to familiarise themselves well enough with the relative positions of menu items. If a participant heard the label “Bookmarks” when trying to target “Help”, for example, they would not know where “Bookmarks” was in the menu (and so not know where the cursor was) and then not know where “Help” was relative to “Bookmarks”. This would require a degree of searching through the menu to then find “Help”. Participants explicitly stated that they often became “lost” within the menu, not knowing where they were or where the target item was, relative to their current position. They also mentioned that they were “distracted” by the ‘warning tone’ and they did not find it useful in knowing that the cursor was close to moving into the next menu item. These problems influenced a redesign of the feedback for Experiment 1, where a spatialised audio design was created to provide a better sense of relative position and movement throughout the pressure space (see Section 3.3.2.1). The warning tone as replaced with a tone indicating the cursor was within the centre of the menu item, as users have been found to aim for the centre of targets during linear targeting [170].

3.5 Experiment 1 – Pressure Input Using Audio Feedback

Following on from the preliminary study, Experiment 1 used a linearised force-sensing resistor to improve control of input and a more useful audio feedback design to provide more information about user input. Control was tested again sitting indoors, to get baseline performance when interacting both visually and non-visually.

3.5.1 Pressure Sensor Used

Stewart *et al.* [228] developed a linearised pressure sensor. They attached an opamp-based current to voltage converter to FSRs (which was then attached to an Arduino[10] interface for A-to-D conversion and output) to produce a good fit to a linear function ($p = 0.0008x + 0.0339$; $R^2 = 0.97$) between pressure applied and the output signal. The authors compared the linear signal to a quadratic mapping (similar to that used by Cechanowicz *et al.* [30]) and found that the linear sensor allowed for a greater degree of control than a non-linear output. Due to these promising results, and the poor accuracy of the N810 screen, Experiment 1 used Stewart *et al.*'s [228] sensor design (see Figure 3-5). The diameter of the sensor pad was 14.7mm.

Figure 3-5. Hardware set up for Experiment 1. FSR is under white adhesive tape and connected to Nokia N810 over USB via microcontroller (black box).

3.5.2 Participants

Seventeen male participants aged between 19 and 35 (mean = 21.5 years) took part in the evaluation, all of which were studying or working in the University of Glasgow. The gender bias was not intentional; a request for volunteers was issued and acceptance was only received from these male participants. Sixteen were right-handed and all were paid £10 for participation, which lasted no more than 90 minutes. None had taken part in the preliminary study.

3.5.3 Experimental Design and Procedure

[10] http://www.arduino.cc

The experimental task was identical to that of the preliminary study with one exception: it followed the design of Ramos *et al.* [192] by comparing only those targets that lie at similar distances. Other studies have also followed this experimental design [30, 209] however, in doing so, the results can only ever examine selection at 4 distances, not the full number of levels stipulated in the interaction (such as 6, 10 or 64). Therefore, Experiment 1 required participants to select all target distances, to see if this would give a clearer picture of pressure control across the entire interaction space. The selection mechanisms used were the same as in the preliminary: Quick Release and 1-second Dwell. The experiment used a within-subjects repeated-measures design with the same 4 independent variables with the exception of *Target Distance*: all 28 distances within the 4, 6, 8 and 10 item menus.

3.5.3.1 Apparatus

The apparatus was set up as seen in Figure 3-5. The Nokia N810 was used to run the experimental software and provide the visual and audio feedback. The FSR was attached to a piece of firm Perspex (under white adhesive tape) to allow for squeezing/pinching action, with the thumb contacting the sensor and the forefinger providing resistance, similar to the action from the preliminary study. The sensor was initially attached to the front, right-hand side of the device body (adjacent to the screen), so that the device could be held in both hands, with the right thumb pressing on the sensor. However, this positioning caused the sensor 'tail' (strip connecting to Arduino) to flex, resulting in random changes in sensor output, so it was necessary to ensure the tail remained stationary. The resulting interaction mechanics, where the sensor is manipulated in a pinch grip between thumb and forefinger, is very similar to the way it would be manipulated if the sensor were attached to the device. Audio feedback was delivered through headphones from the audio jack of the N810.

Due to the linear output, it was possible to accurately calibrate the sensor and so measure the pressure space in Newtons. The sensor could detect a total of approximately 12N, but a pressure space of 3.5 N was used to reduce the potential for fatigue when applying 4 N [170], but to also provide a large enough interaction space that the same number of target levels can be used without significantly reducing their width. Therefore only 30% of the sensor's detectable range was used.

3.5.3.2 Procedure

The procedure for the main evaluation was identical to that of the preliminary other than two details. Because all target distances in a given number of menu items were to be selected (rather than just 4), this resulted in an uneven number of selections for each menu size (4, 6,

8 or 10). This would make any comparison between numbers of items uneven, as, for example, 1 error in the 4-item menu would represent a 25% error rate, but only a 10% error rate in the 10-item menu. Therefore, when comparing performance across number of target items, only the targets with similar distances (identified by Ramos *et al.* [192]) were used. This would mean that, for example, "25% errors" means the same number of errors for a 4-item menu and a 10-item menu (1 incorrect selection). However, it does raise the problem that incorrect selections from outwith these 4 targets will not be counted when comparing the effect of Menu Size. For all other performance analyses, all target distances were considered. Also, selecting each target distance twice (as was the case in the preliminary) would have increased the trial count and task time beyond reasonable levels, considering participant fatigue. Therefore each distance was only acquired once, giving a total of: 17 participants x 2 selection techniques x 2 feedback techniques x 3 blocks x 28 target distances = 5712 trials. Participants completed NASA TLX workload estimation forms after each condition, which included two extra scales titled "Thumb Fatigue" and "Audio Annoyance", the latter for indicating how annoying the participants found the audio feedback design. Experimental instructions and raw data for all measures can be found in Appendix A.

3.5.3.3 Hypotheses

H1: There will be fewer errors in the Dwell conditions than in the Quick Release conditions.

H2: Errors will increase as the number of menu items increases.

H3: That movement time will be lower in the Quick Release conditions than in the Dwell conditions.

H4: That movement time will be lower in the visual conditions compared to the audio conditions.

H5: That the number of crossings will increase as the number of menu items increases.

H6: There will be more crossings in the audio conditions compared to the visual conditions.

3.5.4 Results and Initial Discussion

Outliers were removed from the data set. A trial was considered an outlier if the pressure value (in Newtons) of the selection was more than 2 standard deviations outside of the mean selection value for that target distance. 291 trials were removed, constituting 5% of all trials. This left 2338 data points under each of the Visual, Audio, Dwell and Quick Release conditions (1169 under each feedback-selection technique pair, e.g., Dwell-Visual). All data was analysed using multi-factorial repeated-measures ANOVA, other than the NASA TLX

subjective ratings data, which was analysed using non-parametric Wilcoxon *T* tests. Analysis was carried out using SPSS.

3.5.4.1 Error Rate (ER)

Learning Effects

A 2 x 2 x 3 (selection technique x feedback x block) repeated-measures ANOVA showed no significant effect of block on error rate ($F_{(2,950)} = 0.237$, $p > 0.05$). This suggests there were no learning effects and performance did not change significantly over time.

Selection Technique and Feedback Type

The mean overall error rate (ER) across all conditions was 20.5%. A 2 x 2 x 4 (selection technique x feedback x number of items) repeated-measures ANOVA showed a significant main effect of selection technique on errors ($F_{(1,1168)} = 280.908$, $p < 0.001$), a significant main effect of feedback type on errors ($F_{(1,153)} = 107.070$, $p < 0.001$) and an interaction between selection technique and feedback type ($F_{(1,153)} = 47.798$, $p < 0.001$). The interaction occurred as the difference in ER between Visual and Audio feedback was much larger for Dwell than for Quick Release. Dwell had a lower error rate (11%) than Quick Release (30%), leading to an acceptance of hypothesis H1. Visual feedback had a lower error rate (15%) than Audio feedback (26%). The difference in ER between the Dwell-Visual and Dwell-Audio conditions was much larger than Quick-Visual compared to Quick-Audio, leading to the interaction effect. Figure 3-6 shows the mean error rate for all conditions.

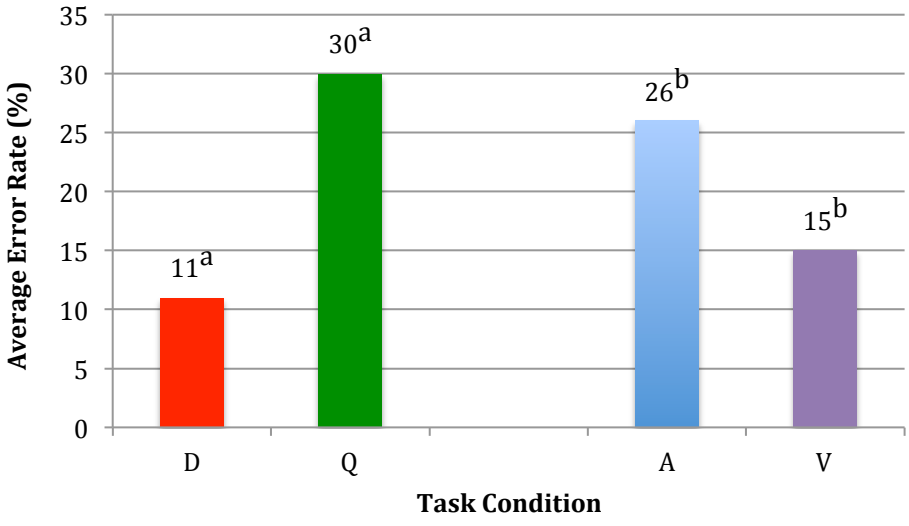

Figure 3-6. Average number of errors for all conditions (D: Dwell; Q: Quick Release; A: Audio; V: Visual). 'a' and 'b' indicate a significant difference $p < 0.001$.

Number of Menu Items

For this comparison only selections from the 4 common-distance targets from each menu size were considered for analysis, similar to Ramos *et al.* [192]. The same 2 x 2 x 4 (selection x feedback x number of items) repeated-measures ANOVA showed a significant main effect of the number of menu items on mean error rate ($F_{(3,327)} = 22.405$, $p < 0.001$) as well as a significant interaction between selection technique and feedback type ($F_{(1,109)} = 12.082$, $p < 0.01$), a significant interaction between selection technique and number of items ($F_{(3,327)} = 2.665$, $p < 0.05$), a significant interaction between feedback and number of items ($F_{(3,327)} = 6.115$, $p < 0.01$) and a 3-way interaction between selection technique, feedback and number of items ($F_{(3,327)} = 2.758$, $p < 0.05$). Mean error rate increased as the number of menu items increased with mean error rates of 9%, 17%, 25% and 26% for 4, 6, 8 and 10 items respectively, leading to an acceptance of H2. *Post hoc* Bonferroni pairwise comparisons revealed that the number of errors differed significantly for all pairs of menu sizes at significance $p < 0.001$ except for 8 x 10 items which was non-significant ($p > 0.05$). Figure 3-7 shows the average number of errors per trial for each menu size.

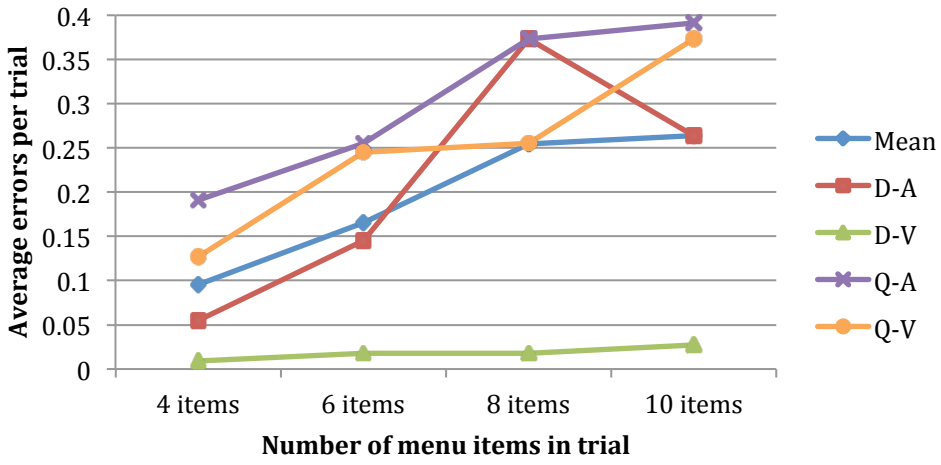

Figure 3-7. Average number of errors per trial for all numbers of menu items. Lines correspond to selection technique-feedback pairs.

The interaction between selection technique and number of items may exist because, in the Dwell conditions, 10 menu items produced fewer errors than 8 items, whereas, in the Quick Release conditions, 10 items produced more errors than 8 items. Error rates for both feedback conditions increased from 4 to 8 items. Upon simple inspection, the feedback x number of items interaction may come from a similar uneven change in error rate from 8 items to 10 items, as it drops from 8 to 10 items under audio feedback but increases from 8 to 10 items under visual feedback. As for the 3-way interaction, error increases with

increased number of items under all selection-feedback pairs except for Dwell-Audio which increases to 8 items before dropping in error rate from 8 to 10 items (see square points in Figure 3-7).

The results here for error rate support the acceptance of hypotheses H1 and H2, and are much more encouraging than in the preliminary study and suggest that near-perfect accuracy is possible in pressure interaction on mobile devices. This is true even with as many as 10 distinct pressure levels (in this case using the Dwell selection technique and visual feedback, triangles/lowermost line in Figure 3-7). It also suggests that non-visual interaction is also highly usable if the number of pressure levels is kept below 8 (again using the Dwell technique, square line in Figure 3-7). Poor performance using the Quick Release technique, however, was quite surprising with this being more evident in the Audio feedback condition.

3.5.4.2 Movement Time (MT)

Selection Technique and Feedback Type

A 2 x 2 repeated measures ANOVA showed a significant main effect of selection technique on movement time ($F_{(1,1427)} = 136.529$, $p < 0.001$) and a significant main effect of feedback type on movement time ($F_{(1,1427)} = 565.253$, $p < 0.001$). Dwell had a higher average movement time (3.4 seconds) compared to Quick Release (2.7 seconds) and Audio had a higher average movement time (3.8 seconds) than Visual (2.2 seconds; see Figure 3-8).

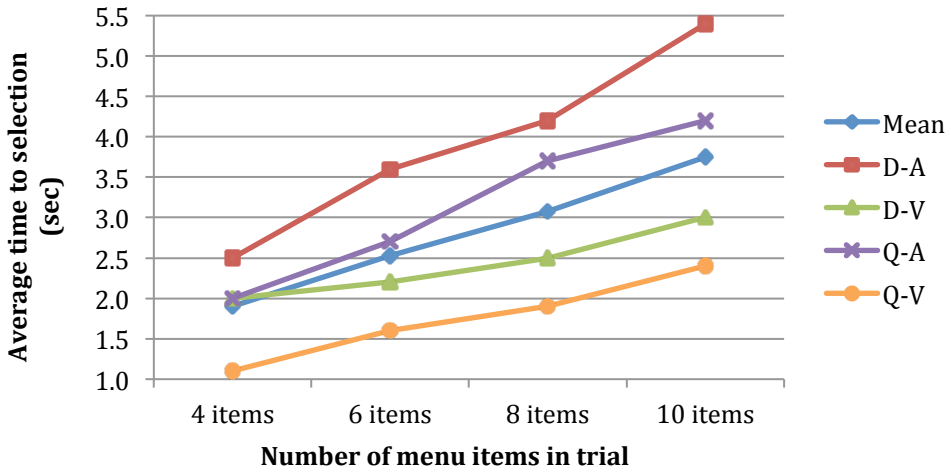

Figure 3-8. Average movement time (MT) per trial in seconds. Lines correspond to selection technique-feedback pairs.

Number of Menu Items

Average movement time increased as the number of items increased, with means of 1.9s,

2.5s, 3.1s and 3.8s for 4, 6, 8 and 10 menu items respectively. Average movement time also increased as target distance increased for all number of items under all conditions. In a similar trend to error rates, the last item frequently had lower MT.

MT results support rejection of the null hypothesis in favour of adopting hypotheses H3 and H4 as Quick Release trials were on average faster than Dwell trials and Visual feedback allowed quicker average selection times than Audio feedback. Audio feedback increases selection time by almost 75%.

3.5.4.3 Number of Crossings (NC)

Selection Technique and Feedback Type

A 2 x 2 x 4 repeated measures ANOVA showed a significant effect of selection technique on number of crossings ($F_{(1,203)} = 72.174$, $p < 0.001$) as well as a significant effect of feedback ($F_{(1,203)} = 59.676$, $p < 0.001$). Dwell had a higher average number of crossings (7.2) compared to Quick Release (4.7), while Audio feedback produced more crossings (7.1) than Visual feedback (4.8).

Number of Menu Items

The same 2 x 2 x 4 repeated measures ANOVA showed a significant main effect of number of menu items on the number of crossings. Mauchly's test indicated a violation in the assumption of sphericity of variance for number of items (chi-square = 105.804, $p < 0.001$), therefore degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity (epsilon = 0.74). Under this correction the significance stood at $F_{(3,609)} = 156.458$, $p < 0.001$.The number of crossings increased as the number of menu items increased with means of 2.2, 4.5, 6.4 and 10.6 crossings for 4, 6, 8 and 10 items respectively. Pairwise comparisons revealed that the number of crossings differed significantly for all pairs of number of menu items at significance $p < 0.001$. Again, the results for NC support acceptance of alternative hypotheses H5 and H6. In a very similar trend to MT, NC also increases as the number of items increases, which suggests that users take more time oscillating back and forth over targets as they become smaller.

3.5.4.4 NASA TLX Workload

Note that low Performance ratings indicate perception of good performance, unlike other scales where higher numbers indicate high levels of the measure. Non-parametric Wilcoxon *T* tests showed that audio feedback produced significantly higher Overall workload (8.73)

than visual feedback (6.74). Quick Release produced significantly higher Overall workload (8.37) than Dwell (7.11).

Measures of Thumb Fatigue were analysed in terms of the order of conditions in which a participant took part, rather than condition type, looking at fatigue from the first condition through to the last condition. A non-parametric Friedman's test showed a significant effect of condition order on Thumb Fatigue (χ^2 (3) = 14.34, p < 0.01). *Post hoc* Wilcoxon *T* tests with a Bonferroni-corrected p-value of 0.0083 showed that the third condition produced significantly higher reports of thumb fatigue than the fourth condition completed (p=0.002). Mean measures were 6.12, 7.76, 8.47 and 5.88 for the 1st, 2nd, 3rd and 4th condition completed. Subjective levels of annoyance produced by the audio design during the Dwell-Audio and Quick-Audio conditions were recorded. There was no significant effect of Control Method on Audio Annoyance, with mean ratings of 7.65 when using Dwell and 6.76 when using Quick Release and an overall rating of 7.2.

3.5.5 Discussion and Initial Conclusions

The results from Experiment 1 suggest that a much better audio feedback design was employed, compared to the preliminary study, and they show that both visual and non-visual pressure-based interaction with a mobile device can be usable and highly accurate. Several accuracy rates shown here are above those found in previous studies using non-linear sensors. Shi *et al.* [209] found 78% accuracy with visual feedback using the Dwell selection technique, whereas Experiment 1 found 83% accuracy with visual feedback and Dwell. In the current study, participants managed 10 levels at 73% accuracy using only audio feedback, almost equalling that of Shi *et al.* [209] when using visual feedback. However, the relatively high accuracy for audio-only interaction came at the cost of significantly higher overall workload, including higher mental and physical demand and perceived effort. The ratings varied from 8.3 for physical demand to 11.8 for effort, out of a maximum of 21, so, although the ratings were significantly higher using audio feedback, no ratings were particularly high.

It would appear from the results that a pressure space of 3.5N allows for good control at up to 10 levels, particularly when using the Dwell selection technique with visual feedback and a linearised sensor. A particularly encouraging set of results is the near-perfect accuracy rates for all numbers of menu items under visual feedback using the Dwell technique. The worst performance was still only at 3% errors for 10 menu items with perfect 0% errors

for 4 items. As was found by other research [30, 192], the Dwell selection technique was more accurate but slower than Quick Release. Quick Release also received significantly higher ratings of overall workload, including significantly higher mental demand and poorer perceived performance. In contrast, Dwell was rated as significantly more frustrating, potentially as maintaining the same level of pressure for the 1-second duration was challenging.

Contrary to the findings of Mizobuchi *et al.* [170], no extreme fatigue was found when acquiring targets at the farthest end of the pressure space. Although errors did increase as the distance increased, as they did in Mizobuchi *et al.*, subjective reports (NASA TLX) of thumb fatigue peaked at 8.5 out of 21, with an average report of 7.2. The pattern of thumb fatigue ratings is peculiar, as they rose, on average, across the first three conditions than a participant took part in, but dropped again for the final condition.

In comparison to both Shi *et al.* [209] and Ramos *et al.* [192], however, the MT and NC results were worse, suggesting that improved accuracy in Experiment 1 came at the cost of the speed of interaction. Both of these measures increased as the number of items increased, but they also tended to increase as the distance to target increased, although this was not apparent across all conditions. Given the very similar increase in both MT and NC this suggests that, rather than deliberately taking more time to carefully orient towards targets, participants are more likely unintentionally moving the cursor back-and-forth over a target in an attempt to pinpoint the small target size. Because only 30% of the sensor's range was used in the study (it could detect up to 12 N), it is possible that there would have been more noise in the output than if the whole range had been used. The sensor can be calibrated to be linear across any pressure range so perhaps reducing the sensor range to 3.5 N would improve control, and consequently MT and NC, even further.

There is a clear difference between the accuracy of selections for Dwell-Visual compared to all other conditions (see Figures 3-6 and 3-7). Examining the selection points (pressure value where selection occurred), almost all selection points fall within the target boundaries for Dwell-Visual, while the majority of all misses in the other conditions occur within a relatively short 'distance' of the lower target boundaries (i.e., the least pressure required to be in the target). The input behaviour was examined to try to determine why this was: why there were so few errors where users have 'overshot' the target (other than the last menu item). For the Dwell trials it seems as though participants simply did not press enough to get to the target. They would take too long to press hard enough and accidentally remain in a non-target item for the 1-second Dwell timer. As most errors occurred in the Dwell-Audio

condition, this hesitancy or lack of speed could come from a lack of familiarity with the order of menu items, as they are not sure where they are relative to other items. However, the errors are spread across all blocks, so they would be expected to have a firmer knowledge of item positioning.

Alternatively, it may be that, as in the preliminary study, there remained insufficient information in the audio feedback to properly facilitate accurate positioning. If this was the case, however, one might expect more errors past the target items as well. Research has shown that, when provided with no external feedback, or impoverished visual feedback, participants often apply less pressure than is required of them [121, 127, 170]. Also, Johansson & Westling [116] found that, when gripping objects, humans apply a small safety margin, or 'just enough' grip-pressure strategy to avoid slippage and not risk damaging the object or unnecessarily over-exerting ourselves. These findings could account for the low levels of pressure, as it seems we may have a natural tendency to err on the side of applying less pressure.

Increasing the length of the Dwell selection timer may reduce the number of accidental selections, resulting from loitering in a lower target for too long, but it would increase the interaction time, and it may be more difficult to maintain pressure accurately for longer periods, increasing the frustration already experienced. Combating a tendency to press less than is required may be difficult. Target boundaries could be dynamic, so, for example, the lower boundary of the next target could move down as the cursor/amount of pressure comes close to the edge. However, it may be that the individual wishes to select the current target, but is simply pressing in the upper extent of the target. Using uneven state-transitions could reduce the number of times the cursor accidentally slips into a lower target level [196]. This method effectively increases the width of the target most recently entered.

3.5.5.1 Quick-Release Performance

Performance using the Quick Release (QR) mechanism was surprising and disappointing. Although QR has been found to be generally more error-prone than other selection mechanisms in other research [14, 30, 192], the difference in performance between QR and Dwell in Experiment 1, particularly when using audio feedback, was much larger than has been found in previous research (see Figures 3-6 and 3-7). In other words, while QR is often found to be less accurate than other mechanisms, it was considerably less accurate here. The same possible factors outlined in relation to the Dwell technique are also relevant to the QR trials. However, looking at the pressure behaviour, one of the primary contributing factors appears to be the QR selection mechanism itself.

Designing an accurate QR mechanism is troublesome because it is difficult to identify a common and clear pattern of sensor behaviour from which user intent can be unambiguously retrieved. For example, a rapid drop in pressure input to 0, or near-0, could simply be the participant deliberately reducing pressure, perhaps to try targeting again from the start, or to target a lower-pressure target after unintentionally over-shooting it. Because the sensor sampled at 52Hz it was almost unavoidable that samples would be taken between lift-off and a 0 reading. The selection method used in Experiment 1 used a simple algorithm comparing where and when samples were taken to decide on the lift-off point. However, looking at the pressure input profiles, it became clear that the algorithm might not always identify the correct lift-off point, instead picking a pressure value somewhere between a higher (genuine) lift-off point and 0. Occasionally, therefore, even if a participant lifted within the target, the algorithm would take a sample outside of that target (on the way back to 0 pressure) instead. To evaluate the effect of this problem, the way in which the algorithm looked for the selection value was improved and the experimental data was re-analysed. The pressure behaviour therefore remained the same, however a different method was used to identify thumb lift-off.

Figure 3-9 shows the 'lift-off' selection points from a Quick-Release-Audio condition during an 8-item menu. The black bars indicate the boundaries of each target item. The left-hand graph shows the original selection points, including the number of erroneous selection points running beneath each target. The right-hand graph shows the same trials, but the selection points are chosen based on the "corrected" QR mechanism. Comparing the two selection distributions highlights what would have been a marked change in recorded error rates, should the corrected mechanism have been used for the main evaluation. The "corrected" selections were much more accurate, decreasing ER rates by up to 50% (of original ER). The ER data was re-analysed using the selection points from the "corrected" selections and, while Dwell was still significantly more accurate (at 11% ER) Quick Release was much more accurate, at 14.7% on average (compared to 30% previously). Accuracy using Visual and Audio feedback individually fell from 28% to 8.3% and from 32% to 21%, respectively. The drawback of this mechanism has always been that it is more error-prone than Dwell, while retaining the benefit of speed. If refinement of the QR mechanism can reduce the error rate as hinted at here, it could become the ideal mechanism.

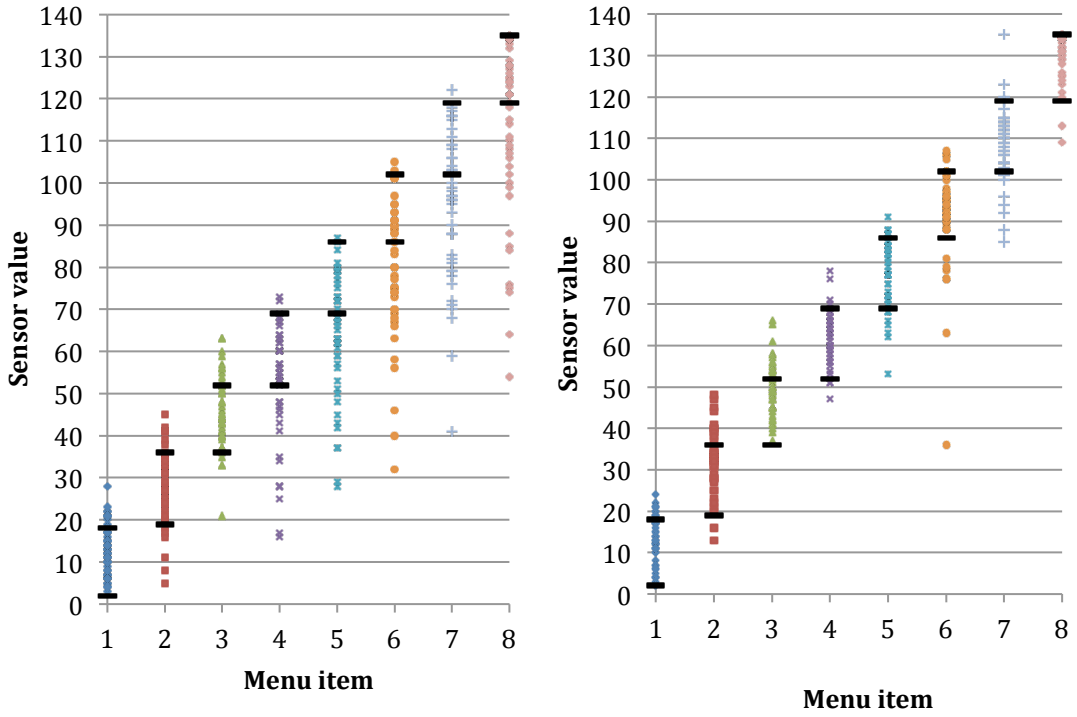

Figure 3-9. Example "corrected" selection point distribution for 8-item Quick-Release-Audio condition (right) compared to original selection distribution (left). Y-axis represents pressure-sensor value at which selection was made. Black horizontal bars indicate target level (menu items, x-axis) boundaries.

3.6 Limitations

This section discusses the main limitations of the research reported in this chapter, which should be considered when interpreting the results. The main limitations of Experiment 1 were: 1) the use of non-equivalent feedback designs and 2) that interaction occurred away from the mobile device.

3.6.1 Non-Equivalent Feedback Designs

Research from both psychophysics and HCI have shown that continuous visual feedback supports optimal performance in applying target levels of pressure and reducing the amount of feedback (i.e., reducing the information available) leads to poorer control [108, 121, 127, 192]. By aiming to develop eyes-free pressure input for mobile devices, it was therefore important to provide as much information as possible in the audio domain. There are several similarities:

1. Both provide a discrete cue regarding which menu item the cursor is currently within (visual: green-border around item and visual label; audio: unique tone that plays continuously and spoken label).
2. Both provide a spatial cue regarding where the cursor is within the menu (visual: location of green-bordered item relative to top and bottom of menu; audio: position of spatialised audio relative to left and right extremes).
3. Both provide a spatial cue regarding where the next target to select is within the menu (visual: location of green-bordered item relative to top and bottom of menu; audio: position of spatialised audio relative to left and right extremes).

However, it is not claimed that the audio feedback and visual feedback designs used here are equivalent, i.e., they do not provide exactly the same information or the same amount of information. For example, the visual feedback shows a continuously moving cursor: the audio feedback provides only a discrete cue regarding cursor position, although this cue plays continuously (constantly). Also, the visual feedback shows the position of all menu items at all times: the participant can always see the label of each item. In the audio design the labels are only heard when the cursor enters the item. Stewart *et al.* [228] found that continuous audio feedback was annoying to participants, so it was important to tailor the audio feedback to be sufficiently informative and yet not annoy or frustrate. As can be seen from the results of Experiments 1-3, in line with the psychophysical and HCI research, using the audio feedback generally resulted in poorer performance than visual feedback, however, in real terms, the audio feedback design still supported highly accurate pressure input and relatively low annoyance levels, rated 6.7-7.7 out of 21.

Another important consideration is the use of visual and audio feedback when in mobile scenarios. Audio feedback was chosen over other non-visual feedback methods such as vibrotactile feedback as it could provide a wider variety of stimuli/information. In Experiments 1-3 participants received the audio feedback through headphones. A future user may be able to receive the audio feedback if they are already wearing headphones, for the purpose of listening to music, or other audio content, for example. However, if they were not, they would have to put the headphones on specifically each time they wished to interact, which could be a hindrance. Non-contact/ambient audio displays, which use speakers to provide audio, could be worn around the head [203], but these have not seen commercialisation. Audio feedback through headphones also blocks out environmental audio sources, such as traffic, voices or warnings, which could be a potential hazard.

3.6.2 Interaction Occurred Away from Device

The purpose of the research was to test control of pressure input *on a mobile device* when using visual and non-visual feedback. The software was run on the N810 and the pressure sensor connected directly to it, showing that the interaction could be incorporated into common, commercial devices. However, as explained in Section 3.5.3.1, the pressure sensor was attached to a piece of Perspex, and not directly to the device, because it led the sensor 'tail' to flex, resulting in errors in sensor output. The intention had been to have the sensor attached to the device, so that the way in which the participant applied pressure and interacted with the device would be a close approximation of how it would be when holding and interacting with a genuinely pressure-sensitive mobile device. However, the pinch-grip interaction on the FSR is very similar to the way it would be manipulated if the sensor were attached to the device and so the interaction remains valid. This issue was rectified in the remaining Experiments 2-5 in the thesis, as the sensors were attached to the device being interacted with.

3.7 Conclusions and Research Question 1

Research Question 1 asked:

"How accurate is pressure-based input on a mobile device when using only audio feedback?"

To answer this question it was first necessary to design non-visual feedback that was useful for pressure-based input. Audio feedback was chosen over vibrotactile feedback, as it was judged capable of conveying more information, even though audio feedback designs had resulted in less accurate pressure input in other research, compared to vibrotactile [228, 232, 233]. The initial preliminary study and Experiment 1 iteratively developed audio feedback that allowed users to control pressure-based linear targeting while sat at a desk. The use of continuous feedback spatialised horizontally through egocentric space resulted in high accuracy (>= 85%) when a Dwell selection technique was used and the pressure space was divided into 4 or 6 levels, more than had previously been controlled using non-visual feedback, and at equal or higher accuracy [228, 232, 233]. Selection time remained relatively high at 2.5-3.5 seconds (including the 1-second Dwell time), but still in line with results from Stewart *et al.* [228] using audio or vibrotactile feedback. Decreasing the size of

pressure levels, by increasing the number of levels, made non-visual interaction more difficult, with both error rates and selection time increasing significantly when 8 or 10 levels were used.

Testing non-visual pressure interaction when mobile is important and Experiment 1 only tested control when the participants were sitting however, in partial answer to RQ 1, it appears that, using the spatialised design, pressure input is highly accurate using only audio feedback when the user is sitting. Therefore, the contribution of the research in this chapter is that non-visual control of a wide range of pressure can be highly accurate when sitting and using a spatialised audio feedback design.

4 Mobile Non-Visual Pressure-Based Input

4.1 Introduction

Pressure-based input has been demonstrated as an accurate means of interacting with desktop systems [30, 170, 192], and some research has begun to show that input can also be accurate on mobile devices [14, 170, 228]. However, only one piece of research has tested control of pressure-based input when the users were actually walking [14], and the interaction used a very limited means of input: two different pressure levels for inputting lowercase or uppercase letters. Walking can significantly hinder the user's ability to accurately [38, 40, 204], and quickly [11, 38, 159, 204] carry out certain tasks on mobile devices. Pressure-based interfaces have made successful use of a much larger input range in desktop settings, so it was necessary to test whether similar ranges can be accurately controlled when walking. Therefore, Research Question (RQ) 2 asks:

"How accurate is pressure-based input when the individual is walking?"

Interaction when walking and using continuous visual feedback would provide the baseline answer to this question, as this form of feedback has facilitated accurate pressure input when seated. To that end, Experiment 2 in Section 4.2 tested control of pressure input when the user was walking and provided with visual feedback. Providing non-visual feedback is important for interaction with mobile devices, and Experiment 1 suggested that eyes-free pressure input could be accurate when seated, using the more stable Dwell selection technique and being provided with spatialised audio. RQ 1 asked:

"How accurate is pressure-based input on a mobile device when using only audio feedback?"

Having designed audio feedback that could provide accurate pressure input when sitting, Experiment 3 in Section 4.3 tested eyes-free pressure input when the user was walking.

The primary purpose of Experiments 2 and 3 was to test control when the user is walking and provided with either visual or audio feedback. Because walking can negatively influence performance in certain tasks, including linear targeting [38], another means of potentially

improving mobile pressure interaction was investigated: the *control method*, or mapping of pressure to input. Pressure-based linear targeting studies, including those in the preliminary and Experiment 1, use what is called Positional control of input, where the position of the cursor is controlled by how much pressure is applied, and is therefore a direct mapping of pressure to input. This movement can be measured and tested to determine the precision of applied pressure. An alternative control method for a pressure-based interface would be Rate-based control, where the velocity of an interaction element, in this case the speed of the cursor movement through the interaction space, is controlled by the amount of pressure being applied. The speed of the cursor increases as the amount of pressure increases. This control method is not as suitable for measuring the precision of applied pressure, as it is an artificial mapping, but it is useful for understanding how the usability of a pressure-based interaction, like targeting or menus, can be improved.

Zhai [260] summarised existing literature on isometric (pressure-based) input devices and concluded that they are better suited to velocity or rate-based input, compared to positional input. Shi *et al.* [210] found that Rate-based control allowed for faster, more precise and less mentally/physically demanding control of pressure-based shape rotation. Outside of pressure interaction, but remaining within linear targeting, Crossan *et al.* [38] found that Rate-based control of cursor movement through head tilting produced more accurate selections than Positional control while the user was walking, with Positional control being faster and more accurate when the user was standing still. As Rate-based control may allow for more accurate or stable control for mobile (and non-visual) interaction, it was decided to compare performance using both control methods while the user was both sitting and walking.

4.1.1 Task

The same linear targeting task as used in Experiment 1 (Section 3.2) was used for both experiments described in this chapter, with the visual feedback shown in Figure 3-1. The same 3.5 N pressure space was used for input. As mentioned in Section 3.3.1, the only difference in visual feedback from Experiment 1 to Experiments 2 and 3 in this chapter was the addition of a highlight for the menu item/target that the cursor was currently in (the target boundary rectangle was displayed in green, rather than black). The task and behaviour of the software was identical to during Experiment 1.

Experiment 2 tested control of the pressure-based targeting task when the participants were sitting, walking and using both the Positional and Rate-based control methods. Experiment 3 used the control method from Experiment 2 that resulted in best task performance when

walking and tested control when sitting and walking and using audio feedback. Section 4.2 describes Experiment 2 and Section 4.3 describes Experiment 3. Section 4.4 compared the results from Experiment 2 and 3 to establish the effect of walking, compared to sitting when provided with visual vs. audio feedback. Section 4.5 discusses the limitations of the research while Section 4.6 summarises and discusses all the results from the chapter.

4.2 Experiment 2 – The Effects of Mobility and Control Method on Pressure-based Input

4.2.1 Apparatus

In Experiment 1, the manipulation method (the finger-thumb pinch against rigid Perspex) was similar to that which would be used when pinching to press a pressure sensor attached to the front (or under the screen) of a mobile device, similar to the "grip" method in Stewart *et al.* [228]. However, the interaction was still different, as the sensor and device were separate and the user did not hold the device in the hands. It was necessary to test control when the user held a device in both hands and interacted directly on/with the device. Attaching a force-sensing resistor to the body of the N810 was problematic, as the sensor 'tail' would flex during use, causing abnormal sensor output behaviour.

Figure 4-1: Interlinks Electronics Force-Sensing Resistor (FSR) model 402 (left) and Samsung UMPC model Q1 (right) with FSR attached (top right).

The experimental software ran on a Samsung Q1 UMPC (see Figure 4-1, right). The Q1 was used for several reasons: 1) it provided a flat surface upon which the sensor could be placed to avoid flexing of the sensor tail; 2) it provided USB input for the sensor, which other mobile devices lack and 3) it had a similar screen resolution, so the spatial movement of the

cursor would be comparable to that on the N810. The apparatus for detecting pressure was different from that used in Experiment 1. An Interlinks Electronics force-sensing resistor (FSR) model 402 (also Figure 4-1, left) was connected to the Q1 over USB via an SAMH Engineering SK7-ExtGPIO01 input/output module, which handled A-to-D conversion and included the same sensor linearization [228] as was used in Experiment 1. The FSR was attached to the front bezel (plastic case surrounding the screen) of the Q1 on the same side as the user's dominant hand, to be operated by the thumb of that hand. This positioning meant that the sensor would be manipulated in a similar manner to the way it was in Experiment 1, pinching the sensor between the thumb (on top) and the fingers behind the device. This same apparatus was used for both experiments.

4.2.2 Control Methods

4.2.2.1 Positional Control

This control method is the same as was used during Experiment 1. The 1-second Dwell selection technique was used, as it provided the best accuracy in Experiment 1.

4.2.2.2 Rate-based Control

In this method the velocity of the cursor's downward motion through the menu was dictated by how hard the participant pressed on the FSR, with no pressure bringing the cursor to a halt (by lifting the thumb from the sensor). This interaction is similar to pushing an object along a smooth surface: how hard you push it dictates how fast it moves, and stopping pushing, stops the object's movement. 'Velocity' in this case refers to the number of pixels (or millimetres) the cursor moves every cycle of the experimental software, which was approximately every 0.03 sec (33Hz). Pilot testing led to the adoption of a maximum speed of 10 pixels (2 mm) per cycle (330 pixels/66 mm per second; see Table 4-1). Initially a maximum speed of 20 pixels (4 mm) per cycle was chosen, balancing speed and control, but this was found to be too fast for accurate control when using audio feedback.

Approx. Pressure (N)	0	0.2	0.4	0.6	0.8	1.0	1.2	1.4	1.6	1.8	2.0
Speed in pixels/sec (mm/sec)	0	33 (6.6)	66 (13.2)	99 (19.8)	132 (26.4)	165 (33)	198 (39.6)	231 (46.2)	264 (52.8)	297 (59.4)	330 (66)

Table 4-1: Rate-based condition speeds in pixels- and millimetres-per-second, based on pressure input in Newtons (N).

This design only allowed for downward motion of the cursor. Although a second FSR could

have been used to allow for upward motion (in the case of overshooting a target), the Positional control method only utilized one sensor, so it was decided to use only one for Rate-based control, to keep the interactions as similar as possible (even though Positional control allows for bi-directional movement). Therefore, if the participant overshot a target, they could push the cursor past the bottom of the menu and it would 'loop' back to the top of the menu and start again. Stopping the cursor within the target item (by lifting off the FSR) and leaving it stationary for 1 second achieved target selection.

4.2.3 Mobility

During the static condition, participants were sat in a padded office chair holding the UMPC in both hands. They were allowed to rest their arms on either their legs or a desk in front of them to provide stability, but could not rest the device or their wrists while interacting. The mobile condition used a similar design to Crossan *et al.* [38] as it requires divided visual attention between task and navigation. Participants were asked to walk in a 4 m x 3 m figure-of-eight route indoors while they interacted with the device (see Figure 4-2). The route was marked by four pieces of paper, one at each corner of the rectangle and users held the device in both hands with no further support.

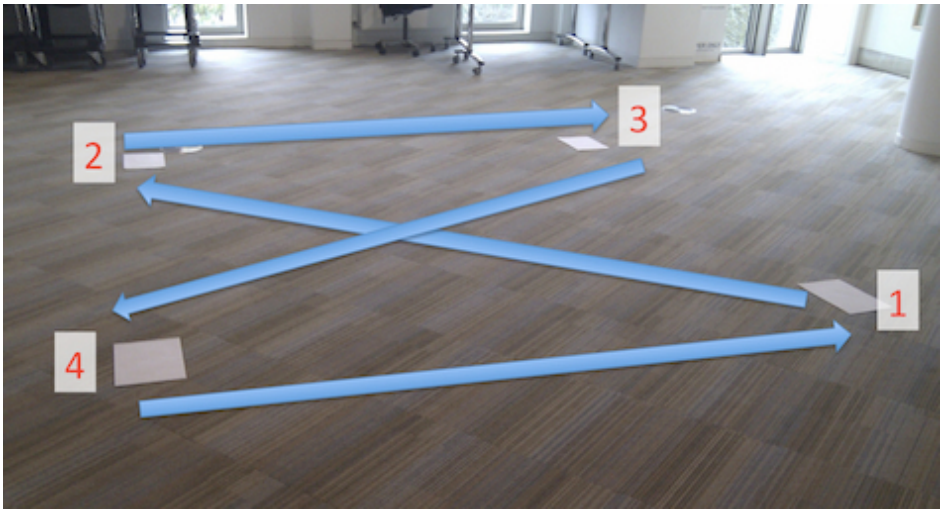

Figure 4-2: Figure-of-eight walking route for Experiment 2 in indoor office space.

4.2.4 Participants and Experimental Procedure

Fourteen participants (11 male, 3 female) aged between 17 and 30 years old (mean 22.8) took part in the evaluation, all of whom were from within the University. Thirteen were

right-handed and all were paid £20 for participation in both Experiment 2 and 3. The experiment was a 2 x 2 within-subjects design (mobility x control) so that participants completed two static and two mobile conditions, using each of the control methods: *Static-Positional, Static-Rate, Mobile-Positional* and *Mobile-Rate*. The order of these four conditions was counterbalanced to avoid order effects. Within each condition every menu item from all four of the menu sizes (4, 6, 8 and 10 items) was to be selected once. The presentation order of menu sizes was randomized, and all targets within that menu were presented in a random order. Each condition began with 10 practice selections and ended with participants completing a NASA TLX workload estimation form. Experimental instructions can be found in Appendix B.

4.2.5 Variables and Measures

There were three Independent Variables: *Control Method* (Positional, Rate-based), *Mobility* (Sitting, Walking) and *Menu Size* (4, 6, 8 or 10 items). Dependent Variables were: *Errors* (ER, % of missed targets), *Movement Time* (MT), *Number of Crossings* (NC, only relevant during Positional control), *Loops* (the number of times the cursor looped to the beginning of the menu after an overshot target in Rate-based control), *Nudges* (the number of discrete presses on the FSR to 'nudge' cursor along during Rate-based control) and *Workload* via the NASA TLX. This gave a total of: 14 participants x 2 Control Methods x 2 Mobility conditions x 28 target distances = 1568 trials. This gave 784 data points for each Control Method and Mobility condition, and 392 data points for each Control Method + Mobility combination condition (e.g., Static-Positional).

NC only applies to Positional control and so a somewhat similar measure, here called *Loops*, was used for Rate-based conditions and measured the number of overshot attempts. A final objective measure recorded during Rate-based conditions was called *Nudges*: the number of press-release cycles the user employs to move the cursor, essentially 'nudging' or 'shunting' it along, as a sort of searching behaviour. This may indicate lower confidence in control over the input.

4.2.6 Results – Experiment 2

The analytical approach was the same as in Experiment 1, so that, when comparing between conditions, the data for every target selected within that condition was used in the analysis. When analysing the potential effect of Menu Size on the variables only the 4 targets of

similar distance from each menu size were compared. Some of the data did not fit a normal distribution, and so non-parametric analyses were used, specifically Wilcoxon *T* test for pairwise comparisons and the Friedman test for non-parametric ANOVA equivalent. Although the use of non-parametric tests increases the validity of results gained from non-normal data, they are limited in their inability to examine interaction effects. Wilcoxon *T* tests were used as *post hoc* pairwise comparisons between levels following significant Friedman's test results, and used the Bonferroni correction on the p-value necessary for statistical significance: $p<0.05/N$, where N is the total number of comparisons. Normality was tested using the Shapiro-Wilk test in SPSS. For normally distributed data, ANOVA was used. Raw data for all measures can be found in Appendix B.

4.2.6.1 Errors

A Wilcoxon pairwise comparison showed a significant effect of mobility on number of errors ($T = 682.50$, $p < 0.01$), as walking (mean = 3.1%) produced more errors than sitting (mean = 1.7%). There was no effect of control method on errors ($T = 1242.50$, $p > 0.05$) as both had ER of 2.4%. Comparing conditions, Wilcoxon *T* pairwise comparisons showed a significant difference between the *Static-Rate* and the *Mobile-Rate* conditions ($T = 170$, $p < 0.05$). All other comparisons were not significant ($p > 0.05$). Error rates for the four conditions were: 1.8% for *Static-Positional* (SD = 0.13), 1.5% for *Static-Rate* (SD = 0.12), 2.9% for *Mobile-Positional* (SD = 0.17) and 3.2% for *Mobile-Rate* (SD = 0.18). Error values are shown in Figure 4-3.

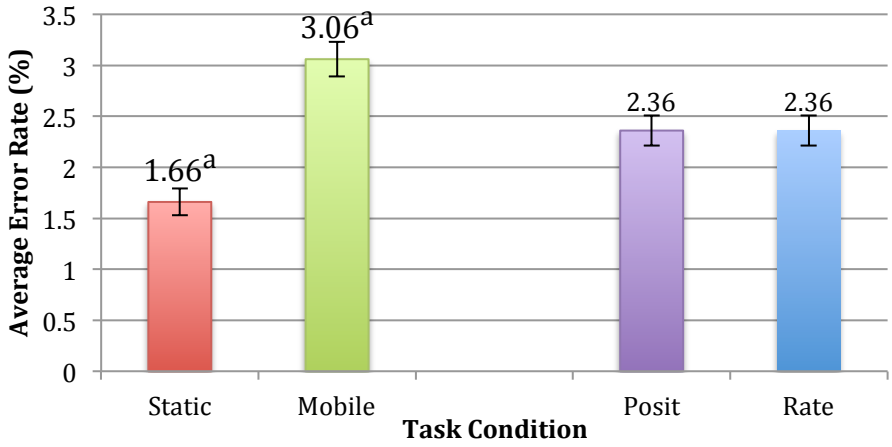

Figure 4-3: Mean error rates for Experiment 2 conditions: Static, Mobile, Positional (Posit) and Rate-based (Rate). Error bars show 1 standard deviation. 'a' indicates significant difference $p < 0.01$.

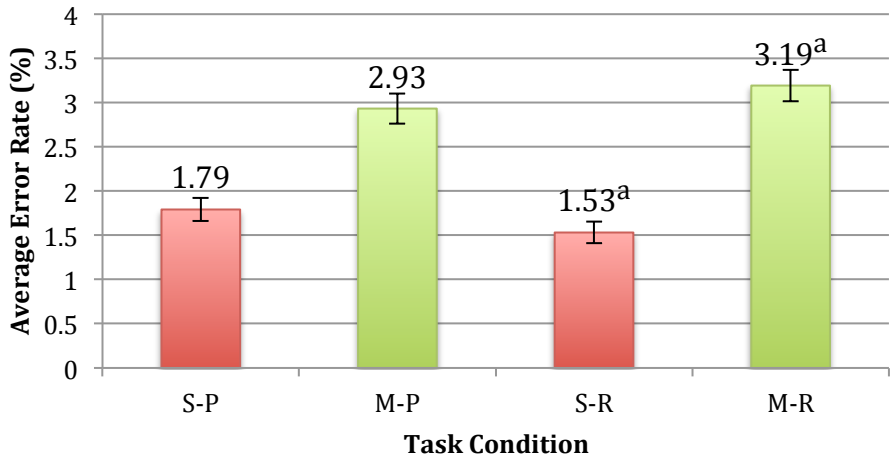

Figure 4-4: Mean error rates for Experiment 2 sub-condition: S = Static, M = Mobile, P = Positional, R = Rate-based. E.g., S-P = Static-Positional condition. Error bars show 1 standard deviation. 'a' indicates significant difference $p < 0.05$.

Friedman's Test showed a significant effect of Menu Size on ER (χ^2 (2) = 9.867, $p < 0.05$). However, no *post hoc* pairwise comparisons reached the required Bonferroni-corrected p-value of 0.0083, using Wilcoxon *T* tests. Mean ER values for each menu size were 1.2%, 2.4%, 0.7% and 3.4% for 4, 6, 8 and 10 item menus.

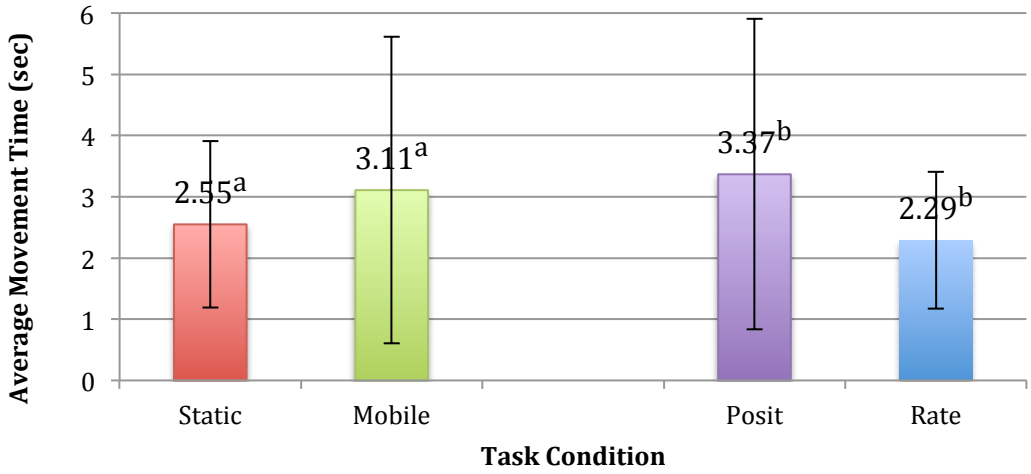

Figure 4-5: Mean target selection times for each condition during Experiment 2: Static, Mobile, Positional (Posit) and Rate-based (Rate). Error bars show 1 standard deviation. 'a' and 'b' indicate significant differences $p < 0.001$.

4.2.6.2 Movement Time

Both mobility (T = 472356.5, $p < 0.001$) and control method (T = 266186.5, $p < 0.001$) significantly affected task time, with Rate-based control (mean = 2.29s) allowing for faster

selections than Positional control (mean = 3.37s) and walking (mean = 3.11s) causing slower selections than sitting (mean = 2.55s). Wilcoxon *T* comparisons found that all conditions were significantly different from each other ($p < 0.001$), with the exception of *Static-Rate* vs. *Mobile-Rate* ($p > 0.05$). Mean movement times for each condition (including the one-second Dwell time) were 2.85s (Static-Positional), 2.24s (Static-Rate), 3.88s (Mobile-Positional) and 2.34s (Mobile Rate). Menu size also had a significant effect on MT ($\chi^2(3) = 307.861$, $p < 0.001$). Wilcoxon comparisons with a Bonferroni-corrected p-value of 0.0083 showed that all menu sizes differed from each other significantly ($p < 0.001$), with mean MT of 2.14s, 2.54s, 2.94 and 3.55s for 4, 6, 8 and 10 item menus respectively.

4.2.6.3 Number of Crossings/Loops

NC only applies to the Positional control method and so only *Static-Positional* and *Mobile-Positional* were compared. Wilcoxon pairwise comparison showed that mobility had a significant effect on the number of crossings ($T = 170$, $p < 0.05$) with mobile selections resulting in more crossings (mean = 6.25) per target than static selections (mean = 3.46). Menu size significantly affected NC (Friedman's $\chi^2(3) = 181.69$, $p < 0.001$) with the NC for each menu size differing significantly from every other one (Bonferroni-corrected Wilcoxon *T,* $p < 0.001$). Mean NC was 1.71, 3.56, 5.19 and 7.69 for 4, 6, 8 and 10 item menus respectively.

Loops only applied to the *Static-Rate* and *Mobile-Rate* conditions and there was a significant effect of mobility found on the number of overshoots (Wilcoxon $T = 39.50$, $p < 0.05$) with mobile selections producing significantly more overshoots per selection (mean = 0.046) than static selections (mean = 0.012). Menu size also had a significant effect on number of Loops (Friedman's $\chi^2(3) = 12.789$, $p < 0.01$) however, no Bonferroni-corrected Wilcoxon *T* comparisons reached the necessary adjusted p-value for significance ($p < 0.0083$). Mean Loops values were 0.01, 0.00, 0.06 and 0.04 for the 4, 6, 8 and 10-item menus, respectively.

4.2.6.4 Subjective Workload

Non-parametric Wilcoxon *T* tests showed that walking significantly increased overall subjective workload (mean = 9.26) compared to sitting (mean = 8.19) ($T = 97.5$, $p < 0.05$) and that Positional (mean = 9.91) control elicited significantly higher overall workload ($T = 22.0$, $p < 0.001$) than Rate-based control (7.54).

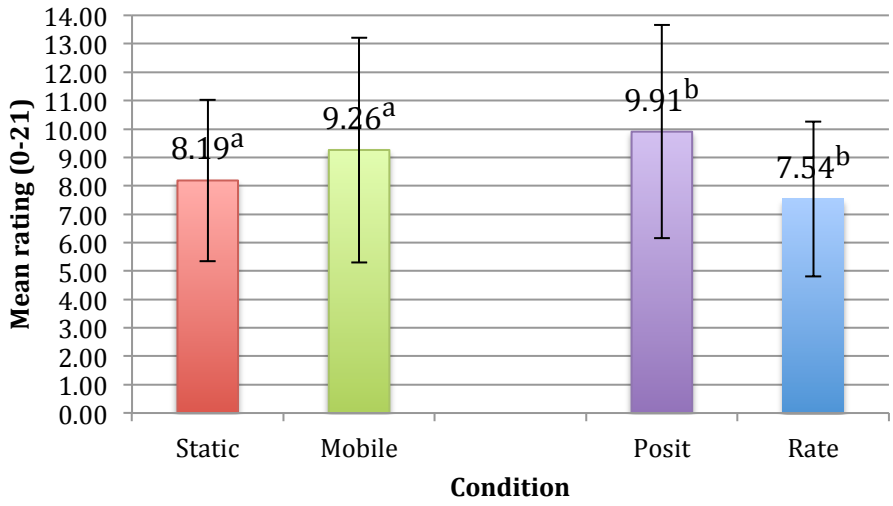

Figure 4-6: Mean overall subjective workload ratings for each condition in Experiment 2: Static, Mobile, Positional (Posit) and Rate-based (Rate). Error bars show 1 standard deviation. 'a' and 'b' indicate significant differences $p < 0.05$ and $p < 0.01$, respectively.

4.2.6.5 Control Tactics

The average number of nudges across both *Static-Rate* and *Mobile-Rate* was 0.22 nudges per selection. A Wilcoxon pairwise comparison showed a significant effect of mobility on the number of nudges (T = 1944, $p < 0.001$) with mobile selections eliciting more nudges per selection (mean = 0.31) than static selections (mean = 0.13). Friedman's Test showed that Menu Size also significantly affected the number of nudges (χ^2 (3) = 51.209, $p < 0.001$). All menu sizes differed significantly from each other ($p < 0.001$) except for 4 vs. 6 items and 8 vs. 10 items, which were not significantly different ($p > 0.05$). Mean number of nudges for each menu size was 0.08, 0.08, 0.37 and 0.35 for 4, 6, 8 and 10 item menus respectively.

4.2.7 Initial Discussion – Experiment 2

Walking had a significant impact on user performance, producing more errors and taking, on average, one second longer per selection. It also greatly increased mental/physical workload levels. Although NC and Loops are not correlate measures, the higher values produced when walking indicate a lower degree of control during mobile selections. Interestingly, mobility appears to have a smaller impact on Rate-based selection time than on Positional selection time, possibly because the influence of unintended changes in input due to bodily movement is stronger for Positional control and so the participants had more difficulty 'homing in' on targets. Walking increased average Positional selection time by 1.02s but only increased it by 0.1s under Rate-based control. Therefore use of Rate-based input may mitigate the

negative effects of mobility to a degree.

The results from Experiment 2 strongly suggested that Rate-based input allows for superior control of pressure-based linear targeting compared to Positional input for both static and mobile interaction. Although both control methods enjoyed equal accuracy, Rate-based selections were significantly faster when both sitting and walking and were rated as significantly less mentally and physically demanding. Therefore this control method appears to be better than the standard method used in linear targeting research, with mobile Rate-based selections even being faster than static Positional ones.

As has been found in many other linear targeting studies, menu size (thus target size) also significantly affected performance, with generally higher ER, MT, NC, Loops and Nudges occurring as the size of targets got smaller (as the number of menu items increased), however ER and Loops did not increase smoothly. The low number of Nudges both overall and even when walking suggests that participants did not engage in 'shunting' or searching behaviour during Rate-based control, even though mobility produced a significantly greater number. Looking at pressure profiles also shows that many users maintained a set speed from start to finish and simply lifted their thumb as soon as the cursor was in the target item.

From these results it appears that mobility negatively influences pressure-based linear targeting but that Rate-based control mitigates these effects to an extent and so is best suited to mobile interaction. Therefore the Rate-based method was chosen for use during Experiment 3, which investigated whether users were able to interact with this application using only audio feedback.

4.3 Experiment 3 – The Effects of Mobility and Feedback Method on Pressure-based Input

In Experiment 1, the use of audio feedback was suggested to represent a more expert usage scenario, as the user would need to be familiar enough with the interface to be able to use it non-visually. Therefore the same participants took part in both experiments 2 and 3, so that they would be more familiar with the interaction when tested using only audio feedback. Thirteen of the same fourteen participants (11 male, 2 female) took part in Experiment 3, as one participant was unable to take part. The second session took place 4-6 weeks after Experiment 2.

4.3.1 Audio Feedback Design

The audio design used here is almost identical to the design used in Experiment 1. However, the audio menu was changed to run from left-to-right, instead of right-to-left (as it was in Experiment 1), as this is the more common directional order for menus and text in English applications (and other Latin languages). Other than this change the audio was identical for the Positional control method.

The Rate-based audio conditions used the same audio design, only with the addition of one additional cue to indicate the speed of cursor movement. This speed cue consisted of a short, light 'tap' sound that played at increasing temporal frequency as cursor speed increased (pressure increased). This was designed to sound like the cursor was being dragged across a sawtooth surface. Pilot testing found this cue to be beneficial. A speed cue was not added to the Positional control feedback, because the speed at which the participant would hear the cursor move through the items (via labels and unique tones) would already provide this information, and the same tap/sawtooth design would not work for such variable changes in speed that occur during Positional control.

4.3.2 Experimental Design

The first condition that all participants engaged in was a walking condition using the Rate-based control with visual feedback, consisting of half as many selections as in other normal conditions. This condition was to familiarise participants with the menu layouts and cursor behaviour after the long break, so they were explicitly told to try and remember the layout/order of labels. Experiment 1 found a lack of familiarity of menu layout led to poorer performance.

After this familiarisation, the main study had three audio-only conditions, presented in a counterbalanced order: *Static-Rate-Audio* (SRA), *Mobile-Rate-Audio* (MRA) and *Mobile-Positional-Audio* (MPA). Although Positional control resulted in poorer performance and higher workload during Experiment 1, MPA was also included in this session to investigate whether the conclusions about Rate-based superiority for mobile non-visual interaction were reliable. The task was identical with every target from all four menu sizes being selected twice at random and in counterbalanced order after 10 practice selections. Audio feedback was presented to participants through stereo headphones connected to the UMPC.

Participants completed a NASA TLX workload estimation form after each condition. The same dependent variables from Experiment 2 were measured during Experiment 3. The independent variables were *Condition* (SRA, MRA and MPA) and *Menu size* (4, 6, 8 and 10 items). There were a total of: 13 participants x 3 Conditions x 28 target distances = 1092 trials. This gave 546 data points for each condition. Experimental instructions and raw data for all measures can be found in Appendix B.

4.3.3 Results – Experiment 3

Note that all *post hoc* Wilcoxon *T* tests used the Bonferroni correction on the p-value necessary for statistical significance: $p<0.05/N$, where N is the total number of comparisons.

4.3.3.1 Errors

Comparing the 3 conditions using a Friedman's Test showed a significant effect of Condition on errors (χ^2 (2) = 62.12, $p < 0.001$). Wilcoxon pairwise comparisons showed that MPA had significantly higher ER than both SRA (T = 609.0, $p < 0.001$) and MRA (T = 704.0, $p < 0.001$). Mean ER stood at 2.5% for SRA, 3.0% for MRA and 12.8% for MPA (see Figure 4-7). There was also a significant effect of Menu Size on ER (χ^2 (3) = 15.26, $p < 0.01$). Wilcoxon pairwise comparisons showed that the 10-item menu produced significantly more errors than the 6-item menu (T = 75.0, $p < 0.001$). Overall mean ER for each menu size was: 4.5% for 4 Items, 2% for 6 Items, 3.9% for 8 Items and 8.7% for 10 Items.

4.3.3.2 Movement Time

Friedman's Test showed a significant effect of Condition on MT (χ^2 (2) = 46.57, $p < 0.001$). Bonferroni-corrected Wilcoxon *T* showed that SRA had lower MT than MPA (T = 50710.50, $p < 0.001$) and MRA had lower MT than MPA (T = 59134.50, $p < 0.001$). Mean MT was 3.67s, 3.96s and 5.08s for SRA, MRA and MPA respectively (see Figure 4-8).

There was also a significant effect of Menu Size on MT (χ^2 (3) = 153.02, $p < 0.001$). Wilcoxon comparisons showed that the MT for all Menu Sizes differed significantly from each other (all $p < 0.001$). MT increased as the Menu Size increased with mean MT of 2.85s, 3.39s, 4.30s and 5.46s for 4, 6, 8 and 10 Item menus respectively.

4.3.3.3 Number of Crossings/Loops

As the number of crossings (NC) only applies to Positional control, analysis here was limited to comparing NC across Menu Sizes. A significant effect of Menu Size on NC was found (χ^2 (3) = 49.627, $p < 0.001$) with Wilcoxon comparisons showing that all sizes differed from each other significantly, apart from 4 items vs. 6 items and 8 items vs. 10 items. Mean NC for each size was 3.44, 4.96, 9.18 and 13.07 for 4, 6, 8 and 10 item menus respectively.

A Wilcoxon comparison of SRA and MRA showed a significant effect of Condition/Mobility on the number of Loops during Rate-based control (T = 1901, $p < 0.001$). Mobile selections produced more Loops (mean = 0.27) per trial than static selections (mean = 0.15). Friedman's Test also showed a significant effect of Menu Size on Loops (χ^2 (3) = 25.553, $p < 0.001$) but no Wilcoxon comparisons reached the necessary level of significance ($p < 0.0083$). Mean number of Loops per trial sat at 0.13, 0.14, 0.25 and 0.33 for 4, 6, 8 and 10 item menus respectively.

4.3.3.4 Movement/Control

Again, we recorded the number of nudges used by participants. The overall average number of nudges across both SRA and MRA was 2.45 nudges per selection. A Wilcoxon pairwise comparison showed a significant effect of Condition/Mobility on the number of nudges (T = 46446.5, $p < 0.001$) with mobile selections producing more nudges per selection (mean = 2.97) than static selections (mean = 1.93). Friedman's Test showed that Menu Size also significantly affected the number of nudges (χ^2 (3) = 49.571, $p < 0.001$), but no Wilcoxon T tests reached the Bonferroni-adjusted level of significance ($p < 0.0083$). Mean number of nudges for each menu size was 1.57, 1.81, 2.50 and 3.07 for 4, 6, 8 and 10 item menus respectively.

4.3.4 Initial Discussion – Experiment 3

Positional input using audio feedback took longer to make selections and was also significantly more error-prone than Rate-based input. This supports the outcome of Experiment 2. Comparing static and mobile audio interaction with Rate-based control showed that walking increased Movement Time as well as the number of both Nudges and Loops. This suggests that being mobile had a similar effect on both audio and visual interaction, only with a stronger negative effect on audio control. However, walking did not affect how accurately targets could be selected when audio feedback was used, as SRA and MRA only differed by 0.5% errors (2.5% and 3.0% respectively). This suggests that pressure

interaction with only audio feedback can reach almost 100% accuracy even while walking, albeit at the expense of task time. As in Experiment 2, mobility increased the number of Nudges from 1.93 to 2.97. As is expanded upon below, these numbers are much higher than Experiment 2, and the difference between them is greater as well.

4.4 Experiments 2 and 3 Compared: The Effect of Feedback

In this section the *Static-Rate* (SRV), *Mobile-Rate* (MRV) and *Mobile-Positional* (MPV) conditions using visual feedback from Experiment 2 were compared to the audio-only equivalent conditions (SRA, MRA and MPA) from Experiment 3. The Independent Variable for this comparison was *Feedback* (Visual, Audio).

4.4.1 Errors

Wilcoxon pairwise comparisons showed a significant difference between SRV and SRA (T = 57.50, $p < 0.05$) with the visual condition having lower ER (mean = 1.5%) than the audio condition (mean = 2.5%). It was also found that MPV (mean = 2.9%) had significantly lower ER than MPA (mean = 12.8%; T = 756.50, $p < 0.001$). As shown in Figure 4-7 MRV (mean = 3.2%) and MRA (mean = 3.0%) were not significantly different ($p > 0.05$).

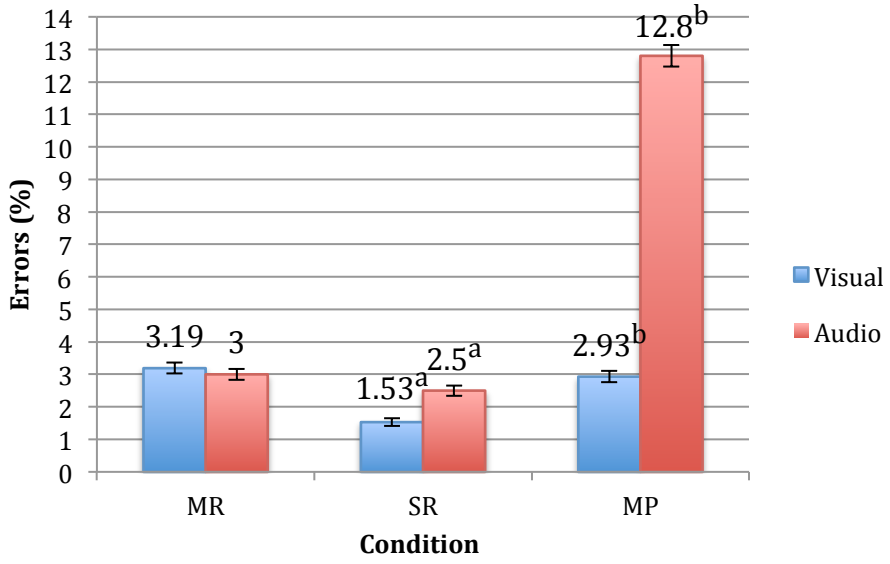

Figure 4-7: Mean error rates for Mobile-Rate, Static-Rate and Mobile-Positional conditions using Visual and Audio feedback. Error bars show 1 standard deviation. 'a' and 'b' indicate significant differences $p < 0.05$ and $p < 0.001$, respectively.

4.4.2 Movement Time

All three visual conditions in Experiment 2 were significantly faster than the audio equivalents in Experiment 3 ($p < 0.001$). Mean SRA MT was 1.44s higher than SRV; MRA was 1.62s slower than mean MRV; and MPA was 1.2s slower than MPV (see Figure 4-8).

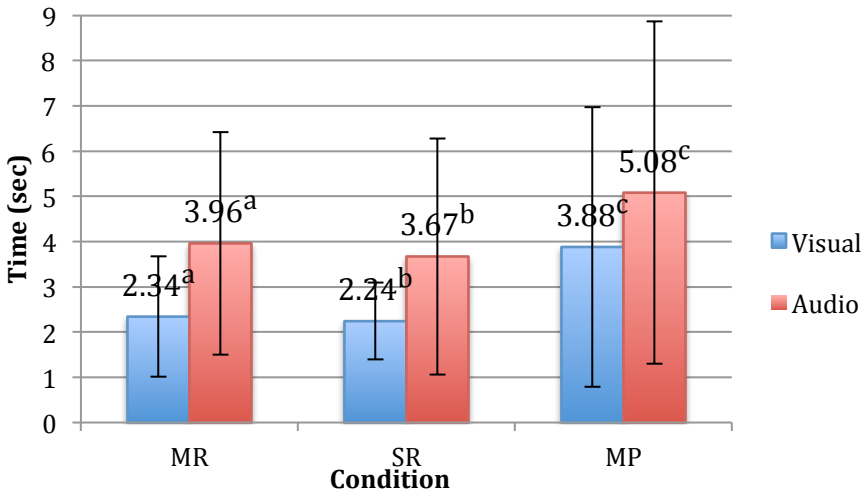

Figure 4-8: Target selection time for Mobile-Rate (MR), Static-Rate (SR) and Mobile-Positional (MP) conditions using Visual and Audio feedback. Error bars show 1 standard deviation. 'a', 'b' and 'c' indicate significant differences, $p < 0.001$.

4.4.3 Crossings/Loops

Wilcoxon comparison of MPV and MPA showed a significant effect of Feedback on NC ($T = 13028.5$, $p < 0.05$) with Audio (MPA) selections producing more crossings (mean = 7.63) than Visual (MPV) selections (mean = 6.25). Wilcoxon comparisons also showed a significant effect of Feedback for both Static (SRV vs. SRA; $T = 194.5$, $p < 0.001$) and Mobile (MRV vs. MRA; $T = 901.5$, $p < 0.001$) selections. In both cases audio selections produced more Loops/overshoots than visual selections.

4.4.4 Movement/Control

Feedback had a significant effect on the number of Nudges for both Static (SRV vs. SRA; $T = 735$, $p < 0.001$) and Mobile (MRV vs. MRA; $T = 2164.5$, $p < 0.001$) selections. For both

conditions more nudges were used during the audio selections.

4.4.5 Initial Discussion – Experiment 2 and 3 Compared

For almost all measures, performance using visual feedback was better than when using audio feedback for all interaction conditions (Mobile, Static, Positional and Rate-based). Therefore audio selections took longer and were more difficult to control. This is perhaps to be expected but there are several interesting results to point out. Firstly, feedback did not affect accuracy while mobile and using Rate-based control. Mobile control with visual feedback (ER = 3.2%) was similarly accurate to mobile control with audio feedback (ER = 3.0%). Secondly, feedback had a much stronger effect on mobile Positional control than Rate-based control, as Positional ER more than quadrupled between visual (2.9%) and audio (12.8%) conditions.

The number of Nudges was vastly different for visual and audio conditions, increasing from a mean of 0.22 Nudges per selection during visual conditions up to a mean of 2.45 Nudges per selection during audio conditions. Therefore it seems as though participants engaged in searching 'shunt' behaviour much more when only audio feedback was provided, rather than continuous, dynamic rate-based control. This may be because they still were not familiar enough with the order and layout of items after the familiarisation condition at the start of Experiment 3. It may also be that they were less confident of their control over the cursor when only audio feedback was available, as the high MT and Loops, combined with low ER, suggests they had lower levels of control and so found it more difficult to correctly acquire the target item quickly.

4.5 Limitations

This section discusses the main limitations of the research reported in this chapter, which should be considered when interpreting the results. The main limitations of Experiments 2 and 3 were: 1) lack of realistic walking route and 2) use of large device.

4.5.1 Lack of Realism in Walking Route

The walking route used for the mobile conditions was based on one used in previous

research [38] and other research has suggested that set walking routes indoors are an acceptably close approximation of the real demands and influences of walking in more realistic settings [4]. Due to ethical considerations, it was not possible to conduct the study outside, nor to introduce obstacles into the course used in Experiments 2 and 3. Therefore, the walking route was simple and required relatively low visual attention from the participants. However, the bodily motion produced by walking was realistic and so the influences of that motion on pressure-based input have been validly tested.

4.5.2 Use of a Large Device

Technological constraints at the time forced the use of the Samsung UMPC for Experiments 2 and 3. While it is a mobile device, in that it is self-contained and can be held in the hands, it is large compared to common mobile devices such as phones and MP3 players. It was also heavier and so had to be held in two hands, with one hand also using the thumb to operate the sensor. This may have had one of two contrasting influences on performance: the extra weight and bulk may have made holding the device and controlling the amount of pressure applied to it more challenging; or the extra weight, and two-handed grip, may have made the device more stable, allowing for more precise input. Without subjective reports from the participants (see Section 8.8.1) it is impossible to know which, if either, was true. However, the interaction dynamics are likely to be different compared to a small device held in only one hand. Because of this, is was necessary to use a small, commercial mobile phone, held in one hand, for Experiments 4 and 5 in Chapter 5.

4.6 Discussion and Conclusions

4.6.1 Research Question 1 – The Use of Audio Feedback

Research Question 1 asked:

"How accurate is pressure-based input on a mobile device when using only audio feedback?"

Experiment 1 had positive results for interaction using only audio feedback when sitting, with relatively high accuracy (>= 85%), but long selection times and less precise input. In

experiments 2 and 3, control using only audio feedback was more error prone (average of 6.1% errors) than using visual feedback (average of 2.5%), however the error rates were much lower than during Experiment 1, even when selecting thinner pressure levels. There are two possible reasons for the large difference in performance between Experiment 1 and Experiments 2 and 3. The first was mentioned in Section 3.5.5, that only a third of the previous sensor's range was used, which may have introduced noise into the sensor behaviour, potentially leading to more erratic cursor motion. This would make targets more difficult to target precisely. The second possible reason may be the slight difference in visual gain between the visual feedback shown on the Nokia N810 compared to the Samsung Q1. The devices have the same resolution of 800 x 480, but the Samsung screen measures 7 inches diagonally, while the N810 measured 4.3 inches. This would provide a higher vertical resolution of millimetres per Newton on the Samsung, and increasing the gain of visual feedback has been shown to improve control of applied pressure [108, 177, 214]. Unfortunately, however, the vertical resolution in past research has been reported in pixels rather than millimetres. Therefore, while both devices used in experiments 1 to 3 had a vertical resolution of 100 p/N (pixels per Newton), close to the 128 p/N that produce good control of pressure, it is not known what physical dimensions these recommended 128 pixels inhabited, and so how the 100 p/N relate to that.

Previous research had only made use of 3 pressure levels in non-visual interactions, and they resulted in considerably lower accuracy than was found for up to 10 levels in Experiments 1 to 3 [228, 232, 233], even when using linearised sensor output [228]. The control method used for input significantly impacted performance. The highest number of errors from Experiments 2 and 3 was only 12.8%, and that was for Positional control while walking and using audio feedback. Errors for Rate-based control peaked at only 3.2%, while sitting and while walking was only 3.0%. These error rates were achieved when the user was walking and are considerably less than those from Experiment 1 when the user was sitting. They are also comparable or lower to those from other studies using seated visual interactions [30, 170, 209]. Therefore, purely in terms of targeting accuracy, mobile and non-visual interaction from Experiments 2 and 3 matched or out-performed previous pressure interfaces which were static and using visual feedback. The amount of time needed to complete a selection increased when using audio feedback and remained higher than other research, and so it seems that, while non-visual accuracy remains high, it is at the cost of speed.

Given the results from Experiments 1 to 3, the answer to Research Questions 1 is that pressure-based interaction on a mobile device is highly accurate when using only audio feedback, when the user is both sitting and walking.

4.6.2 Research Question 2 – The Influence of Walking

Research Question 2 asked:

"How accurate is pressure-based input when the individual is walking?"

There are two aspects to this question: 1) how accurate is the control of applied pressure when walking and 2) how accurate is pressure-based interaction when walking. Both have been answered by Experiments 1 to 3, but in different places. Positional input is a direct measure of the pressure applied to the sensor, and so the participants' ability to accurately select target pressure levels using this control method answer aspect 1. While Positional control can also be used as a means of interacting with a pressure-based interface, the Rate-based control method is superior for linear targeting/menu-based interactions. Therefore the results from this control method can better answer aspect 2.

Regarding aspect 1, walking negatively affected the participants' ability to accurately control the amount of pressure applied to the device. Being mobile significantly increased the number of Errors, selection time and Number of Crossings, suggesting input was more variable and led to more unintended incorrect selections. While the error rate increased 64% when walking (compared to sitting), it only increased to 2.93%, when presented with visual feedback, which is still very low. However, the time required to make a selection increased by 1.02 seconds (to 3.88s) when walking and the number of target crossings increased by 80%, which suggest that maintaining that high accuracy was challenging. The combination of walking and using audio feedback had a stronger negative impact on precision of input, as errors greatly increased to 12.8%, more than 4x the errors when walking with visual feedback, and selection time increased by a further 1.2 seconds. In partial answer to Research Question 2, the application of pressure is significantly less precise when walking, both in terms of target selection as well as variability of input. Participants clearly find it challenging to maintain very precise levels of pressure when in motion, especially when only audio feedback is available.

Regarding aspect 2, walking significantly decreased overall accuracy during Rate-based control, but, interestingly, only when using visual feedback. It had no effect on overall accuracy when only audio feedback was used: static and mobile non-visual Rate-based interaction was equally accurate. Selection time increased and overall control, measured in Loops and Nudges, degraded under both visual and audio Rate-based walking

conditions, however. Mobile interaction was also significantly more mentally/physically demanding. This effect of mobility on accuracy for visual interaction but not audio interaction is intriguing, but may be due to the movement of the device or, more importantly, its screen. If it was solely the negative influence of walking-induced motion, then the effect should arise for both feedback conditions. During static-visual selections the screen of the device is stationary so cursor movement can be tracked easily. During mobile-visual selections, however, the screen of the device is moving, as is the head of the participant, making it potentially harder to track the movement of the cursor. Audio feedback should not have been affected in the same way by bodily motion.

In answer to Research Question 2, accuracy of pressure-based input was very good, at only 3% errors, and selection time was also no longer compared to sitting. Using audio feedback did not make mobile accuracy any worse, but it did make it seemingly more difficult, as selection time increased by 1.6 seconds.

The contributions of the research in this chapter are:

1. Walking significantly degrades control of applied pressure but linear targeting performance across a wide range of pressure input remains as good as in previous research that tested control when sitting in desktop interactions scenarios.
2. Non-visual control of a wide range of pressure, using only spatialised audio feedback, can be highly accurate when walking.

4.6.3 Context and Limitations

An explanation for the poorer precision of input during Positional control may come from the support and stability afforded by the apparatus. The lowest ER, MT and NC from previous research were achieved via either desktop stylus input or FSRs attached to a computer mouse. A stylus grip generally consists of a thumb and two fingers providing opposing and stabilizing forces. Part of the hand and arm are also resting on the table. Those studies using a mouse for linear targeting did not require any x-y movement of the mouse for the interaction so it could remain stationary. Again, the hand and arm would be resting on the table providing stability, with multiple fingers gripping the mouse. These factors provide more stable interactions than in experiments 2 and 3, where the user applies pressure through the thumb or through a thumb-finger pinch while the hands also hold up the device. Although the participant could rest their arms on their knees or the table while sitting, the wrists were unsupported and there was no extra support when mobile. Stewart *et al.* [227] have subsequently found that walking introduces unintended increases and variations in

pressure applied to a mobile device, and this was also found by Crossan *et al.* [38] during head-tilt based linear targeting. This excess movement may have lead to a higher number of Crossings and Loops, which were correlated with higher MT.

Because the way in which pressure is applied to a device, including the choice and number of digits used, could have a potential influence on control, the experiments reported in the next chapter investigated the use of multiple digits for pressure-based input on mobile devices.

5 Multi-Digit Pressure Input on a Mobile Device

5.1 Introduction

Experiments 2 and 3 in Chapter 4 established that pressure-based input could be highly accurate when applied to a mobile device while the user was walking and provided with either visual or audio feedback. However, there were three important limitations of these studies:

1. The device used was large
2. It required two-hands to carry and operate
3. Only one digit/pressure sensor was used for input

The device used in the experiments in Chapter 4 was an Ultra Mobile PC (UMPC), measuring 228 x 139 x 25 mm and weighing approximately 800 grams. The participants had to hold the device in both hands (one at either side) and press with a single thumb on one Force-Sensing Resistor (FSR) attached to the front of the UMPC. Most mobile devices, such as phones and mp3 players, are considerably smaller, such as the popular Apple iPhone, which measures 115 x 58.6 x 9.3 mm and weighs 140g. These kinds of device are small and light enough to be held and interacted with one-handed and, because all 5 digits of the hand may be in contact with the device at one time, there is an opportunity for pressure-based input to either come from different digits/locations at different times, or from multiple different digits simultaneously. Further, by providing multiple pressure inputs, more complex mobile interactions may be controlled with a single hand.

Most HCI research on pressure-based input using FSRs has used either the thumb or index finger (or gripping/pinching with both) for input [14, 30, 162, 228], with some research also using the middle finger [30, 210]. Tang and colleagues studied the formation of three-digit pressure "chords", with each of the first three fingers applying one of three levels of pressure [232, 233]. Each chord required 2 to 2.5 seconds to form, with an error rate of just 13-14%, but only three distinct levels of pressure were used. Psychophysical research suggests that some digits are better able to apply pressure than others [138] and the number of digits used

to apply pressure influences the accuracy with which pressure is applied [177]. To judge if alternative digits or multiple digits in combination can be used for input on a mobile device, it was necessary to understand how well each digit can control applied pressure, both in isolation, as well as in combination with other digits. Because the aim was to provide input on smaller mobile devices, control had to be judged when holding and pressing on a mobile device form-factor. Therefore, Research Question 3 asked:

"How accurate is pressure-based input when multiple digits apply pressure to a mobile device?"

If multiple different digits can provide accurate input then they could be used to provide multiple different inputs to the system and facilitate more complex interactions than are available when using only a single sensor. Sections 5.3 and 5.4 describe two experiments that tested control of applied pressure from each individual digit of the right hand, as well as in a number of combinations. 7 FSRs were attached to the sides, back and top of a smartphone and the participant held and squeezed the device with one hand. After identifying which digits, and combinations thereof, provided the most accurate input, one-handed pressure-based input was compared with two-handed multitouch gestures in a zooming and rotating map application, to judge if one-handed, multi-digit pressure input could potentially substitute for two-handed touch, leaving the second hand free.

A final consideration was the size of the pressure-space. Newell & McDonald [177] found that increasing the number of digits (or "degrees of freedom"), used to synchronously squeeze on a dynamometer, altered the precision of pressure output. Specifically, precision at low levels of pressure degraded as more digits were added but precision at higher pressure levels improved. This means that our ability to control pressure depends on both 1) how many fingers we are using and 2) the level of pressure being applied. This presents something of a double-edged sword for pressure-based interactions. It suggests that we may need to tailor the size of the pressure space of an interaction to the number of digits used, so that the levels of pressure required suit the number of digits. For example, if a user pinches his/her phone with a thumb and forefinger, he/she will be able to precisely apply lower pressure levels than if the phone was squeezed with the whole hand. However, it may also mean that interactions can exploit larger spaces, or varying sizes of space, with the addition (or subtraction) of digits for different purposes. For example, squeezing with two fingers could move through a small list of recent contacts, but squeezing with the whole hand could move through a larger list of contacts. Therefore, Experiment 4 varied the size of the pressure space and the total amount of pressure involved in the interaction, to judge how it

influenced precision of control.

5.2 Sensor Positions: Choice and Rationale

5.2.1 Finger Positions

The aim was to investigate not only the use of multiple digits and varying grips on precision of pressure application on a mobile phone but also the use of digits not commonly used in pressure input. Therefore, the sensors needed to be placed in locations around the device that were easily reachable without repositioning the hand or device. They also needed to be in positions that could provide opposing forces so that the phone could be held and squeezed freely with the same hand: one hand needed to be able to hold the device, interact with the sensors and provide opposing forces. Therefore, the most logical positions for the sensors were locations around the device near to where users naturally place fingers when holding a mobile touchscreen phone. A brief survey of these holding grips was conducted with users of touchscreen devices around the University. Users were asked if they used their phone one-handed and, if they did, they were asked to hold the device as they would when interacting with it (in the hand they would naturally use). The most common form of grip/holding pattern is shown to the left of Figure 5-1, in the right-handed variant. By placing the sensors around the body of the device, away from the screen, the visual content is fully visible and not obscured by "fat fingers".

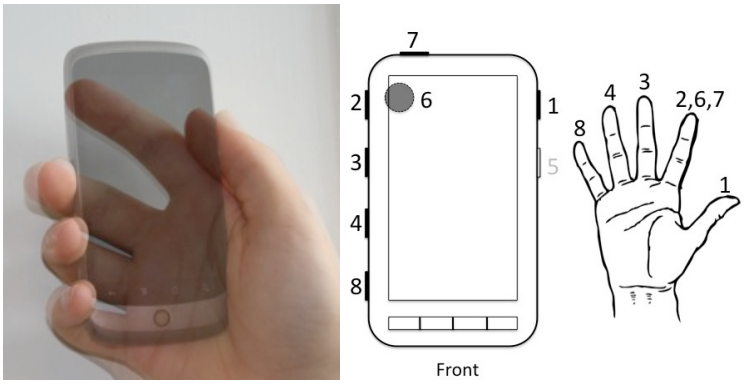

Figure 5-1: Common one-handed touchscreen device grip (left) and the sensor locations used for Experiment 4 (right).

The middle (<M>), ring (<R>) and little (<L>) finger reach round and clasp the lower left-hand side of the device, pushing it against the palm for grip and stability, leaving the thumb

(<T>) free to interact with the screen. The index finger (<I>) rests along the back of the device, providing further balance. From this common grip, the numbered locations shown to the right of Figure 5-1 were chosen for the pressure sensors (the apparatus itself is shown in Figure 5-2). The numbers above each digit of the hand indicate which sensor/s that digit pressed. It was decided that the sensor locations would remain the same for each participant, rather than re-positioning them based on where each individual participant may rest their fingers. This was done because detaching and re-attaching the sensors to suit each individual would be time-consuming and problematic, as the regular repositioning could potentially damage the sensors. However, the disadvantage of this approach is that the positioning may be sub-optimal for some participants.

Sensors numbered 3, 4 and 8 (in Figure 5-1) were in positions similar to those of <M>, <R> and <L> (respectively) in the common grip, with these forces being opposed by the palm. Sensors 1 and 5 were placed as alternatives for <T> (thumb), however sensor 5 was not used in the limited number of grips chosen for the study (see Section 5.2.2). Sensors 2, 6 and 7 are alternatives for <I>. Sensor 2 allowed for input from a five-digit grip along the same plane, along with sensors 1, 3, 4 and 8. Sensor 6 was near the resting point of <I> in the common grip (see Figure 5-1, left), providing input from the back of the device. Although some research has looked at touch input from the back of mobile devices [242, 254] pressure input from the back is still relatively novel [228]. Finally, sensor 7 introduced input from a second novel position, namely the top of the device. These locations allowed for pressure input along three different dimensions.

5.2.2 Grips

Due to time constraints it was not practical to test all 31 possible combinations of fingers, and so 14 grip configurations, including each digit individually, were tested to keep the task time manageable but also provide a good range of grips. The 14 grips are described in Table 5-1 in terms of the sensors and digits used in the grip, with the sensor numbers corresponding to those in Figure 5-1. They are referred to here by their grip ID number. Grips G1-G6 gave an indication of how precisely each individual digit can apply pressure. As mentioned above, sensors 6 and 7 (in G4 and G5 respectively) introduce pressure input from the back and top of the device, positions that are not commonly used for input, and involve pressing along different axes than the sensors down the sides of the device. These gave an indication of how precise pressure application is when applied from different orientations.

Grip ID	Digits Used	Sensors Used	Grip ID	Digits Used	Sensors Used
G1	Thumb(T)	1	**G8**	I, L	2, 8
G2	Middle(M)	3	**G9**	R, L	4, 8
G3	Ring(R)	4	**G10**	T, I, M	1, 2, 3
G4	Index(I)	6	**G11**	M, R, L	3, 4, 8
G5	Index(I)	7	**G12**	T, I, M, R	1, 2, 3, 4
G6	Little(L)	8	**G13**	I, M, R, L	2, 3, 4, 8
G7	I, M	2, 3	**G14**	T, I, M, R, L	1, 2, 3, 4, 8

Table 5-1: Grip configurations used in the evaluation, described in terms of the fingers and sensors used.

From the ten possible combinations of two-digit grips, G7-G9 were selected. Combinations using the thumb and one other digit were avoided for particular reasons: <T> + <I> is a very similar grip to <T> or <I> individually, as one digit opposes the other. Pressing with <T> and either <R> or <L> results in slight rotation of the device, as <T> pushes the top of the device to the left (for a right-handed grip) and <R>/<L> pushes the bottom to the right. In this case the sensors to be pressed are being pushed away from the digits pressing on them, making control more difficult. From the remaining possible choices, G7 and G9 provide grips using more adept (<I> + <M>) and less adept (<R> + <L>) fingers respectively, and G8 uses a combination of adept (<I>) and less adept (<L>) fingers.

The thumb is the most precise in matching target pressures compared to other digits [138]. Therefore, grips that make use of the thumb for active control of targeting may be more precise than those that do not. Therefore, of the ten possible three-digit combinations, one grip with <T> (G10) and one without (G11) were chosen to see if this is the case. G11 also provides input from the grasping fingers in the common touchscreen grip. For the same reason one four-digit grip (G12) used <T> and the other (G13) did not. Although alternative combinations of five-digit grips are possible (for example using sensors 6 or 7 instead of 2) G14 used only those sensors opposing along the same plane. Using sensors 6 or 7 would introduce tangential force into the grip (pressure pushing perpendicular to the other fingers/thumb), which could result in compensatory increases in the normal, gripping, pressure [116]. Therefore introducing tangential force may cause unintentionally higher pressure output along the normal plane.

Each sensor was only pressed on by the specific single digit enumerated in Figure 5-1. Also, for each grip, the whole hand was in contact with the phone, but only those sensors listed in Table 5-1 for the relevant grip provided input to the experimental software.

5.3 Experiment 4 – The Effect of Grip and Pressure Space on Precision of Pressure Applied to a Mobile Phone

This experiment tested control of pressure applied to a mobile device from the various grips described in Section 5.2.2, when the size of the underlying pressure space (range of pressure) was varied. The experimental task was essentially identical to that used in Experiments 1 to 3, but with certain aesthetic changes imposed by the use of a smaller device with a different operating system (Google Android).

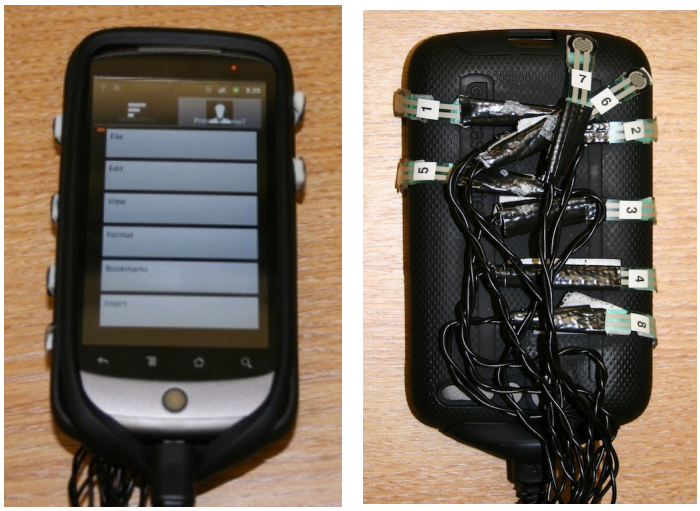

Figure 5-2: Nexus One phone encased in a Tough Case (left) and sensor positions around the device (right).

5.3.1 Apparatus

Pressure input was taken from seven Force-Sensing Resistors (Interlink Electronics model 400FSR with sensor pad diameter of 5.1mm). These were connected to two SAMH Engineering SK7-ExtGPIO01 I/O modules for analogue-to-digital conversion and sensor output linearization [228], four sensors per board. The two I/O modules were then connected to a MacBook Pro via USB for signal processing, which forwarded the sensor output over USB to an HTC Nexus One Android mobile phone (see Figure 5-2, left) to present the application GUI. The sensors were attached to a CaseMate© Tough Case (www.case-mate.com) by double-sided adhesive pads. A case was used because the FSR's response is less variable when pressed against a flat surface. The Tough Case was chosen specifically

over other cases because it was more rigid, with flat edges, providing better opposition to pressure exerted on it. The FSRs were attached in the configuration shown in Figure 5-2, right, for right-handed input.

5.3.2 Pressure Spaces

Two different pressure spaces were chosen for the study, referred to here as Fixed and Incremental. The input from each individual digit was measured and recorded separately (to judge the relative contribution of each digit), but the input from the sensors to the software was cumulative, so the total amount of pressure applied across all the sensors is taken as input. For example applying 1 N to each of sensors 1, 2 and 3 (in G10) gives a total input to the system of 3 N.

5.3.2.1 Fixed Pressure Space

Previous research on pressure input on mobile devices has used a pressure space of approximately 1.5-4 N when one point of pressure input was used (one digit or one sensor), and the good performance from Experiments 1 to 3 in Chapters 3 and 4 suggested that 3.5 N provided a good range, without inducing fatigue. Therefore, the Fixed pressure space size was set at approximately 3.5 N: regardless of which fingers, or how many fingers, were used in a given grip, the maximum pressure detected in the task was always 3.5 N. The performance results from this condition would provide comparisons between grips/fingers as well as comparing back to previous research using the same pressure space to better understand the effects of number of fingers.

5.3.2.2 Incremental Pressure Space

If the use of multiple digits shifts precision from lower levels of pressure to higher levels [177] then increasing the size of the pressure space may improve control. As 3.5 N is suitable for use when one digit is in use, the Incremental pressure space increased by 3.5 N with the addition of each digit, meaning 7 N for two digits, 10.5 N for three, 14 N for four and 17.5 N for all five digits.

5.3.3 Experimental Task

The experimental task was a linear targeting task very similar to that used in Experiments 1 to 3, however the pressure space was only ever divided into 6 levels or menu items.

Requiring participants to select each target from 4-10 levels, at 2 pressure spaces across 14 grips would have led to unreasonable task time and potential fatigue and boredom. 6 levels were chosen (over 4, 8 or 10) so that accurate performance might be easier than using 8 or 10 levels, but not so easy as when using 4 to introduce ceiling effects. The 6 levels were again visualized onscreen as a vertical menu of 6 items running from top-to-bottom, measuring 465 x 600 pixels (45 x 63mm; see Figure 5-3, left). Each menu item had the same labels as the previous experiments: *File, Edit, View, Format, Bookmarks* and *Insert*. Positional control was used again for this experiment, as it is necessary for measuring input precision. The position of an onscreen cursor, displayed to the left of the menu, indicated the total level of pressure being applied (see Figure 5-3). Each trial involved selecting a single target item by applying a target level of pressure and the 1-second Dwell selection technique was used.

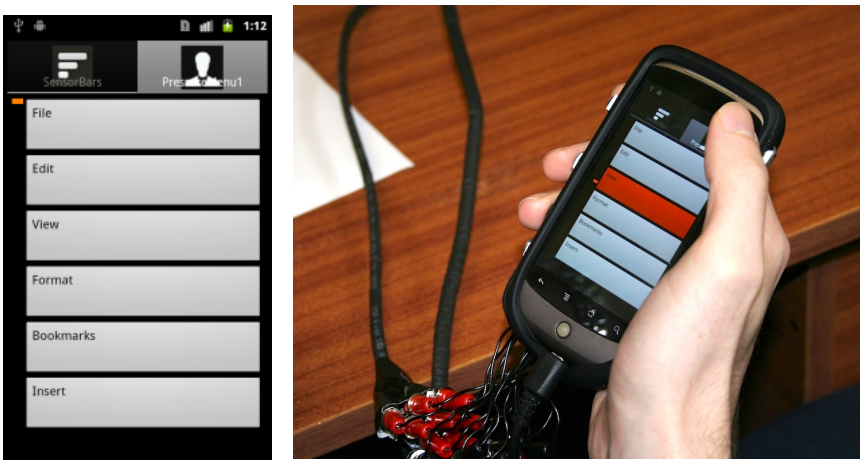

Figure 5-3: Experimental software showing target menu items (left) and participant using the apparatus (right).

5.3.4 Participants and Experimental Procedure

Thirteen participants (6 male, 7 female) aged between 21 and 63 (mean 29.18) took part, all from within the University. None had taken part in Experiments 1 to 3. Due to the positioning of the sensors, all participants were required to be right-handed and each was paid £10 for participation, which took approximately 90 minutes.

The study was a 2 x 14 (Pressure Space x Grip) within-subjects design, where all participants performed all grips under both the Fixed and Incremental pressure spaces. The order of both Grip and Pressure Space was varied via Latin square to avoid ordering effects. The experiment was divided up by Pressure Space: all grips were done within one Pressure Space

size before moving on to the next Pressure Space. Within each grip condition every menu item was selected twice in a random order. At the start of each trial the item to be selected was highlighted in green for one second before returning to the common grey. Each Pressure Space was started with 6 practice selections and ended with participants completing a NASA TLX workload estimation form. There were a total of: 13 participants x 2 Pressure Space x 14 Grips x 6 target distances x 2 selections = 4368 trials. This gave 2184 data points for each Pressure Space, 312 data points for each Grip and 156 data points for each Pressure Space + Grip combination condition (e.g., Fixed + Grip G1).

For the entire session, participants sat in a padded office chair holding the apparatus in their right hand. They were allowed to rest their arm on either the desk in front of them or on their lap; however, the wrist and hand remained unsupported. They were able to put the phone down in between Pressure Spaces, to allow them to rest. Experimental instructions can be found in Appendix C.

5.3.4.1 Variables and Measures

There were two Independent Variables: *Grip* (14 variations) and *Pressure Space* (Fixed and Incremental). Dependent Variables were the same as Experiments 1-3: *Errors* (ER, % of missed targets)*, Movement Time* (MT, from first non-0 pressure reading to target selection including 1-second Dwell time), *Number of Crossings* (NC, number of times cursor crosses target boundary) and *Subjective Workload* ratings, via NASA TLX.

5.3.5 Hypotheses

H1: The Incremental pressure space will result in lower ER, MT, NC and Workload than the Fixed pressure space.
H2: There will be differences in how accurately each digit can apply pressure.

5.3.6 Results – Experiment 4

5.3.6.1 Notes on the Analysis

The data for ER, MT and NC did not fit a normal distribution (via Shapiro-Wilk test) and so normally non-parametric analyses would be necessary. However, these do not allow for *post hoc* comparisons or the testing of interaction effects in multi-factorial analyses. Therefore, the advice of Wobbrock *et al.* [253] was followed and their Aligned Rank Transform (ART) was used to reformat the data for use in traditional factorial analysis. For all three measures a

2 x 14 Mixed Model REML (Restricted Maximum Likelihood) analysis was carried out, with 'participant' as a random factor. It should be noted that the analysis uses ranked data, and the non-normal data distribution means that measures of central tendency are less informative and representative than during parametric analyses, but means are presented here for illustration. TLX data were normally distributed. Raw data for all measures can be found in Appendix C.

5.3.6.2 Pressure Space

The analysis found no effect of pressure space on ER ($F_{(1,155)} = 3.162$, $p > 0.05$), with means of 5.8% for the Fixed space and 4.7% for the Incremental space. There was a significant effect of pressure space on movement time (MT). The Incremental space allowed for significantly faster selections (mean = 2.30s) than the Fixed space (mean = 2.76s; $F_{(1,155)} = 61.379$, $p < 0.001$). Mean values for ER and MT are shown in Figure 5-4. A significant effect of pressure space was also seen for measures of crossings (NC). The Incremental space resulted in significantly fewer crossings per selection (mean = 1.97) than the Fixed space (mean = 3.24; $F_{(1,155)} = 57.773$, $p < 0.001$). There was no effect of pressure space on overall subjective workload ratings using a non-parametric Wilcoxon *T* test. The Fixed Pressure Space produced a mean Overall Workload rating of 8.22, while the Incremental Pressure Space produced a mean rating of 6.85. While MT and NC were significantly lower for the Incremental pressure space, and Incremental ER was lower than Fixed, the results still call for a rejection of hypothesis H1, as the difference in ER was not significant and there was no significant difference in Workload either.

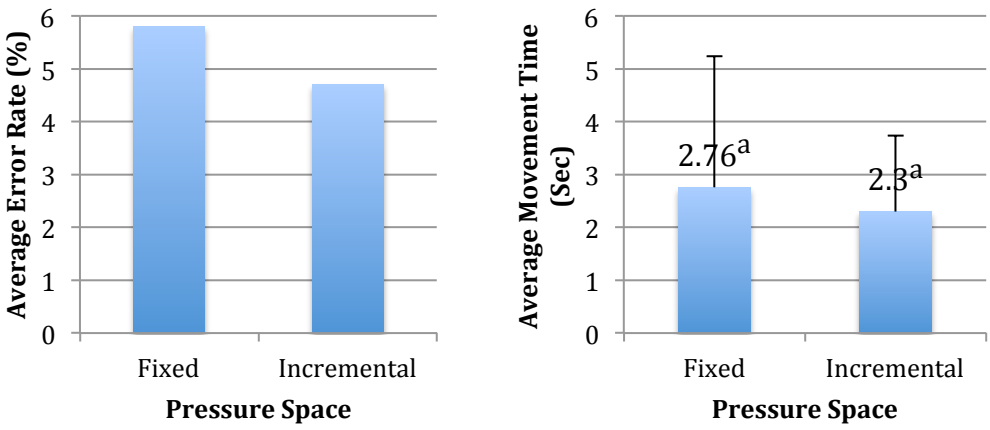

Figure 5-4: Mean Errors (left) and Movement Time (right) for both pressure spaces compared in Experiment 4. Error bars show 1 standard deviation. 'a' indicates significant difference $p < 0.001$.

5.3.6.3 Grip

Mean ER, MT and NC values for each grip are shown in Figures 5-5 to 5-7. The Mixed Model analysis showed a significant effect of grip on ER ($F_{(13,2015)}$ = 2.775, p=0.001). Bonferroni corrected pairwise comparisons indicated a significant difference between G4 and each of G3 ($p < 0.01$), G7 ($p < 0.05$), G9 ($p < 0.01$) and G10 ($p < 0.05$). In all cases G4 (<I> from back) had higher ER than the other grips. There were no other significant differences.

A significant effect of grip was also found on MT ($F_{(13,2015)}$ = 3.651, $p < 0.001$). Bonferroni pairwise comparisons of grip showed that G4 was significantly faster than each of G6 ($p < 0.05$), G10 ($p < 0.05$), G13 ($p < 0.01$) and G14 ($p < 0.05$). G5 was significantly faster than G10 ($p < 0.05$), G13 ($p < 0.01$) and G14 ($p < 0.05$). G9 was significantly faster than G13 ($p < 0.05$).

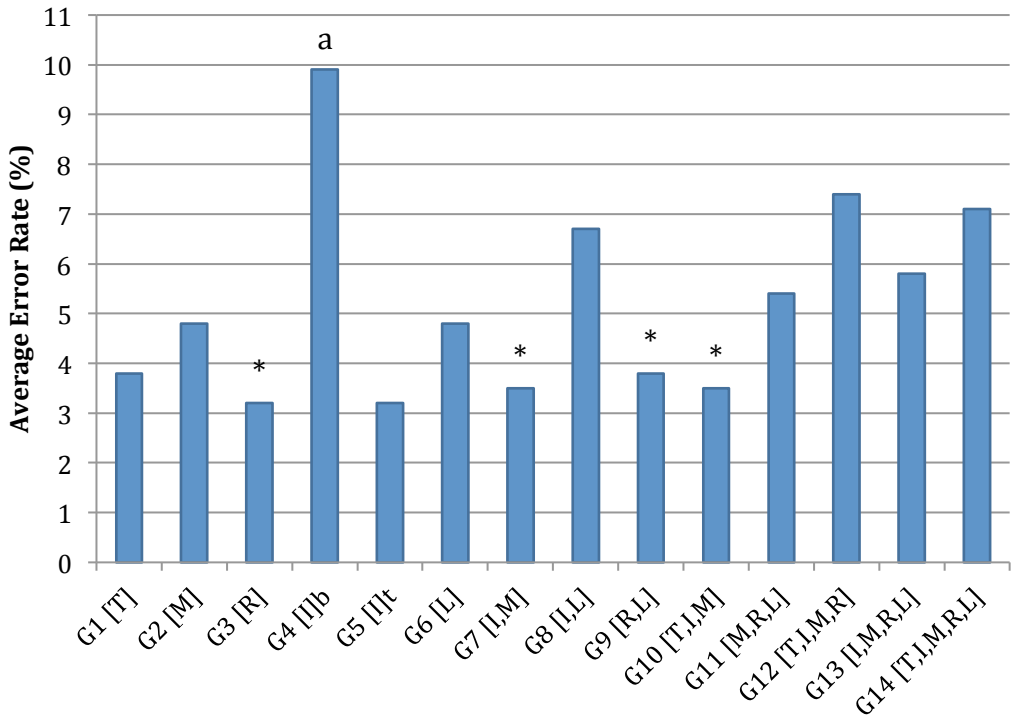

Figure 5-5: Mean Errors (ER) for all Grips used in Experiment 4. Digits: T=Thumb, I=Index (b=back position, t=top position), M=Middle, R=Ring and L=Little. 'a' = significantly higher than G3, G7, G9 & G10 (marked with *).

NC was significantly affected by Grip ($F_{(13,2015)}$ = 10.627, $p < 0.001$). Bonferroni comparisons showed that G4 had significantly fewer crossings than all other grips (p <=

0.05) except for G5 and G11, from which G4 was not significantly different. Similarly G5 also had significantly fewer crossings than all other grips ($p <= 0.01$) other than G4, G9 and G11. G11 had significantly fewer crossings than G6 ($p < 0.05$) and G10 ($p < 0.01$). G9 had significantly fewer crossings than G10 ($p < 0.01$). Because of these results, hypothesis H2 was accepted.

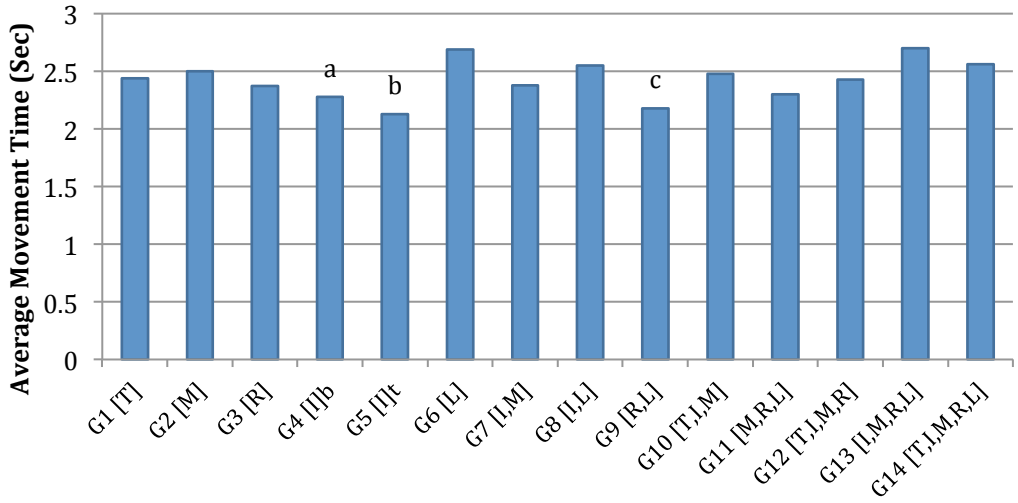

Figure 5-6: Mean Movement Time (MT) per trial for each Grip in Experiment 4. Digits: T=Thumb, I=Index (b=back position, t=top position), M=Middle, R=Ring and L=Little. 'a' = significantly less than G6, G10, G13 & G14; 'b' = significantly less than G10, G13 & G14; 'c' = significantly less than G13.

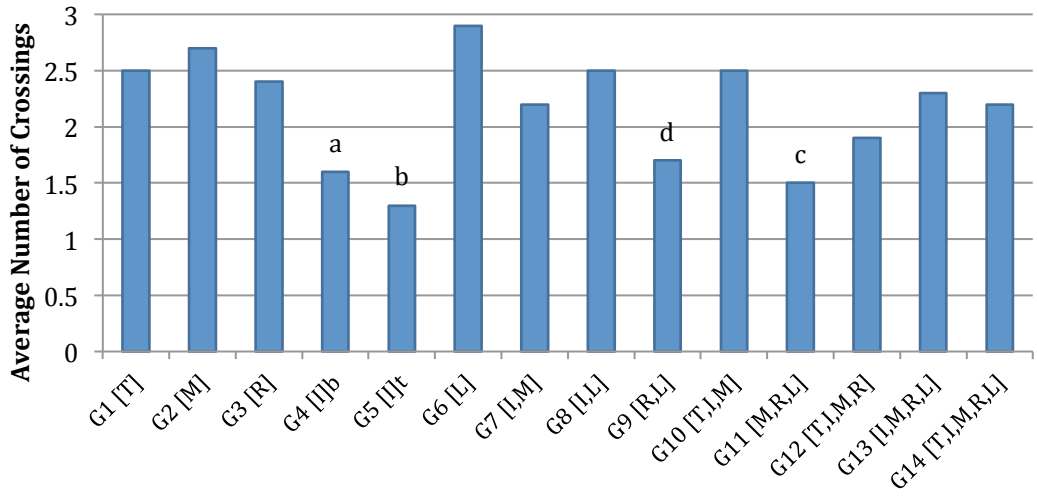

Figure 5-7: Mean Number of Crossings (NC) per trial for each Grip in Experiment 4. Digits: T=Thumb, I=Index (b=back position, t=top position), M=Middle, R=Ring and L=Little. 'a' = significantly less than all Grips, except G5 & G11; 'b' = significantly less than all Grips except G4, G9 & G11; 'c' = significantly less than G6 & G10; 'd' = significantly less than G10.

5.3.6.4 Pressure Space * Grip

The Mixed Model analysis found a significant interaction between pressure space and grip for all three measures: ER ($F_{(13,2015)} = 23.887$, $p < 0.001$), MT ($F_{(13,2015)} = 5.627$, $p < 0.001$) and NC ($F_{(13,2015)} = 57.773$, $p < 0.001$). The trends for both MT and NC are clear, with both measures generally decreasing as the number of digits increase under the Incremental space, while they both increase with more digits under the Fixed space (see Figures 5-9 and 5-10). The ER interaction is less clear, although ER generally seems to increase with more digits under the Fixed space while it remains fairly constant across grips under the Incremental space (although certain 1-digit (Grips G1 and G2) and 3+ digit grips (G11, G12) appear worse; see Figures 5-5 to 5-7).

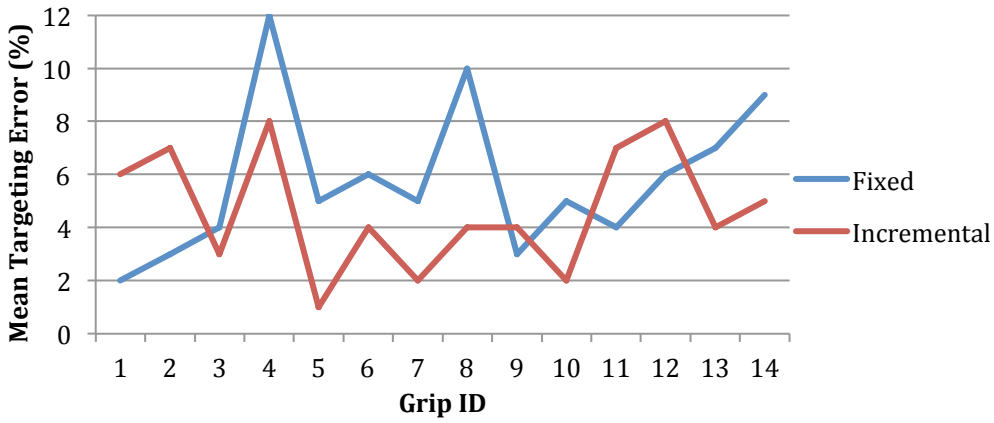

Figure 5-8: Interaction between Grip and Pressure Space on Error (ER).

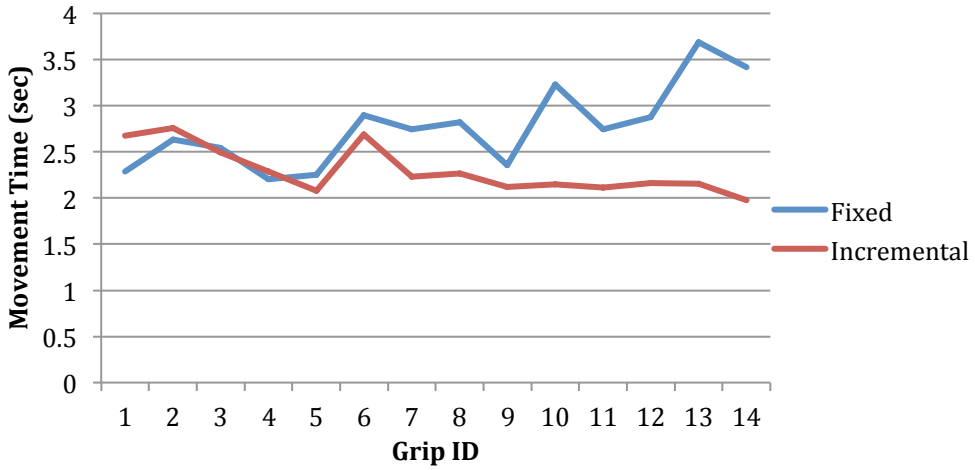

Figure 5-9: Interaction between Grip and Pressure Space on Movement Time (MT).

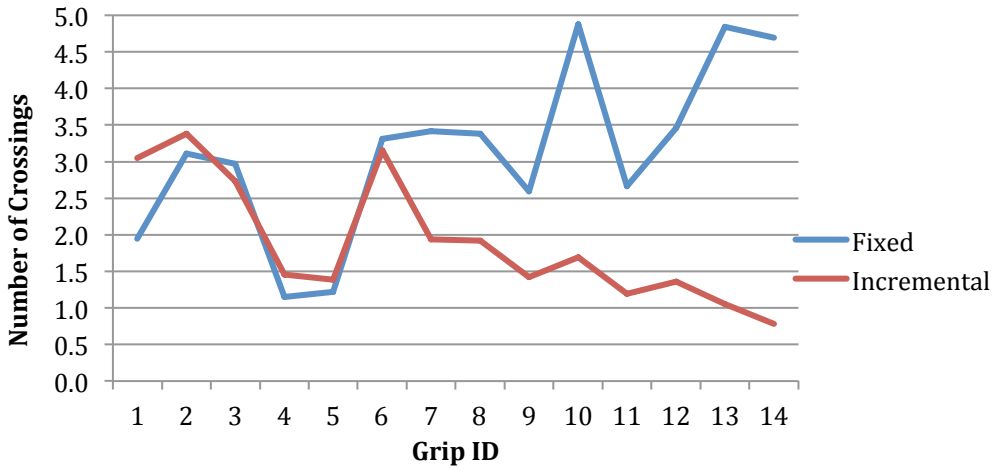

Figure 5-10: Interaction between Grip and Pressure Space on Number of Crossings (NC).

5.3.7 Initial Discussion – Experiment 4

5.3.7.1 Pressure Space

Generally the Fixed pressure space was worse than the Incremental space: although it had a comparable error rate to the Incremental pressure space, Fixed selections took significantly longer and input was significantly less well controlled. Accuracy, in terms of menu items correctly selected, was very high for both pressure spaces, with overall means of 95% and 96% for Fixed and Incremental respectively. Therefore, purely in terms of task success both pressure spaces were equally good. The high accuracy here may be in part due to the use of only 6 pressure levels, as they would be wider and easier to select than a higher number of divisions. These accuracy figures are comparable to those found by some previous research using single pressure input points [170, 192] but higher than those in other research [30, 209]. They are also very similar to the 96-97% accuracy found in Experiments 1 and 2, using Positional control and visual feedback. These results provide the contribution of one-handed input on a mobile device across multiple pressure points, suggesting that multi-digit pressure interaction can be highly accurate.

Despite the high accuracy, Fixed selections took on average 0.56 seconds longer than Incremental selections, at 2.86 sec compared to 2.30 sec respectively. 0.56 seconds is not long in real world terms but it constitutes a 24% duration increase per selection. Note that these figures include the one-second Dwell selection time. These times are comparable to those found in Experiments 2 and 3 using pressure input from only the thumb, however they remain slightly longer than other research [30, 192, 209]. Increased selection time was accompanied by an increase in the number of inadvertent target crossings (NC) during the

Fixed condition, as a result of less controlled cursor movement. The Incremental space had a mean NC of 1.97 crossings per target, with an average of 3.24 for the Fixed pressure space. It seems that participants had more difficulty controlling the cursor during the Fixed condition, resulting in more unintended movement and longer selection times. As can be seen from Figures 5-8 to 5-10, performance using the Fixed space got worse as more digits were used to apply pressure. Newell and McDonald [177] found that the error and variance of applied pressure increased when lower levels of pressure were applied, especially when more than two digits (index finger + thumb) were used. The results in Experiment 4 echo this finding.

Overall, the results support the hypothesis that the size of the pressure space should scale in relation to how many digits will be used in the interaction. The Incremental pressure space increased the total interaction space by 3.5 N per digit. As is explained in Section 5.3.7.2, performance was poorer when more than three digits were used, however, interaction designers may be able to increase the interaction space as much as 3x (compared to one digit) without significant drop in performance.

5.3.7.2 Grip

Figures 5-5 to 5-7 show the mean performance for each grip averaged across both pressure spaces. Overall, performance was relatively good, with accuracy of 90%+ for all grips, selection times of less than 3 seconds (including the 1-second Dwell time) and similar numbers of crossings as were seen in Experiments 1 to 3. The results suggest that every individual digit can be used for pressure input during one-handed use, as can the combinations chosen for testing in Experiment 4, at least when selecting up to 6 pressure levels. Therefore it may be possible to provide multiple different pressure-based inputs to a mobile device from one hand, which would leave the other hand free for other tasks, and the screen clear of obstructions. However, not all digits and grips performed equally well.

In general, error rates were highest when using four or five digits, more so those involving <I> with at least two other digits. However, single-digit G4 had the highest number of errors, providing input from the back of the device using <I>. Targets (menu items) 3, 5 and 6 had equally high error frequency when using G4. Looking at the pressure input profiles for these errors, it appears as though participants simply did not press hard enough to reach the given target in time. G4 introduced tangential force (from the back) to the normal forces gripping the device (from the sides, through the thumb and other digits). As the digits could not directly oppose this force, pushing too hard from the back could effectively push the phone away from the hand, reducing the maximum amount of pressure that could be applied. This

may explain the error for targets 5 and 6, however target 3 required relatively low levels of pressure (1.7-2.3 N) so would be less likely to be affected by this. Also, the fact that there were no errors selecting target 4, suggests insufficient pressing may not have been the issue with target 3 errors. There were only 3 errors selecting this target, across 26 total selections, so the number is low and may not be due to a systematic issue, but a possible explanation is still unknown.

In contrast to the poor performance of <I> from the back in G4, <I> using the top sensor in G5 was the best performing grip, even though it also introduces force along a different plane to the gripping digits along the sides. The reason for the difference in performance may be because the bottom of the device could be pressed into the lower part of the palm, as seen in Figure 5-1, providing opposing forces. <I> + <L> together (G8) were more error-prone than other one- and two-digit grips, although the majority of errors came during the Fixed pressure space. This result was slightly surprising, as other grips using <I> (G7) and <L> (G9) were much more accurate. Combining certain digits, such as <M> or <L>, with other digits provided more precise input than when these fingers were used alone. For example, <L> was poor individually and yet was part of three of the five best grips.

Because the muscles controlling thumb flexion are particularly good at matching target levels of pressure [138], grips using <T> were hypothesised to provide better control of input than those without. The results indicated that the grips using <T> as well as 3 or 4 other digits performed poorly (3-digit G10 using <T>, <I>, <M> was better). The common grip in Figure 5-1 shows where <I> usually rests when holding a touchscreen device one-handed, and few people in our survey wrapped <I> round the device to the same side as <M>, <R> and <L>. From participant observation, data analysis and anecdotal feedback from participants, it appears that it was occasionally awkward to properly orient the hand and fingers to make sufficient contact with as many as four or five sensors. This would lead to low activation of the awkwardly positioned sensors and so decreased input to the system. An exception to this is G10 (using <T>, <I>, <M>) which performed well, possibly because <R> and <L> could be positioned anywhere that was comfortable for the participant. Therefore, the poor performance when using <T> and <I> with other digits is perhaps due to the positioning of the sensors rather than the digits being used. The deliberate decision was made to keep the sensor positions constant across participants rather than move them to suit individual differences, in order to avoid damage to the sensors, keep task time low and for consistency. Tailoring the apparatus to suit individual differences may have improved performance.

The hypothesised benefit of the Incremental pressure space was to improve control when using multiple digits. As can be seen from Figures 5-9 and 5-10, MT and NC dropped as more digits were used under the Incremental space, but ER was poorer for all 3, 4 and 5-digit grips, other than G10 (<T> + <I> + <M>). ER rates were not significantly different in Experiment 4, but with better-positioned sensors, the ER rates for 3- to 5-digit grips may improve, potentially showing a greater performance difference between Pressure Spaces.

Digit Contributions to Each Grip

The varied results, where digits performed differently when used alone to when used in combination, prompted an analysis of how much pressure each digit contributed. For example, <L> performed poorly by itself, but two of the grips it was involved in performed well (G9 and G11). The better performance may be because there are other, more stable digits, to provide the majority of control. However, this does not necessarily appear to be the case.

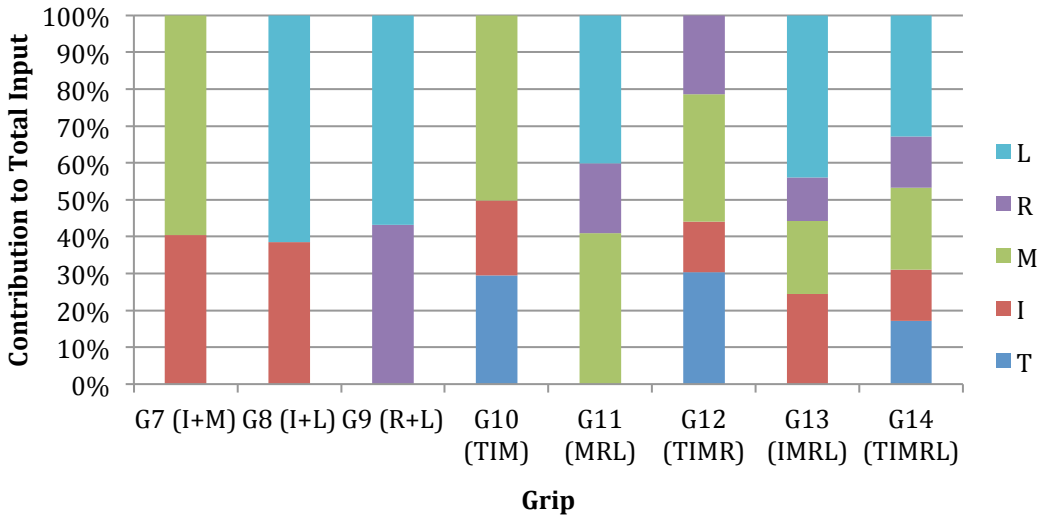

Figure 5-11: Averaged contributions of each digit to total pressure during Fixed pressure space grips: Thumb (T), Index (I), Middle (M), Ring (R) & Little (L).

While there were large individual differences, the averaged, relative contributions each digit made to each 2- to 5-digit grip are shown in Figure 5-11, during the Fixed pressure space and Figure 5-12, during the Incremental space. What is apparent is that, overall, the digits in each individual grip did not contribute equally to the input of that grip. Interestingly, each digit's contribution to a given grip varied depending on the pressure space. During the Fixed pressure space, <M> and <L> generally contributed the largest share of input to their grips, as can be seen by the green and light blue sections in Figure 5-11. In contrast, <I> and <R>

contributed a relatively small amount. Individually, both <M> and <L> performed less well than <T>, <R> and <I> (from the top), suggesting they are less adept at controlling the amount of pressure they apply, yet they dominated Fixed space grips. This may have contributed to the poorer overall Fixed performance. <I> may have been used less because of the awkward sensor positioning, but the meagre use of <R> (outside of G9) is slightly surprising.

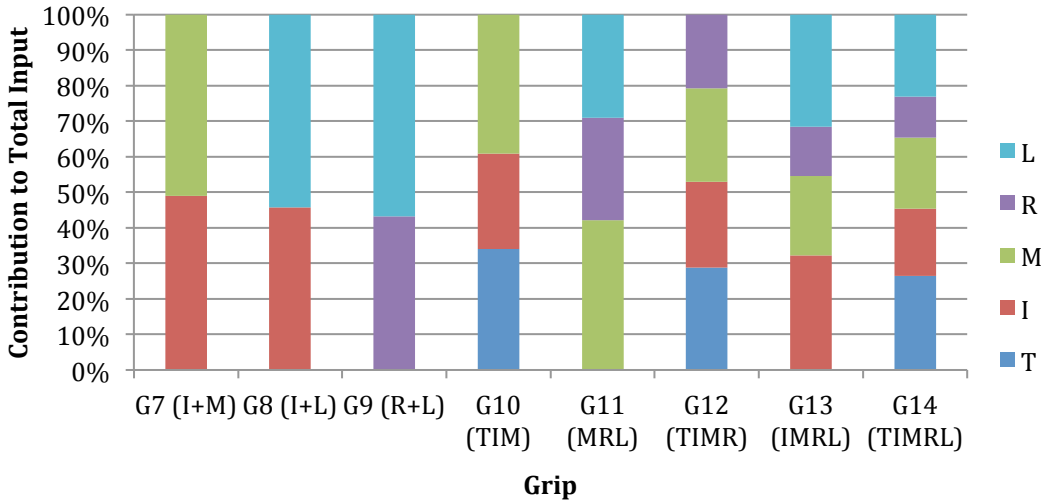

Figure 5-12: Averaged contributions of each digit to the total pressure during Incremental pressure space grips: Thumb (T), Index (I), Middle (M), Ring (R) & Little (L).

The contribution pattern was slightly different during the Incremental pressure space (Figure 5-12). <M> still makes a significant contribution, although slightly smaller than before, and <R> contributes only slightly more. However <I> and <T> contribute much more to Incremental grips, and <L>'s contribution drops. The Incremental pressure space created larger target levels of pressure, and so <T>, <I> and <M> may be more reliable for applying higher levels.

5.3.7.3 Best Grips

In order to compare overall grip performance, a simple scoring system was used. For each measure (ER, MT and NC) the grips were given a 'score' of between 14 (best) and 1 (worst). The highest score possible was therefore 42 points (14 x 3 measures). The top five grips were G5 (42pts), G9 (34pts), G3, G7 and G11 (all 30pts). The lowest scores went to G8 (13pts), G13 (13pts), G14 (14pts) and G2 (15pts).

Up to this point, Experiments 1 to 4 have all used the same task of linear targeting, in the form of menu selection. While it has been useful to properly judge the precision of input, it provides limited scope for appraising the more general usefulness of pressure-based input on mobile devices. Having identified which digits, and combinations thereof, can accurately provide input, Experiment 5 aimed to expand the use of pressure to common interactions with modern touchscreen devices. One potential advantage of having multiple pressure inputs available is that more complex interactions may be controlled. Also, by using input from sensors placed around the outside of the device, the screen remains entirely free of obstructing fingers. Multitouch gestures on touchscreen devices require two hands to use and the gesturing fingers obstruct the very content with which the user is interacting. Multi-digit pressure input could provide two-handed input with only one hand, all the while leaving the screen fully visible. Therefore, based on the results from Experiment 4, the Incremental pressure space and some of the best grips were used to compare one-handed pressure input to two-handed touchscreen input to determine if two-handed interaction can be mapped to one hand.

5.4 Experiment 5 – Comparing One-Handed Multi-Digit Pressure Input to Two-Handed Multitouch

Common touchscreen gestures include one-finger swiping (for example, to pan content larger than the screen viewfinder) and two-finger rotating and zooming. The latter two are often used on touchscreen mobile devices in map or photo applications, for example. Both are bi-directional, single axis tasks, but ones that can be done simultaneously on touchscreen devices. Because translation (vertical and horizontal movement) is also possible while zooming and rotating, it is in fact possible to control 4 axes simultaneously. This has made them very popular and a wide range of different touchscreen devices now use them. To carry out these tasks, however, two hands are needed: one to hold the device and the other to gesture on the screen. This can make the interaction awkward if the individual requires the use of one hand for other tasks, such as carrying things, opening doors, or holding onto handles on public transport. Also, the two fingers gesturing on the screen obscure what is presented underneath, making it hard to see labels on maps, for example. Being able to carry out these functions with just one hand would free up the second hand for other tasks and potentially make the interactions easier when mobile.

Pressure-based input allows for continuous control over a single axis, or one axis per

pressure sensor, and so is well suited for zooming/rotation. However, existing mobile pressure interfaces only tend to allow control over a single axis or task at a time, such as zooming [168], linear targeting [228] or scrolling [90]. To test whether one-handed, multi-digit pressure input can substitute for two-handed touch input, a map-browsing application was developed that included two multitouch gestures: pinch to zoom and rotation. In both cases, one hand holds the device and the other performs the gestures. Pressure-based alternatives were also generated for these actions, all performed with one hand. It should be noted here that this comparison is not between one-handed input and two-handed input, as the second hand in the multitouch interaction is only used to hold the device, and not to provide input. In traditional two-handed input methods, both hands provide some input to the system [129].

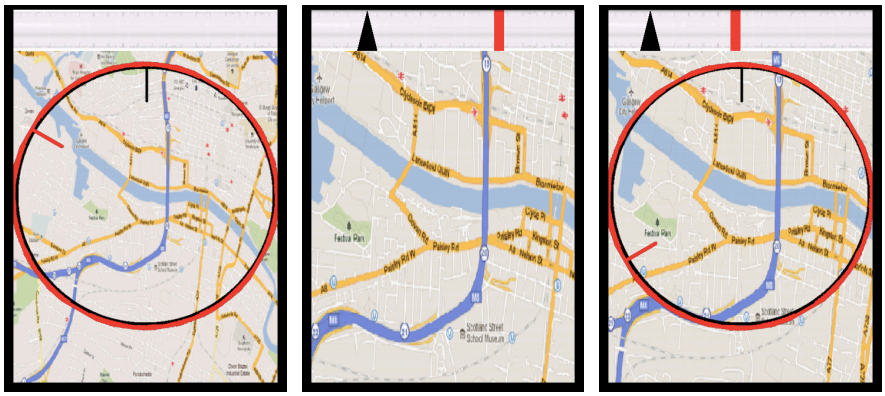

Figure 5-13: Experimental software showing rotating (left), zooming (middle) and combined zooming & rotating (right) tasks.

The study asked participants to engage in three simple tasks: *Rotation* (rotating the map to a target angle), *Zooming* (zooming the map to a target zoom level) and both zooming and rotating combined (*Combination*). The images in Figure 5-13 show the GUI for each task. The image on the left shows the rotating task. Overlaid on top of the map are one black and one red circle, each with a line pointing towards the centre. The black circle rotates with the map, and the line indicates north. The line on the red circle is the target angle and the task is to match the black and red lines by turning the map. The image in the middle of Figure 5-13 shows the zooming task. Above the map is a black arrow pointer indicating the current level of zoom and a red target zoom level marker. The task is to match the current zoom level to the target zoom level. The combined task simply includes both of these and is shown to the right of Figure 5-13. Participants used the controls explained below to alter the angle and/or zoom level to the target angle/level. Selection of the set angle or level was confirmed by leaving the map stationary for three seconds. Participants were expressly instructed to focus on being fast rather than precise during the task, as very high precision was considered less

important during zooming and rotating in real map applications.

5.4.1 Grip Controls and Pressure space

The sensors were left in the same positions as during Experiment 4, to compare their use during more realistic task performance with the targeting/control task. The Incremental Pressure Space was used for Experiment 5, but the choice of digits/grips was not straightforward. The decision could have been based on which sensor positions would make the most logical sense in the context of the task, such as using positions around the device relative to their function (such as moving in, out, left and right), meanwhile ignoring how well we could apply pressure to those positions. The other alternative is using the digits/grips that performed best, even if their positions may not be the most logical concerning their functions in the task. A middle ground was chosen, using four of the best grips from Experiment 4 while basing some decisions on providing more easily understandable controls. It needed to be possible to perform both rotating and zooming at the same time, as it is in multitouch gestures, and so the same digit could not be used for both zooming and rotating.

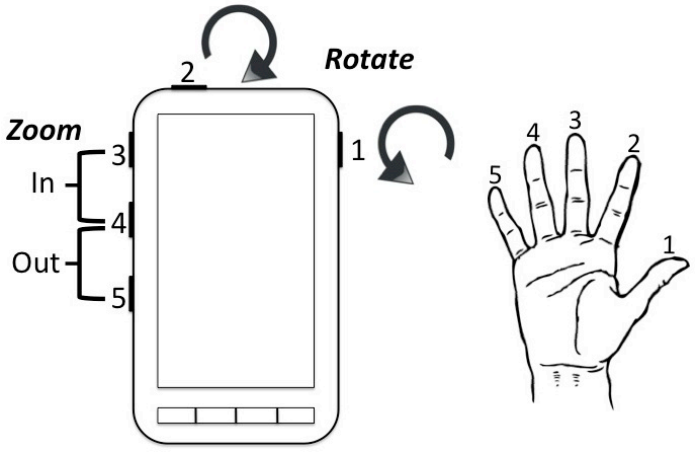

Figure 5-14: Sensors used for pressure-based controls with relative function and digit used for input.

The pressure-based alternative controls are shown in Figure 5-14, and their formation went as follows. The five best-performing grips overall were G5, G9, G3, G7 and G11 (in descending order). G5 and G9 use different digits (<I> from the top and <R> & <L>, respectively) and so were suitable for use. However G3 (<R>) and G7 (<I> + <M>) both use digits from G5 (<I>) and G9 (<R> + <L>). With <I> already used and pressing from the top of the device, it would make G7 a very different grip: either both fingers would push from

the top or <I> would push from the top and <M> from the side, and neither of these was tested in Experiment 4. G11 includes the used <R> and <L> from G9, however <M> is free and so an alternative of this grip was used, consisting of <M> and <R>. <T> was used for the final input, as it performed moderately well in Experiment 4 and would provide somewhat compartmentalized controls, with the rotation and zooming inputs being physically separate from each other.

5.4.2 Controls

The pressure-based alternative controls are shown in Figure 5-14. In each case Rate-based input was used, as it provided better task performance in Experiments 2 and 3: the speed at which the map rotated or zoomed was controlled by the amount of pressure applied. Mappings of pressure-to-movement were based on initial pilot testing.

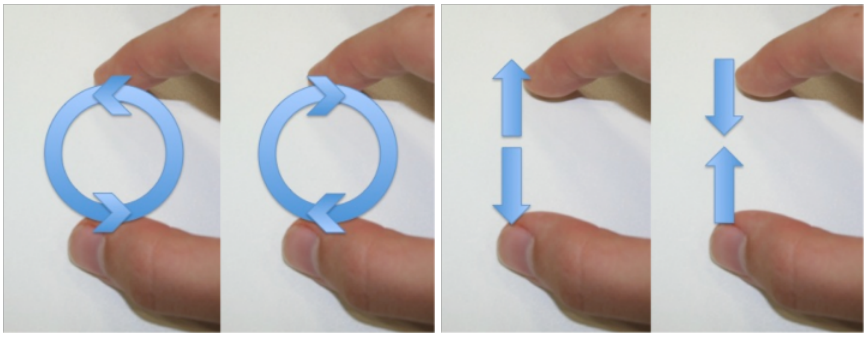

Figure 5-15: Common multitouch touchscreen gestures for rotating (left) and zooming (right).

5.4.2.1 Rotation

For standard touchscreen gestures, elements onscreen can be rotated through the gesture shown in the left of Figure 5-15, where two fingers rotate in unison. For the pressure-based controls, <T> and <I> provided bi-directional rotation: pressing on sensor 1 with <T> rotated the map anticlockwise and pressing on sensor 2 with <I> rotated clockwise. In both cases, the speed of rotation increased as the amount of pressure applied increased, at a rate of 2-200°/sec (equivalent of 57.14° per N, per second).

5.4.2.2 Zooming

Zooming is commonly achieved through inward, or outward, pinch gestures with two fingers, as shown to the right in Figure 5-15**,** and the same gestures were used here. For the

pressure-based controls, pressing sensors 3 & 4 together (with <M> and <R> respectively) zoomed in and pressing sensors 4 & 5 (with <R> and <L>) zoomed out. Increasing the amount of pressure increased the speed of zooming. Experiment 4 showed that increasing the size of the pressure space when more than one digit is used provides better control and so 3.5 N was measured from each digit (total of 7 N). The pressure from both digits was summed, giving a zoom increase/decrease rate of 1-20% per second (2.8% per N, per second).

5.4.3 Participants and Procedure

The same apparatus from Experiment 4 was used for the pressure-based controls. Issues with multitouch detection on the Nexus One led to the use of an HTC Desire S (see Figure 5-16) for the touchscreen condition. It is a very similar size to the Nexus One, with similar specifications.

Figure 5-16: HTC Desire S mobile phone used for multitouch input in Experiment 5.[11]

Twelve of the thirteen participants from Experiment 4 took part in Experiment 5 a week later, and were again sat in a padded office chair. Before the study began, participants filled in a brief questionnaire indicating their familiarity with touchscreen devices as well as multitouch gestures. The within-subjects study was split into two by Control Method: one half using the pressure-based controls first and one half using touch controls first, with the order counterbalanced. In the pressure control condition, participants held the apparatus in their right hand, while during the touch condition they held the Desire S in their left hand and gestured on the screen with the right hand. Within each Control Method condition, each participant completed two blocks, each consisting of twelve rotation tasks, twelve zooming tasks and twelve combined rotating and zooming tasks, in a random order, giving a total of

[11] Image from http://www.techdigest.tv/htc-desire-s.jpg

24 of each task. The first block of each Control Method started with nine practice tasks, three of each type. Experimental instructions can be found in Appendix C.

The Independent Variables were *Control Method* (Multitouch, Pressure-based), *Block* and *Task* (Rotating, Zooming, Combined). The Dependent Variables were: *Error* (distance of zoom level/angle to target level/angle, expressed as % of zoom or rotation space), *MT* (time from first movement to last movement in each task) and *Subjective Workload* (NASA TLX).

There were a total of: 12 participants x 2 Control Methods x 3 Tasks x 2 Blocks x 12 Task trials = 1728 trials. This gave 864 data points for each Control Method, 576 data points for each Task and 288 data points for each Control Method + Task combination condition (e.g., Pressure + Zooming).

5.4.4 Results – Experiment 5

The data for Experiment 5 also violated the normality assumption and so an Aligned Rank (ART) transformation of the data was carried out for analysis using a 2 x 2 x 3 Mixed Model REML with participant as a random factor. Note that all MT figures *do not* include the three-second stationary confirmation time and therefore represent time from first movement to last movement. Seven of the twelve participants owned touchscreen devices and, of those, four used multitouch gestures "often" or "every day", with the other three using them "seldom" or "occasionally". Four others had used touchscreen devices in the past, of which two had used multitouch gestures and two had not. Only one participant neither owned nor had used a touchscreen device. Raw data for all measures can be found in Appendix C.

5.4.4.1 Block/Learning Effects

Comparing blocks one and two showed no learning effect in terms of Error ($F_{(1,131)} = 1.804$, $p > 0.05$) with mean values of 6.78% and 6.47% for block one and block two respectively. There was, however, a learning effect in terms of MT ($F_{(1,131)} = 32.091$, $p < 0.001$) with means of 6.77 seconds for block one and 5.66 seconds for block two. A significant interaction effect between Block and Task was also found for both Error ($F_{(2,262)} = 3.798$, $p < 0.05$) and MT ($F_{(2,262)} = 10.429$, $p < 0.001$). ER dropped from Block 1 to Block 2 for both the rotating and combination tasks, but zooming ER increased. MT dropped more between Block 1 and Block 2 for rotating and combination than for zooming.

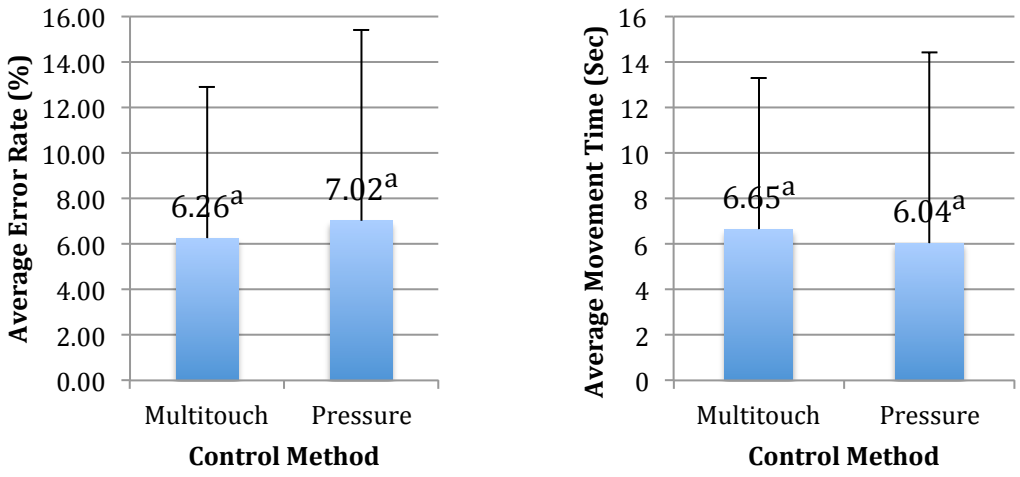

Figure 5-17: Mean targeting Error (distance from target) and Movement Time for each Control Method used in Experiment 5. Error bars show 1 standard deviation. 'a' indicates a significant difference, p < 0.001.

5.4.4.2 Control Method

Error, MT and workload values for each control method can be seen in Figure 5-17**.** There was a significant effect of Control method on Error ($F_{(1,131)}$ = 11.743, p=0.001) with the Multitouch controls producing lower average Error (smaller distance to target angel/level) than the Pressure-based controls (means of 6.24% for Multitouch and 7.01% for Pressure). There was also a significant effect of Control Method on MT ($F_{(1,131)}$ = 36.111, p < 0.001). In this case, Pressure-based controls produced faster targeting times (mean = 5.95s) than Multitouch controls (mean = 6.49s). An interaction effect was found for Control Method and Block on MT ($F_{(1,131)}$ = 10.826, p=0.001), as MT dropped more from Block 1 to Block 2 under the Multitouch controls (7.39s to 5.9s), compared to Pressure (6.42s to 5.67s). A non-parametric Wilcoxon *T* test on the NASA TLX workload ratings showed there was no effect of Control Method on either overall subjective workload, with mean ratings of 8.97 and 7.05 for Multitouch and Pressure input, respectively.

5.4.4.3 Task

Error and MT values for each task can be seen in Figures 5-18 and 5-19**.** There was a significant effect of Task on both Error ($F_{(2,262)}$ =1492.021, p < 0.001) and MT ($F_{(2,262)}$ = 696.252, p < 0.001). In both cases, each task differed significantly from the other (p < 0.001). The Rotation task had the lowest error (mean = 0.44%) followed by the Combination task (mean = 7.91%) and the Zooming task (mean = 11.52%). Concerning MT, the Zooming task was fastest (mean = 2.95s) followed by the Rotation task (mean = 4.85s) and

Combination task (mean = 10.85s).

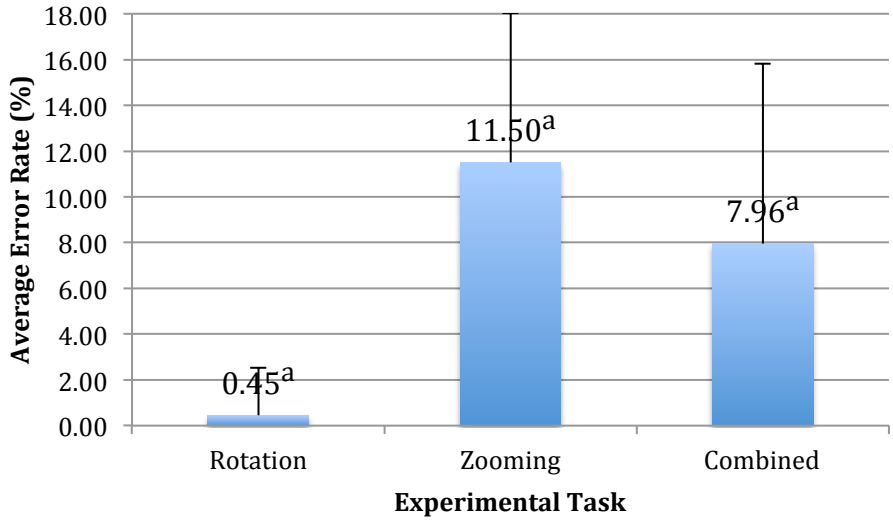

Figure 5-18: Mean Error Rate for each Experimental Task in Experiment 5. Error bars show 1 standard deviation. 'a' indicates significant difference, p < 0.001.

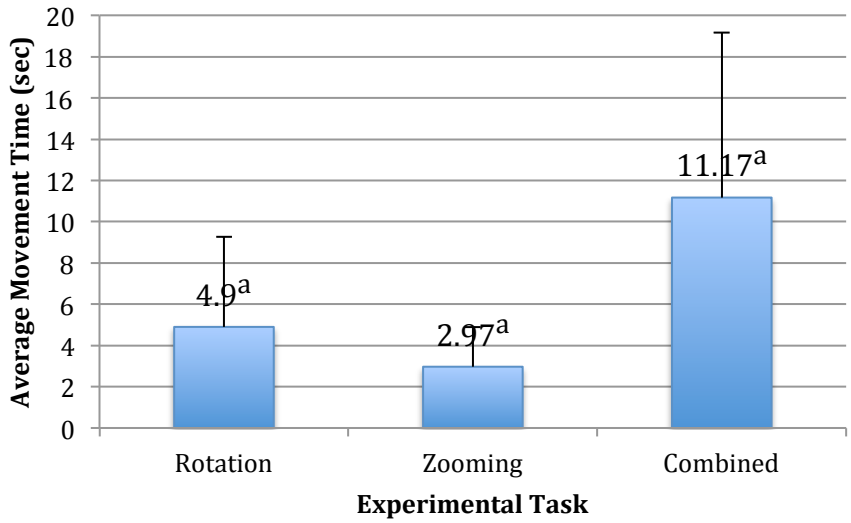

Figure 5-19: Mean Movement Time for each Experimental Task in Experiment 5. Error bars show 1 standard deviation. 'a' indicates significant difference, p < 0.001.

5.4.4.4 Control * Task

A significant interaction effect was found between Control Method and Task for both Error ($F_{(2,262)}$ = 13.526, p < 0.001) and MT ($F_{(2,262)}$ = 26.877, p < 0.001). For ER, Multitouch performed better for both the rotation and combination tasks, but pressure performed better

for the zooming task. For MT, Multitouch was faster for rotation and zooming, but Pressure was faster for the combination task. Also, a three-way interaction between Control method, Block and Task was found for both Error ($F_{(2,262)} = 7.318$, $p = 0.001$) and MT ($F_{(2,262)} = 7.709$, $p < 0.001$).

5.4.5 Initial Discussion – Experiment 5

The results from Experiment 5 are highly encouraging and indicate that one-handed pressure-based input could provide concurrent control over two axes or tasks at one time, in this case two-handed zooming and rotating touchscreen gestures. Eleven of the twelve participants either owned a touchscreen device or had used one in the past, with nine of those eleven having experience with multitouch gestures, four of whom used gestures "often" or "every day". Despite this high level of familiarity, pressure-based input still performed similarly to the standard multitouch gestures. There was a small but significant difference between the control methods in their targeting Error, with the multitouch controls producing better accuracy than the pressure-based controls. However, the real world difference between the means was less than 1% in absolute terms. In contrast, the pressure controls allowed for significantly faster task completion times than Multitouch but, again, with a relatively small mean difference of around 0.6 seconds per task (a relative increase of 9%). A further encouraging result is that there were no significant differences in subjective workload ratings between the two control methods, although Pressure workload ratings were slightly lower overall.

So, whereas pressure input may not have been quite as precise as Multitouch, it allowed for faster input with only one hand and with no significant increase in workload demand on the user. The Pressure controls provided more accurate (but slower) control over zooming but provided faster, and less accurate, control over the combination task. Rotation performance was comparable. Therefore, multi-digit pressure input may be a suitable means of controlling multiple axes or tasks with only one hand on mobile devices, without the fingers obscuring the screen and leaving the other hand free for other tasks. The large difference in average Error between the Zooming and Rotating tasks is likely due to the smaller total zooming space, compared to the rotation space. The high overall interaction times indicate that participants took time to be as accurate as possible. In future, real-world use of these tasks should be studied to test how precision and task time are affected.

The interaction between Task and Block is interesting, as zooming ER increased from Block 1 to 2, and zooming MT reduced less, compared to the other tasks, from Block 1 to

Block 2. The total zoom range was smaller than the total rotation range, which may have made smaller movements more difficult. The *increase* in ER, however, is surprising. With greater familiarity with the task the participants seemed to zoom faster, but at the sacrifice of accuracy. They were instructed to be fast rather than accurate during the task, so this behaviour may have become more apparent during the second block.

5.5 Discussion

Experiment 4 looked at the factors influencing our ability to control one-handed pressure interaction, while Experiment 5 applied this understanding to the design of an interface aimed at utilising one-handed, multi-digit input. There are four important outcomes from the results summarized here: the influence of digit choice, the influence of digit number, the influence of pressure space and the positioning of sensors.

5.5.1 The Influence of Digit Choice and Number

In general, using more than three digits reduced the precision with which pressure could be applied to the mobile phone, even when increasing the target levels of pressure to suit the number of digits, however this may be partly due to awkward positioning of some of the sensors. In contrast, performance using one, two or three digits can be highly precise in various combinations (depending on the digits used). Some digits perform better in combination with other digits, compared to on their own (such as <M> and <L>) while others are precise both alone and in combination (<I>, <R>). Therefore, interfaces may be more easily controlled if <M>, <R> and <L> are used in 2- or 3-digit combinations, rather than alone. For example, the three digits that are commonly used to hold touchscreen devices (<M>, <R> <L>; Figure 5-1) performed relatively well, better than the three-, four- and five-digit grips using the thumb. Finally, the top of the device using <I> may be an optimal location for pressure input on mobile devices.

5.5.2 The Influence of Pressure Space

As was hypothesized, increasing the target pressure levels when using more digits produced better performance in terms of speed and input stability, even though error levels remained similarly low for both pressure space sizes. For interaction designers, it means that the

interaction would benefit from tailoring the size of the interaction space (pressure space) to the number of digits being used for input. When comparing the relative contributions each digit made to the total pressure input in the Fixed and Incremental spaces, it appeared that different digits dominated, depending on the target pressure level. Lower target pressure, during the Fixed space, led to <M> and <L> providing the majority of input, whereas when targeting higher pressure during the Incremental space, <T> and <I> provide a much larger proportion of input, along with <M>.

The thumb, index and middle fingers are relied upon to generate more pressure, and so interfaces that make use of pressure spaces larger than 3.5 N should facilitate input from these digits. Using more digits makes targeting lower levels of pressure more difficult [177], and it may be that using stronger digits, such as <T>, <I> and <M>, especially in combination, also make targeting lower pressure more difficult. <L> and <R> were used much more than the other digits (i.e., their relative contribution was higher) during the Fixed space, which suggests participants relied upon these digits for targeting lower pressure levels. Therefore, interfaces (based on input down the sides of a mobile device) which use small pressure spaces of 3.5 N or less, should perhaps allow for input to come from these digits (<L>, <R> and <M>).

5.6 Limitations

This section discusses the main limitations of the research reported in this chapter, which should be considered when interpreting the results. The main limitations of Experiments 4 and 5 were: 1) potentially easy targeting task, 2) the influence of sensor location and 3) the lack of mobility.

5.6.1 Potentially Easy Targeting Task

The targeting task in Experiments 1-3 split the pressure space into as many as 10 target levels to test accuracy in targeting a narrow range of pressure. As discussed in Section 5.3.3, 6 levels were chosen to balance task time and task difficulty. However, given the very high accuracy across Grips and Pressure Spaces, the task may have been too easy, leading to a ceiling effect. In this case, the performance differences between Grips/Pressure Spaces would not be as apparent, compared to a more difficult task. The task would have benefited from either dividing the space into different numbers of target levels (6, 8 and 10) or using

only 8 or 10 levels, instead of only 6. However, as they stand, the results still show that the same number of pressure levels (6) can be as accurately selected using each individual digit (and select combinations) as when sitting at a desktop interacting with a hand-held stylus [192].

5.6.2 The Influence of Sensor Location

The sensor positioning was partly based on where the fingers commonly hold mobile devices during one-handed use, with the addition of three alternative locations for the index finger. One of those (top) performed well and one did not (back). The third location (sensor 2, Figure 5-1) was used to provide full-hand grip around both sides of the device. However, the grips using sensor 2 with at least two other sensors (G10, G12, G13 and G14) did not perform well, particularly G13 and G14, which used all four sensors down the left-hand side. During these grips it was difficult to orient or arrange the fingers to make sufficient contact with the relevant sensor. This conclusion is further supported by the good performance of G7 using just <I> and <M>, and of G9 and G11 using <M>, <R> and <L>. Therefore the fixed positions of the sensors may have made input more difficult, as individual differences in hand size, span and finger length would make the ideal locations different. It will be necessary to design apparatus that can take pressure input from all digits regardless of where they are placed around the device, in order to assess whether input precision using four or five digits is better when input from all digits is accurately measurable.

5.6.3 The Lack of Mobility

When judging the usefulness of pressure input for interaction with mobile devices, understanding the influence of walking is highly important. Experiments 2 and 3 tested control of pressure input when walking, but only using a single digit (thumb) and sensor for input on an unrealistically large mobile device. While Experiment 5 tested control on a small, commercial mobile device using all digits available, it did so only sitting down. Therefore, it is unknown how walking influences interaction with a realistic device, when interacting with each digit. Also, one of the main advantages of using one-handed pressure input over two-handed multitouch is that the second hand is free to hold objects or interact with objects in the mobile environment, such as doors or handles. Comparing task performance of pressure and multitouch while the user is walking and/or holding an object would provide a clearer picture of any benefits that might come from one-handed, multi-digit pressure input.

5.7 Conclusions and Research Question 3

Research Question 3 asked:

"How accurate is pressure-based input when multiple digits apply pressure to a mobile device?"

In answer to this question, Experiment 4 found that all individual digits could be used for precise input on a mobile device (albeit not all equally well) even while simultaneously holding the device. The 14 digit combinations tested also performed well, and testing the other sixteen possible grips may reveal yet more usable inputs. Therefore, purely from the perspective of pressure-matching precision, pressure input on a mobile device using multiple digits is highly accurate. However, input was also accurate from an interaction perspective. An interaction using four input channels was developed and mapped common two-handed touchscreen interactions to one-handed pressure interactions. In Experiment 5, this setup was shown to provide a very similar level of performance compared to traditional multitouch input, but with the advantages of using only one hand and leaving the screen fully visible. Therefore, multi-digit pressure input is not only precise (Experiment 4) it can also provide accurate input to common mobile interactions (Experiment 5). One-handed pressure input could also be used for other common touchscreen interactions such as scrolling (with scroll speed controlled by amount of pressure) or menu interaction (with different menus mapped to different digits).

The contributions of the research in this chapter are:

- One-handed, multi-digit pressure input can provide comparable control over bi-directional zooming and rotating as two-handed multitouch input, with the advantages of requiring only one hand and leaving the screen fully visible.
- Each individual digit of one hand, and various combinations, can apply force accurately but not equally so, when both holding and squeezing a mobile phone.
- Increasing the range of detectable force, relative to the number of digits used, improves precision of applied pressure.

6 Identifying Detectable and Comfortable Thermal Feedback Parameters

6.1 Introduction

One aim of this research was to investigate the feasibility of using thermal stimulation as a means of conveying information non-visually in mobile interaction scenarios. Specifically, the aim was to investigate the design of structured thermal feedback that can convey multidimensional information, in a similar manner to how Earcons [10, 16, 18, 165] and Tactons [12, 20, 21] convey information in the audio and tactile domains respectively. Thermal sensation is a vital, inherent aspect of human touch, which continually provides information about our environment and each object that we touch, and so it naturally conveys information.

While both Earcons and Tactons have been shown to be effective in conveying information in mobile scenarios, there are environments in which audio and vibrotactile feedback may not be suitable or appropriate. These may include very loud or very quiet places unsuitable for audio, or simultaneously loud and bumpy environments, such as public transport, that are unsuitable for both. In very quiet environments, such as religious buildings, museums or meetings, vibrations may still be heard or felt by others. Thermal feedback is entirely silent and so may be suitable for these settings. Further, user preference for when a feedback modality is desired varies by location and situation [98, 101, 102], and thermal feedback may provide a third alternative. If thermal feedback cannot convey information sufficiently in isolation, it may instead provide an extra, complimentary layer on top of existing visual and non-visual feedback channels. Very little research has been done on designing thermal feedback for everyday usage scenarios, including mobile interaction. No HCI research has taken into account the possible influences of walking and being outdoors on the perception of thermal feedback, and so no research exists on how best to design thermal feedback to be suitable in a range of environments. This research aimed to address these issues.

In Earcons and Tactons, information is encoded in the parameters of the sound (for example, rhythm, pitch, timbre) and vibration (rhythm, waveform), with each parameter conveying one piece of information. By creating two or three perceptually distinct variations, or levels,

each parameter can provide one of two or three bits of information. Through Earcons, for example, McGookin & Brewster [165] used timbre, rhythm and register to represent the type, intensity and cost, respectively, of a theme park ride. Three instruments represented three types of rollercoaster, three rhythms indicated three levels of intensity and three pitches (at least an octave apart) represented three cost values. Using this design, three pieces of information could be presented concurrently, with a range of 27 possible combinations. For example, a trumpet timbre playing a 6-note rhythm in a high pitch would indicate the ride is 1) a rollercoaster of 2) high intensity and 3) high cost. In order to use thermal stimulation as a feedback channel in the same manner, it is necessary to identify which features of thermal stimulation create different sensations in the individual, and so could be used as feedback parameters. Then levels of those parameters could be chosen based on the perceptual distinctiveness of points along each parameter's spectrum.

Therefore, the first necessary step was to identify which features of thermal stimulation would be suitable for use as feedback parameters, and then identify which variations or levels of those parameters can reliably produce sensations that are both:

- *Salient* – reliably detectable;
- *Comfortable* – the stimulus does not cause an undesirable or painful sensation.

Parameter levels that are salient but uncomfortable would not be acceptable in an interface, and levels that are comfortable, but less reliably detected, would be of little use. Therefore, Research Question (RQ) 4 asks:

"What parameters of thermal stimulation are most salient and comfortable when using equipment designed for mobile interaction?"

Chapter 2 reviewed the wealth of psychophysical research on human thermal perception, and described how the intensity and quality (comfort, pain, prickling) of thermal sensation changes as the stimulus varies in terms of the extent of thermal change, the rate of change or the area of stimulation. These results formed the basis for identifying suitable thermal feedback parameters; however, the process was not straightforward: there are fundamental differences between how the psychophysical results were gained and how thermal feedback would be presented in mobile HCI. Firstly, these psychophysical results were obtained in highly controlled laboratory conditions, often from participants who had engaged in many hours of training and testing. Mobile interaction scenarios, and even static indoor scenarios, are far more variable and users receive no prolonged training on identifying feedback.

Therefore, the way each parameter influences perception in the literature may not necessarily apply in realistic HCI scenarios. Perception needs to be tested in these environments. Secondly, the psychophysical apparatus used to produce the thermal sensations were often large, complex and specific to the location being tested. Perception of thermal stimuli needs to be tested when the apparatus used to produce the sensations has been designed specifically for mobile interaction: compact, light and efficient and is sufficiently generic in design that it can be attached to a variety of devices.

6.1.1 Hardware Design

RQ 4 specifically focuses on what thermal stimuli are perceivable through more limited stimulating apparatus: hardware that has been designed for use in 'realistic' mobile HCI scenarios. Rather than refer to one or more specific usage scenarios, the aim was to have hardware designed that could be used in a number of scenarios, but that had particular characteristics that would make it relatively easy to transport or wear, and that could feasibly be integrated into mobile device form factors, primarily mobile phone forms. The apparatus needed to be small enough to fit within the width and height of a typical mobile phone (approximately 11 cm x 6 cm), and be light enough to be carried (either attached to a mobile device, or the user or in a pocket) without encumbering the individual. Part of the feasibility consideration was in terms of power requirements. Far less power is available from compact battery technology (typically Lithium-ion in mobile devices) compared to mains electricity. While some psychophysical research has used relatively low voltage apparatus to produce low rates of thermal change, some have required very high rates of change and so required much higher voltage. Jones & Berris [128] recommend that Virtual Reality (VR) thermal interfaces employ maximum rates of change of 10-20°C/sec and multiple stimulating elements, which would also require large amounts of power. Being limited to battery power would mean being limited to lower rates of change (due to lower voltage) and shorter duration (due to capacity).

Based on functional and design requirements specified, SAMH Engineering was commissioned to construct the hardware for presenting thermal stimuli. A full description of the requirements and the hardware design itself is included in Appendix D. The functional requirements specified that the hardware had to be capable of producing and manipulating the suitable parameters of thermal stimulation outlined in Section 6.2 of this chapter. However, because of certain unforeseen technical limitations, the precise levels of each parameter that could be produced were restricted. The details of these limitations are described in the section relevant to the parameter.

Section 6.2 summarises the main parameters and influences of thermal stimulation outlined by psychophysical science, and explains the process behind choosing which parameters are potentially suitable for conveying information in mobile HCI and which are currently considered unsuitable. Section 6.3 includes two evaluations, which tested perception of each of these parameters using the proprietary hardware design. The experiments focused on the salience and comfort of different parameter levels. It includes an evaluation conducted when participants were seated indoors, and so formed a baseline level of perception. This is followed by a second identical evaluation carried out when participants were walking a pre-defined route indoors, to test the influence of walking on the perception and comfort of the chosen thermal stimuli. Section 6.4 gives overall discussion and conclusions of the research while Section 6.5 describes the limitations of the research. Final design recommendations for thermal feedback in HCI are outlined in Section 6.6.

6.2 Choosing Parameters and Parameter Levels

Psychophysical research has identified a number of features of thermal stimulation that have a profound effect on the related internal subjective sensation. Each of the main features is addressed individually in this section, including a judgement on that feature's suitability as a parameter in structured thermal feedback design.

6.2.1 Important Influences: Skin Temperature and Stimulus Magnitude

As discussed in Section 2.3 in Chapter 2, thermal perception is focused on the extent or magnitude of stimulation (the subjective intensity) [43, 135, 220, 223], and the subjective intensity of the stimulation is relative to the individual's skin temperature [53, 71, 77, 184]. The preoccupation with intensity means the sensations that can be aroused by thermal feedback may be limited primarily to those of varying intensity. In attempting to identify usable parameters for thermal feedback, a similar path to that of Brown and colleagues [20-22] in designing Tactons was followed.

The primary design challenge was the influence of skin temperature on our sensitivity to thermal changes, particularly thermal thresholds: the smallest amount of change required to produce a perceivable sensation. It is discussed in detail in Section 2.3; in short, faster changes are more salient when the skin is resting in the 'neutral zone' of 28-36°C [128, 223]

and, when outside the neutral zone, smaller changes are more salient when they warm or cool the skin further towards the pain thresholds. Once a stimulus has been detected, continuing to change temperature increases the intensity of the sensation [184, 185, 223].

The challenge for thermal feedback in mobile HCI is that the rate of change and extent of change need to be controlled relative to skin temperature, so that the stimulus is detectable, but not so fast or large as to be uncomfortable. While the individual's skin temperature is a largely uncontrollable influence, a thermal stimulator could influence the temperature of a local area of skin that it is in contact with.

6.2.2 Suitable Thermal Feedback Parameters and Levels

6.2.2.1 Direction

"Direction" is used here to refer to warming and cooling from a baseline skin temperature. Unlike extent of change (6.2.2.2) and rate of change (6.2.2.3), direction does not, by itself, result in sensations of varying subjective intensity. It is a more qualitative feature. There are considerably more cold receptors than warm receptors around the body [223], and cold thresholds are often smaller than warm thresholds under the same circumstances [61, 133], suggesting a higher sensitivity to cooling changes. At 45°C, heat pain thresholds (the point at which sensations of warmth become predominantly sensations of pain) are generally closer to the skin's resting temperature of 28-36°C, than cold pain thresholds at around 13-15°C [118, 223]. This asymmetry may mean that the extent by which a stimulator changes may have to fall within a more limited range of <= 12°C changes (from skin temperature) when warming, compared to <= 20°C when cooling. This would result in an asymmetric interaction/design space.

However, we remain highly sensitive to changes in both directions [133, 223], as both warming and cooling are integral aspects of thermal perception and information transmission in everyday life. Because it is one of few qualitative parameters, it is important to include it in the design of structured thermal feedback, so that sensations can vary on more than the single dimension of subjective intensity.

Levels

1. Warming
2. Cooling

In order to both warm up and cool down it is necessary to have a 'starting' temperature. Starting from within the skin's resting 'neutral zone' is most logical, as adaptation to the starting temperature can occur and no thermal sensation will arise from the starting temperature itself [128, 223]. Therefore, 32°C was chosen for the starting temperature; it has been used as such in other studies [61, 133]. Warming increased the stimulating temperature above 32°C and cooling decreased temperature below 32°C.

6.2.2.2 Extent of Change

The extent of change is the difference, in degrees Celsius (°C), between the individual's current skin temperature (or a pre-set starting temperature) and the temperature of the stimulator, and so indicates the extent of warming or cooling of the skin. As discussed in Chapter 2, most psychophysical research on extent of change has focused on thermal thresholds. While the salience of a thermal change is inherently tied to its rate of change (ROC), it can still be isolated and used separately to influence sensation. Research that has examined supra-threshold extents of change, beyond minimal threshold, has found that, when ROC is kept constant, greater extents of change are perceived as having greater intensity or strength [53, 66, 81, 161, 229].

Levels

1. ± 1°C
2. ± 3°C
3. ± 6°C

Because the starting temperature of 32°C is within the range of neutral resting temperatures, changes should be relatively large to maximise their salience. Therefore, three different extents of change were used: ± 1°C, ± 3°C and ± 6°C. From previous research, 1°C changes were detectable at rates of change equal to, as well as below, those chosen for use in this research, but these results were gained in highly controlled psychophysical laboratory conditions [36, 133, 185]. 3°C changes, at ~1°C/sec, were perceivable by most of the participants in a much less controlled desktop HCI experiment [174]. Wettach *et al.* [239] and Lee & Lim [151] suggested that users may be able to differentiate varying degrees of warmth. Therefore, a stronger change of 6°C was also used to investigate user responses to a wider range of changes. These would give set-point temperatures of:

- 26°C (-6°C), 29°C (-3°C), 31°C (-1°C)
- 33°C (+1°C), 35°C (+3°C), 38°C (+6°C)

None of these temperatures is near to the pain thresholds [43]; however, the asymmetric space does mean that 38°C is closer to heat pain (~45°C) than 26°C is to cold pain (~13°C). This could potentially make the warming changes less comfortable than cooling changes; on the other hand, sensitivity to changes improves as the stimulating temperature moves away from neutral [133, 136], so it may also mean that participants will be more sensitive to the warming changes.

6.2.2.3 Rate of Change

As outlined in Chapter 2, the relationship between rate of change (ROC) of thermal stimulation and thermal threshold is roughly U-shaped. This may be related to the phenomenon of temporal summation (see Section 2.3.3.2), which would result in faster changes feeling more intense [223]. Therefore, varying the ROC could produce varying subjective intensities of stimulation, even when using the same extent of change.

Changes of ~0.2°C from resting skin temperature can be detected at ROCs as low as 0.1°C/sec when in ideal laboratory conditions, and after many hours of testing [133]. In contrast, VR thermal interfaces are recommended to have temporal resolutions as high as 20°C/sec [128] to mimic the thermal conductivity of various materials. The very low ROC may not be suitable for real world use, and the very high rate of change may not be necessary to produce salient changes in feedback designed for mobile HCI.

Levels

1. 1°C/sec
2. 3°C/sec

No research has yet looked at the influence of ROC in the context of feedback for HCI, so it was important to include this parameter. Two different rates of stimulus change were used: 1°C/sec and 3°C/sec. Research has shown that 1°C/sec should be adequate to produce detectable sensations in ideal situations [36, 133, 185]. But other work has suggested it may not always be large enough [174], and the influences of walking and being outdoors may make detection more difficult, so a higher rate of change was also used. The specific rates were limited by the capabilities of the hardware design, specifically the current and battery capacity available (details are available in Appendix D). This meant that ROCs above 3°C/sec would not be feasible using a battery supply.

6.2.2.4 All Stimuli

Three features were chosen as potential parameters in thermal feedback design:

1. Direction of Change: warm and cool
2. Extent of Change: 1°C, 3°C and 6°C
3. Rate of Change: 1°C/sec and 3°C/sec

Employing two directions of change, three extents of change and two rates of change (ROC) gave a total of 12 stimuli/thermal changes, which are shown in Table 6-1. A single stimulus consisted of warming or cooling from 32°C by a set extent (1°C, 3°C or 6°C) at one ROC (1°C/sec or 3°C/sec), for example, warming by 3°C (to 35°C) at 1°C/sec. When first chosen, the sensations that would result from these stimuli were not known, but they were hypothesised to vary only in their subjective intensity, and whether they felt warm or cold. The use of different extents of change and different rates of change meant that different subjective intensities might be produced in several different ways. Although the same extents of change were used for both ROCs, the effect of ROC on perception meant that the same extent might feel differently intense at one rate compared to the other. The precise magnitude of these subjective intensities was not known; nor was how salient or comfortable each parameter and level was. Therefore, the two experimental studies reported in this chapter were carried out to test the detection, comfort and subjective intensity of each chosen thermal feedback parameter and level. This would then indicate the range of detectable and comfortable intensities that can be produced from this limited selection of stimuli.

	Warm		**Cool**	
	1°C/sec	**3°C/sec**	**1°C/sec**	**3°C/sec**
1°C	33°C	33°C	31°C	31°C
3°C	35°C	35°C	29°C	29°C
6°C	38°C	38°C	26°C	26°C

Table 6-1: Stimuli by intensity, direction and ROC.

6.2.3 Excluded Candidate Parameters

6.2.3.1 Area

Due to spatial summation, stimulating a larger area using the same temperature would result in a stronger subjective sensation [25, 135, 220, 223, 225, 226]. Using this phenomenon, varying degrees of subjective intensity could be produced by varying only the area of

stimulation, for example by activating/deactivating physically adjacent stimulators placed in an array. Due to poor spatial resolution, it is less likely that the areal extent of stimulation itself could be a perceivable parameter (for example, small/medium/large), unless the changes in area were large or spread out, which would be less suitable for integrating into mobile devices. Spatial pattern recognition is also poor [202]. While varying the intensity through the area of stimulation would be useful, the extent of spatial summation reduces as the extent of change increases [226], to the point where no spatial summation occurs at the pain thresholds [223]. Finally, the intention was to design thermal feedback that would be suited to mobile interaction, and so the space available for placing multiple stimulators around a small mobile device would be limited. For these various reasons, area was not used as a parameter for feedback.

6.2.3.2 Rhythm

Rhythm is an integral parameter in the design of both Earcons and Tactons, created by using multiple notes of different length, or on-off pulses of varying length. Uniquely identifiable rhythms can then convey a piece of information during interaction. Rhythmic thermal changes have not been employed in either psychophysics or HCI. Therefore, it is open to interpretation how a rhythm would be formed. While Earcon and Tacton rhythms are effectively unidirectional (they are either on or off), thermal rhythms could be bidirectional by using both warmth and cold. Thermal 'pulses' or 'waves', which warm or cool the skin from neutral by a set extent, before returning to neutral/skin temperature, could be placed in sequence, varying the length of the pulse, the inter-pulse length or both. Alternatively, pulses could include both warm and cold, transitioning from neutral-warm-neutral-cold-neutral, for example.

Certain issues might make rhythm a difficult parameter to use in feedback design, some of which are perceptual, while others are technological. Because the thermal sense is focused on the overall extent/intensity of stimulation, rather than precisely where or what temperature is being presented, it is more of a comparative sense: comparing current body and skin temperature to stimulatory temperature. Providing a stimulation of 35°C to skin that is 32°C may produce a sensation of warmth. But following this with a return to 32°C may not simply feel like 'returning to neutral' or the removal of warmth, it may actively feel like cooling. This may be because the skin can adapt to (change to and remain at) temperatures within the 6-8°C range of the neutral zone [223], and so the initial warming change may slightly increase the skin temperature, meaning the 'return' to 32°C is actually cooling the skin again. The net result may be that, should the user not perceive the initial direction of change (which may occur if the change is too small, too slow or they are distracted) it may

be difficult to determine if the rhythm is comprised of warming pulses (neutral-warm-neutral-warm) or cooling pulses (neutral-cool-neutral-cool). The "neutral" starting temperature might be within the skin's neutral zone, but the skin itself has no set baseline temperature to 'return to' as it changes fluidly within the range. Because of this, as the rhythms warm/cool, the skin will slightly warm/cool with them, with the net result being that the comparative sensation of warm and cool rhythms is the same: they both include warming and cooling changes.

Another issue might be how fast the thermal interface hardware can change the stimulating temperature. Even if it is limited to an ROC that is relatively fast, in terms of what is needed for perception (for example, 1-3°C/sec), then a single neutral-warm-neutral (or neutral-cold-neutral) 'round-trip' may take several seconds. If different rhythms (made up of differing pulse or inter-pulse durations) were to be used to convey different meanings, the user would need to wait several seconds for each pulse in order to distinguish rhythms. Feedback identification could be very time-consuming. This case assumes that all the stimulators being used to produce the sensation are producing the same temperature. If an array is large enough, for example a 2 x 2 grid, then more rapid changes in temperature can be produced, by alternating warmth/cooling of adjacent elements [202]. As mentioned before, thermal rhythm is a new idea in HCI and so warrants investigation to judge its merits as a feedback source. However, a time-consuming feedback channel is likely to be less useful than one of more immediacy. Because of this, and the potential perceptual problems mentioned above, rhythm was not used as a feedback parameter.

6.2.3.3 Bodily Location/Spatial Pattern

Spatial location has also been considered as a parameter for Tacton design [21], partly due to the more limited range of vibrotactile parameters compared to audio parameters in Earcons. As there are also limited ways of manipulating thermal sensation, spatial location may be of use going forward. As discussed in Chapter 2, thermal perception varies in sensitivity around the body. With the exception of high sensitivity in the palm and fingertips, sensitivity is generally better on the head and trunk and worse towards the extremities [36, 226]. In order to use body location or spatial location as a parameter, the locations would best be on the same part of the body, for example the forearm or the upper arm. Were they to be on different areas of variant perceptual fidelity, the same feedback stimuli may be perceived differently, such as feeling less or more intense. Thermal perception varies even between the front and back of the trunk [25] so perception of four stimulators placed at cardinal points around the waist, as used in Tactons by Hoggan & Brewster [98], may not be equal. Although the temperatures provided by the stimulators could be calibrated to feel similar, the

ability to detect and interpret relative changes would be different. In these situations, it may be possible to use very simple feedback, such as only presenting a simple warm or cold pulse; finer differences, such as judging varying extent of change, may not be perceivable.

The two-point threshold (TPT, the minimum distance at which two points contacting the skin are felt as two points rather than one) for thermal perception is much larger when the stimulation is radiant [25] compared to when in contact with the skin [221, 222]. TPT/thermal spatial localisation improves as the two contact points become warmer, cooler or more different [221, 222]. Spatial summation may make spatial location a difficult parameter to use. However, with sufficient distance between points (8cm or more on the forearm [150]) it may be a usable parameter. However, spatial location was not chosen for use; it is a less feasible or realistic parameter for use in everyday mobile HCI scenarios, as it would require the user to wear additional technology on the relevant body part. It was important to fit the hardware into a mobile phone-sized form factor.

6.3 Experiments 6 and 7 – Testing Detection and Comfort of Thermal Parameters

The first section of this chapter discussed which parameters of thermal stimulation might be suitable for conveying information in structured thermal feedback for mobile HCI. From these parameters three were chosen for further testing, to establish their suitability for use as feedback parameters: *direction of change*, *extent of change* and *rate of change (ROC)*. Note that, from here on, the display of these three terms in italics denotes the experimental factors and specific stimuli used in the research, while display of the terms in normal text denotes the use of the terms in a general sense, such as the act of varying extent or rate of change. This section includes two user studies that were conducted to test detection, and the participants' subjective perception, of these thermal parameters. Experiment 6, described in Section 6.3.3, was conducted while the participants were seated, in order to provide baseline results. Because walking can negatively affect performance in certain tasks [11, 38, 204], and environmental temperatures can influence thermal perception [92, 229], both mobility and environmental temperature were considered separate variables to control and look at individually. Mobility was chosen as the first variable to study, and so the second study, Experiment 7, reported in Section 6.3.4, was conducted while participants were walking a set route in an indoor environment. Having two indoor studies would then provide ideal/baseline results that could be compared to results when participants are sitting and walking outdoors,

to establish the effects of both a) outdoor environments and b) walking in outdoor environments.

6.3.1 Experimental task

The experimental design employed by both studies is very similar to many psychophysical studies on identifying thermal thresholds [67, 71, 133, 224]. In these experiments, stimuli that vary in their direction and rate of change[12] are presented to the individual and they are tasked with responding as soon as they detect a change in thermal sensation. In the traditional psychophysical studies, the extent by which the stimulator has changed temperature when the individual responds is recorded to show the minimum extent of change that can be detected. A smaller amount of detectable change indicates a more salient stimulus. However, identifying the smallest amount of detectable change was not of interest for the research described in this chapter, as the purpose was to determine how salient and comfortable set levels of each parameter are, to establish which can be used for feedback. Specifically, it was important to establish how subjectively intense each stimulus felt, to understand what range of intensities could be used to represent different information. Therefore, the salience of each stimulus was determined by 1) whether or not the change was detected (detection rate across all participants), 2) how quickly (after the stimulus began) the participant responded that a change was felt and 3) the participants' subjective rating of how intense/strong the stimulus felt. Higher detection rates, lower detection times and higher ratings of intensity would be considered indicators of higher salience.

Stimuli were presented to four body locations that could be associated with holding, transporting or interacting with a mobile device. Thermal perception varies around the body [36, 71, 81, 166, 226] and mobile devices can be held or placed in different locations on the person. Therefore, it was important to consider how perception of each of the potential thermal feedback parameters might vary when presented to different bodily locations. Note that location is not being used as a feedback parameter, as perception is being tested at different locations. The fingers and palm of the hand are the most logical choices, as mobile phones are held against the palm and gripped with the fingers. The thenar eminence (the bulbous area of skin adjoining the thumb) was chosen specifically over the central palm due to its apparent increased sensitivity to thermal stimuli [118]. The dorsal surface (hairy skin) of the forearm was chosen partly as it has differing thermal sensitivity to both the thenar

[12] There are generally no set extents of change used, the stimulator simply continues warming or cooling until a response is received from the participants indicating that they have detected a change in temperature.

eminence and the fingertips [66], but also as it is conceivable that a watch or wrist band containing thermal elements could be worn, which is connected to the user's phone or the network wirelessly. Finally, mobile devices, such as phones and MP3 players, are commonly held against the upper arm by elastic straps while engaged in sporting activities such as running, cycling and weight lifting, and so the dorsal surface of the upper arm, around the bicep, was used as well. Such a device could include thermal elements in the same way as one worn around the wrist. While the breast pocket and trouser pockets are also commonly used, there are layers of fabric through which any thermal stimulus would have to pass to reach the skin. Subsequent research has suggested that the thermal conductivity of intermediary fabrics has a strong influence on the perception of thermal stimuli [75], and as the immediate concern was with identifying the baseline of perception and comfort, direct skin contact was important.

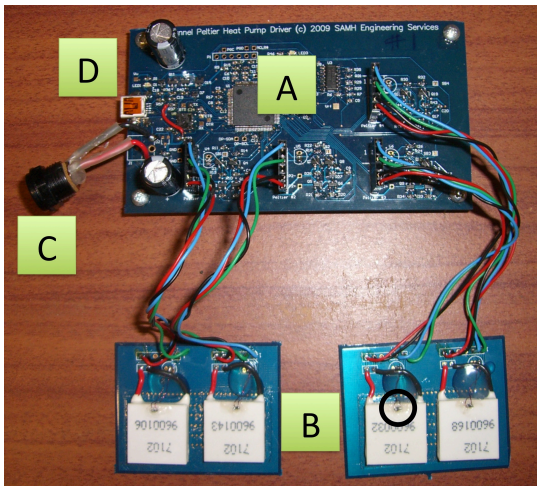

Figure 6-1: Hardware used to produce thermal stimuli: Microcontroller (A), 4 Peltier modules (B), battery connector (C) and USB connector (D). Thermistor for measuring Peltier temperature is ringed in black.

6.3.2 Apparatus

The hardware used to produce the thermal stimuli is described in detail in Appendix D, and is shown in Figure 6-1. It consisted of a microcontroller board, which received commands from a host PC over USB, and which controlled the temperature output of two 2 cm^2 Peltier modules. The Peltiers could output a temperature from -20°C to +45°C, at a resolution of 0.1°C. While the board was capable of controlling four Peltier modules, only two Peltiers were used for the experiments, to reduce size and better fit into a compact mobile form. The controller board was powered by four 1.2V AA NiMH batteries (2650mAh). The board was

controlled from a MacBook Pro running a Pygame application, which controlled presentation of stimuli and recorded responses from participants. A 22" external monitor was used to show the Pygame user interface (shown in Figure 6-5) and user input was received via a mouse.

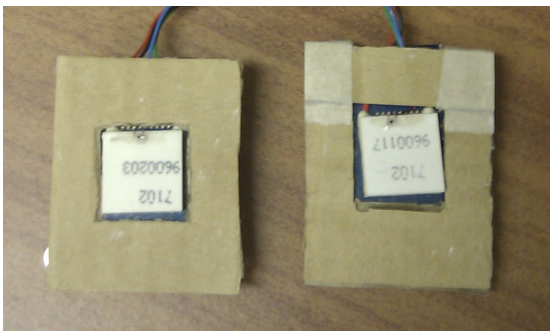

Figure 6-2: Peltier modules with cardboard cover as barrier between potentially warm circuit board and participants' skin.

To produce cold sensations on the exposed side of the Peltiers, heat is transferred to the side bonded to the circuit board. While heat sinks were attached to the underside of the circuit board to dissipate heat, some heat spread onto the circuit board itself, causing it to warm slightly. As it was possible for the board to contact the skin during the experiment, this may have presented the participant with either a larger area of warmth during warm stimuli, or conflicting cold and warm sensations during cold stimuli. As area of stimulation [223] and conflicting presentation of warm and cold [63, 64] influence thermal perception, thin cardboard sheets were placed over the exposed circuit boards as a barrier between skin and board (see Figure 6-2). The thermal conductivity, *k*, of cardboard is quite low, at 0.21 $W \cdot m^{-1} \cdot K^{-1}$ (watts (W) per meter Kelvin (K)), so the warmth would not pass through quickly or easily.

6.3.3 Experiment 6 – Seated Indoor Evaluation

The initial investigation looked at how well users could detect thermal stimuli while seated at a desk in an indoor usability lab. Participants sat in an office chair, with the apparatus lying on the desk in front of them along with the MacBook Pro, monitor and mouse.

6.3.3.1 Participants and Experimental Procedure

Fourteen participants (9 male, 5 female) aged 21 to 57 (mean = 29.2 years) took part in the evaluation, all studying or working at the University of Glasgow. All were right-handed and

were paid £6 for participation, which lasted just over an hour. The stimuli used in Experiments 6 and 7 are described above in Section 6.2.2.4. The skin was adapted to the 32°C neutral temperature for 1 minute at the beginning of each location condition and for 20 seconds between each stimulus presentation. Each stimulus in this set was delivered twice in a random order, giving a total of 24 stimuli (3 extents of change x 2 directions x 2 rates x 2 presentations) presented at each of the four bodily locations. Experimental instructions and data for all measures can be found in Appendix E.

	1	2	3	4	5	6	7
Intensity	Very Cold	Cold	Cool	Neutral	Warm	Hot	Very Hot
Comfort	Very Uncomfortable	Uncomfortable	Slightly Uncomfortable	Neutral	Slightly Comfortable	Comfortable	Very Comfortable

Table 6-2: Likert scales for subjective reports of stimulus intensity and comfort.

The Independent Variables were: *Direction of Change* (warm or cool), *Rate of Change* (1°C/sec or 3°C/sec), *Extent of Change* (1°C, 3°C or 6°C) and *Body Location* (fingertip, thenar eminence, dorsal forearm, dorsal upper arm).

The Dependent Variables were: *Number of Detections* (if stimuli were perceived), *Detection Time* (how long after the initiation of a stimulus that it was detected), *Threshold Size* (extent of stimulator temperature change, in °C, from neutral at the time the stimulus was felt), *Subjective Intensity of Stimulus* and *Subjective Comfort of Stimulus.* Table 6-2 shows the 7-point Likert scales used to record subjective reports of stimulus intensity and comfort, as used in similar research [32, 53].

Environmental Consideration

Environmental temperature influences skin temperature [92, 128], as high (>25°C) or low (< = 15°C) environmental temperatures can cause the skin temperature to shift from the neutral zone [92, 229]. Climate-controlled facilities were not available, so room temperature and humidity were recorded throughout the study, with a view to incorporating these into the analysis of results. It should be noted, however, that the stimulated skin under the Peltier was always adapted to the neutral 32°C between trials.

Procedure

The task was split up into four conditions based on body location, and was a within-subjects design: all participants took part in all four conditions in a counterbalanced order, and the

presentation order of all stimuli was randomised. The participant was seated at a desk upon which there was a computer monitor and mouse (see Figure 6-3). For the fingertip and thenar eminence conditions, the Peltier stimulator lay on the desk in front of the seated participant, facing up so that the users could lay their finger/hand on the stimulator, supported by a padded rest (see Figure 6-4 left and centre). For the forearm and upper arm conditions, the stimulator was held against the arm with an elastic fabric strip secured with Velcro pads. The stimulator was held between this strip and the skin (see Figure 6-4 right). Every effort was made to position the hands, stimulator and elastic strip in such a way as to fully contact the skin, to ensure maximum and equal areal stimulation across participants.

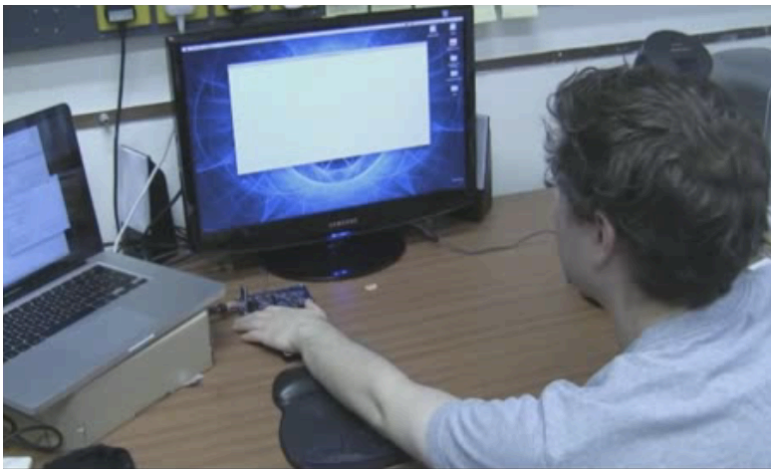

Figure 6-3: Setup for Experiment 6. Participant is resting his left thenar eminence on the Peltiers, supported by a padded rest (interface shown on screen (see Figure 6-6)).

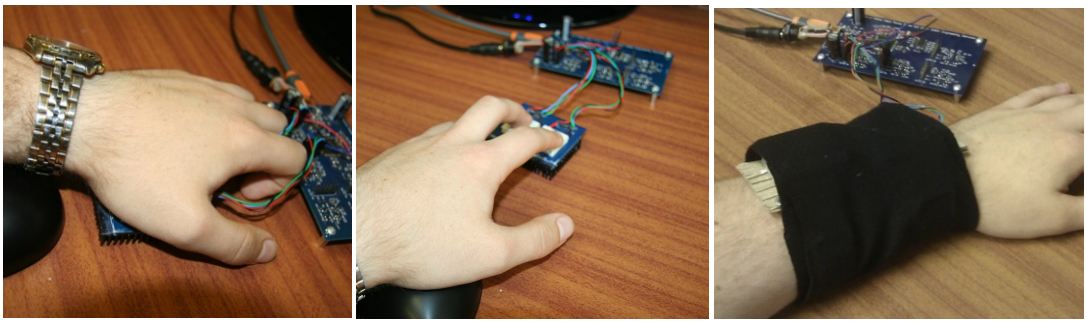

Figure 6-4: Stimulator sites for the thenar (left), fingertips (centre) and forearm (right) conditions.

The stimulator remained in contact with the skin of the non-dominant hand/arm for the duration of that condition. Green [66] found that participants reported higher intensity perceptions when they were in contact with a stimulator between successive stimuli, compared to removing their hand from the stimulator in between trials. At the start of each

condition the stimulators were set to the neutral starting temperature of 32°C for one minute to adapt the skin to this temperature. After the adaptation period, all 24 stimuli were presented in a random order.

A stimulus presentation comprised of 10 seconds of stimulus, followed by a return to the neutral temperature and 20 seconds of adaptation. There were no visual or auditory cues as to when stimuli were presented. Participants were instructed to click the right mouse button as soon as they felt a change in thermal stimulation, in any direction and at any intensity. Once this occurred, the temperature of the Peltiers was taken as the temperature that was detected (the threshold would be calculated by subtracting the 32°C starting temperature), and the time elapsed since the initiation of the stimulus was taken as the time-to-detection. At this point, 2 Likert scales appeared on screen (see Figure 6-5) asking the participant to rate how the stimulus felt, in terms of "intensity" (from "very cold" to "very hot") and "comfort" (from "very uncomfortable" to "very comfortable"). They then clicked on a submit button and another stimulus was presented after 20 seconds of adaptation at 32°C. As soon as the participant clicked the mouse button to register a change in sensation, the Peltiers were immediately returned to 32°C and the rating scales were presented.

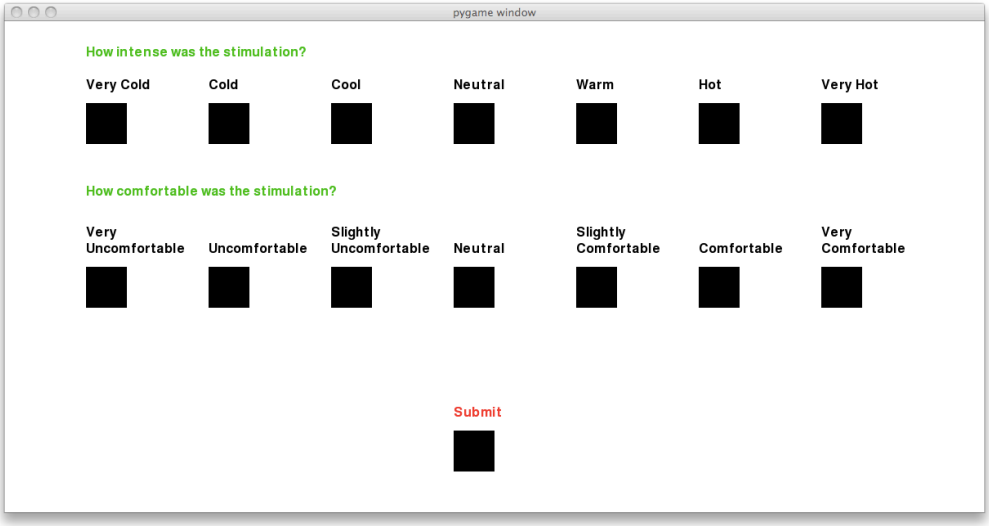

Figure 6-5: GUI screen used to get user subjective reports of stimulus intensity (top row) and comfort (bottom row).

If no response was received from the participant within 10 seconds of a stimulus being initiated, it was considered "missed" (not detected), at which point the miss was logged in the software and the Peltiers were set to 32°C in preparation for the next stimulus to be presented. If a stimulus was missed, *Detection Time, Threshold Size, Subjective Comfort and Subjective Intensity* data points were not recorded for that trial. This led to an uneven number

of data points for each condition. There were a total of: 14 participants x 4 Body Locations x 2 Directions x 2 ROC x 3 Extents of Change x 2 presentations = 1344 trials. This gave 1344 total data points for *Number of Detections*, as missed stimuli were still counted: 336 data points for each *Body Location*, 672 for each *Direction* and *ROC* and 448 for each *Extent.* Following the removal of data points following missed stimuli, the number of data points for the remaining Dependent Variable were: 290 for each *Body Location*, 288 for each *Direction* and *ROC* and 154 for each *Extent*, due to a high number of missed ±1°C stimuli.

6.3.3.2 Results – Experiment 6

The data for all measures violated the assumption of a normal distribution, following Shapiro-Wilk tests. Therefore non-parametric analyses were carried out. Effects of *Body location* and *Extent of Change* were analysed using a Friedman's test (non-parametric one-way ANOVA equivalent) while all pairwise comparisons (including for effects of *ROC* and *Direction of Change*) were conducted using Wilcoxon *T* tests. When Wilcoxon tests were used as *post hoc* pairwise comparisons following a significant Friedman's test, the necessary p-value for significance was adjusted using the Bonferroni correction, which dictates that $p = 0.05/N$, where N is the number of comparisons being made. For example, when comparing all combination pairs of the four Body locations, the necessary significance level is $p = 0.05/6 = 0.0083$.

Environmental Temperature

During the experiment, room temperature ranged from 20.5-25.3°C with an average of 23.6°C. Humidity ranged from 42-67% with an average of 48.5%. Perceptual research has shown that, within this range of room temperatures, skin temperature sits at neutral temperatures of 30-35°C [92, 229]. The neutral starting temperature used here would have been close to overall skin temperature, producing no sensation of warmth or cold. Therefore, room temperature and humidity were not considered when interpreting the results.

Number of Detections

A Friedman's test showed a significant main effect of *body location* on the number of thermal stimuli detected (χ^2 (3) = 17.56, $p < 0.01$). *Post hoc* Wilcoxon *T*-tests with a Bonferroni-corrected p-value of 0.0083 showed the thenar eminence (mean = 87.5%) produced significantly more detections than the fingertips (mean = 75.5%) and the upper arm (mean = 78.5%), but non-significantly more than the forearm (mean = 79%; see Figure 6-6).

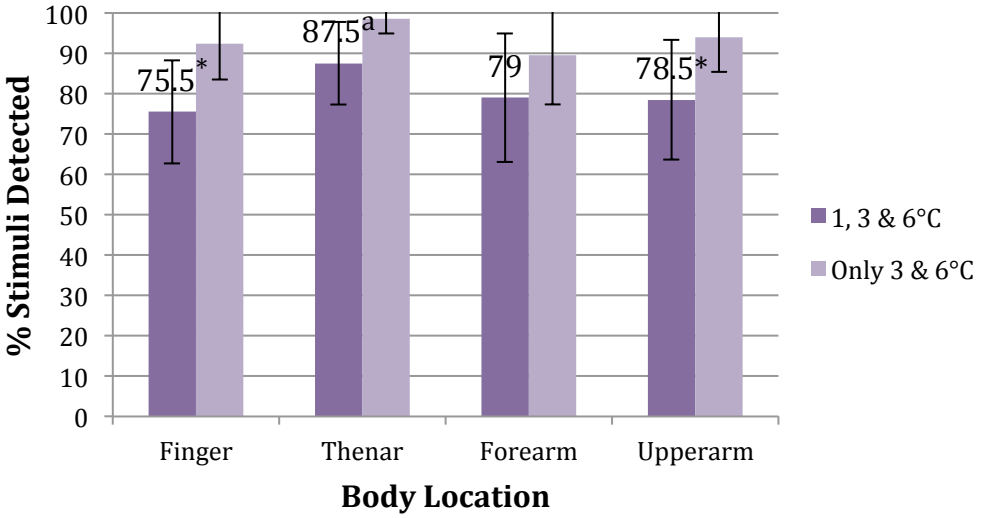

Figure 6-6: Mean detection rate of stimuli at the four body locations. Error bars show 1 standard deviation. 'a' indicates significantly higher value than '*', p <= 0.0083.

There was also a significant effect of stimulus *extent of change* on the number of detected stimuli (χ^2 (2) = 206.23, p < 0.001). Bonferroni-corrected Wilcoxon *T*-tests with an adjusted p-value of 0.0167 showed a significant difference between the numbers detected from all *extents of change*. The number increased as the *extent of change* increased, with means of 53%, 90.5% and 97% for 1°C, 3°C and 6°C extents (see Figure 6-7 and Figure 6-14, 'S' columns). There was no significant effect of *ROC* or *direction of change* on the number of stimuli detected. Both warm (mean = 79%) and cool (mean = 81.5%) stimuli produced similar numbers of detections. There were no significant interactions.

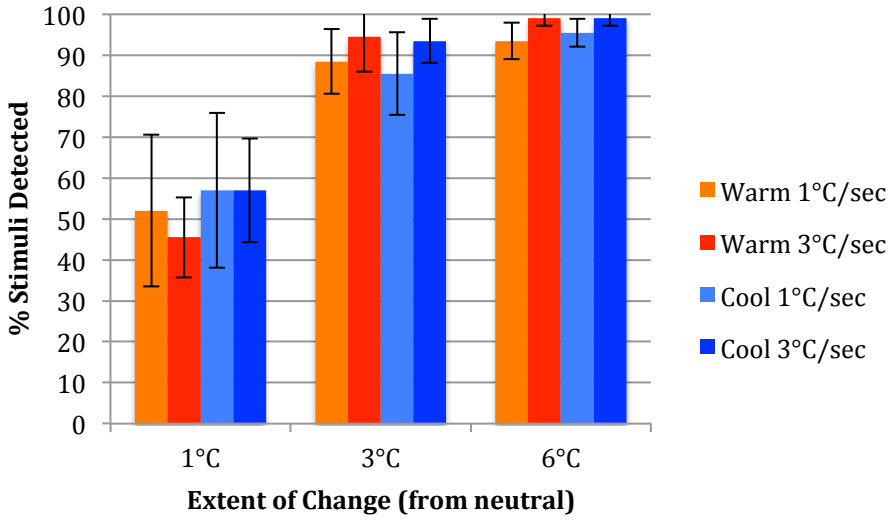

Figure 6-7: Mean detection rate at each extent of change and rate of stimulus change. Error bars show 1 standard deviation.

Time-to-Detection and Threshold Size

The time-to-detection and size of threshold are directly related and so are considered together. Friedman's test showed location had a significant effect on both time-to-detection (χ^2 (3) = 24.71, p < 0.001) and threshold size (χ^2 (3) = 41.65, p < 0.001). Wilcoxon *T* comparisons with an adjusted p-value of 0.0083 showed that the finger produced significantly longer times (median = 3.52s) than the thenar (median = 3.04s; *T* = 3672, p=0.001), forearm (median = 3.02s; *T* = 3205, p < 0.001) and upper arm (median = 2.87s; *T* = 2988, p < 0.001). Median times are shown in Figure 6-8.

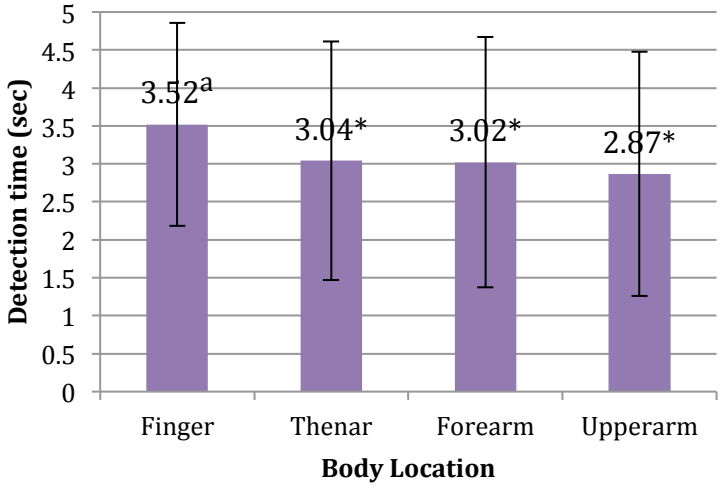

Figure 6-8: Median time-to-detection at each body location. Error bars show 1 standard deviation. 'a' indicates significantly higher value than '*', p <= 0.0083.

Similarly, Wilcoxon pair-wise comparisons with an adjusted p-value of 0.0083 showed that the finger (2.9°C) had a significantly larger threshold size compared to thenar (1.9°C), forearm (2.2°C) and upper arm (2.25°C). Significant effects of *direction of change* showed that warming stimuli were detected significantly more slowly (median = 2.91s) than cooling stimuli (median = 2.46s) (*T* = 15,643, p < 0.001) and, consequently, warm threshold size was significantly larger (median = 2.80°C) than cold threshold size (median = 1.85°C; *T* = 8327.50, p < 0.001).

Stimulus *extent of change* also had a significant effect on time-to-detection (χ^2 (2) = 63.01, p < 0.001; see Figure 6-9). Wilcoxon *T* comparisons with an adjusted p-value of 0.0167 showed significant differences in the time-to-detection between 1°C and 3°C (*T* = 2906, p < 0.001) and between 1°C and 6°C (*T* = 1638, p < 0.001) but not 3°C vs. 6°C. The amount of time decreased as the *extent of change* increased with median values of 3.67s, 2.59s and 2.30s for 1°C, 3°C and 6°C changes respectively. Friedman's analysis also showed a

significant effect of *extent of change* on threshold size (χ^2 (2) = 261.27, p < 0.001). Wilcoxon comparisons with an adjusted p-value of 0.0167 showed that all *extents of change* were significantly different from each other (p < 0.001) with median thresholds of 1°C, 2.7°C and 3.75°C for 1°C, 3°C and 6°C respectively.

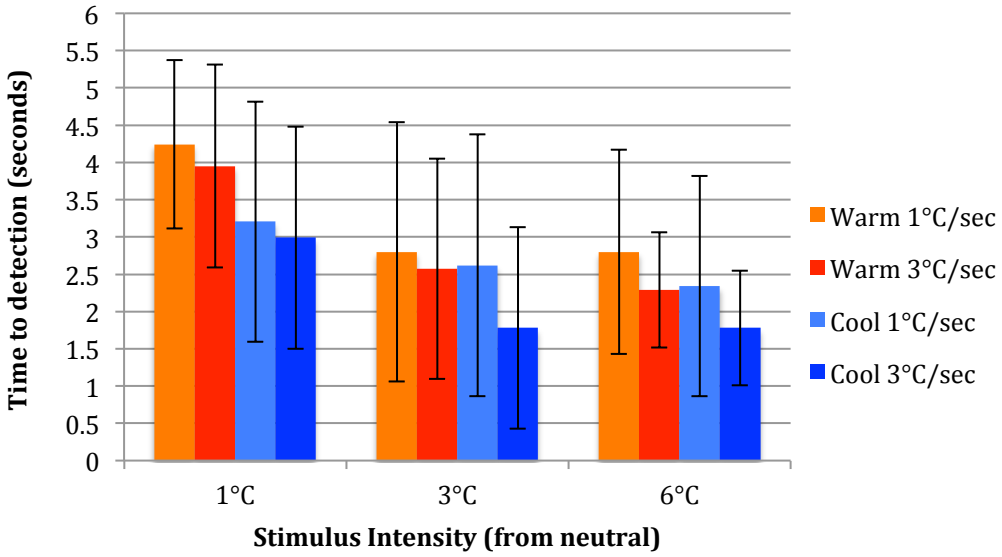

Figure 6-9: Median time-to-detection at each extent of change and rate of stimulus change. Error bars show 1 standard deviation.

Finally, *rate of change* significantly affected the time-to-detection (*T* = 12,461, p < .001) and threshold size (*T* = 8213.50, p < 0.001). The higher *ROC* (3°C/sec) produced a significantly lower time (median = 2.43s) than the lower *ROC* (1°C/sec; median = 3.04s) while the lower *ROC* had significantly lower threshold size (median = 1.90°C) than the higher *ROC* (median = 3.00°C).

Subjective Stimulus Intensity

Intensity ratings ranged from 0 to 3, where 0 denoted "Neutral" and 3 denoted "Very Intense". A Wilcoxon *T* test showed a significant effect of *ROC* on subjective stimulus intensity (*T* = 3935, p < 0.001). The higher *ROC* produced significantly higher ratings of intensity (median = 1.5) than the low rate of change (median = 1.00). There was also a significant effect of *direction of change* (*T* = 7263.50, p < 0.001), as warm stimuli were rated as significantly more intense (median = 1.5) than cold stimuli (median = 1.0).

A Friedman's test showed a significant effect of stimulus *extent of change* on perceived stimulus intensity (χ^2 (2) = 160.742, p < 0.001). Wilcoxon *T* comparisons showed that all

three *extents of change* were significantly different from each other ($p < 0.001$) with median ratings of 1, 1.5 and 2 for 1°C, 3°C and 6°C changes respectively (see Figure 6-10).

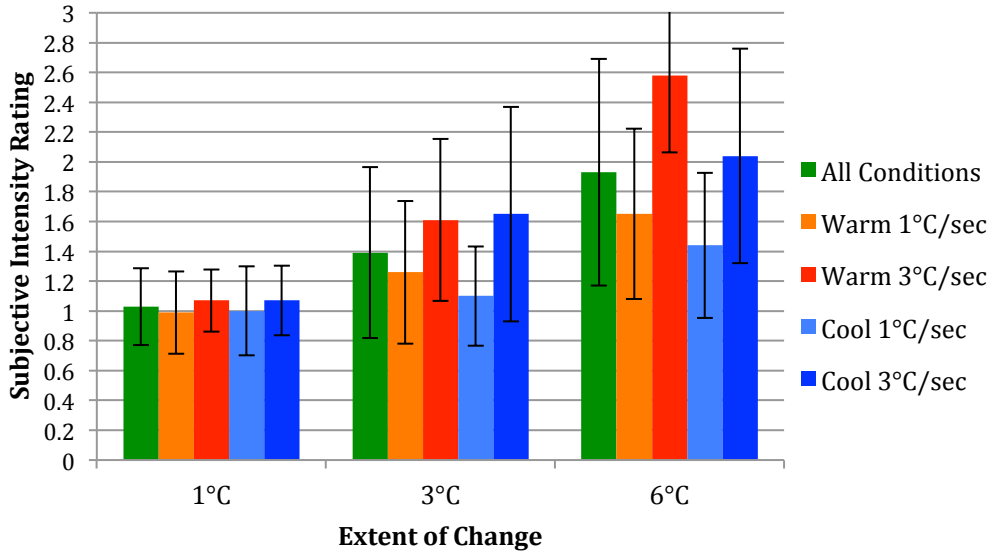

Figure 6-10: Median subjective ratings of stimulus intensity across each extent of change and each rate of change. Error bars show 1 standard deviation.

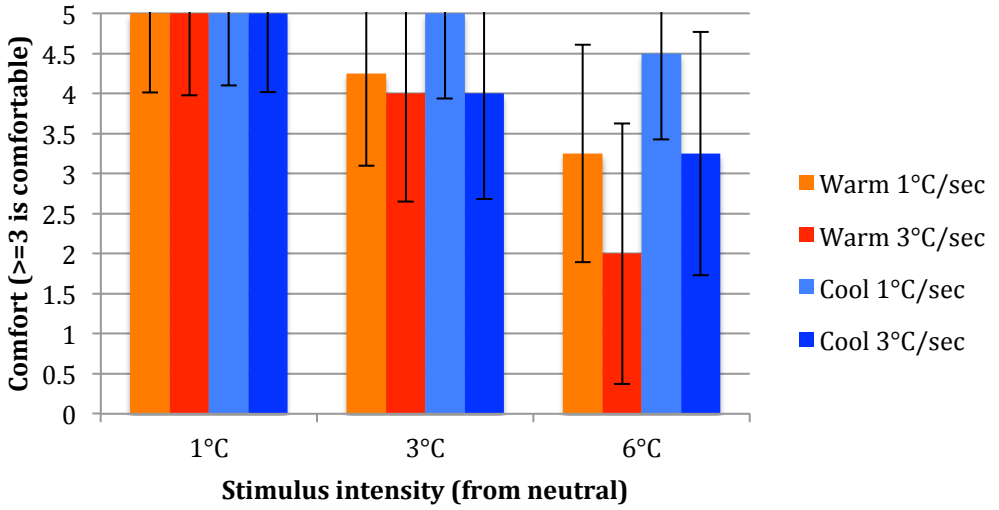

Figure 6-11: Median subjective comfort ratings at each extent of change and rate of change. Rating of >=3 indicates comfort.

Subjective Stimulus Comfort

ROC had a significant effect on subjective comfort ($T = 8775$, $p < 0.001$) with the higher *rate* (3°C/sec) producing significantly lower ratings of comfort (median = 4.0) than the lower *rate* (1°C/sec, median = 4.50). A Wilcoxon *T* also showed a significant effect of

direction of change on reports of comfort (T = 8922.50, p < 0.001). Warm stimuli had a significantly lower average comfort rating (median = 4.0) than cool (median = 4.5), see Figure 6-11. Friedman's analysis of variance by ranks showed that *extent of change* also had a significant effect on subjective comfort (χ^2 (2) = 48.46, p < 0.001). Wilcoxon T comparisons with an adjusted p-value of 0.0167 showed significant differences in subjective comfort between all *extents*: 1°C vs. 3°C (T = 2107.50, p < 0.001), 1°C vs. 6°C (T = 1500.50, p < 0.001) and 3°C vs. 6°C (T = 4343.50, p < 0.001). Ratings of comfort decreased as the *extent of change* increased, with median values of 5.0, 4.0 and 3.0 for 1°C, 3°C and 6°C respectively.

6.3.3.3 Initial Discussion – Experiment 6

The results of Experiment 6 suggest that all of the chosen thermal parameters, direction of change, extent of change and rate of change (ROC) can produce salient and comfortable stimuli and so are suitable for use in thermal feedback design. However, not all levels of each parameter were as salient or as comfortable. 1°C changes from neutral had relatively low detection rates (53%) and long detection times (3.67s), even when changing at the faster *ROC*. This result also came from a relatively controlled and seated indoor environment; so if 1°C changes are difficult to detect in these situations, they may be even more difficult to detect when walking or outdoors. Therefore, 1°C changes may not be suitable for thermal feedback, at least when the starting temperature is in the middle of the neutral zone of skin temperatures (in this case 32°C). In contrast, both 3°C and 6°C changes were reliably detected (90.5% and 97% respectively) and relatively fast, at 1.8-2.8s, suggesting they are suitable for inclusion in thermal feedback.

Both warming and cooling had equal detection rates, at ~80% overall, but cooling changes were detected 0.5s faster, felt less intense and more comfortable. The finding that cooling changes were more quickly detected is in line with other research [223]. That cooling changes were reported as more comfortable may fit with the asymmetrical perception space, where the heat-pain threshold is nearer to the skins resting temperature than cold pain. Because of this, cooling temperatures would have to change more to become more uncomfortable or painful, and because the three *extents of change* used here were the same for both warming and cooling, the warming changes were less comfortable. The finding that cooling changes were less intense than their warm counterparts is interesting but may also be related to the asymmetry, because they were further from thermal discomfort. Despite the large differences, it would appear that both warm and cold changes are suitable for use in designing thermal feedback; however, the asymmetry may need to be taken into account. If the intention of feedback designers is to create both warming and cooling changes of equal

subjective intensity, then either cooling changes may need to change *more* or *faster*, or warming changes may need to change *less* or *slower*. User comfort also needs to be taken into account, as it may be easier to arouse discomfort when warming than when cooling.

Both *ROC* were reliably detectable, with the faster *ROC* being detected 0.6s more quickly but, due to changing faster, the stimulator had changed more before the participant recorded feeling a change. Stimulus comfort was rated from 0 ("Very Uncomfortable") to 6 ("Very Comfortable"), with 3 being "Neutral". Therefore, ratings of 3 or above were considered to indicate that a stimulus was considered "comfortable". The only stimulus to be rated below 3, on average, was warming by 6°C at 3°C/sec which had a median rating of 2 ("slightly uncomfortable"). Warming by 6°C at the lower *ROC* of 1°C/sec elicited an average comfort rating of 3.25, above the comfort level, which would suggest that the discomfort arose from the speed at which the stimulator changed. Faster thermal changes are often felt as more intense [185, 223], and more intense stimuli are often reported as less comfortable [53, 185].

The 38°C set-point temperature is far enough away from heat-pain threshold so the temperature itself was not dangerous, and the above-comfort rating of the slower 6°C change also supports its innocuousness. However, from the perspective of the thermal sense, the skin is receiving a fast warming change that is rapidly heading towards pain, so alarm bells may be rung within the perceptual system to notify the individual that the stimulus is potentially dangerous. The asymmetry of perception is also apparent here, as cooling by 6°C at 3°C/sec was rated as more comfortable than the same extent of warming. ROC and extent of change are inextricably linked, so while *ROCs* of both 1°C/sec and 3°C/sec appear to be suitable for use in thermal feedback, this is dependent on the *extent of change*. 1°C changes were too small for either *ROC*, 3°C changes were well detected at both *ROC*, and while 6°C changes were also well detected at both *ROC*, they were most comfortable at the slower 1°C/sec. Feedback designs should carefully consider how much to warm by at 3°C/sec changes.

In terms of the sensitivity of different bodily locations, the results of Experiment 6 show that the thenar is the most sensitive area, in line with previous psychophysical research [71]. This judgement was based on the highest overall detection rate and fast detection time. The arm locations had slightly lower detection rates, but comparable detection times to the thenar, while the fingers had both the lowest detection rates and the longest detection time. However, while the fingertips performed worst out of the locations chosen in this study, they still performed acceptably, with a detection rate only 3-4% lower than the arm locations (12% lower than the thenar) and a detection time only 0.5s longer than the other locations. Also, as can be seen from Figure 6-6, if only the more reliably salient 3°C and 6°C are

considered, the fingertips had a very high detection rate of 92.5%, slightly better than the forearm (89.5%), but still slightly worse than the upper arm (94%). This bodes well for thermal feedback for mobile devices, as the device therefore may not necessarily be held or be in contact with the hand for feedback to be perceived reliably.

From the subjective ratings of intensity, it was hoped that an impression of how different each stimulus felt, in terms of its overall strength, would form. As the thermal sense is focused primarily on the subjective extent or intensity of change, stimuli that are reported as feeling differently intense would be better for use in thermal feedback as they are more likely to be perceptually distinct. Stimuli that are more perceptually distinct may then be more likely to be differentiable and so able to convey different bits of information. The subjective ratings for each *extent of change* at each *ROC* are shown in Figure 6-10. The Likert scale used to extract the ratings ran from "very cold" (rating of 0) to "very hot" (rating of 6) with "neutral" in the middle. As the stimuli were bipolar, the ratings were converted into two 4-point scales, from 0 to 3, as shown in Table 6-3**.**

	0	**1**	**2**	**3**
Warming	Neutral	Warm	Hot	Very Hot
Cooling	Neutral	Cool	Cold	Very Cold

Table 6-3: Adjusted Subjective Intensity scales.

What is clear is that, while subjective intensity generally increases as either *extent of change* or *ROC* increases, *ROC* has a larger effect than *extent*; an interaction between the two, changing both *extent* and *ROC* together, produces the largest changes in subjective intensity, larger than the combined changes of both parameters individually. Interestingly, increasing *ROC* individually strongly affects intensity ratings of only 3°C or 6°C changes. As can be seen in Figure 6-10, changing from 1°C/sec to 3°C/sec barely affected intensity ratings for 1°C changes (ratings increased from 0.99 up to 1.07 respectively), but increasing the *extent of change* to 3°C (remaining at 1°C/sec) increased intensity ratings from 0.99 for 1°C change to 1.26. Increasing both *ROC* and *extent* (from 1°C change at 1°C/sec to 3°C change at 3°C/sec) resulted in intensity ratings of 1.61. This suggests that two stimuli may feel perceptually distinct if they vary by a few degrees Celsius (at least 2°C or 3°C) or in the rate that they change, but they will feel more different if both features vary simultaneously. Thermal feedback would benefit from employing such tactics as a way of producing distinct types of stimulus. This evaluation did not ask participants to make comparative judgements of stimulus intensity, nor judge whether they feel differently intense. Therefore, no conclusions can be drawn regarding what values of subjective intensity actually indicate

perceptual distinctiveness.

Conclusions

Research Question 4 asked:

"What parameters of thermal stimulation are most salient and comfortable when using equipment designed for mobile interaction?"

Experiment 6 was designed to test perception of thermal stimuli that varied along three parameters chosen for their potential to produce detectable changes in sensation. The purpose was to test perception using hardware designed for realistic mobile interaction, in order to judge which parameters, and which levels within those parameters, could produce reliably detectable and comfortable sensations, and so be used in designing structured thermal feedback for mobile HCI. In general, all three chosen parameters could provide salient and comfortable sensations, provided changes were large enough. Also, the thenar eminence was the most sensitive location for feedback. The research aim is to develop thermal feedback for mobile interaction and so it was necessary to understand how being mobile influences perception. By conducting Experiment 6 sitting indoors, a baseline level of perceptual acuity was established, which can be used in comparison to results when walking. Therefore, a second study was carried out while participants were walking a predefined route indoors.

6.3.4 Experiment 7 – Mobile Indoor Evaluation

6.3.4.1 Experimental Design

The experimental design and procedure were almost identical to that of Experiment 6. The Independent and Dependent Variables were the same, as were the stimuli, again starting from 32°C neutral. Even though 1°C changes were unreliably detected in Experiment 6, it was important to leave them in for Experiment 7 to see if detection varied when the participant was walking. Similarly, although warming 6°C at 3°C/sec was rated as slightly uncomfortable by participants in Experiment 6, it was necessary to also include it here to see if mobility affects this perception.

The only ways in which the experimental design diverged from Experiment 6 was in the apparatus used and the body locations. For Experiment 7, the Peltier microcontroller was connected over USB to a Samsung Q1 Ultra Mobile Personal Computer (UMPC) running

Windows XP, which was carried by the participant in a backpack (see Figure 6-13). The software ran on an HTC Nexus One Android mobile phone (Figure 6-12), which sent commands to the UMPC over Bluetooth to control the behaviour of the Peltiers. The phone also took input from the user regarding his or her subjective reports of stimulus comfort and intensity. The mobile phone was held in the participant's dominant hand, as all stimuli were presented to the non-dominant hand.

Figure 6-12: Nexus One mobile phone used during Experiment 7 for receiving participant responses and subjective reports of intensity and comfort.[13]

The other variant from Experiment 6 was that, in this experiment, only three *body locations* were used. The same dorsal positions on the forearm and the upper arm were used in this experiment. However, the Peltier devices were placed more in the centre of the palm of the hand, rather than the thenar specifically, to more realistically simulate the location that might be stimulated when holding a mobile phone. The fingertips had the poorest performance in Experiment 6, and securely attaching the Peltiers to the fingertips would have been problematic, and so that condition was not included in Experiment 7.

Participants and Procedure

Fourteen new participants (10 Male, 4 Female) aged between 23 and 41 (mean = 30.2 years) took part in the evaluation, and all were studying or working at the University of Glasgow. All were right-handed and were paid £10 for participation, which lasted just over an hour.

The task was identical to Experiment 6, other than being split into three conditions based on the location of stimulation. A within-subjects design was employed, with all participants taking part in all three conditions in a counterbalanced order and the presentation order of all stimuli was randomised. Each participant walked around a triangular route in an indoor

[13] Image from http://www.androidguys.com/wp-content/uploads/2010/03/nexus_one.jpg

office environment. For all conditions, the stimulator was held against the arm with the same elastic fabric strip secured with Velcro pads used in Experiment 6. Participants were instructed to press a virtual button on the phone screen when they felt a change in thermal stimulation, in any direction and at any extent of change. Once this occurred, the temperature of the Peltiers (threshold) and the time elapsed since the initiation of the stimulus was recorded. At this point the same 2 Likert scales appeared on the screen of the mobile phone to receive the participants subjective ratings of intensity and comfort of the detected stimulus.

There were a total of: 14 participants x 3 Body Locations x 2 Directions x 2 ROC x 3 Extents of Change x 2 presentations = 1008 trials. Unexpected, and unpredictable, issues arose with the external thermistors on the Peltier devices; moisture from skin lead to erroneous readings by the thermistors of the Peltiers' temperature, which occasionally led to the safety mechanism engaging and the Peltiers would stop responding. Therefore, not every trial of every condition was completed by every user. The issue was remedied by placing insulating tape over the external thermistors, allowing them to function as normal. As a result, 662 out of the 1008 stimulations were submitted to the participants. As was done in Experiment 6, data points for all Dependent Variables, other than *Number of Detections* were not used for analysis if a stimulus was missed. The number of data points for each condition and Dependent Variable are shown in Table 6-4.

Condition	**Dependent Variable**	**Number of Data Points Per Level of Condition**
Body Location	• Number of Detections • All others	• 183 • 121
Direction of Change	• Number of Detections • All others	• 224 • 133
ROC	• Number of Detections • All others	• 326 • 196
Extent of Change	• Number of Detections • All others	• 215 • 68

Table 6-4: Table showing total number of data points used for analysis, for each level of each experimental condition (Independent Variable).

Figure 6-13: Stimulator locations for forearm (left) and upper arm (right) conditions.

6.3.4.2 Results – Experiment 7

As in Experiment 6, data for all measures were not normally distributed, and so non-parametric analyses were used: Friedman's test and Wilcoxon *T* pairwise comparison (with Bonferroni-corrected p-value as appropriate).

Environmental Temperature

Room temperature ranged from 20.2-25.9°C, with an average of 21.8°C, and humidity ranged from 39-69%, with an average of 48.9%. As with the static experiment, the neutral starting temperature was close to overall skin temperature, producing no sensation of warmth or cold and so humidity and temperature were not taken into account when conducting the analysis.

Number of Detections

Friedman's analysis found no effect of *body location* on the number of stimuli detected, with mean detection rates of 65%, 63.2% and 58.9% for the palm, forearm and upper arm, respectively. There was a significant effect of *extent of change* on number of detections (χ^2 (2) = 134.105, p <0.001). *Post hoc* Wilcoxon *T* comparisons with an adjusted p-value of 0.0083 showed significant differences in the number of detections between all *extents of change*: 1°C vs. 3°C (T = 6837.0, p < 0.001), 1°C vs. 6°C (T = 9100.0, p < 0.001) and 3°C vs. 6°C (T = 2251.5, p<0.001). Median detection rates were 28.44%, 69.68% and 85.78% for 1°C, 3°C and 6°C, respectively. There were no significant main effects of either *ROC* or *direction of change* on number of detections. Mean detection rates for *ROC* were 59.5% for 1°C/sec and 65% for 3°C/sec, while means for each *direction* were 58.9% for warming and 62.3% for cooling.

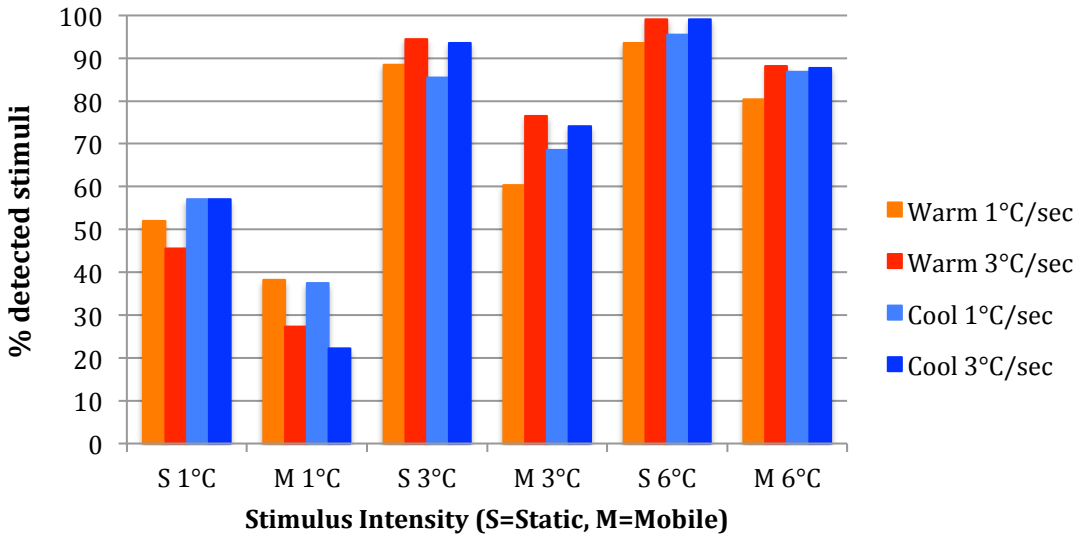

Figure 6-14: Detection rates at each extent of change and rate of change for both Static (S) and Mobile (M) studies.

Time-to-Detection and Threshold Size

Friedman's analysis of variance by ranks was used to analyse the effect of *location* and stimulus *extent of change* on the average time-to-detection. *Location* did not have a significant effect on time-to-detection, with median times of 3.15s, 3.28s and 3.04s for the palm, forearm and upper arm, respectively. *Extent of change* did have a significant effect on time-to-detection (χ^2 (2) = 36.102, $p < 0.001$). Wilcoxon *T* comparisons with an adjusted p-value of 0.0167 showed significant differences in the time-to-detection between the smallest *extent of change* and the two larger *extents*: 1°C vs. 3°C ($T = 1687.5$, $p < 0.001$), 1°C vs. 6°C ($T = 1844.5$, $p < 0.001$), but not between 3°C and 6°C. The amount of time decreased as the *extent of change* increased, with median values of 3.74s, 3.07s and 2.95s for 1°C, 3°C and 6°C extents, respectively. Wilcoxon *T* tests were used to analyse the effect of *ROC* and *direction* on time-to-detection. *ROC* had a significant effect on time ($T = 12{,}472$, $p < 0.001$): the higher *ROC* (3°C/sec) produced a significantly lower time (median = 2.96s) than the lower *ROC* (1°C/sec; median = 3.35s). A significant effect of *direction of change* ($T = 6105.5$, $p < 0.001$) showed that warming stimuli were detected significantly more slowly (median = 3.28s) than cooling stimuli (median = 2.94s).

As stated earlier, time-to-detection is linked directly to the size of threshold. A Friedman's test showed a significant effect of *extent of change* on the size of threshold (χ^2 (2) = 79.993, $p < 0.001$). *Post hoc* Wilcoxon pair-wise comparisons with an adjusted p-value of 0.0167 showed that all *extents* differed significantly from each other ($p < 0.001$). Median threshold

sizes were 1°C, 2.3°C and 4°C for the 1°C, 3°C and 6°C *extents of change*, respectively. Friedman's analysis of variance by ranks showed no significant effect of *location* on the size of threshold, with medians of 2.6°C, 2.7°C and 2.3°C for the palm, forearm and upper arm, respectively. Wilcoxon *T* comparisons showed a significant effect of *ROC* on threshold size ($T = 14047.5$, $p < 0.001$): 1°C/sec had significantly lower threshold size (median = 1.90°C) than 3°C/sec (median = 3.00°C). A further Wilcoxon *T* also showed a significant effect of *direction of change* on threshold size ($T = 6820$, $p<0.001$). Warm stimuli produced significantly larger thresholds (median = 3.1°C) than cold stimuli (median = 2.1°C).

Subjective Stimulus Intensity

As with Experiment 6, intensity ratings ranged from 0 to 3, where 0 denoted "Neutral" and 3 denoted "Very Intense". A Friedman's test showed a significant effect of *extent of change* on perceived stimulus intensity ($\chi^2 (2) = 160.742$, $p < 0.001$). Wilcoxon *T* comparisons with an adjusted p-value of 0.0083 showed that all three *extents* were significantly different from each other ($p < 0.001$) with median/mean ratings of 1/1.09, 1/1.47 and 2/1.95 for extents of 1°C, 3°C and 6°C respectively. A Friedman's test showed no significant effect for *body location* on ratings of stimulus intensity, with mean ratings of 1.52, 1.63 and 1.68 for the palm, forearm and upper arm, respectively. Similarly, Wilcoxon *T* tests revealed no significant effect for either *direction* or *ROC*. Mean intensity ratings for each direction were 1.74 for warming and 1.57 for cooling, and mean ratings for each *ROC* were 1.47 for 1°C/sec and 1.76 for 3°C/sec.

Subjective Stimulus Comfort

A Wilcoxon *T* test revealed that *ROC* had a significant effect on subjective comfort ($T = 7364.5$, $p < 0.005$) with the higher ROC (3°C/sec) producing significantly lower ratings of comfort (median = 4, mean = 3.63) than the lower *ROC* (1°C/sec, median = 4, mean = 3.99). A Wilcoxon *T* test also showed a significant effect of *direction of change* on reports of comfort ($T = 6408.5$, $p < 0.001$). Warm stimuli had a significantly lower average comfort rating (median = 4, mean = 3.52) than cool stimuli (median = 4, mean = 3.89). Friedman's analysis of variance by ranks showed that *extent of change* also had a significant effect on subjective comfort ($\chi^2 (2) = 412.703$, $p < 0.005$). Wilcoxon *T* comparisons with an adjusted p-value of 0.0167 showed significant differences in subjective comfort between 3°C vs. 6°C ($T = 5712$, $p < 0.001$). Ratings of comfort decreased as the *extent of change* increased, with median values of 4 (mean = 4.1), 4 (mean = 4.06) and 3 (mean = 3.4) for 1°C, 3°C and 6°C respectively. A Friedman's analysis of *body location* also revealed a significant effect on subjective comfort ($\chi^2 (2) = 9.968$, $p < 0.01$). Pairwise Wilcoxon T tests with an adjusted p-

value of 0.0083 revealed significant differences between the forearm and upper arm (T = 39911.5, p = 0.005); there were no differences between the upper arm and palm of the hand. The median values of the subjective comfort were 4 (mean = 3.63), 4 (mean = 3.95) and 4 (mean = 3.71) for the upper arm, forearm and palm respectively.

6.3.4.3 Perception When Sitting (Experiment 6) vs. Walking (Experiment 7)

The detection rates, time-to-detection and threshold results were compared between sitting down in Experiment 6 and walking in Experiment 7. Mann-Whitney *U* comparisons were used for analysis.

The participants were not able to detect as many stimuli when walking as when sitting but the patterns of detection were approximately the same. Analysis showed that mobility significantly affected the number of stimulus detections (U = 313515.5, Z = 12.049, $p < 0.001$), with higher detection rate when sitting (86.2%) compared to when walking (62.1%).

Time-to-detection was significantly affected by mobility (U = 33452, Z = 7.330, $p < 0.001$), with walking detections (3.69s) taking significantly longer than sitting detections (3.07s). As threshold size and time-to-detection are interlinked, the size of threshold was also significantly affected by mobility (U = 69784.0, Z = 11.440, $p < 0.001$), again with a greater threshold size for walking (2.88°C) in comparison with sitting (2.53°C).

6.3.5 Initial Discussion – Experiment 7, Comparison to Experiment 6

The results from Experiment 7 were very similar to those in Experiment 6, and as such they go some way to validating both the results themselves and the feedback recommendations that arise from them. In general, stimuli were more difficult to detect when participants were walking; however, similar patterns emerged. As might have been expected, 1°C changes were poorly detected when walking, even more so than when sitting, with a detection rate of only 28.44% on average and long detection time (3.74s). Detection rates for 3°C and 6°C changes were much higher, at 69.68% and 85.78% respectively, but these values were still lower than 90.5% and 97% when sitting. Therefore, similar conclusions can be drawn about the suitability of different *extents of change* for thermal feedback across both studies: 1°C changes are unsuitable, while 3°C and 6°C changes are more suitable, albeit with slightly less confidence in 3°C changes. These larger changes were detected more slowly when walking than when sitting, however, at 3.07s (3°C) and 2.95s (6°C). Warming and cooling changes were equally salient when walking, at around 60% accurate, as they were when

sitting, albeit at a lower overall rate. This reinforces the conclusion that both directions are equally suitable for thermal feedback designs. The lack of an effect of ROC on detection rate also echoes the results when sitting.

Another prominent difference between the two experiments is the minimal effect *body location* had on all experimental measures during Experiment 7. During Experiment 6, *body location* significantly affected the number of stimuli detected, the time-to-detection and the size of threshold, but not subjective comfort or intensity. One reason for this difference may be the omission of the finger condition in the Experiment 7 walking study as, comparing the results for the three non-finger conditions in Experiment 6, measures of both number of detections and time-to-detection were very similar. The threshold sizes for all locations were significantly different; however, the fingertip still performed the worst of all areas. Therefore, it may have been the fingertip performance that skewed the sitting results. The other important difference between the walking and sitting environments was the finding that neither direction of change nor rate of change significantly affected subjective intensity while walking, whereas they both did while stationary. This is discussed in more detail below. Aside from these findings, all other variable relationships/effects were the same across both mobile and static studies and are discussed below.

6.4 Discussion and Conclusions

In order to put these results into context for thermal feedback design for HCI, this section will discuss each parameter's significance in terms of suitability for use in feedback, as well as how the relative sensitivity of the four arm locations could impact on presentation of stimuli. Because of the similar pattern of results between the static and mobile studies, unless otherwise made clear, all relationships/effects refer to both mobile and static results.

6.4.1 Comparison of Perception When Sitting and Walking

There were some large differences in how well stimuli were detected when walking, compared to sitting. The environmental thermal conditions were very similar in both the sitting and walking studies, and so the drop in identification rates is most likely to have come from the act of walking itself. This could lead to several issues: the stimulators may have made slightly uneven contact with the skin over time, reducing the strength of the signal; the participant would have to pay some attention to the environment to avoid obstacles and

walls, and stay within the route, removing attention from the thermal stimulator; walking may also act to slightly elevate skin temperature, which would influence how well the stimuli could be detected.

Overall, stimuli were much more difficult to detect when the user was walking. Therefore the same stimuli cannot necessarily be relied on to be equally salient when in different interaction environments. This means that stronger feedback is necessary to increase the likelihood of detection, especially when walking. This could be achieved by using faster rates of change, larger extents of change, or combination of both. If it were possible, larger stimulators would also increase the strength of the stimulus, increasing its salience. For the stimuli that were detected, detection times were also significantly slower when walking. The same methods for making stimuli more salient would likely act to make them detectable more quickly. In particular, increasing the rate at which the stimulator changes temperature would reduce reaction time.

Interestingly, stimuli were rated as equally intense when sitting and walking and were generally rated as being similarly comfortable when walking. If "stimulus intensity" is to be used as a parameter for feedback design (for example using low, moderate and strong intensities to represent proximity to a travelling destination) this result might suggest that the same feedback stimuli could be suitable for use both when sitting (for example, on a train) and walking (for example, around a building). However, despite being equally intense as when sitting, the stimuli were less likely to be detected when walking. Increasing the strength of a stimulus, as is suggested to make them more reliably detectable, may result in the stimuli feeling differently intense when sitting and walking, requiring adaptive feedback design for static and mobile interaction environments.

6.4.2 Stimulus Perception

Direction of change, extent of change and rate of change were identified as features of thermal stimulation that might be suitable for use in thermal feedback design. The purpose of the two evaluations in this chapter was to determine three aspects of several parameter levels, namely each level's:

- Salience
- Comfort
- Subjective intensity

The measures that relate most to judgements of 'salience' here are the number of detections, the time-to-detection and, to a lesser extent, the size of just noticeable difference (threshold). A higher number of detected stimuli, a faster detection or a smaller threshold size are all indications of higher salience. Parameter levels that are salient but uncomfortable to perceive would not be acceptable in an interface, and levels that are comfortable when detected, but less reliably detected, would be of little use. Therefore, stimulus comfort is also paramount. Finally, because the thermal sense is based primarily on the overall magnitude of sensation, rather than precisely where the sensation is or how long it lasts [223], the primary means of conveying information in thermal feedback may be through varying the subjective intensity of the stimulation. Therefore, it was important to measure how subjectively intense the different parameter levels felt, to get an understanding of what range of intensities could be produced. Recall that intensity values here ran from 0 (neutral) to 3 (very strong) and comfort ratings ranged from 0 (very uncomfortable) up to 6 (very comfortable) with a neutral value at 3. We therefore, considered any rating of 3 or above as indicating acceptable comfort levels for feedback design.

6.4.2.1 Rate of Stimulus Change

Previous research has suggested that stimulus detection is heavily influenced by the rate of stimulus change, with higher rates of change producing more salient stimuli [36, 61, 81, 133]. In this chapter, salience has been measured by a combination of the detection rate/likelihood and the time-to-detection. With these measures in mind, the position that increasing ROC increases salience was only partly supported by the results here, as, although the high *ROC* (3°C/sec) produced a significantly faster time-to-detection, both rates produced equal numbers of detections. Given the lower detection times, it seems that ROCs over 1°C/sec do not affect if a stimulus is detected as much as when it is detected. Indeed, stimuli are perceivable relatively quickly even at ROCs much lower than 1°C/sec (~0.1°C/sec), at least in laboratory conditions [133]. The faster time at least suggests that faster-changing stimuli are more immediately salient. Perhaps thermal interactions that are time-critical would be recommended to use faster ROC, to bring attention to the event more quickly.

The *high ROC* also produced significantly larger threshold sizes. Threshold size is inherently linked to the time taken to respond to a change so, while stimuli changing at the *high ROC* were detected faster than those at the *low ROC*, the high speed itself will have meant that the stimulator changed more by the time a response was recorded. Some research has shown that threshold size either stays relatively constant or slightly decreases as ROC increases up to about 3°C/sec at similar body locations [36], while other research has shown threshold to

increase as ROC increases from 1.4-3.9°C/sec [185]. These results were gained with different methods and apparatus, so it is not necessarily the case that precisely the same behaviour will be observed across experimental designs.

The *high ROC* was significantly less comfortable than the *low ROC*; however, the ratings were very similar, at 4-4.5 (out of 6). Ratings of 3 or above were considered 'comfortable' so, from a comfort point of view, both *ROCs* are suitable for use in feedback design. Perhaps a cause of the lower comfort, the *high ROC* also felt subjectively more intense than the *low ROC*, but only in the static setting, not while mobile. Temporal summation would cause faster changes in a small amount of time to be perceived as being more intense [223], which would explain both the higher intensity ratings while stationary as well as the lower comfort ratings (due to higher intensity). Why the *high ROC* was not considered more intense and yet still less comfortable in the mobile condition is unclear. As pure speculation, perhaps the distraction of walking and/or body movement meant that perception was limited to more noticeable aspects of the sensation, namely comfort/discomfort.

In terms of suitability for feedback design, it appears that increasing the rate of change brings no benefit to stimulus detectability, and it can reduce the subjective comfort of the stimulus. However, it can be used to produce faster-detected changes and, importantly for feedback design, it is a reliable means of strongly influencing the subjective intensity of the stimulus.

6.4.2.2 Direction of Change

In almost all measures, cold stimuli were more salient than warm stimuli. Although they produced a roughly equal number of detections, cold stimuli were faster to detect, produced smaller thresholds and were more comfortable. Cold perception has been found to be faster than warm [128, 185], producing smaller thresholds [133], and this was also found here. These three factors (time, threshold and comfort) make cold stimuli particularly appealing for feedback design as they require less power to produce a detectable amount of change and are detectable sooner, compared to warm stimuli. Warm stimuli were reported as subjectively more intense than cold during the static condition but not while mobile, a similar trend to ratings for ROC. Therefore, as the warm stimuli here were closer to painful levels, it may well be that these felt more intense than cold stimuli. But then, why this wouldn't be true when participants are mobile is unknown.

While both warming and cooling are suitable for feedback design, they do not afford equal capabilities. Care must be taken not to use too hot stimuli, or warm stimuli that change very

quickly, as these may be too intense and so less comfortable. Warming and cooling by the same *extent* produces sensations of unequal subjective intensity, meaning that, if intensity is used to convey information, different extents of change will need to be used to produce equal subjective intensities.

6.4.2.3 Extent of Change

Although most research on thermal perception focuses on the rate of change rather than the 'end-point' extent of change, extent was controlled and manipulated here as well, to identify set stimulus characteristics, or levels, that could produce detectable stimuli and so be used in structured feedback. Previous psychophysical work has suggested that 1°C stimuli are highly detectable at lower rates of change than those used here [61, 133], although it was hypothesised that such a low intensity may not be suitable for less controlled experimental situations with less highly trained users, which seems to be the case here. The results showed that increasing *extent of change* significantly increased the number of stimuli detected, and, as can be seen from Figure 6-14, detection rates for 1°C stimuli were very low, compared to those of 3°C and 6°C. Therefore, it seems that 1°C changes are not suitable for thermal feedback due to the low detection rate. Increasing *extent of change* also significantly increased threshold size and significantly reduced time-to-detection.

This result is, at first, slightly confusing. The *high ROC* led to higher threshold yet lower detection time, due to the rate at which the temperature changed, so the stimulator had changed more by the time a response came. But larger *extents of change* should not cause the same issue, as *ROC* remains constant across them. The issue may instead be down to reaction time. Median threshold increased from 1°C to 2.7°C to 3.75°C when sitting, and from 1°C to 2.3°C to 4°C when walking, for the 1°C, 3°C and 6°C changes respectively. It is apparent that the stimulator commonly reached, or almost reached, its full *extent of change* before a response was recorded when producing the 1°C and 3°C changes. These changes would take approximately 1 sec. and 3 sec. at the *slow ROC*, or 0.3 sec. and 1 sec. at the *high ROC*. Median time-to-detection when sitting was 3.67s for 1°C and 2.59s for 3°C, and when walking the detection times were 3.74s for 1°C and 3.07s for 3°C. So, while the time-to-detection dropped, the total time was still over, or close to, the total time needed to change to the full *extent of change*.

Perhaps unsurprisingly, each *extent of change* felt significantly more intense than those beneath it, because, as just illustrated, the stimulator changed considerably more by the time participants tended to respond to a change. This is a useful finding as it suggests that, in agreement with Wettach *et al.* [239], differing levels of warmth or cold feel perceptually

different, and so thermal feedback need not be simply a dichotomy of ‘warming’ or ‘cooling’. This is an important result for feedback design as it suggests extent of change can be used as a design parameter in order to manipulate the subjective intensity of a stimulus.

6.4.3 Location Sensitivity

The three measures used for determining stimulus salience are also used to determine location sensitivity. Considering these measures only, the thenar eminence was shown to be the most sensitive location. In all measures in Experiment 6 it performed outright best or equal best with high detection rates, low detection speeds and small threshold sizes. For Experiment 7, the palm was most sensitive, in terms of number of stimuli detected. The thenar eminence is said to have slightly higher thermal sensitivity than the centre of the palm [118], which was stimulated in Experiment 6. The detection rate dropped by 22.5% from 87.5% for the thenar when sitting, to 65% on the palm when walking. This is similar to the average difference in detection rate between sitting and walking in general (24.1%), and so, taking the effect of walking into account, the palm does not appear to be significantly less sensitive than the thenar.

Detection measures for both arm locations were very similarly to each other, with slightly lower detection rates and larger thresholds than the thenar and palm, but having roughly equal detection speeds. The fingers performed worst out of all locations in all three of these measures. The epidermis of glabrous skin, particularly that on the pads of the fingers, can be up to 5 times thicker than that of hairy skin (for example on top of the arm), which increases delay in the thermal stimulus reaching thermoreceptors [240].

From these results the thenar eminence is recommended as the optimal location for thermal feedback out of the four locations tested; however, in mobile situations where contact with precisely the thenar is not practical, the palm is recommended and non-glabrous arm locations are also suitable. As the volar/glabrous skin on the forearm is suggested to have comparable sensitivity to the thenar [71, 118], this location may also be highly suitable. While the fingertips performed worst out of the four locations, they still performed relatively well, with a detection rate only 3-4% lower than the arm locations (12% lower than the thenar) and a detection time only 0.5s longer than the other locations. 1°C changes have been found to be unsuitable across the board, so if only the more reliably salient 3°C and 6°C are considered, the fingertips had a very high detection rate comparable to, or even higher than, other locations. The higher detection times may still be an issue for time-critical applications.

6.5 Limitations

This section discusses the main limitations of the research reported in this chapter, which should be considered when interpreting the results. The main limitations of Experiments 6 and 7 were: 1) the uneven number of data points for each condition, 2) the regularity of stimulus presentation and 3) the realism of the walking route.

6.5.1 Uneven number of data points.

As discussed in Section 6.3.3.1, if a stimulus was not detected by the participant (i.e., they did not click the mouse button to indicate they had felt a change), the data point for all measures other than *Number of Detections* was not recorded or used in the analysis, as there would be no data point for the subjective rating scales, and the threshold *Time* and *Size* would be invalid. This led to an uneven number of data points for different levels of each condition. The fingertips were the least sensitive area, and so had fewer data points than the other locations. Similarly, the slow *ROC* and the 1°C *Extent* were less well detected, leaving those levels with fewer data points than the others. When analysing the data in SPSS, the number of data points used to compare all levels of a condition is the smallest number of data points among the levels. For example, if one level, such as the 1°C *Extent*, has only 100 data points while the 3°C and 6°C *Extents* have 120, only the first 100 points from the 3°C and 6°C levels are used in the analysis, ignoring the remaining 20 in each. This means that the analysis does not always include all the data, giving a slightly misrepresentative view of the results. This was not an issue with the *Number of Detections* measure, as all data points were included for analysis.

6.5.2 Regularity of Stimulus Presentation

The stimuli were presented at regular intervals: every 30 seconds a stimulus was produced, with no variation in delay, and no 'fake' trials, where no stimulus was presented, were used. This was done so that all stimuli (apart from the very first) were preceded by the same length of adaptation at 32°C, to minimise any uneven influence of this factor on the detection of stimuli. However, participants were not told that stimuli would be produced at regular intervals, and no time frame (i.e., 30 seconds) was given for the time between each stimulus. Participants were merely told that "periodically the stimulator will heat up or cool down". It

is possible that participants could have come to a realisation of regular or rhythmic stimulus presentation, and so have begun to expect a stimulus at regular points, or become primed to detect one. While this may be true, a large number of stimuli were not detected, which suggests that, if present, expectation was not pervasive, and these missed stimuli would have resulted in varying and longer inter-stimulus delays, which could have acted against the habituation to a regular stimulus presentation cycle.

6.5.3 Lack of Realism in Walking Route

The walking route used for Experiment 7 was a simple triangular route around Computing Science offices. In the same way as influenced the walking route for Experiments 2 and 3, ethical considerations meant that obstacles had to be avoided. Therefore, unfortunately, the walking route required relatively low visual attention from the participants, meaning they could effectively focus all attention on any potentially arising thermal sensation. Walking and being outdoors were considered separate influences to test individually, so the lack of outdoor testing was not considered a limitation at this point. However, a more demanding course would have tested attentional limitations and tested whether a more demanding perceptual environment might have influenced perception of thermal changes.

6.6 Design Recommendations and Research Question 4

From the results of the first two evaluations, several interesting and important factors have been highlighted that should be considered when designing thermal feedback both in general and for mobile devices.

1) The thenar eminence is the optimal location for feedback, but non-glabrous arm locations are also suitable.

In measures of 'number of stimuli detected', 'time-to-detection' and 'size of threshold' the thenar eminence either performed outright best or equal best, showing it to be the most sensitive area. Similarly, the palm was superior when mobile. Although the forearm and upper arm suffered lower overall detection rates than the thenar/palm, they both performed well on measures of time-to-detection and threshold size, indicating suitability. The fingers performed poorly on all measures, showing them to be slower and more inaccurate in thermal perception.

Research Question 4 asked:

"What parameters of thermal stimulation are most salient and comfortable when using equipment designed for mobile interaction?"

The following guidelines were developed based on which parameters were found to be both sufficiently salient and comfortable when detected in both sitting and walking conditions indoors.

2) 1°C/sec and 3°C/sec changes are both suitable, but power requirements should be considered.

Both *rates of change* produced approximately equal numbers of detectable stimuli, with the best detection rates when using 3°C and 6°C stimuli. Each has its own advantages, however. 1°C/sec changes are slower to detect and require a larger change to be reliably detectable, but feel less intense and so feel more comfortable. Therefore, low rates of change may be best suited to ambient displays. 3°C/sec changes, however, are much faster to detect but sacrifice a degree of comfort without any benefit in likelihood of detection. Faster changes may be necessary when mobile, as they will increase likelihood of detection.

3) Warm and cold stimuli are both suitable for use.

Although both warm and cold stimuli are equally detectable, cold stimuli are faster to detect, require less change to detect and are more comfortable as they feel less intense. Warm stimuli should be used carefully or more subtly as they are generally less comfortable and feel more intense. A potentially problematic effect of using Peltier-based apparatus for producing cold feedback is that of heat sinking. As the skin side cools, the other warms, which could then potentially increase the temperature within the housing/body of the device. Drawing heat away from this side of the Peltier would be necessary for safe and effective use.

4) Extent of change or 'end-point' can be used as a parameter for feedback design as different intensities appear perceptually different.

This would allow for many levels of warm and cold to be used for event semantics, such as the priority of received messages, proximity to destinations during navigation or the progress

of file downloads. 1°C extents are best avoided despite their low power costs and high comfort level, due to high levels of missed stimuli and slow detection speed. Finally, larger extents of change are best used at lower rates of change, as this will minimise discomfort.

The contributions of the research in this chapter are:

- It constitutes the first detailed investigation into the perception and design of thermal feedback for mobile HCI. It has identified which features of thermal stimulation might be suitable for thermal feedback in mobile HCI
- Thermal stimuli are less salient in realistic HCI scenarios than in controlled psychophysical lab studies, and walking reduces stimulus salience further.

7 Conveying Multi-Dimensional Information Thermally

7.1 Introduction

Thermal stimulation has been used to convey information in Virtual Reality (VR) simulations, specifically the thermal conductance characteristics of various materials. However, this is a very high-resolution thermal 'display' with very fine thermal resolution (changes as small as 0.001°C), high rates of change (10-20°C/sec) and a large number of stimulating elements (8-10) [128]. Research suggests that users may be able to identify materials uniquely based only on temperature change [111] but this was achieved with fixed-choice tests, a large thermal stimulator (16 cm^2) and large power supply, which is currently unrealistic for mobile interaction. Simpler feedback designs may be necessary to transfer information thermally in mobile HCI. A few attempts have been made to convey simpler information, but they are perhaps too simple, often merely conveying "warmth" or "cold" [142, 143, 174, 176], sometimes as a means of conveying emotional information [52, 58, 112]. While these simple designs may be more feasibly recreated in mobile interaction scenarios, they convey little information. Earcons [10, 16] and Tactons [20, 21] can convey two or three pieces of information non-visually, as a way of freeing visual attention from cramped mobile (or desktop) screens. They do so by using structured sounds or vibrations, where the form of the sound or vibration conveys information. Research Question (RQ) 5 asks:

"Can thermal stimulation be manipulated to convey multi-dimensional information?"

In order to convey information, the stimulation must first be reliably detectable, and for the feedback to be usable, the stimulation must be comfortable as well. Chapter 6 described the process through which forms of thermal stimulation were identified as being suitable for use in the design of structured thermal feedback. These were the forms of thermal stimulation that were most frequently detected by participants and that were considered comfortable to perceive, when the individual was both sitting and walking in an indoor testing environment. This research produced a set of suitable stimuli that could be used in feedback designs:

- Two *directions of change*: warming and cooling
- Two *extents of change*: 3°C and 6°C
- Two *rates of change*: 1°C/sec and 3°C/sec

Although there are three different features of thermal stimulation, they are only used to alter the sensation along two axes/spectra: *direction* (warm or cool) and *subjective intensity*. As discussed in Chapter 6, altering either the *extent* or *rate of change* will alter the perception of the stimulus, making it either more intense (by increasing the *extent* or *rate*) or less intense (by decreasing *extent* or *rate*). Having identified a suitable set of stimuli, actually conveying information requires applying structure to the range of stimuli, so that unique forms of stimulation can be identified in isolation.

Information is conveyed by the attachment of meaning or significance to these unique forms. In Earcons [10, 16] and Tactons [21, 95, 96], two or three feedback parameters (such as rhythm and timbre/roughness) convey one piece of information each. Varying the quality of a single parameter alters the specific meaning of that piece of information. For example, the timbre of an Earcon may indicate the type of message being received, with a string instrument indicating an email and a horn instrument indicating a text message. In a similar way, using two thermal axes/spectra as feedback parameters, it may be possible to convey two pieces of information: one from the *direction of change* and one from the *subjective intensity*. Then, using two different *directions* and two or more *subjective intensities*, the specific meaning can be varied. From this position, three more steps were necessary in order to sufficiently answer RQ 5:

1. Create a set of two-dimensional thermal icons from the set of suitable stimuli
2. Test identification of the icons in a controlled (ideal) indoor setting
3. Test identification of the icons in more realistic outdoor and mobile settings

This chapter addresses each of these steps and, in doing so, answers RQ 5. The research studies in Chapter 6 indicated that walking has a negative effect on perception, and psychophysical research suggests that the humidity [56] and temperature [36, 92, 184] of our surrounding environment influences perception of thermal stimuli. This may have severe ramifications for the use of thermal feedback in general, but especially for mobile HCI, as perception of thermal cues may diminish and so the information conveyed by the feedback may be missed or misinterpreted. Therefore, because movement and environmental conditions may both influence perception and interpretation of thermal icons, the initial test

was carried out in an indoor setting whilst seated. Then, the same test was run when different participants were both sitting and walking outside. These two tests would provide an indication of the influence of 1) walking, 2) being outdoors and 3) walking outdoors on thermal icon identification.

As well as being potentially suitable for conveying information by itself, thermal feedback may also be a useful augmentation of existing non-visual feedback methods, such as Tactons. Spatial location is a highly identifiable Tacton parameter [21, 98], however, it may not be as feasible for everyday mobile interaction, as it would require that the user wear or attach extra equipment and carry it while going about their business. Roughness is a relatively reliable parameter, and more feasible for everyday use, but it has lower identification rates than either rhythm or spatial location [98]. One or more thermal feedback parameters/spectra could potentially be used to replace these less feasible (location) or reliable (roughness) Tacton parameters, to improve tactile information transmission. As thermal stimulation is said to potentially possess an inherent hedonic quality, it may also mean that a richer, more multi-faceted tactile experience could be created. Previous research has shown that sensory magnitude of a vibrotactile stimulus is influenced by skin temperature [46, 62, 256]. It was therefore considered useful to test identification of combined vibrotactile and thermal structured feedback, or 'intramodal' icons, as well as purely thermal icons. The purpose was to test if presenting information to two different tactile feedback channels simultaneously would impair perception/identification of one or both, either from a sensory perspective (interference of signals) or a cognitive perspective (attentional limitations).

Section 7.2 describes the design and creation of the set of thermal icons, as well as the set of intramodal icons, while the meanings attached to the icons for the identification study are described in Section 7.3. Section 7.4 describes the experimental task and the apparatus used during it. Section 7.5 describes Experiment 8, the first experimental study on identification of thermal icons and intramodal icons, in a controlled indoor location while sitting in a chair. Section 7.6 includes the follow-up Experiment 9, looking at identification of thermal icons when the participants were sitting outdoors and walking outdoors. Intramodal icons were not included in Experiment 9. Section 7.7 presents discussion on the research in the chapter, Section 7.8 discusses the limitations of the research and Section 7.9 presents overall conclusions.

7.2 Thermal Icon Design

With two parameters it would be possible to convey two pieces of information, one through the *direction* of the stimulation, and one through the *subjective intensity* of the stimulation. There are only two directions of change, warm and cool (from 32°C), and both were equally detectable in the first experiment (albeit with unequal intensity, comfort and detection time) so the two 'levels' of this parameter to be used were very straightforward. The main challenge was choosing appropriate *extents* and *rates of change* to produce stimuli of distinct *subjective intensity*. The perceptual study in Chapter 6 measured participants' subjective ratings of how "intense" each stimulus felt. This was recorded on a 4-point Likert scale (see Table 6-3) from 0-3, where 0 indicated a "neutral" intensity (very weak) and 3 indicated a very strong intensity. The subjective intensity ratings were significantly different (statistically) between each *extent of change* (medians of 1, 1.5 and 2 for the 1°C, 3°C and 6°C changes) as well as between both *rates of change* (1 and 1.5 for the 1°C/sec and 3°C/sec rates). However, no explicit comparison of stimulus intensity was conducted so, while these ratings were statistically different, they may not represent ratings of stimuli that are perceptually distinct, merely ones that are slightly different.

Figure 6-10 (in Chapter 6) shows the subjective intensity ratings for each of the *extents of change* at each *rate of change* from the perceptual study. As discussed in Chapter 6, increasing either one increases the subjective intensity, but increasing both at the same time increases intensity more than each individual increment combined. To increase the likelihood that stimuli will feel differently intense it was, therefore, better to use 3°C changes at 1°C/sec and 6°C changes at 3°C/sec to create two different levels of *subjective intensity*. 3°C change at 1°C/sec received median subjective intensity ratings of 1.5 and 1.0 for warming and cooling respectively, which were somewhere between the labels of "warm"/"cool" (rating of 1) and "hot"/"cold" (rating of 2). 6°C changes at 3°C/sec received ratings of 2.58 and 2.04 for warming and cooling respectively, between the labels of "hot"/"cold" and "very hot"/"very cold" (rating of 3).[14]

The two levels of the *subjective intensity* parameter were labelled as "moderate" for 3°C at 1°C/sec and "strong" for 6°C at 3°C/sec. Therefore, each icon parameter had two levels: 'warming' and 'cooling' for *direction of change* and 'moderate' and 'strong' for *subjective intensity*, giving four thermal icons:

[14] Note that the same change results in different intensity ratings depending on whether it is warming or cooling, a finding discussed in Section 6.3.3.3 in Chapter 6.

1. **Moderate Warmth**: warming by 3°C (32°C → 35°C) at 1°C/sec
2. **Strong Warmth**: warming by 6°C (32°C → 36°C) at 3°C/sec
3. **Moderate Cooling**: cooling by 3°C (32°C → 29°C) at 1°C/sec
4. **Strong Cooling**: cooling by 6°C (32°C → 26°C) at 3°C/sec

7.2.1 Intramodal Icon Design

Some research has shown that skin temperature can have an effect on the tactile perception of texture and vibrotactile stimuli. Green *et al.* [69] found that the magnitude of perceived roughness (created by grooves cut in an aluminium sheet) decreased as skin temperature dropped below 32°C. They also found that apparent roughness increased as skin temperature increased *above* 32°C, but this effect was smaller than the dulling effect of cooling. Green [62] and Yang *et al.* [255] both found that concurrent changes in thermal and vibratory stimulation influenced vibrotactile perception. Green [62] found a U-shaped relationship between skin temperature and the threshold amplitude of vibrations (the smallest amplitude that creates a detectable vibration). Greatest sensitivity (or smallest threshold amplitude) was at a skin temperature of 34°C, and sensitivity decreased (threshold amplitude increased) as skin temperature either dropped (to 24°C or 20°C) or increased (to 37°C, 40°C or 42°C). Also indicating a lowering of vibrotactile sensitivity, Yang *et al.* [255] found that cooling the skin from 32°C to 25°C reduced the perceived intensity of vibrations with a frequency of 150Hz or 250Hz (lower frequency vibrations of 30Hz were not affected). Therefore, skin temperature seems to have a strong influence on the perception of tactile stimuli, which is an important consideration in using thermal stimulation as an augmentation of vibrotactile Tactons, where warmth/cold and vibration would be presented simultaneously.

At this point it is necessary to talk briefly about feedback "modalities" and the terminology used to describe interfaces. Structured thermal+vibrotactile feedback is referred to as "intramodal icons", as both thermal and vibrotactile stimulation are from within the same tactile modality. The definition of what constitutes a "modality" in HCI varies. Some consider the boundary to generally fall between the main human senses, perhaps in addition to a small number of other channels such as physical gestures, and this is the approach followed here. From this perspective, modalities include visual, audio and tactile/haptic. Others believe that different channels within the same sense can be considered a modality, so that thermal and vibrotactile feedback would be considered two different modalities. "Multimodal" interaction tends to refer to an interface that uses more than one modality (within that author's definition of "modality"), although the modalities need not necessarily

be used simultaneously. Multimodal interfaces may simply provide alternative feedback methods to visual feedback. The simultaneous use of two different modalities in an interface is sometimes called "bimodal" (or even "trimodal" for three modalities) where, for example visual feedback is presented together with complimentary audio feedback. The term "crossmodal" has been used to describe audio and vibrotactile feedback share the same perceptual properties in both audio and tactile domains, such as the same rhythms, timbre/textural quality or spatial location [98, 99, 101, 102]. Icons can be learned in one modality and recognized in the other [98]. There is no accepted definition for modality, or an accepted set of them, and so the reasoning behind the naming of intramodal icons is made clear. They are "intra" because tactile feedback is considered one modality, and the two feedback channels are from within the same tactile modality. For a similar reason they are neither multimodal nor bimodal as, although they are presented at the same time, they are, again, from one modality.

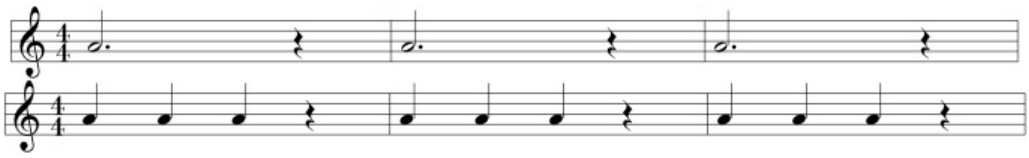

Figure 7-1: Vibrotactile rhythms used in the intramodal icons.

As there has been little research on the presentation of intramodal stimuli (compared to unimodal Tactons and Earcons, or bimodal stimuli combining two of audio, visual and tactile), there were no HCI guidelines on best practice for the design of intramodal icons. The thermal parameter chosen was *direction of change* as it is particularly salient, while the vibrotactile parameter chosen was *rhythm*, as this is one of the most easily identifiable parameters of Tactons [21]. Again, warming and cooling were used for the two levels of the *direction* parameter, in this case warming or cooling by 6°C at 3°C/sec, the large *extent of change* from the thermal icons. As identification of different vibrotactile rhythms is easier when they contain different numbers of notes [21], two 3-second rhythms were used, one made up of 3 long notes and one made up of 9 short notes. The *rhythms* in musical notation are shown in Figure 7-1. The four intramodal icons were:

1. Warming + Slow rhythm
2. Warming + Fast rhythm
3. Cooling + Slow rhythm
4. Cooling + Fast rhythm

The 3-second length was chosen as this is approximately the length of time before the full extent of warming or cooling occurs, with both the vibrotactile *rhythm* and the thermal change being initiated at the same time. This was done to test how well the thermal and vibrotactile stimuli can be processed when presented simultaneously, as skin temperature influences vibrotactile perception [62, 255].

7.3 Assigning Meaning to Icons

In order to test information transmission through thermal and intramodal icons, meaning had to be attached to each icon. A common means of conveying information on mobile devices is through non-visual notifications: sounds or vibrations that alert the user that an event has occurred and requires attention. This scenario has also been used for testing the information transmission capabilities of Earcons [165] and Tactons [20, 21]. Modern mobile phones allow the user to assign unique ringtones to individual contacts so that the user can know who is calling from the ringtone. This provides more information than simpler devices, which use the same ringtone for all calls, but it is still limited. Therefore, the thermal and intramodal icons were used to represent the arrival of an SMS or email, and convey two pieces of information about the message, namely the "Source" of the message and the "Importance" of the message, in a similar manner to Brown *et al.* [21]. These were then mapped to the two parameters of each icon type. As both types of icon have two levels for each of the two parameters, they can provide two possible meanings for the Source and Importance. The Source (i.e., the sender) could either be from a "Personal" source, such as a friend or family member, or from a "Work" source, such as a colleague or associate. The Importance (or priority) could either be "Standard" or "Important". This provided four message types:

1. Standard Personal
2. Important Personal
3. Standard Work
4. Important Work

The mappings between icon and message type are shown in Table 7-1. The two thermal icon parameters were *direction of change* and *subjective intensity*. *Direction of change* began at a neutral skin temperature of 32°C [128], warming and cooling from there. Warmth was chosen to represent the Personal messages, as there is evidence of an innate association

between physical warmth and interpersonal warmth, trust or closeness [58, 243]. Work messages were mapped to cold changes. The two levels of *subjective intensity* were labelled "Moderate" and Strong" and so the Important messages were mapped to subjectively stronger changes, as these larger changes have been found to be more attention-grabbing [223]. For the intramodal icons, *direction of thermal change* and *rhythm* were used as parameters. Because of the association between warmth and interpersonal closeness, *direction* was retained as representing the Source of the message, with warmth indicating Personal message and cold indicating Work messages. The Slow *rhythm* was used to indicate Standard messages, and the Fast *rhythm* was used to indicate Importance messages.

	Thermal	**Intramodal**
"Source" Parameter	*Direction of Change*	*Direction of Change*
Personal	Warm	Warm
Work	Cool	Cool
"Importance" Parameter	*Subjective Intensity*	*Tactile Rhythm*
Standard	Moderate	Slow Rhythm
Important	Strong	Fast Rhythm

Table 7-1: Mappings of thermal and tactile parameters to type of message received.

7.4 Experimental Task and Apparatus

The task used to test absolute identification of thermal and intramodal icons closely resembled that used by Brown *et al.* [21] for identifying multidimensional Tactons. The task required that participants identify which type of message has arrived, by interpreting the feedback presented to them and reporting which message they thought the feedback represented. Each message type was presented to the participant four times in a random order and the task was preceded by a training session where participants had the opportunity to learn the mappings of feedback-to-message type. Full details and procedures for the static and outdoor experiments is included in the relevant sections below.

The hardware used to present the thermal stimuli was the updated Design 2 of the hardware created by SAMH. It is described in full detail in Appendix D, and is shown in Figures 7-2 and 7-3. In short, the new design was considerably more compact and portable than the initial hardware design. The microcontroller board was much smaller and could connect to both mobile and desktop devices over Bluetooth, rather than USB. Two Peltiers were used to provide the stimuli, the same number and size as in the perceptual experiment in Chapter 6

and they had cardboard sheaths over the exposed heat sinks to avoid extra contact with the hand. 4 x AA batteries powered the microcontroller and Peltiers. For the vibrotactile feedback, an EAI C2 Tactor was used (www.eaiinfo.com), driven by audio files which are converted to vibration in the same way as Brown *et al.* [21] and Hoggan *et al*. [98]. The C2 is shown in Figure 7-2, right.

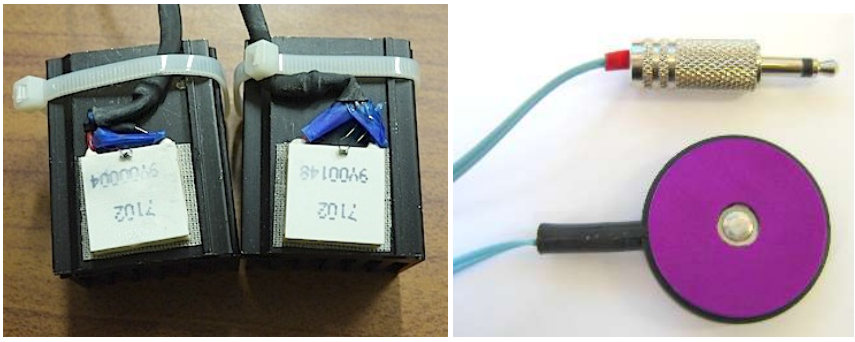

Figure 7-2: Peltier modules used to produce thermal stimuli (left) and the EAI C2 Tactor vibrotactile actuator (right).

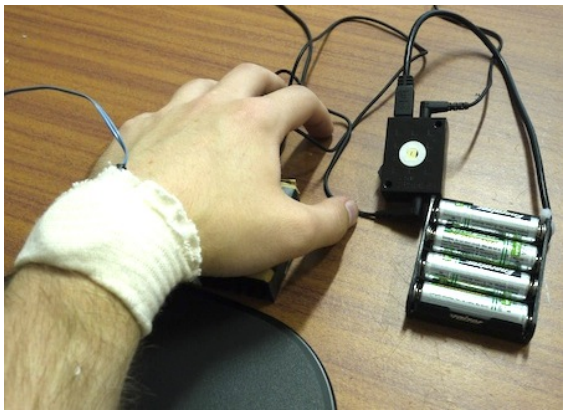

Figure 7-3: Experiment 8 apparatus, with Peltiers under palm and C2 under white elastic strap.

7.5 Experiment 8 – Identification of Thermal Icons When Sitting Indoors

The first study, Experiment 8, looked at identification of thermal and intramodal icons when the participants were sat indoors at a desk. The Peltier microcontroller was controlled by a MacBook Pro running Windows 7, and a GUI was presented on the laptop screen, which the participants interacted with via a PC mouse using their dominant hand. The feedback was presented to the non-dominant hand, in the manner described below.

7.5.1 Design and Procedure

Twelve participants took part (7 male, 5 female), aged from 18 to 43 (mean 25.08), and who were all from within the University of Glasgow and paid £6 for participation. The evaluation had two conditions: one identifying thermal-only icons (Thermal condition) and one identifying intramodal thermal+vibrotactile icons (Intramodal condition). It was decided not to include a vibrotactile-only condition, as identification of multi-dimensional Tactons has already been conducted successfully while sitting indoors in other research [19-21]. The procedure was the same for both conditions and participants took part in both, with the order counterbalanced. For both conditions, the thenar/lower palm was chosen as the site for thermal stimuli as it was the most sensitive of the locations in the perceptual study. Participants laid the thenar/lower palm of their non-dominant hand on top of two Peltiers, with the arm supported by a padded rest. During the intramodal condition, they also had the C2 contacting with the top of the non-dominant wrist, secured by the elastic strap. Figure 7-3 shows the experimental setup. Although it was important to have both stimuli (thermal and tactile) presented close to the same location (as both would theoretically be presented from the mobile device itself), it was not feasible to have both presented to the palm of the hand in this case, due to the size and placement of the Peltiers.

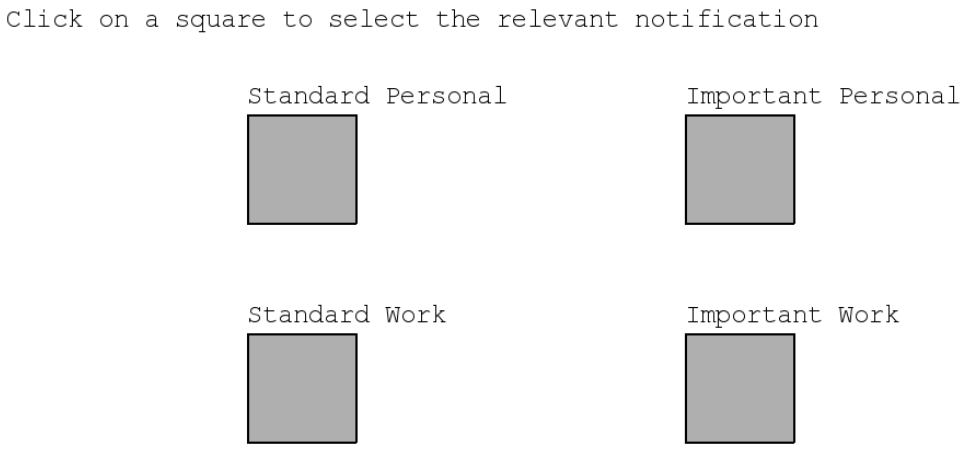

Figure 7-4: GUI shown during the training session and the main experimental task.

Each individual condition started with 60 seconds of adaptation, where the palm was rested on the Peltiers and they were set to the neutral starting temperature of 32°C. Participants first completed a 10-minute training session. During this time they were at liberty to feel each feedback/icon a number of times in order to learn the mappings of icon-to-message type. Four buttons were shown on the PC screen with the label of the corresponding message type (see Figure 7-4). When clicked, the relevant icon feedback was produced. In both conditions, the Peltiers changed to the relevant temperature and remained there for 10 seconds,

before returning to neutral for 30 seconds to ensure the skin was back to the neutral 32°C for a sufficient time before the next stimulus was presented. During the intramodal condition the tactile rhythm and thermal change began simultaneously. The participant was made to wear a set of headphones to make any vibrations from the C2 inaudible.

For the full task, in each condition all four stimuli/message types were presented four times in a random order, giving 16 icons per condition. The same interface screen as during the learning period was shown as soon as the icon was initiated and the participants were asked to click the button corresponding to which message type they interpreted the icon as representing. As soon as a button was clicked, the Peltiers were returned to neutral for 30 seconds, after which the next random icon was presented. This repeated until all icons were judged four times.

The Independent Variables were Icon *Modality* (thermal, intramodal) and *Icon Type* (Standard Personal, Important Personal, Standard Work, Important Work). The Dependent Variables were: *Accuracy* (whether the right message type was identified) and *Identification Time* (IDT, the time taken to choose a message type). *Identification time* will give an indication of how long it takes participants to become confident in their identification of the icon. There were a total of: 12 participants x 2 Modalities x 4 Icon Types x 4 presentations = 384 trials. This gave 192 data points for each Modality, 96 data points for each Icon Type and 48 data points for each Modality + Icon Type combination condition (e.g., Thermal + Standard Personal). Experimental instructions and raw data for all measures can be found in Appendix F.

7.5.2 Results – Experiment 8

Shapiro-Wilk tests showed that the data from Experiment 8 violated the assumption of a normal distribution, and so non-parametric tests were used, specifically Friedman's test for the effect of Icon Type, and Wilcoxon *T* tests for the effect of Modality. *Post hoc* pairwise comparisons following a significant Friedman's test were conducted using Wilcoxon *T* tests, using the Bonferroni-correction of $p=0.05/N$, where N is the total number of comparisons. Pearson's product-moment correlation coefficient was used to investigate any relationship between trial number and Accuracy to identify learning effects. While a moderate positive relationship was found, it was not significant (r (14) = 0.457, $p > 0.05$). The overall mean Accuracy for two-parameter thermal icons was 82.8% (SD = 37.8). Mean Accuracy for the two thermal parameters individually was 85.4% for *subjective intensity* and 97.4% for *direction of change*. Mean Accuracy is shown in Figures 7-5 and 7-6 and the confusion

matrix for thermal icons is shown in Table 7-2. There was a significant effect of Icon on Accuracy ($\chi^2(3) = 8.730$, $p < 0.05$). Wilcoxon *T* tests, with a Bonferroni-adjusted p-value of 0.0083, showed a significant difference between Moderate Warmth (MW) and Strong Warmth (SW) ($T = 104$, $p = 0.005$). Mean Accuracy was 95.8% (SD = 20.2), 72.9% (SD = 44.9), 79.2% (SD = 41.0) and 83.3% (SD = 37.6) for the MW, SW, Moderate Cooling (MC) and Strong Cooling (SC) icons, respectively.

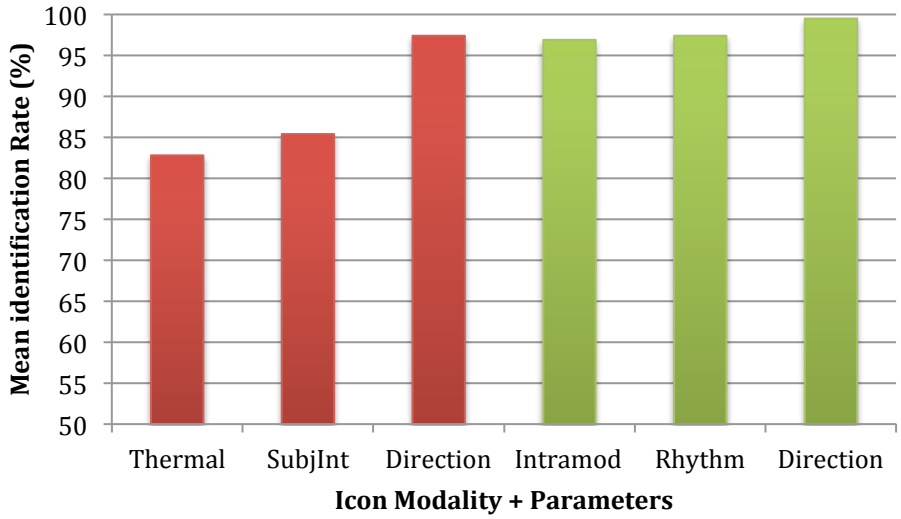

Figure 7-5: Mean Identification Accuracy for each Icon Modality (thermal, intramodal) as well as each individual icon parameter (SubjInt = Subjective Intensity).

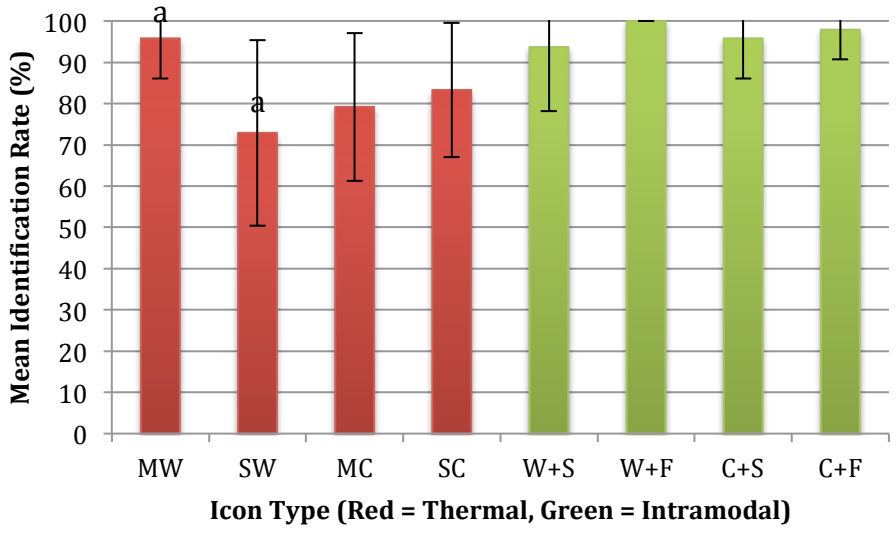

Figure 7-6: Mean Identification Accuracy for each Icon. Thermal icons Moderate Warmth (MW), Strong Warmth (SW), Moderate Cooling (MC) and Strong Cooling (SC) and intramodal icons Warm+Slow (W+S), Warm+Fast (W+F), Cold+Slow (C+S) and Cold+Fast (C+S). Error bars show 1 standard deviation. 'a' indicates significant difference, p = 0.005.

		Perceived Icon				
		Mod Warm	Strong Warm	Mod Cool	Strong Cool	
Actual Icon	**Mod Warm**	46	1	1	0	95.8%*
	Strong Warm	11	35	1	1	72.9%*
	Mod Cool	0	1	38	9	79.2%
	Strong Cool	0	1	7	40	83.3%

Table 7-2: Confusion matrix for the thermal icons. Right-hand column shows mean Accuracy for each icon; * = p< 0.0083.

The median IDT for each thermal icon was 5.40s (SD = 4.23), 5.34s (SD = 3.76), 5.61s (SD = 4.13) and 4.29s (SD = 4.33) for the MW, SW, MC and SC icons respectively. A Friedman's analysis indicated a significant effect of icon on Identification Time (χ^2 (3) = 11.075, p < 0.05), with MC and SC being significantly different from each other (Wilcoxon T = 854, p = 0.006) with an adjusted alpha of 0.0083. IDT are shown in Figure 7-7.

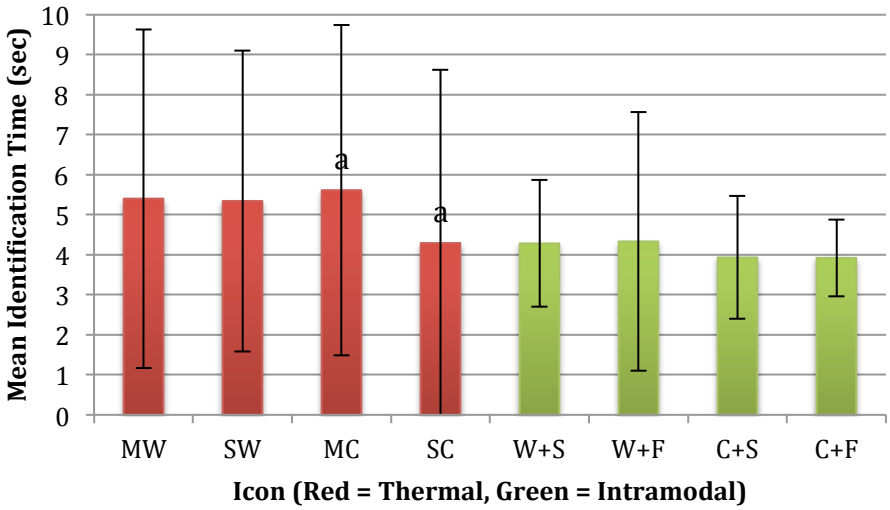

Figure 7-7: Mean Identification Time for each Icon. Thermal icons Moderate Warmth (MW), Strong Warmth, Moderate Cooling (MC) and Strong Cooling (SC) and intramodal icons Warm+Slow (W+S), Warm+Fast (W+F), Cold+Slow (C+S) and Cold+Fast (C+S). Error bars show 1 standard deviation. 'a' indicates significant difference, p = 0.006

The overall mean Accuracy for two-parameter intramodal icons was 96.9% (SD = 17.4). Mean Accuracy for the two intramodal parameters was 97.4% for *rhythm* and 99.5% for

direction of change (Figures 7-5 and 7-6). A Wilcoxon *T* test showed a significant effect of Icon Modality, as participants identified significantly more intramodal icons than thermal icons ($T = 558.0$, $p < 0.001$). Friedman's test found no effect of Icon on Accuracy ($\chi^2(3) = 3.750$, $p > 0.05$), with means of 93.7% (SD = 24.5), 100% (SD = 0), 95.8% (SD = 20.2) and 97.9% (SD = 14.4) for Warm+Slow (W+S), Warm+Fast (W+F), Cold+Slow (C+S) and Cold+Fast (C+F) icons, respectively. The median IDT for each intramodal icon was 4.12s (SD = 1.58), 3.77s (SD = 3.23), 3.63s (SD = 1.53) and 3.91s (SD = 0.96) for the W+S, W+F, C+S, and C+F icons respectively. A Friedman's analysis of the data indicated a significant effect of icon on IDT ($\chi^2(3) = 11.275$, $p < 0.05$), but no Wilcoxon *T* tests reached the adjusted level of significance ($p = 0.008$).

7.5.3 Initial Discussion – Experiment 8

The mean identification rate for thermal icons is high, with 82.9% Accuracy in identifying two pieces of information, suggesting thermal icons are a promising method of conveying information. This figure is higher than the 71% identification rate for two-parameter Tactons [21], although only two levels per parameter were used and Brown *et al.* used three. While the results are promising, future revisions should address the slightly more error-prone parameter of *subjective intensity (SI)*. Of the 33 errors, 28 confused the *SI* of the icon, resulting in 85% Accuracy for that parameter. The other 5 (of 33) errors confused warm for cold or *vice versa*, giving *direction of change* a much higher Accuracy of 97%. As seen in Table 7-2, a roughly equal number of *SI* errors occurred within both warming (12) and cooling (16), however, the pattern of confusion was different. All but 1 of the warm confusions felt subjectively less intense than was intended (where Important Personal was interpreted as Standard Personal) leading to a low Accuracy of 73% for the important Personal icon. In contrast, roughly equal numbers of cold confusions were perceived as less (7) or more (9) intense than intended.

Participants reported believing that either they became "less sensitive" to changes over time or simply that the stimuli became harder to differentiate. Looking at the frequency of errors over the course of the Thermal condition showed no pattern of increasing error with time, as a Pearson's product-moment correlation showed no significant correlation between trial number and Accuracy (in fact the coefficient was positive (0.457) suggesting an increase in performance over time). Therefore, even if the subjective ability to identify the *SI* got worse, it did not do so enough to make them indistinguishable. One way of increasing the subjective difference between stimuli would be to decrease the extent of change of the moderate

warmth/cold and increase the extent of strong cold. Increasing the strong warmth may move too close to the pain threshold, so this is not recommended. Another way might be to increase the area of stimulated skin [223].

The intramodal icons had a significantly higher mean identification rate than the thermal icons, at 97%, with 99% Accuracy for *direction of change* and 97% Accuracy for *rhythm*, a figure similar to rhythm identification in Tactons (96.9%) [21]. Therefore, presenting thermal and vibrotactile stimuli together does not appear to significantly hinder interpretation of either and so thermal changes may be a useful additional parameter to one- or two-parameter Tactons. The Identification Time (IDT) of thermal icons showed that, in line with previous research [74], the coldest icon (strong cool/Important Work) was the fastest to be identified, but only compared to the moderate cool, otherwise the times were comparable. However, overall the IDT were quite high, at 5-6 seconds. Because the two *subjective intensities* varied in their rate of change, they took different lengths of time to reach their extents. Participants may have waited a length of time to see how far the Peltiers changed temperature. In contrast, the amount of thermal change was irrelevant in the intramodal icons, and so identification was 1-2 seconds faster than thermal.

Experiment 8 examined thermal icon and intramodal icon identification when the individual was sitting indoors. The results were encouraging, as identifying two pieces of information from thermal icons was relatively easy and could be improved with adjusted designs. Therefore, thermal icons appeared to be similarly effective in conveying information as Earcons and Tactons. Also, presenting two different tactile stimuli simultaneously does not seem to produce confusion. However, the rationale for designing structured thermal feedback was for use in mobile interaction. A necessary next step was testing identification while the user is walking and in outdoor environments.

7.6 Experiment 9 – Identification of Thermal Icons When Sitting and Walking Outdoors

The second study looked at identification of thermal icons and purely vibrotactile Tactons when the participants sat and walked in an outdoor environment. While intramodal icons were identified well in Experiment 8, the decision was taken to compare thermal icons to a more established non-visual feedback method for mobile HCI. As mobility and

environmental temperature are considered two different influences, identification of thermal icons and Tactons was tested in two different scenarios: sitting outside and walking outside. If thermal icons are to be judged in terms of potential usefulness for mobile HCI it is necessary to know how well they can be identified in these more realistic scenarios. The apparatus used to produce the thermal stimulation was the same as in Experiment 8 (see Appendix D). However, to facilitate mobility, all the software was run on a Google Nexus One Android mobile phone (see Figure 7-8). The Nexus One communicated with the Peltier apparatus over Bluetooth and the Peltiers themselves were physically attached to the back of the Nexus One in order to make contact with the palm of the left hand, which held the device (see Figure 7-8). The thermal hardware was powered by 4 x AA batteries and both the microcontroller box and the battery pack were placed in a small travel bag that the participant carried over their right shoulder, across the body to the left side.

Figure 7-8: Experimental software ran on a Google Nexus One Android mobile phone (right). The Peltier apparatus was attached to the back of the Nexus one, to contact the palm of the hand holding the device (left).

Vibrations for the Tactons were produced by the EAI C2 Tactor (see Figure 7-2), which was attached to the top of the left wrist by an elastic strap, identical to that used in Experiment 8. In Experiment 8 the audio files for the vibrations were played to the C2 through the headphone socket of the MacBook Pro, which could produce a loud enough volume/amplitude to make the vibrations easily perceivable. Unfortunately, the Nexus One could not produce a suitable volume to make the Tactons perceivable and so a small headphone amplifier had to be used, which was placed in the carry bag with its input connected to the Nexus One by a 3.5mm headphone cable, and the C2 connected to its output. No headphones were worn, as the vibrations were not audible in the outdoor environment. Therefore, the participant held the Nexus One + Peltiers in the left hand, which were connected to the microcontroller board and amplifier sitting in the carry bag, which rested behind the left hand, over the shoulder (see Figures 7-11 and 7-12). The software produced a simple GUI screen for interaction (shown in Figure 7-13) and the participants

interacted by simply pressing on-screen buttons.

7.6.1 Icons

The thermal icons were identical to those from Experiment 8, described in Section 7.2. Although participants had some difficulty differentiating the moderate and strong warmth icons when sitting indoors, it was decided not to change the design for Experiment 9, to provide a more direct comparison with previous results and gain an idea of how sitting and walking outside influence interpretation of thermal icons. The Tactons used were a subset of those created by Brown *et al.* [21].

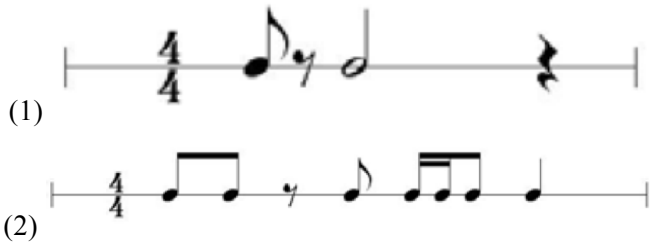

Figure 7-9: Tacton rhythms used in Experiment 9: 2-note Rhythm 1 for Personal messages and 7-note Rhythm 2 for Work messages. From Brown *et al.* [21].

As the thermal feedback had four icons in a 2 x 2 design (two-parameters with two levels per parameter), the same number/design of Tactons was used. *Rhythm* was used as one parameter, with the two *rhythms* shown in musical notation in Figure 7-9. Like the *rhythm* design in the intramodal icons, different numbers of notes were used to make the icons more distinguishable [21]. One *rhythm* consisted of two notes, one very short followed by one long note. The second *rhythm* consisted of seven notes: six short notes and one slightly longer note. The second parameter was vibration roughness, with a "smooth" 250Hz sine wave, and a "rough" 50Hz amplitude-modulated 250Hz sine wave. This gave four Tactons:

1. Rhythm 1, Smooth
2. Rhythm 1, Rough
3. Rhythm 2, Smooth
4. Rhythm 2, Rough

In Experiment 8, the icons were used to indicate the Source (Personal or Work) and Importance (Standard or Important) of a hypothetical message, and the same was done in Experiment 9. The mappings of thermal icons to message types were the same as Experiment 8, shown in Table 7-1. The mappings of Tacton to message type are shown in

Table 7-3. *Rhythm* was used to indicate the Source, with the two-note rhythm (rhythm 1) used to indicate a Personal message and the seven-note rhythm (rhythm 2) used to indicate Work messages. *Roughness* was used to indicate the Importance of the message, with a smooth vibration indicating a Standard message and a rough vibration indicating an Important message.

	Tacton
"Source" Parameter	*Rhythm*
Personal	2-note rhythm 1
Work	7-note rhythm 2
"Importance" Parameter	*Roughness*
Standard	Smooth
Important	Rough

Table 7-3: Mapping of Tacton parameters to message information.

7.6.2 Location

The process of choosing where to run the study was not entirely straightforward. Because participants would have to walk around, the area needed to be safe and free from obstacles that could potentially cause injury. However, it would also ideally be a location with a reasonable amount of activity and environmental noise, to realistically reflect a public location. Urban volume levels vary wildly depending on time and location [55]. Different parts of a town or city will have very different noise levels at the same time and one location will vary in noise over the course of a day [55]. Particularly busy and noisy areas are those near large bodies of motorised traffic, such as major roads, however, these areas are not safe for running a study due to the proximity of cars and the number of people on the pavements lining the roads. Further, the training sessions for both thermal icon and Tacton conditions needed to be carried out indoors to provide more stable and ideal conditions, and so the experimental location needed to have an enclosed space nearby.

A suitable location was identified adjacent to a University building (Fraser Building), and is shown in Figure 7-10. It was located approximately 70 meters from University Avenue, a busy road through Glasgow's West End (behind the trees in Figure 7-10) and had a public area inside the Fraser Building where training could take place. There were also benches that could be used for the sitting condition. While the outdoor area was public, it had little foot traffic, and so there were few human obstacles for participants to potentially collide with. In

many ways this location was highly suitable. Unfortunately, however, it was necessary to change locations after two participants had taken part.

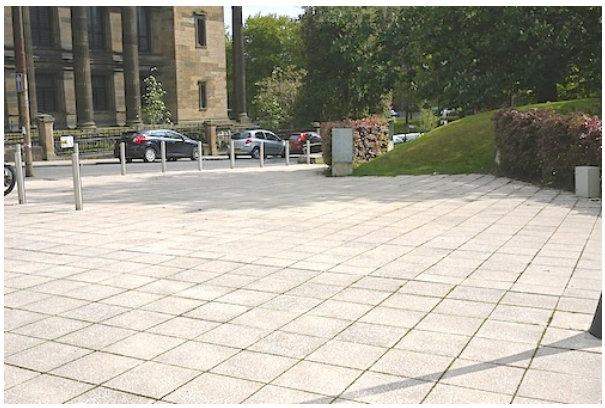

Figure 7-10: Initial experimental location, adjacent to Glasgow University's Fraser building. The location was abandoned due to high temperatures and participant discomfort.

The topography of the experimental location led to it being a suntrap, with high walls or buildings on 3 sides and little to no wind. This caused the temperature to rise to uncomfortable levels. Participants had to walk around the area for two conditions of approximately 10 minutes, and sit for a further two 10-minute conditions. The heat, long exposure and walking time resulted in notable participant discomfort, and so it was necessary to change the location. As mentioned above, environmental temperature influences thermal perception [36, 92, 184]. By deliberately choosing a cooler location, it means that only the influence of a smaller range of temperatures could be tested on thermal icon identification. While it is important to understand how warm and hot temperatures influence identification, participant comfort was of paramount importance.

Figure 7-11: Experimental location inside Glasgow University Quadrangle.

The new location was in the Western courtyard of the University Quadrangle (see Figure 7-11). This location was also enclosed, like the first location, but it provided large shaded areas, which would be cooler. The main issue with this area was that it was quieter, as there was no nearby traffic, although there was a degree of footfall from students, staff and tourists. The Hunterian Museum is inside the Quadrangle, and the entrance to it has a bench on which the training could take place and there are benches outside where the sitting conditions could take place. The area was flat and so free of potentially dangerous obstacles. The data from the first two participants was kept, to provide data from a wider range of environmental temperatures.

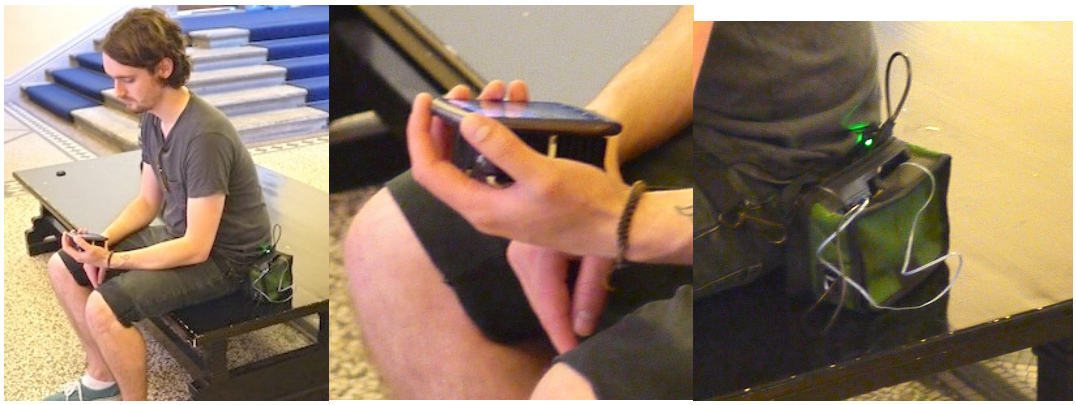

Figure 7-12: Training location (left), close-up of apparatus held in left hand (centre) and carry bag holding microcontroller, battery pack and C2 amplifier (right). C2 not shown, but was attached to the back of left wrist.

7.6.3 Design and Procedure

Twelve participants took part (10 male, 2 female), aged from 22 to 31 (mean 25.61), and were paid £6 for participation. Using two icon modalities (thermal, vibrotactile) and two mobility conditions (sitting, walking) gave four conditions:

1. Sitting + Thermal icons
2. Walking + Thermal icons
3. Sitting + Tactons
4. Walking + Tactons

The procedure was the same for all conditions and participants took part in all conditions. The Nexus One + Peltiers was held in the left hand, so that the Peltiers made contact with the palm of the hand, and the carry bag was slung over the right shoulder, across the body (see

Figures 7-11 and 7-12). During the Tacton conditions the C2 tactor was held against the back of the left wrist by the elastic fabric strap. The experiment was split into two halves by modality, one involving thermal icons and one involving Tactons. Within each half the participant took part in one condition whilst seated on a bench outside and one while walking outside. The order in which the participant ran through the modalities was counterbalanced, as was the order in which they were seated/walking. Each half started with a 10-minute training session seated in the indoor location (Figure 7-12).

The thermal training session started with 60 seconds of adaptation, where the Peltiers were set to the neutral starting temperature of 32°C. The Tacton training session had no adaptation. After this time the participants were given 10 minutes to feel each icon and learn the mapping of icon-to-message type. The software running on the phone showed the screen in Figure 7-13, where participants could press on the radio button next to a message type and press "Submit" to feel the icon that represents that message type. They could do this as many times as they liked for the 10 minutes. After this the participants took part in the first mobility condition, followed by the second mobility condition, before changing modality and repeating.

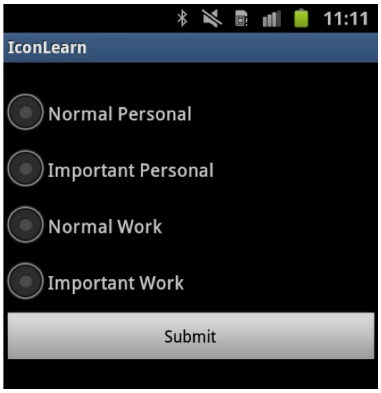

Figure 7-13: Experimental software for both training and testing sessions in Experiment 9.

During the sitting conditions participants sat on a bench while holding the Nexus One and interacting with it. During the walking conditions, participants were asked to walk in a simple loop around the courtyard at their normal walking pace. Other than these details the procedure for the sitting and walking conditions was the same. Each icon was presented to the participant four times in a random order, with 30 seconds in between each icon. This amount of time was chosen initially to give sufficient time for the Peltiers to adapt the skin back to 32°C neutral. While no such adaptation is necessary for Tactons, the same gap was used in the Tacton conditions, to keep the frequency of stimuli consistent. Unlike during

Experiment 8, the screen (shown in Figure 7-13) was shown at all times and so there was no cue as to when icons were presented. Participants were instructed that, whenever they felt an icon, they were to press on the radio button relevant to that icon and press "Submit" to report the message type they interpreted the icon as representing. At this point in the thermal icons condition, the Peltiers were immediately set back to 32°C. In both conditions, the report was recorded and another icon was presented at random after 30 seconds. Participants were also told that, if the system received no input from them within 20 seconds of an icon being presented, it would interpret this as the participant 'missing' the icon (not detecting it). If this occurred, the event was logged, the Peltiers were returned to 32°C (in the thermal conditions) and another icon was presented at random after 30 seconds. This continued until all four icons had been presented four times each. Experimental instructions and raw data for all measures can be found in Appendix F.

The Independent Variables were *mobility* (sitting and walking), *modality* (thermal and tactile) and *icon* (four icons). The Dependent Variables were: *Accuracy* (whether the right message type was identified) and *Identification Time* (the time taken to choose a message type). There were a total of: 12 participants x 2 Mobility conditions x 2 Modalities x 4 Icon Types x 4 presentations = 768 trials. In the same way as Experiments 6 and 7, if an icon was not detected the data point for *Identification Time* for that trial was ignored (as it was not detected the data point was included in the *Accuracy* analysis). Therefore, including both Thermal Icons and Tactons together, there were a total of 768 data points for *Accuracy*: 384 for each Mobility and Modality and 192 data points for each Icon Type and Mobility + Modality combination condition (e.g., Walking + Thermal). No Tactons were missed, while 18 Thermal Icons were missed, leaving 178 data points for each Mobility Condition, and 84 for each Icon Type, when analysing Thermal Icons.

The environmental temperature, humidity and noise level was also recorded throughout the study. The temperature and humidity was recorded at the beginning of each of the four conditions, and the ambient noise was recorded constantly throughout each condition using an ambient noise meter. Performance results could then be analysed in relation to temperature/noise to determine if these factors influence icon identification.

7.6.4 Results – Experiment 9

7.6.4.1 Influence of Environmental Temperature, Humidity and Noise

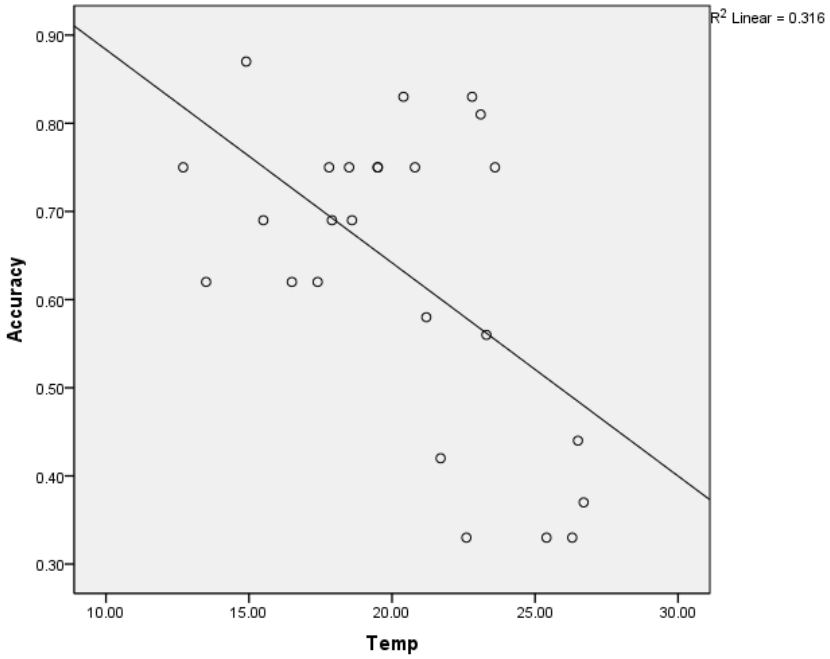

Figure 7-14: Significant negative correlation of thermal icon Accuracy (% correct) with Environmental Temperature (Temp, °C). Pearson coefficient *r* = -0.562.

Temperature

The mean temperature across all conditions was 20.27°C (SD = 3.99), with a minimum of 12.7°C and a maximum of 27.4°C. The potential relationship between environmental Temperature and thermal icon/Tacton Accuracy was investigated using Pearson's product-moment correlation coefficient. Temperature and thermal icon Accuracy had a significant negative correlation (r (23) = -0.562, $p < 0.01$), with Accuracy decreasing as Temperature increased (see Figure 7-14). There was a positive relationship between Temperature and Tacton Accuracy, although the correlation was only just outside of significance (r (20) = 0.414, $p > 0.05$). The relationship between Temperature and Humidity was also tested and found that Humidity had a significant negative correlation with Temperature (r (22) = -0.872, $p < 0.01$; see Figure 7-15).

Humidity

Humidity ranged from a minimum or 46.6%, to a maximum of 87.1%, with a mean of 64.22% (SD = 9.29). Again, Person's correlation coefficient was used to examine the relationship between Humidity and Accuracy. There was a positive but non-significant relationship between Humidity and thermal icon Accuracy (r (23) = 0.381, $p > 0.05$), however, there was a significant negative correlation between Humidity and Tacton Accuracy (r (20) = -0.453, $p < 0.05$), as Accuracy decreased as Humidity increased (see

Figure 7-16).

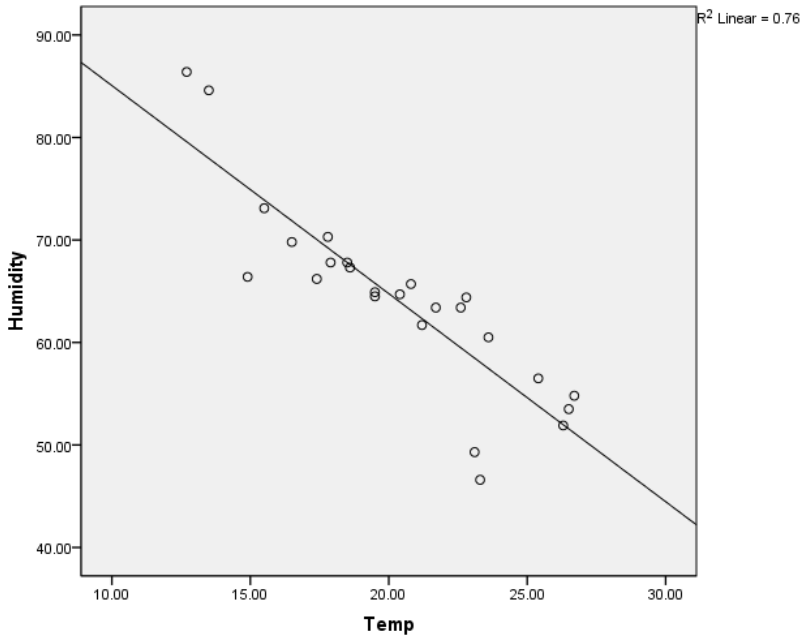

Figure 7-15: Significant negative correlation of Humidity (%) with Environmental Temperature (Temp, °C). Pearson coefficient *r* = -0.872.

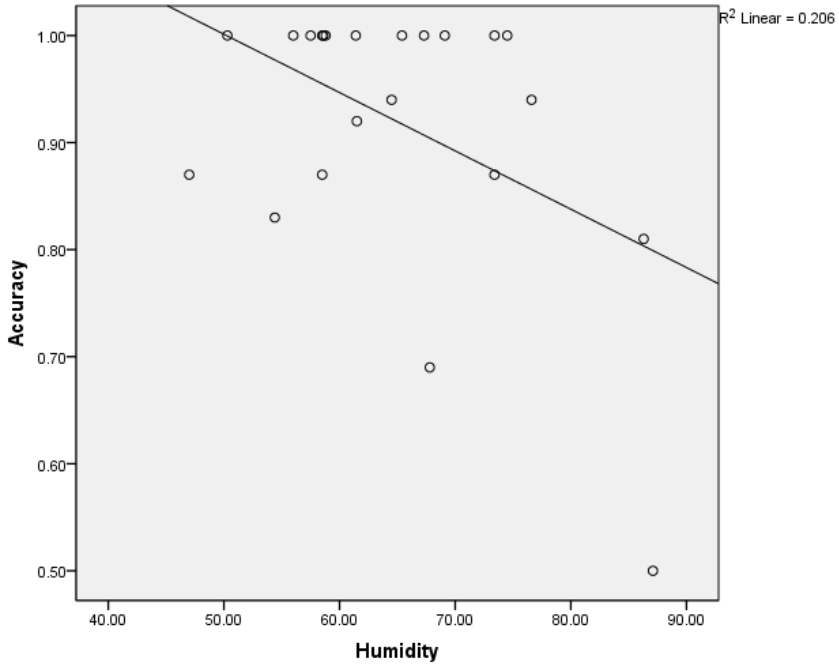

Figure 7-16: Significant negative correlation of Tacton Accuracy (% correct) with Humidity (%). Pearson coefficient *r* = -0.453.

Environmental Noise

Environmental noise varied from 38-88dB, with a mean of 51.87dB (SD = 3.64). No correlation was found between Noise and either thermal icon Accuracy (r (19) = -0.370, p > 0.05) or Tacton Accuracy (r (18) = 0.139, p > 0.05).

7.6.4.2 Thermal Icons

The data for Thermal Icons was normally distributed, and so 2 x 4 (Mobility x Icon Type) Repeated-measures ANOVA were used to analyse both *Accuracy* and *Identification Time,* using SPSS. *Post hoc* pairwise comparisons were run within the SPSS analysis, using the Bonferroni correction.

Accuracy

Pearson's correlation coefficient was used to test for learning effects by investigating the relationship between trial number and Accuracy. No significant relationship was observed (r (30) = 0.149, p > 0.05) suggesting no learning took place over the course of the study and training was sufficient.

		Perceived As					
		Mod Warm	**Strong Warm**	**Mod Cool**	**Strong Cool**	**Missed**	**Accuracy**
Actual Icon	**Mod Warm**	70	10	3	1	10	74%
	Strong Warm	31	57	1	3	2	61%
	Mod Cool	0	3	54	33	4	57%
	Strong Cool	2	1	27	62	2	66%

Table 7-4: Thermal icon confusion matrix, showing each icon presented (Actual Icon) and the number of each icon it was interpreted as (Perceived Icon) and how many were not detected (Missed).

The overall identification rate for the two-parameter thermal icons was 64.6% (SD = 47.75). The mean Accuracy for each individual thermal parameter was 96.3% for *direction of change* and 73.1% for *subjective intensity*, with 18 missed thermal icons (4.79% of all icons). The confusion matrix for all thermal icons is shown in Table 7-4. A 2 x 4 repeated-measures ANOVA (Mobility x Icon) found no effect of Mobility on Accuracy ($F_{(1,46)}$ = 3.783, p > 0.05) as identification rates were similar when walking (mean = 61%, SD = 49.0) to when sitting (mean = 69%, SD = 46.5). The ANOVA found no effect of Icon on identification rate either ($F_{(3,138)}$ = 2.205, p > 0.05), with mean Accuracy of 74% (SD =

43.8), 61% (SD = 49.1), 57% (SD = 49.7) and 66% (47.6) for the moderate warm, strong warm, moderate cool and strong cool icons respectively. Finally the ANOVA found no interaction between Mobility and Icon ($F_{(3,138)}$ = 2.423, p > 0.05). Identification rates are shown in Figures 7-17 and 7-18.

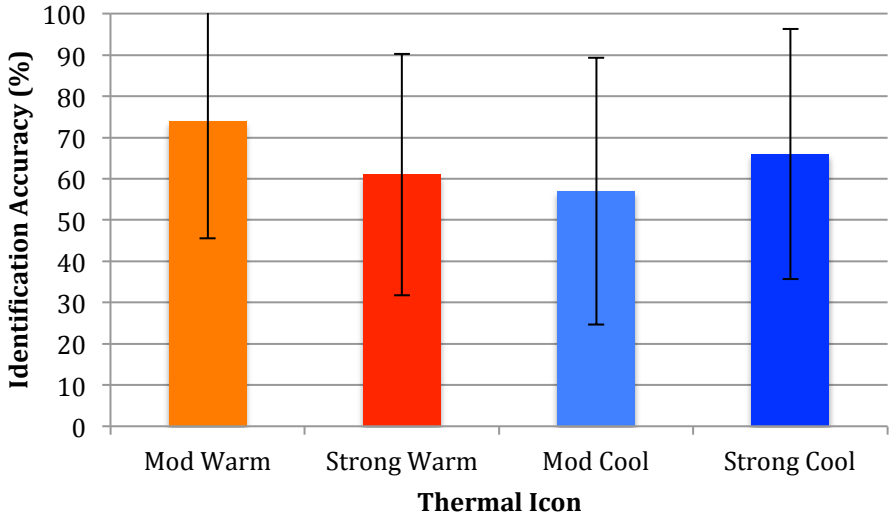

Figure 7-17: Mean identification rates for each thermal icon. Error bars show 1 standard deviation. There were no significant differences.

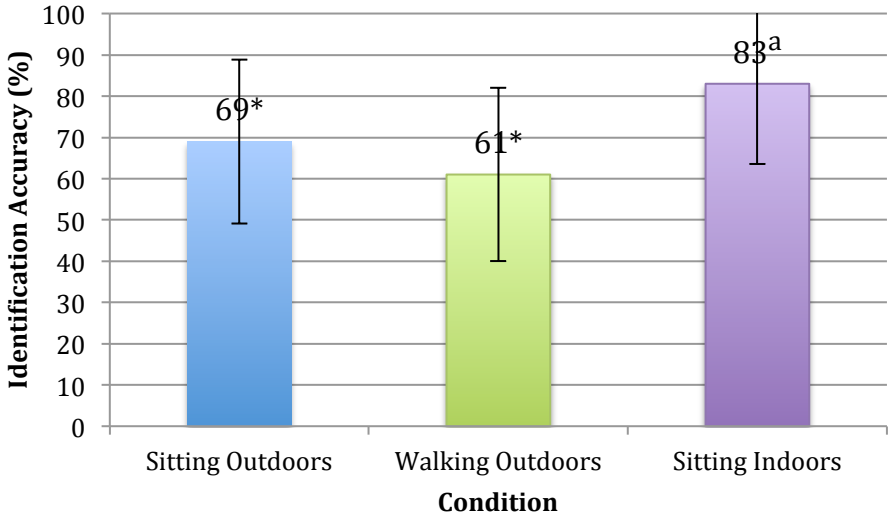

Figure 7-18: Mean thermal icon identification rates for the two mobility conditions in Experiment 9 (Sitting Outdoors and Walking Outdoors), as well as the results from Experiment 8 (Sitting Indoors). Error bars show 1 standard deviation. 'a' indicates significantly higher than '*', p<= 0.01.

Statistical comparisons of identification rates were carried out between the static indoor

thermal icon condition from Experiment 8 and both static and walking thermal icon conditions from Experiment 9. Two Mann-Whitney *U* tests showed a significant difference between the first indoor condition and both the static outdoor condition ($U = 15486$, $p < 0.01$) and the walking outdoor condition ($U = 14046$, $p < 0.001$). In both cases, the identification rate was significantly higher in the indoor study from Experiment 8 (82.8%; see Figure 7-18).

Late Corrections, False Positives and Neutral Returns

If a participant submitted a response indicating that he or she had detected an icon when, in fact, no icon had been presented since the last response, this event was recorded as a False Positive (FP). 23 FPs were recorded across all participants. These did not contribute to the Accuracy scores, as they occurred when no icon was presented for judgement. However, looking through the data actually suggests that not all responses were technically False Positives (where the participant detected something when no icon was presented). All events were put into the following categories:

1. Late Corrections
2. Late and Wrong
3. Neutral Returns
4. False Positives

The software had a set timer, which started once an icon had been initiated. If no response from the participant had been received within 20 seconds, the icon was considered "missed", the Peltiers were returned to 32°C neutral and an error was logged next to that icon, before another icon was queued to be presented after a further 30 seconds. From the event logs it appears as though this time was perhaps not quite long enough in a very small number of occasions, as responses from participants were actually received within just a few seconds of the timer running out. This situation arose seven times, leading to three Late Corrections and four Late and Wrong events. Late Corrections (LC) refer to when a participant submitted the correct response to the presented icon, however, only doing so after the 20 second timer had elapsed. An error was logged with the icon and an FP was logged for the correct response. Two LC arose from Moderate Warm icons and one LC arose from a Strong Cool icon. Late and Wrong (LW) events arose from the same timing as LC events, however, the wrong response was submitted shortly after the timer ended. An event was considered LC or LW if it happened within 10 seconds of the 20-second timer ending.

Neutral Returns (NR) refer to the situation where the initial thermal change from neutral to

the *extent of change* is not detected, however, the change back to 32°C from the given *extent* (after the 20-second missed icon timer has elapsed) was detected. Therefore, a response was received indicating the detection of a warming icon, when in fact the Peltiers are returning to neutral from a cold icon, or a cold icon, when returning to neutral from a warm icon. This occurred on ten occasions, eight when returning to neutral from warm icons and two when returning from cold icons. A breakdown of which icons were wrongly interpreted is shown in Table 7-5. An event was considered an NR if a response was received between 10 and 20 seconds after the 20-second timer ended.

Finally, there were only six true False Positives (FP), which were categorised if they occurred more than 20 seconds after either the end of the 20-second timer, or any time if the previous icon had been responded to, correctly or incorrectly. Two FPs reported feeling the Moderate Warm icon, when none was presented while four reported feeling the Moderate Cool icon, when none was presented.

Missed Icon	Return to Neutral Perceived As	Number of Events
Moderate Warmth	Moderate Cool	4
Moderate Warmth	Strong Cool	2
Strong Warmth	Moderate Cool	2
Moderate Cool	Moderate Warmth	2

Table 7-5: Number and makeup of Neutral Return events, where the initial icon is not detected, but the subsequent return to 32°C neutral is.

Identification Time

Pearson's correlation coefficient was used to investigate the relationship between trial number and Identification time. The two factors correlated positively and significantly ($r(30) = 0.423$, $p < 0.05$) with Identification Time increasing as the number of completed trials increased (see Figure 7-19).

A 2 x 4 repeated-measures ANOVA (Mobility x Icon) was run on the Identification Times (IDT). A significant effect of Mobility on IDT was found ($F_{(1,38)} = 9.460$, $p < 0.01$), as identification took significantly longer when sitting (mean = 9.20s, SD = 3.88s) than when walking (mean = 8.57s, SD = 3.78s). There was no effect of Icon on IDT, with means of 8.33s (SD = 3.27), 8.83s (SD = 3.77), 9.00s (SD = 3.80) and 9.48s (SD = 4.36) for Moderate Warm, Strong Warm, Moderate Cool and Strong Cool icons respectively. A significant interaction effect was found ($F_{(3,114)} = 3.664$, $p < 0.05$). Mauchly's test of Sphericity was significant, and so a Greenhouse-Geisser correction was applied. Mean Identification Times

are shown in Figures 7-20 and 7-21.

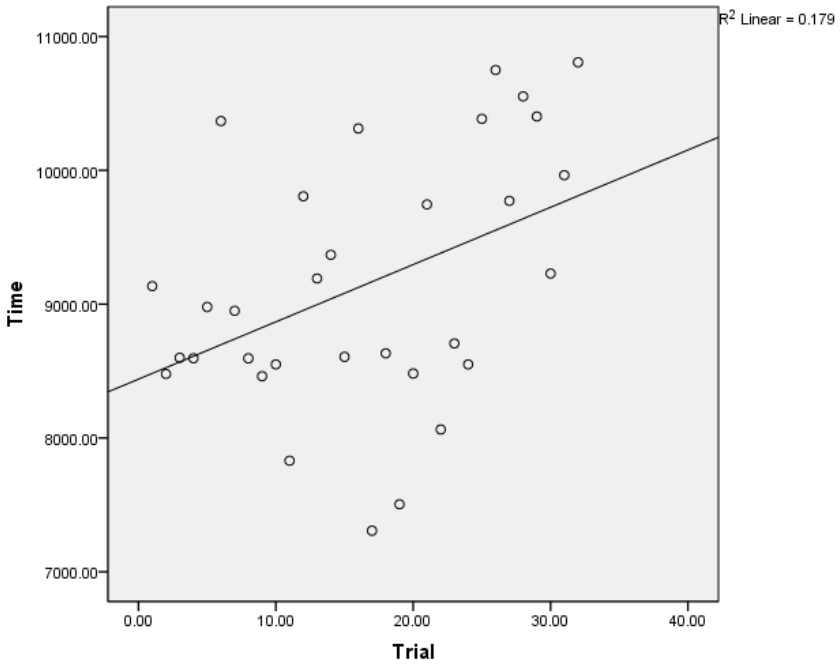

Figure 7-19: Significant positive correlation of thermal icon Identification Time (milliseconds) with number of completed Trials. Pearson coefficient $r = 0.423$.

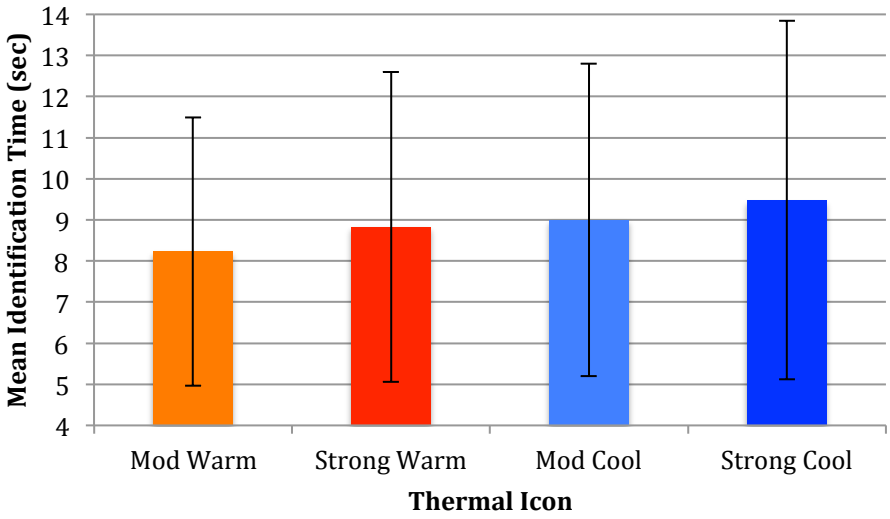

Figure 7-20: Mean Identification Times for each thermal icon in the outdoor study. Error bars show 1 standard deviation. There were no significant differences.

Statistical comparisons of Identification Times were carried out between the static indoor thermal icon condition from Experiment 8 and both static and walking thermal icon conditions from Experiment 9. Two Mann-Whitney *U* tests showed a significant difference between the first indoor condition and both the static outdoor condition ($U = 8734.5$, $p <$

0.001) and the walking outdoor condition (U = 9720.5, $p < 0.001$). In both cases, the Identification Time was significantly lower/faster in the Experiment 8 indoor study (mean = 6.57s; see Figure 7-21).

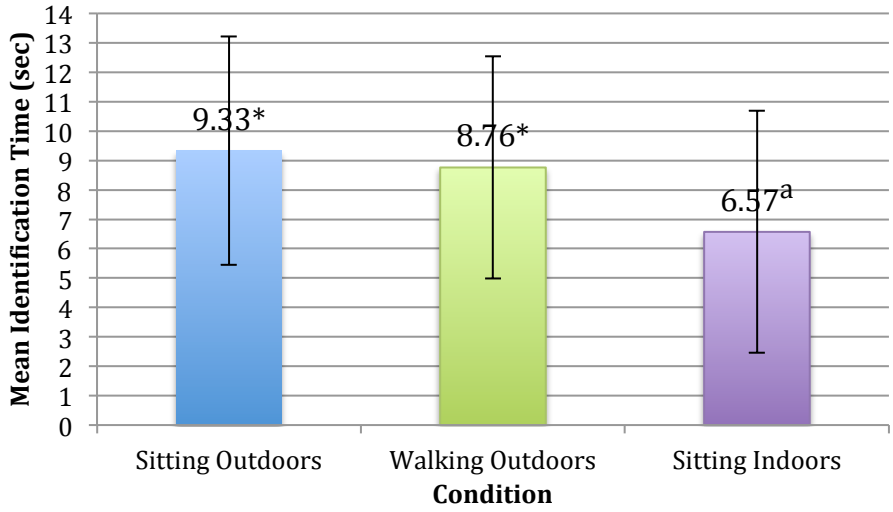

Figure 7-21: Mean thermal icon Identification Times for the two mobility conditions in the second study (Sitting Outdoors and Walking Outdoors), as well as the results from Experiment 8 (Sitting Indoors). Error bars show 1 standard deviation. 'a' indicates significantly lower than '*', p < 0.001.

7.6.4.3 Tactons

The data for Tactons violated the assumption of a normal distribution, and so non-parametric tests were used, specifically Friedman's test for the effect of Icon Type, and Wilcoxon *T* tests for the effect of Modality. *Post hoc* pairwise comparisons following a significant Friedman's test were conducted using Wilcoxon *T* tests, using the Bonferroni-correction of *p=0.05/N*, where N is the total number of comparisons.

Accuracy

The mean overall identification rate for the two-parameter Tactons was 91.6% (SD = 27.75). The mean Accuracy for each individual Tacton parameter was 96.6% for *rhythm* and 96.25% for *roughness*. The confusion matrix for all Tactons is shown in Table 7-6 and identification Accuracy is shown in Figure 7-24. A Wilcoxon *T* test found no effect of Mobility on Accuracy (T = 120, $p > 0.05$), as mean Accuracy was similar when walking (mean = 90%, SD = 30.1) to when sitting (mean = 93%, SD = 25.4).

A Friedman's test found a significant effect of Icon on Accuracy ($\chi^2(3) = 18.231$, $p < 0.001$). *Post hoc* Wilcoxon *T* tests, with a Bonferroni-adjusted p-value of 0.0083, found a

significant difference between Important Personal (two-note *rhythm 1*, *rough*) and Important Work (seven-note *rhythm 2*, *rough*) and between Standard Work (seven-note *rhythm 2*, *smooth*) and Important Work. In both cases, the mean Accuracy of Important Work was significantly higher (mean = 100%, SD = 0) than the Accuracy of either Important Personal (mean = 91%, SD = 28.4) or Standard Work (mean = 82%, SD = 38.2). Mean Accuracy for Standard Personal (two-note *rhythm 1*, *smooth*) was 92% (SD = 26.5).

		Perceived Icon					
		Rhythm 1 Smooth	**Rhythm 1 Rough**	**Rhythm 2 Smooth**	**Rhythm 2 Rough**	**Missed**	**Accuracy**
Actual Icon	**Rhythm 1 Smooth**	75	5	0	0	1	92%
	Rhythm 1 Rough	2	73	4	0	1	91%‡
	Rhythm 2 Smooth	0	7	66	5	2	82%*
	Rhythm 2 Rough	0	0	0	80	0	100%*‡

Table 7-6: Tacton confusion matrix, showing each icon presented (Actual Icon) and the number of each icon it was interpreted as (Perceived Icon) and how many were not detected (Missed). * and ‡ = p<0.0083.

Identification Time

A Wilcoxon *T* test found no effect of Mobility on Identification Time (IDT) for Tactons, with similar times when walking (mean = 4.67s, SD = 1.36) to when sitting (mean = 4.80s, SD = 2.05). A Friedman's test found a significant effect of Icon on IDT ($\chi^2(3) = 49.492$, $p < 0.001$). *Post hoc* Wilcoxon *T* tests, with a Bonferroni-adjusted p-value of 0.0083, found significant differences between each pair of Icons, other than Standard Personal (two-note *rhythm 1*, *smooth*) and Important Work (seven-note *rhythm 2*, *rough*). Mean IDT were 4.84s (SD = 2.45), 4.41s (SD = 1.34), 5.14s (SD = 1.67) and 4.55s (SD = 1.11) for Standard Personal, Important Personal, Standard Work and Important Work respectively.

7.6.4.4 Thermal Icons vs. Tactons vs. Intramodal Icons

The comparison of thermal icons when sitting indoors to intramodal icons was reported in Section 7.5.2 and is not repeated here. This section compares Accuracy and Identification Time of thermal icons when sitting and walking outdoors, Tactons when sitting and walking outdoors and intramodal icons when sitting indoors.

Accuracy

A Wilcoxon *T* test was run on the combined outdoor sitting and walking data and found a significant difference between thermal icon and Tacton Accuracy ($T = 7326$, $p < 0.001$), with Tactons having significantly higher Accuracy (92%, SD = 27.8) than thermal icons (65%, SD = 47.9). Further Wilcoxon *T* tests compared each individual mobility condition (mobile-thermal, static-thermal, mobile-Tacton and static-Tacton), with a Bonferroni-adjusted p-value of 0.0083. They found significant differences between each combination of mobility condition (all $p < 0.001$), apart from the two within-modality conditions (i.e., mobile-thermal vs. static-thermal and mobile-Tacton vs. static-Tacton). Mean identification rates for each condition are shown in Figure 7-24.

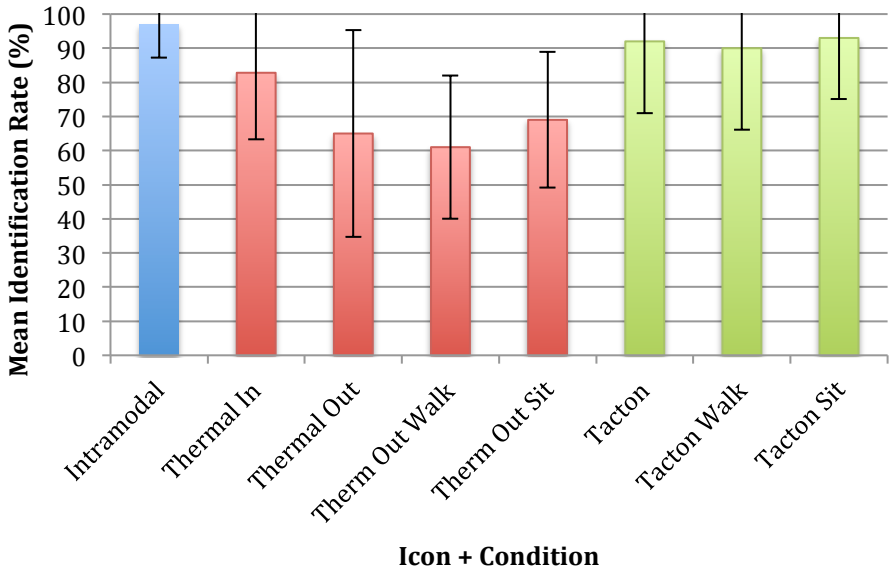

Figure 7-22: Mean Identification rates for all icon types: intramodal icons (indoors); thermal icons when sitting indoors (Thermal In); thermal icons when outdoors (Thermal Out) as well as walking and sitting conditions; and Tactons (outdoors) including walking and sitting conditions. Error bars show 1 standard deviation.

Mann-Whitney *U* tests were run comparing intramodal icon Accuracy to mobile-thermal, static-thermal, mobile-Tacton and static-Tacton Accuracy, with an adjusted p-value of 0.0125. Significant differences were found between intramodal icon Accuracy and mobile-thermal ($U = 11508$, $p < 0.001$), static-thermal ($U = 12948$, $p < 0.001$) and mobile-Tacton ($U = 14304$, $p < 0.01$), but not between intramodal icons and static-Tactons ($U = 14784$, $p > 0.05$). In each case, bar the non-significant difference, intramodal icons had significantly higher identification Accuracy than the thermal icons or Tactons.

Identification Time

The same analysis was carried out on the IDT data. A Wilcoxon *T* test found a significant difference between thermal icon and Tacton IDT (T = 47830.5, $p < 0.001$), with Tactons having significantly lower/faster IDT (4.73s, SD = 1.73) than thermal icons (9.04s, SD = 3.83). Further Wilcoxon *T* tests compared each individual mobility condition (mobile-thermal, static-thermal, mobile-Tacton and static-Tacton), with a Bonferroni-adjusted p-value of 0.0083. They found significant differences between each combination of mobility condition (all $p < 0.001$), apart from the two within-modality conditions (i.e., mobile-thermal vs. static-thermal and mobile-Tacton vs. static-Tacton). Mean IDT for each condition are shown in Figure 7-25.

Mann-Whitney *U* tests (with adjusted p-value of 0.0125) found significant differences between intramodal icon Accuracy and all conditions from Experiment 9 ($p < 0.001$). In all cases, intramodal icons had significantly lower/faster IDT than the thermal icons or Tactons.

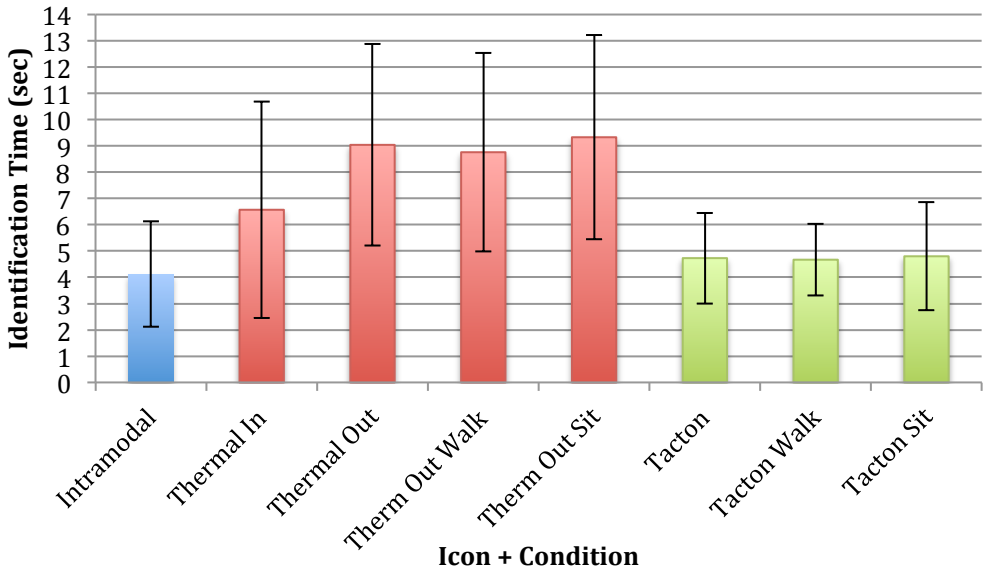

Figure 7-23: Mean Identification Times in seconds for all icon types: intramodal icons (indoors); thermal icons when sitting indoors (Thermal In); thermal icons when outdoors (Thermal Out) as well as walking and sitting conditions; and Tactons (outdoors) including walking and sitting conditions. Error bars show 1 standard deviation.

7.7 Discussion

7.7.1 Thermal icons

There were some highly encouraging results from both Experiments, including the 83% identification rate when sitting indoors, the reliability of *direction of change* when outdoors, and the lack of any significant drop in identification accuracy when walking, compared to sitting, when outdoors. However, some significant issues were encountered which have major implications for the use of thermal feedback for mobile interaction and so this section discusses recommendations and potential obstacles in the design of thermal feedback.

7.7.1.1 Indoor accuracy and the Salience of Direction of Change Outdoors

When sitting in a stable indoor environment, identification of two pieces of information was good, at 83% Accuracy and Identification Times of 4-6 seconds. Direction of change was particularly salient, but the overall accuracy suggests thermal feedback could be a potential alternative to other non-visual feedback methods, if only one or two pieces of information are to be conveyed. Even when sitting and walking outdoors, the results show that *direction of change* was extremely well identified, at 96%, which was as accurate as both individual Tacton parameters. Therefore, basic warming and cooling thermal stimulation is highly salient, even when walking outdoors, and both warm and cold stimuli can be felt using simple, compact apparatus. The low-bandwidth affective feedback designs that simply warm or cool to provide information are therefore likely to be useful even when walking outdoors.

Experiment 8 found that identification of both Intramodal parameters simultaneously was highly accurate, at ~97%, suggesting thermal feedback may be a suitable augmentation of existing Tacton design. Given the high identification of *direction* when walking outdoors, it may be a suitable replacement for more problematic parameters during mobile interaction (as long as only two parameter levels are required). While identification rates were lower than *direction of change*, *subjective intensity* still performed fairly well when outdoors, at 73.1% accuracy. Means of potentially improving this value are discussed in Section 7.7.1.4 but this result, and the high reliability of *direction* still suggest that multi-dimensional thermal feedback is a promising and feasible means of conveying simple information.

Recommendations

Direction of change is a useful parameter for thermal feedback in mobile environments. If only a single piece of information, with two alternatives, is to be conveyed, then *direction of change* is a suitable alternative to vibrotactile feedback. Based on the small number of missed icons in Experiments 8 and 9, a thermal change of at least 3°C is recommended.

7.7.1.2 Walking Does not Significantly Impair Identification

Identification rate was 69% when sitting outdoors, while walking outside reduced it slightly (but not significantly) further to 61%, and Thermal Icons were identified significantly faster when walking. Interestingly, the initial study into perception of thermal parameters in Chapter 6 showed that walking significantly reduced the number of thermal stimuli detected when indoors (compared to sitting), yet walking outdoors did not significantly affect identification Accuracy of thermal icons, compared to sitting outdoors. Given the negative effects of walking and environmental temperatures [74] individually, it might have been expected to see a more pronounced drop in identification when the two influences acted together. In Experiment 7, the negative effect of walking indoors acted on all three *extents of change*, even the larger 3°C and 6°C changes that were used in thermal icon designs, so it might have been expected to see a more pronounced drop in identification when the two influences of environmental temperature and walking acted together. An important point to note is that the participants were cued as to when an icon was being presented during Experiment 8, as the GUI screen, which recorded responses, appeared as the icon was presented. In the outdoor studies there was no cue. Cueing participants in Experiment 8 may have artificially inflated the Accuracy scores slightly, as a small number of thermal icons were not detected (at least within the 20-second response timer) when sitting or walking outdoors. A small number (7) of these missed icons were in fact simply late responses, three of which were correct responses (the other four were incorrect), so a lack of cueing only appears to have led to participants to be outright oblivious to 11 icons, out of a total of 376. While the Accuracy for sitting and walking outside is quite low, it is encouraging that there appears to be only a small interaction cost when walking.

A surprising finding, however, was that the Identification Time was significantly (0.6 seconds) slower when sitting than when walking. In contrast, during the perceptual study, walking led to longer detection times. 0.6 seconds is perhaps not a large difference in real-world terms, but it is difficult to explain why this would be. Perhaps the need to pay visual attention to where the participant was walking led to a little more haste in responding.

Recommendations

The results suggest that thermal feedback may be as suitable for use when walking outdoors as when sitting, however, future research should test identification in a wider range of mobile scenarios, including on transport and when walking more cognitively demanding routes.

7.7.1.3 Environmental Influences

The influence of outdoor environmental temperatures significantly reduced identification of thermal icons, as a strong negative correlation was observed between temperature and identification rate, and Identification Times increased to 8-10 seconds. Even within the small range of temperatures recorded during the experiment (13-27°C), thermal icons became significantly less identifiable as temperature increased. Further, a correlation was found between number of trials and Identification Time, suggesting that participants took longer identifying icons as time went on. No significant correlation was found between trial number and Accuracy, however, so it seems that extra time was taken to maintain Accuracy.

Average identification accuracy for both bits of information was 64%, ranging from a high of 87% for one participant (during 14.9°C outdoor temperature) down to 33% (at 25-26°C outdoor temperature). The overall value is markedly lower than the 83% accuracy found for the same thermal icons when sitting indoors, although the high value of 87% is slightly higher. The individual differences are worthy of note, however, as one participant managed only 62% Accuracy at 13.5°C outdoor temperature and another managed 83% at 23°C, so, while the environment clearly influences accuracy, individual differences should also be monitored.

Of the ten Neutral Return (NR) events recorded, six occurred within the first four participants, who took part under four of the highest average environmental temperatures. Interestingly, all six NR events were identified returning to neutral from warm icons, where the warming changes were not detected, but the cooling changes to neutral were. Of the other four NR events (which occurred during cooler trials), only two were missed warming icons, while the other two were missed cooling icons. Environmental temperature is known to influence skin temperature and, in turn, skin temperature influences thermal perception [92, 223, 229]. Increases in environmental temperature will increase the skin temperature (the physical activity of walking may also act to raise it), and so it is possible that the skin was above the 32°C neutral starting temperature when these warming icons were presented. The skin directly underneath the Peltiers should have been close to 32°C due to the 30 seconds of inter-icon adaptation, however, the immediately surrounding skin may not have been, potentially leading to referral or domination [63]. Warmth from surrounding areas on the skin may have led to erroneous interpretation of greater warmth at the stimulation site. Therefore, the true extent of thermal change from neutral to the warming *extent of change* (up to 35°C or 38°C) may have been smaller (or been perceived as smaller) because of the warming of the skin, while the return change to neutral may have been larger (or perceived as larger), leading it to be more salient.

Recommendations

As environmental temperature has such a strong impact on identification of thermal icons, feedback designs will have to adapt to the environment and adjust the starting temperature and the extent/rate of thermal change (see Section 7.7.1.4) to make the feedback more salient. An example might be to match the starting temperature to current skin temperature, by monitoring changes, and then warm and cool from that point [201], similar to adaptive brightness relative to environmental light. If this is the case, more in depth research is needed to understand how well individuals can detect thermal changes of various magnitudes, starting from a variety of starting temperatures. Finally, research needs to be done in a wider range of environmental temperatures and conditions, to assess their impact on perception and identification of thermal feedback.

7.7.1.4 Issues with Subjective Intensity

The main source of error came from the *subjective intensity* (*SI*) parameter. During Experiment 8, 85.4% of *SIs* were correctly identified, which is promising, but in Experiment 9, 73.1% were identified. Analysing the two confusion matrices provides interesting insights into where the problems lie. In both the indoor and outdoor studies, more cold *SIs* were confused than warm (16 vs. 12 indoors, 60 vs. 41 outdoors respectively). This is perhaps unexpected, as cold stimuli were significantly better detected during Experiments 6 and 7; however, they were also reported by participants as feeling significantly less intense than warm stimuli, which may mean that they were more difficult to tell apart. The Strong Warm thermal icon is closer to (yet safely distant from) the heat pain threshold than the Strong Cold icon is to the cold pain threshold, while both Moderate icons were relatively mild. This may have lent the Strong Warm icon a unique, more intense, quality, making it easier to tell apart from the Moderate Warmth compared to the two cold icons. If this were true, however, it did not prevent errors occurring within warm icons. Whereas slightly more cold *SI* were perceived as being more intense (42) than less intense (34) than intended, the vast majority of warm *SI* were perceived as less intense than intended (42, compared to 11 feeling more intense) across all three conditions (sitting indoors, sitting outdoors and walking outdoors).

However, given the relatively poor performance of *subjective intensity*, the range of thermal stimuli that can be used to convey information may be limited. Because the thermal sense is focused on the overall magnitude, or extent, of stimulation, it was hypothesised that *subjective intensity* would be a suitable parameter for thermal feedback design. If it is a suitable parameter, it may be that the stimuli chosen to produce the intensities here were not

appropriate. The design was based on research that suggests faster and larger changes feel subjectively more intense than slower, smaller changes [36, 81, 185, 229, 249]. The rates and extents of change used here may simply not have been fast or large enough to tell apart.

Recommendations

As discussed in Chapters 2 and 6, there are many different ways of varying the subjective intensity of a thermal stimulus and only a small subset of those were used in the design of thermal icons. For example, the difference between *extents of change* should be larger than the 3°C used here, or the rates at which the temperature is changed should be more different than 1°C/sec vs. 3°C/sec. Only 2 extents of change and 2 rates of change were used in the design of thermal icons. A more thorough examination of the different possibilities could yield stimuli that are more reliably perceivable and perceptually distinct, compared to those used here.

Finally, several participants reported that they felt desensitization over the course of both Experiments 8 and 9, as the icons felt quite different at the beginning but became more difficult to tell apart as time went on. There was, in fact, a moderate *positive* correlation between trial number and Accuracy; however, this was not significant. The subjective interpretation may still have had an influence on perception. This may be an artefact of the experiment, as icons were presented at fairly regular intervals (2 every minute), and for up to 20 seconds at a time. More research should be done into how presentation frequency influences thermal perception over time. What this may mean for thermal feedback design is that the feedback may be more effective, and easily identifiable, if stimuli are provided less frequently. Constant and/or regular vibrotactile stimulation has been found to lead to adaptation/habitualisation to the stimulus, which makes an individual less sensitive over time [235]. However, there have been a number of psychophysical studies that have involved extended periods of testing where participants receive thermal changes regularly, and habitualisation has not been reported as a problem. One possible reason for the discrepancy may be that the changes used in Experiments 8 and 9 were larger than other psychophysical research, which often focuses on identifying the smallest changes detectable [43, 133, 136]. Regular, and large, shifts to both warm and cold temperatures may have led to a more pronounced influence on the thermal sense.

7.7.1.5 Feedback Alternatives

If *subjective intensity* remains an unreliable parameter, a replacement would need to be sought. *Area of stimulation, rhythm* and *spatial location* were discounted as parameters for

this research for several reasons, but research should be conducted to test their suitability. While discounted for being less practical, *spatial location* may be the more suitable parameter of the three (see Section 6.2.3.3), and it has been used successfully in Tacton presentation. However, varying the *area* of stimulation also varies the *subjective intensity* of the sensation, which this chapter has shown to be a problematic means of conveying information outdoors. Using a larger stimulator for all stimuli may make the feedback more salient, however: only two 2cm^2 Peltiers were used in this research. The design challenges for thermal *Rhythms* were discussed in Section 6.2.3.2.

Another possible use for thermal feedback would be as an augmentation of other tactile cues. Experiment 8 found that presenting thermal and vibrotactile feedback together did not hinder interpretation of these intramodal icons. Identification of intramodal icons was not tested in the second study; instead thermal icons were compared with more established non-visual feedback in the form of two-parameter Tactons. Experiment 9 found no significant effect of walking on identification of Tactons, which is at slight odds with other research [179], although three levels per parameter were used instead of only two in the second study here. Simple warming or cooling thermal changes could hypothetically be added to the two-, or even three-parameter Tactons to increase the available bandwidth, or replace less practical Tacton parameters such as spatial location [21].

7.7.2 Tactons

In slight contrast to thermal icons, Tactons were well identified as a whole (92%), and both *rhythm* and *roughness* were identified accurately individually. They were also identified much faster than thermal icons, at 4-5 seconds. Not all the different Tactons were equally clear, however, as the Tacton representing the Important Work message (seven-note *rhythm* with a *rough* texture) had 100% Accuracy, with no errors, yet the Standard Work Tacton (the same *rhythm* with a *smooth* texture) had the highest error rate, at 18%, as well as the longest Identification Time. Roughness has been a slightly problematic parameter in Tacton design [21, 179], as it can be difficult to tell different textures apart. However, if participants had issues in identifying the roughness, then it might be expected that they would confuse each importance level with the other. This was partly the case for the Personal messages (the two-note *rhythm*), as seven *roughness* confusions occurred. While participants wrongly identified the *smooth* Work *rhythm* as the *rough* Work *rhythm* on five occasions, the reverse was never true: the *rough* Work Tacton was never interpreted as *smooth.*

What is more surprising is that, on 11 occasions, the Important Personal Tacton (two-note *rhythm, rough*) and the Standard Work Tacton (seven-note *rhythm, smooth*) were confused. This is a confusion of both *rhythm* and *roughness* at the same time, and this error was more common than inter-*roughness* confusions. This result may be due to the design of the Tactons. The Tactons used were the same as created by Brown *et al.* [21] based on their own research on which forms of vibration felt subjectively "rougher" to the participants. Several participants described the "rough" vibration as feeling "broken up", essentially consisting of many pulses. The may have made the *rough* two-note *rhythm* feel like it consisted of more than two notes, making it subjectively similar to the *smooth* seven-note *rhythm*. A deliberate decision was made to use a different number of notes in the two *rhythms* to make them more distinguishable. However, this choice may have inadvertently introduced confusion when combined with *roughness*.

7.8 Limitations

This section discusses the main limitations of the research reported in this chapter, which should be considered when interpreting the results. The main limitations of Experiments 8 and 9 were: 1) attentional priming in Experiment 8, 2) basic environmental influences in Experiment 9 and 3) lack of skin temperature measure.

7.8.1 Attentional Priming During Experiment 8

In Experiment 8, the screen that asked participants to indicate which icon they interpreted the feedback as meaning was presented as soon as the icon feedback was presented, no input from participants was required to first indicate that they had detected a stimulus. This design choice was based on the same experimental design used in testing identification of Tactons [21]. Also, the stimuli used to create the Thermal Icons were those that were found to be most salient in Experiments 6 and 7, and so were least likely to be missed (i.e., not detected). However, the appearance of the response screen may have primed participants, indicating to them that an icon is being presented, potentially before they felt a change. Participants may, therefore, have been expecting an icon, which may have unintentionally improved detection and identification rates, as well as identification times. A better design would have been to require a response form participants indicating that they have felt a stimulus before presenting the response screen, as was done in Experiments 6 and 7.

7.8.2 Basic Environmental Influences

The choice of outdoor location for Experiment 9 was based on 1) participant comfort and 2) necessary amenities nearby, including flat open space, a place to sit and a nearby indoor area. Ethical considerations meant that the study could not be conducted in more representative locations, which might have a variety of noise levels and obstacles. Therefore, the location was very quiet and effectively free of obstacles. The environmental temperature and wind would have been a realistic influence on identification of thermal icons, but attentional demands would have been low. As was discussed concerning Experiment 7, the low visual and auditory attentional demands may have meant that participants could focus more closely on the changes in thermal stimulation, more so than they would in more realistic situations. This may have slightly improved identification performance.

7.8.3 Lack of Skin Temperature Measure

Skin temperature varies with environmental temperature, and skin temperature influences thermal perception [92, 223, 229], but the specific temperatures that the skin will change to will vary from person to person [118, 128, 223]. While the changes in environmental temperature were measured during Experiments 8 and 9, it was not possible to measure precise skin temperature of each participant. Due to the significant influence environmental temperature had on identification of thermal icons, it would have been beneficial to know skin temperature, to compare identification to skin temperature directly, rather than through the proxy of environmental temperature.

7.9 Conclusions

The first study, Experiment 8, tested identification of two-dimensional thermal icons and two-dimensional intramodal icons presented to the non-dominant hand when the participants were sat at an office desk indoors. Overall, identification Accuracy was quite high at 83%, suggesting that structured thermal feedback is a feasible means of conveying two pieces of information, if the application is not highly time-critical, as identification required 4-6 seconds. The *subjective intensity* parameter was more error-prone, at 85%, than the *direction of change* parameter at 97%. In particular the Strong Warmth icon was confused relatively frequently with the Moderate Warmth icon (but rarely *vice versa*), while both Cold icons were confused equally. This result, coupled with subjective opinions of reducing

sensitivity to thermal changes, suggested that *subjective intensity* might be a problematic parameter for feedback design. Intramodal icon identification Accuracy was significantly higher than that of thermal icons, at 97%: *rhythm* was identified with 97% Accuracy and *direction of change* was correctly identified at 99%. Intramodal identification was also 1-2 seconds faster. These results gave an indication that thermal feedback may be a suitable augmentation of vibrotactile feedback, as presenting two different tactile feedback channels together (thermal + vibrotactile) did not hinder interpretation of either.

The second study, Experiment 9, expanded upon the first to test identification of two-parameter thermal icons and two-parameter vibrotactile Tactons presented from a mobile phone to the left hand while participants sat and walked outdoors. The results were not as promising for thermal icons, as successful identification of both pieces of information dropped to 69% when sitting and 61% when walking, and Identification Time increased to 8-9 seconds. As during Experiment 8, the *subjective intensity* parameter was the least accurate, with only 73% Accuracy, compared to 96% for *direction of change*. Accuracy negatively correlated with increasing outdoor environmental temperatures, potentially because skin temperature changed with environmental temperatures, influencing the delta (difference) between skin temperature and Peltier stimulator temperature. The *extents of change* used to influence the *subjective intensity* would therefore, not be as far from skin temperature as they were intended to be. It also meant that, occasionally, icons were not detected at all, and yet the return of the stimulator to the neutral starting temperature *was* detected. This strong effect of external temperature suggested that thermal feedback should adapt to the environment, shifting the starting temperature to better match skin temperature, as well as vary how much it changes by to make it more perceivable. While the environment strongly influenced identification, there was no effect of mobility on identification, so walking outdoors did not markedly impact ability over sitting outdoors.

Thermal icon identification was compared to Tacton identification in Experiment 9 to judge the usefulness of thermal icons in relation to more established non-visual feedback method. Tacton identification remained high throughout the study (at 92%), significantly higher than thermal icons and identified significantly faster. Therefore, they remain the better means of conveying multidimensional information when walking and in outdoor environments. One important issue that arose, however, was the confusion of the *rough* vibrotactile texture, which felt subjectively "broken up" and the seven-note *smooth rhythm*. The "broken" nature of the sensation apparently felt subjectively similar to the multi-note *rhythm*, leading to confusion between the two. While having different numbers of notes in different Tacton *rhythms* can improve identification of those *rhythms*, it may be that the subjective texture of

the vibration should also be considered so that other confusions are not introduced.

7.9.1 Research Question 5

Research Question 5 asked:

"Can thermal stimulation be manipulated to convey multi-dimensional information?"

In general terms, the research in this chapter confirmed that thermal stimulation can be manipulated to convey multi-dimensional information, however, this is best achieved in a more thermally stable indoor environment. The research also suggested several ways of improving the design of thermal feedback to make individual icons unique and differentiable, such as increasing the difference in *subjective intensity*, using larger area for stimulation, using different spatial locations and dynamically adjusting the start and end-point temperatures used in the icons to improve salience and perceptual distinctiveness for outdoor environments. While thermal feedback is a feasible means of conveying information itself, the research in this chapter has also shown that thermal stimulation may successfully augment vibrotactile feedback, such as adding extra information bandwidth to Tactons. The two tactile channels were processed simultaneously with high Accuracy, and Tactons were well identified when outdoors and walking. The *direction of thermal change* was well perceived when outdoors and walking, so simple warming or cooling changes might successfully substitute less feasible Tacton parameters such as spatial location.

The contributions of the research in this chapter are:

- Structured, two-dimensional thermal stimulation that vary in their *direction of change* (warming/cooling) and *subjective intensity* (moderate/strong) can convey two pieces of information.
- Outdoor environmental temperature significantly hinders identification of structured thermal feedback.

8 Discussion and Conclusions

The hand is a multifaceted investigative and manipulative tool. Certain aspects of manual touch and haptic interaction have been well researched within HCI, including spatial gestures (such as tilting or drawing shapes) [42, 76, 160, 178, 193, 195], force-feedback [28, 110, 163], vibrations and textures [1, 21, 41, 139, 152], finger orientation [199, 200] and tangible devices [34, 54, 86, 106, 107, 172, 186, 252]. The application of pressure and thermal perception are two other inherent aspects of human touch, and human-object interaction that have an enormous contribution to make to interaction with mobile devices. However, their relative merits in designing mobile interfaces have not been fully explored. The aim of this research was to test the feasibility and usability of pressure input and thermal feedback for use in mobile HCI. The thesis statement read as follows:

The hand is a multifaceted investigative and manipulative tool. The application of pressure and the reception of thermal feedback are inherent aspects of manual touch and provide new opportunities to broaden the input and output capabilities for mobile device interaction. Pressure input on mobile devices is highly accurate when walking, when provided with either visual or audio feedback and when applying pressure from multiple digits, both individually and in combination. Individuals can detect a range of thermal stimuli produced from limited hardware designed for mobile interaction when both sitting and walking indoors. Using these stimuli, structured thermal icons can be created to convey two pieces of information to users when in both indoor and outdoor environments.

The following Research Questions have been addressed by the research in this thesis:

> **RQ1:** How accurate is pressure-based input on a mobile device when using only audio feedback?
>
> **RQ2:** How accurate is pressure-based input through the fingers when the individual is walking?
>
> **RQ3:** How accurate is pressure-based input when multiple digits apply pressure to a mobile device?

RQ4: What parameters of thermal stimulation are most detectable and comfortable for use in mobile interaction?

RQ5: Can thermal stimulation be manipulated to convey multi-dimensional information?

These five questions have been addressed by a review of the literature on the control of pressure applied through the fingers and the fundamental influences on human thermal perception, as well as a series of empirical research studies undertaken within HCI paradigms. These focused on applying the understanding of human capabilities to the design of interfaces utilising pressure input and thermal feedback. This chapter summarises the work reported in this thesis and discusses how the findings answer the five research questions. Limitations of the research conducted are discussed along with potential avenues for future research on both topics of pressure input and thermal feedback. The chapter ends with a final summary with general conclusions and the contributions of this thesis research.

8.1 Thesis Summary

Chapter 2 reviewed the literature on the application of pressure and thermal perception, from both a perceptual/psychophysical as well as an HCI perspective. The review described the factors influencing precise application of pressure that could impact the use of pressure as an input channel for mobile HCI. These included the number of digits used, the levels of pressure required and the influence of varying the feedback available to the individual. Chapter 2 also described several features of thermal stimulation that could be used to create unique and identifiable thermal feedback for mobile interaction, as well as how environmental factors could influence perception of thermal stimuli.

Chapter 3 reported an experiment that tested the precision of pressure input applied to a single pressure sensor during a linear targeting task, when participants were sat in a chair and provided with only visual feedback and only audio feedback. Previous research had suggested that non-visual pressure interaction was feasible, but a limited range of pressure, and more basic audio feedback designs, had been used. This experiment improved upon previous interfaces by facilitating pressure-based input across a larger range of pressure, at equal or better precision, using a semi-continuous audio feedback design. This chapter partly answered Research Question (RQ) 1 by showing non-visual pressure interaction can be

accurate when the individual is sitting.

Chapter 4 included two experiments that extended the findings from Chapter 3 to test precision of pressure input during linear targeting when walking a route indoors and provided with either visual or audio feedback. Two alternative input control methods, Positional and Rate-based, were also tested. Walking significantly impacted how accurately participants could control the pressure they applied, with incorrect selections, selection time and input variation (number of target crossings) all increasing when walking, compared to sitting. In the context of the linear targeting/menu interaction task, the use of a Rate-based control method improved task performance when walking. Using Rate-based control, mobile pressure-based input using only audio feedback was just as accurate as sitting and using visual feedback, albeit at the expense of longer selection times. This chapter completed the answer to RQ 1, and answered RQ 2, by showing that pressure-based input can be accurate when walking, even when using non-visual feedback.

The experiments in Chapter 3 and 4 only used a single pressure sensor pressed by the thumb. Chapter 5 described two experiments using pressure input from all five digits of one hand, both individually and in various combinations. Sensors for pressure input were placed on the sides, back and top of a mobile phone. The first experiment tested how precisely pressure could be applied to the device by each digit and various combinations of digits (grips), while the same hand also held the device. The results showed that each individual digit could accurately select 6 levels of pressure in a menu interaction task, but not all were equally well controlled. Grips varied in how precisely they could control applied pressure, with 2- and 3-digit grips performing well, but 4- or 5-digits performed poorly, possibly because of awkward sensor placement. The second experiment used the best-performing digits/grips and compared one-handed, multi-digit pressure input with two-handed multitouch input during a zooming and rotating map task. Performance was very close between the two methods, with multitouch input being 1% more accurate but pressure input being 0.5 sec faster, but the results suggest that concurrent multi-digit pressure input is accurate and capable of replacing two-handed touch input. Chapter 5 answered RQ 3, demonstrating how accurately each individual digit, and various combinations, can apply pressure to a mobile device.

Chapter 6 described two experiments that tested detection of thermal stimuli that varied in their *direction* (warm or cold, starting at 32°C), *rate of change* (1°C/sec or 3°C/sec) and *extent of change* (1°C, 3°C or 6°C) when sitting and walking indoors. The purpose was to identify which forms of feedback might be suitable for use in designing structured thermal

feedback for mobile interaction. The stimuli chosen for testing were based on the perceptual and psychophysical research in Chapter 2 describing the influence of the three factors on thermal sensation. The results showed that, while equal numbers of warm and cold stimuli were detected, cooling stimuli were detected faster and were rated as subjectively less intense and more comfortable than warming stimuli. Faster changes, and larger extents of change (from skin temperature) were more likely to be detected, and detected faster, but were consequently rated as more intense and less comfortable than slower/smaller changes. Walking significantly reduced the number of stimuli detected, although the same patterns of detection/intensity/comfort remained. The chapter ended with guidelines for the creation of thermal feedback, based on the results. This chapter answered RQ 4.

Chapter 7 included two experiments that tested absolute identification of unique, structured thermal stimuli called 'thermal icons' when the individual was a) sitting indoors, b) sitting outdoors and c) walking outdoors. The design of the icons was based on the thermal stimuli from Chapter 6 that were reliably detected and rated as comfortable. Icons were two-dimensional, using *direction of change* (warm and cold) and *subjective intensity of change* (moderate and strong), to convey the Source (Personal or Work) and Importance (Standard or Important) of a hypothetical message. Both pieces of information could be identified through only thermal stimulation with 83% accuracy when sitting indoors. Accuracy rates for both pieces of information dropped to 69% and 61% when sitting and walking outside, respectively. Environmental temperature had a large impact on identification rates, with a strong negative correlation between accuracy and temperature. Chapter 7 answered RQ 5, confirming that thermal stimulation can be manipulated to convey multi-dimensional information.

8.2 Research Question 1

How accurate is pressure-based input on a mobile device when using only audio feedback?

Research Question 1 was answered in Chapters 3 and 4. The first step was to create a suitable audio feedback design, and establish its usefulness. Previous research had suggested that continuous external feedback was necessary to facilitate precise application of pressure, but existing non-visual feedback designs for pressure input had used only discrete feedback, generally indicating transition between set pressure levels. The preliminary study and Experiment 1 in Chapter 3 iteratively designed a horizontally spatialised, egocentric audio

display consisting of both synthetic speech and musical components and providing both continuous and discrete cues. Using a linear targeting task, Chapter 3 compared control of pressure-based input when the participant was provided with this audio feedback design to when he/she was provided with traditional continuous visual feedback, to judge the efficacy of the design. This was done while the participants were sitting down, to provide an ideal, baseline level of performance. The results showed that overall accuracy using the audio feedback was poorer (74%) than using visual feedback (85%), but that accuracy was higher (>= 85.5%) when a smaller number of wider targets were used (4 or 6 levels) and a Dwell selection technique was employed. This suggested that pressure-based interaction using only audio feedback can be accurate, at least when sitting.

Chapter 4 also addressed RQ 1, by comparing performance in the linear targeting task when the participants were walking and sitting, and provided with audio or visual feedback. It also compared two different control methods: Positional and Rate-based. Using only audio feedback resulted in poorer control of applied pressure (Positional input) when walking, reducing accuracy from 97% when using visual feedback to 87% when using audio, and increasing selection times by 1.2 seconds. However, using the Rate-based control, targeting accuracy (97%) and subjective ratings of workload were not negatively affected by using audio feedback, although selection time still increased by 1.4-1.6 seconds.

While pressure input (using Positional control) was less precise when walking and using audio feedback, targeting accuracy remained relatively high at 87%. Also, using Rate-based input improved targeting task accuracy to 97%. These accuracy rates are equal to, or above, those of other research looking at pressure-based input when sitting at a desktop and provided with visual feedback. Therefore, overall, this thesis concludes that pressure-based input can be highly accurate when using only audio feedback.

8.3 Research Question 2

How accurate is pressure-based input when the individual is walking?

There are two aspects to this question: 1) how accurate is the *control of applied pressure* when walking and 2) how accurate is *pressure-based interaction* when walking. Both aspects were answered by Experiments 1 to 3 in Chapters 3 and 4. Positional control is a measure of applied pressure, and so performance when using this method while walking during Experiments 2 and 3 in Chapter 4 addressed the first aspect. Rate-based control

allowed for superior linear targeting task performance, and so the Rate-based results from Experiments 2 and 3 addressed the second aspect.

Regarding aspect 1, walking negatively affected the participants' ability to accurately control the amount of pressure applied to the device. Therefore input was more variable and led to more unintended incorrect selections. Measures of subjective workload were also higher. While the error rate increased by 64% when walking (compared to sitting), it only increased to 2.93% when presented with visual feedback, which is still very low. The combination of walking and using audio feedback had a stronger negative impact on precision of input, as errors increased to 12.8%, more than 4x the errors when walking with visual feedback, and selection time increased by a further 1.2 seconds. Therefore, in answer to aspect 1 of Research Question 2, the application of pressure is significantly less precise when walking, both in terms of target selection as well as variability of input. Participants clearly find it challenging to maintain very precise levels of pressure when in motion, especially when only audio feedback is available.

Regarding aspect 2, walking significantly decreased overall accuracy during Rate-based control, but, interestingly, only when using visual feedback. It had no effect on overall accuracy when only audio feedback was used: static and mobile non-visual Rate-based interaction was equally accurate. While selection time increased when walking and using audio feedback (compared to sitting) sitting and walking selection times were equal when using visual feedback, so walking had no effect on task time when using visual feedback. Mobile interaction was significantly more mentally/physically demanding, however. Therefore, in answer to aspect 2 of Research Question 2, accuracy of pressure-based input was very good, at only 3% errors, and selection time was also no longer compared to sitting, when using visual feedback. Overall, the research suggests that control of applied pressure is good when walking, and pressure-based interaction, through the use of Rate-based input, can be as highly accurate and fast when walking as it is when sitting.

8.4 Research Question 3

How accurate is pressure-based input when multiple digits apply pressure to a mobile device?

In a similar manner to Research Question 2, RQ 3 also has two aspects to it: a) the accuracy of pressure applied to a mobile device by multiple digits and b) the accuracy of a multi-digit

pressure interface. Both aspects were answered by Experiments 4 and 5 in Chapter 5. For aspect 1, seven force-sensing resistors were attached to the sides, back and top of a mobile phone to provide input from each digit of the one hand holding the phone. Control of the pressure applied by each digit individually, and 14 different digit combinations (grips), was tested while the participants were sitting down, again using linear targeting. Experiment 4 found that all individual digits could be used for precise input on a mobile device (albeit not all equally well) even while simultaneously holding the device. The 14 digit combinations tested also performed well in general, but there were large differences between grips. The results therefore suggested that multi-digit pressure input from different positions on a mobile phone can be highly accurate.

Addressing the second aspect, Experiment 5 used some of the best performing grips, and a zooming and rotating map interaction was developed using four concurrent pressure-based input channels, which mapped common two-handed touchscreen interactions to one-handed pressure interaction. This setup was shown to provide a similarly high level of accuracy and task time compared to traditional multitouch input, but with the advantages of requiring only one hand and leaving the screen fully visible. Therefore, in answer to RQ 3, multi-digit pressure input is not only precise in isolation (Experiment 4) it can also provide accurate input to common mobile interactions (Experiment 5).

8.5 Research Question 4

What parameters of thermal stimulation are most salient and comfortable when using equipment designed for mobile interaction?

Research Question 4 was addressed firstly in Chapter 2, where the literature on human thermal perception was discussed. The features of thermal stimulation that influence thermal sensations, and how they do so, were of particular interest, as the intention was to create feedback that produced sensations which varied in systematic ways, to make them uniquely identifiable. Research showed that the spatial location of stimulation (around the body), the rate of simulation change and the extent of stimulation change all influenced the qualitative thermal sensation. Therefore, the first experiment in Chapter 4 tested the identification of stimuli that warmed or cooled by either 1°C, 3°C or 6°C at either 1°C/sec or 3°C/sec, presented to the fingertips, thenar eminence/palm of the hand, top of the forearm or top of the upper arm. Stimuli were presented when both sitting and walking indoors. The purpose was to gain measures of how reliably detectable the stimuli were and how subjectively

intense and comfortable they felt. From the results, several guidelines were developed on which stimuli would be suitable for use in thermal feedback. The guidelines answered RQ 4:

1) The thenar eminence is the optimal location for feedback, but non-glabrous arm locations are also suitable.
2) 1°C/sec and 3°C/sec changes are both suitable (but power-requirements should be considered).
3) Warm and cold stimuli are both suitable for use.
4) Extent of change can be used as a parameter for feedback design as different intensities are perceptually different, however 1°C changes are not suitable.

8.6 Research Question 5

Can thermal stimulation be manipulated to convey multi-dimensional information?

This Research Question was answered by two experiments in Chapter 7. Following the guidelines from Chapter 6, two-dimensional structured thermal icons were developed, based on the design of multi-dimensional Earcons and Tactons. They were intended to convey two pieces of information, one through the *direction of thermal change* and one through the *subjective intensity of change*. Using two *directions* (warm and cold) and two *intensities* (moderate and strong) produced four icons: moderate warmth, strong warmth, moderate cool, strong cool.

When sitting indoors, participants could correctly identify both the *direction* and *intensity* together with 83% accuracy, and each individual parameter with 97% and 85% accuracy, respectively. When sitting outdoors, identification accuracy for both parameters dropped to 69%, and walking reduced correct identification to 61%. Overall, *direction of change* was correctly identified 96% of the time, and *subjective intensity* was correctly identified with 73.1% accuracy. In general terms, Experiments 8 and 9 in Chapter 7 confirmed that thermal stimulation *can* be manipulated to convey multi-dimensional information, however, this is best achieved in an indoor environment and can currently only convey two pieces of information. The research also suggested how to potentially make icons more easily identifiable outdoors, such as increasing the difference in *subjective intensity*, using larger area for stimulation, using different spatial locations and dynamically adjusting the start and end-point temperatures used in the icons to improve salience and perceptual distinctiveness for outdoor environments.

8.7 Contributions

This section summarises the novel contributions of the research in this thesis, on Pressure-based input and thermal feedback.

8.7.1 Pressure Input on Mobile Devices

- Walking significantly degrades control of applied pressure but linear targeting performance across a wide range of pressure input remains as good as in previous research that tested control when sitting in desktop interactions scenarios.
- Non-visual control of a wide range of pressure, using only audio feedback, can be highly accurate when sitting and walking.
- One-handed, multi-digit pressure input can provide comparable control over bi-directional zooming and rotating as two-handed multitouch input, with the advantages of requiring only one hand and leaving the screen fully visible.
- Each individual digit of one hand, and various combinations, can apply force accurately but not equally so, when both holding and squeezing a mobile phone.
- Increasing the range of detectable force, relative to the number of digits used, improves precision of applied pressure.

8.7.2 Thermal Feedback for Mobile Devices

- This thesis research has included the first detailed investigation into the perception and design of thermal feedback for mobile HCI. It has identified which features of thermal stimulation might be suitable for thermal feedback in mobile HCI
- Thermal stimuli are less salient in realistic HCI scenarios than in controlled psychophysical lab studies, and walking reduces stimulus salience further.
- Structured, two-dimensional thermal stimulation that vary in their *direction of change* (warming/cooling) and *subjective intensity* (moderate/strong) can convey two pieces of information.
- Outdoor environmental temperature significantly hinders identification of structured thermal feedback.

8.8 Limitations and Future Work

8.8.1 Lack of Qualitative Data

The NASA TLX gave basic subjective reports about the interaction, but no longer form qualitative data was gathered, such as via interviews, concerning more specific opinions participants might have had about the feedback designs or the interface. The objective measures of error rates, selection times and crossings only inform about task performance, and not whether the participants liked or disliked the design choices of the interaction in general and so whether they would want to use such an interface in reality. The relatively low ratings of subjective workload, and annoyance at the audio design, suggest that attitudes towards the interaction and the interface were not overly negative, but details as to what was or was not appreciated are missing.

8.8.2 Environmental Considerations

The walking routes used during both the pressure research (Experiments 2 and 3) and thermal feedback research (Experiments 7 and 9) were simple loops, used to introduce the potential influence of bodily movement on control of pressure or perception/identification of thermal stimuli. However, they did not necessarily introduce the cognitive demands or attentional distractions that might occur in everyday walking scenarios, such as monitoring and adjusting unpredictable routes or paying visual attention to potential hazards. These factors are difficult to realistically test, for safety and ethical concerns, but it would be important to understand their potential impact on applied pressure or thermal perception.

8.8.3 Pressure input

8.8.3.1 Vibrotactile Feedback Design

While audio feedback successfully facilitated eyes-free mobile pressure input, audio feedback is not always useful or desired [97, 101]. In Experiments 1 to 3, headphones were needed to present the stereo sound and users may not always have, or want to use, them. Using vibrotactile feedback may be an alternative, and would allow for in-pocket interaction. Measuring the precision of pressure input when providing only vibrotactile feedback has been researched, but only using a small range of pressure, and a small number of (wide) targets levels [228, 232, 233]. Some research has used a wide, expressive range of pressure

input and vibrotactile feedback to create messages/convey emotions [83, 193, 230], but user ability to accurately control the amount of pressure applied was not measured. In these cases discrete vibration pulses or rhythms were used to indicate the transition from one pressure level to another. Future work should look at alternative feedback designs that might facilitate interaction with a wider range of pressure input.

8.8.3.2 Personalisation

The poorer performance using 4- and 5-digit grips in Experiment 4 may be partly due to the positioning of the sensors around the device. Ideally a device that is being used to detect multiple pressure inputs should, therefore, detect pressure across its entire body, to ensure input is detected regardless of finger location. If this is not possible, the positioning of sensors should be carried out on an individual basis, suited to the hand and finger size, and preferred finger placement of each user. Then, further research should ascertain whether multi-digit grips perform better when input can be provided without obstacle, or, perhaps, whether the individual differences in digit placement are so wide that few general assumptions or conclusions can be drawn.

There are wide individual differences in how much pressure a person can physically or comfortably apply [33, 109, 214], and Li *et al.* [153] recommended tailoring the pressure space to the individual. The Incremental pressure space improved control of pressure in Experiment 4, increasing the total space by 3.5 N per digit used. However, further research should be done into personalising the Incremental pressure space, including the smallest single-digit space, to test the influence on accuracy of pressure input.

8.8.3.3 One-handed Pressure Input when Walking and Encumbered

Experiments 3 and 4 tested single-digit pressure input when walking, and Experiment 5 showed that multi-digit pressure input when sitting could match multitouch input, with the benefit of only using a single hand and leaving the screen clear. Future research should compare one-handed, multi-digit pressure input to multitouch for a wide range of tasks when the user is walking and possibly encumbered, for example holding something in a second hand. Example tasks include zooming, scrolling, rotating and OS navigation. This research would show whether or not having the screen visible and only needing one hand for input improves task performance (accuracy and speed), keeps walking speed closer to optimal or reduces physical demand.

8.8.3.4 Alternative Multi-digit Functions

Multi-digit input was only tested for menu interaction and map zooming/rotation. Future work should research other uses for multiple concurrent and continuous inputs. Examples could include:

- More complex pressure chords than those of Tang and colleagues [232, 233].
- As suggested in Chapter 5, the use of different numbers, or constellations, of digits could enact different functions. Squeezing with two digits could scroll a small list of common contacts, or several digits could scroll a larger list. Alternatively, using audio feedback, squeezing on one sensor lightly could inform the user how many emails have been received, pressing moderately reads out the sender and title of the most recent mail and pressing strongly activates a synthetic speech reading of the latest email. Pressing a second sensor alters the speed at which the email is read. Different sensors/positions could access different tasks, such as email, SMS, social network feeds in this manner.

8.8.4 Thermal Feedback

8.8.4.1 Alternative Hardware

The hardware commissioned from SAMH Engineering used Peltier modules to provide the thermal sensations, as they can be controlled precisely and produce both warm and cold stimuli. However, they require relatively large amounts of power to change temperature quickly and regularly. There may be other, more energy-efficient means of producing thermal sensations, that should be investigated and tested for what range of stimuli they can produce, and how accurately.

Spatial Location was not used as a parameter in the design of thermal feedback, as it was deemed less useful for realistic use, requiring the individual to attach or wear specific additional hardware around the body to get feedback. However, there is a growing trend towards 'wearable computing', where commonly used or worn devices, such as watches[15][16] or glasses[17], augment existing perfunctory functions with additional digital capabilities. Thermal elements could feasibly be incorporated into these, and other[18], wearable devices to provide thermal feedback around the body in a less obtrusive and encumbering fashion.

[15] http://www.sonymobile.com/gb/products/accessories/smartwatch/

[16] http://getpebble.com/

[17] http://www.google.com/glass/start/

[18] https://getmyo.com/

Being directly in contact with the body at all times would also make detection of feedback easier, rather than relying on a mobile device to be held in the hand.

8.8.4.2 Wider Range of Stimuli

The range of stimuli tested for salience and comfort in Experiments 6 and 7 was partly limited by the hardware being used, as the available rates of change (ROC) were limited to only 1°C/sec and 3°C/sec. A small number of extents of change were chosen to keep task time reasonable. Those that were chosen were away from pain thresholds and were hypothesised to be perceptually distinct but also reliably detectable in mobile interaction scenarios. However, future research should conduct further tests on other ROC and other extents, to identify other stimuli that might be usable for thermal feedback.

Similarly, research should be conducted into the usefulness of those parameters considered unsuitable for this thesis research: area, rhythm and spatial location/spatial pattern. While they were legitimately discounted for the purposes of this research, they may still provide uses or benefits for thermal feedback, and their merits should be investigated.

Finally, with the use of other stimuli or parameters, it may be possible to convey more than the two pieces of information achieved in the research in Chapter 7. While the poor identification accuracy when outdoors suggests that the information transmission capacity for thermal feedback may be limited, alternative feedback designs may prove successful.

8.8.4.3 Adapt Thermal Icons to Environmental Temperature

The results from Experiment 9 showed that the surrounding environmental temperature heavily influenced the identification of thermal icons, with identification accuracy becoming worse as temperatures rose. Environmental temperature is known to influence skin temperature, which, in turn, influences perception of thermal stimuli [92, 223, 229]. A deliberate decision was made to keep the starting temperature for thermal icons constant at 32°C for several reasons. Firstly, while skin temperature varies across individuals, and within the same individual, over time, 32°C is likely to be within most, if not all, individuals' natural range of resting skin temperatures [128, 223], making it an appropriately representative of skin temperature. Secondly, 32°C had been used as starting temperature in previous psychophysical research [61, 133], making comparisons with controlled laboratory studies concerning stimulus salience/detection easier, to highlight the influence of being in more realistic HCI scenarios. Finally, it was the intention to create a set of clearly structured thermal icons that can be identified in a variety of environments. The subjective intensity of

a given thermal stimulus varies depending on current skin temperature, such that cooling changes feel stronger at cold skin temperatures while warming changes feel weaker, and vice versa for warming changes/warm skin temperature. This inherent variability would have made defining a clear, and generalizable, set of icons, including a set of useable parameters for icon construction, much more challenging.

Because 32°C rests within the natural range of skin temperatures, and because the Peltiers could be set to adapt the skin to 32°C, it was hypothesised that the influence of (non-extreme) environmental temperatures could be mitigated, allowing the icons to be more clearly perceived and identified. However, Experiment 9 suggested that this is not the case. To use thermal icons in mobile environments, the mobile device must be able to directly measure (through a thermometer), or calculate from (potentially online) information sources the current environmental conditions, and dynamically adapt the design of thermal icons to suit resulting skin temperature. This would potentially involve changing: 1) the starting temperature, 2) the extent of change and 3) the rate of change, based on the desired magnitude of subjective intensity, relative to skin temperature. This could make the icons a) more salient and b) more clearly identifiable. Future research should investigate how best to do this and then whether the icons can be better identified.

8.8.4.4 Inherent Responses or Semantic Associations of Thermal Feedback

There are inherent meanings to some thermal sensations, such as “danger” to extreme hot or cold, or emotional responses to others, such as a cold breeze on a hot day or the warm touch of another person. Unfortunately, there are also associations to be made between electronic devices and heat, as it is often a sign of intense processor/cellular activity or even a fault within the device. Warming changes were used to represent Personal messages in Experiments 8 and 9, as there may be an inherent associations between physical warmth and interpersonal warmth [243]. But do these associations hold when the stimulus comes from an electronic device? What does the user automatically think, or feel, when his/her mobile device warms up or cools down? Does the extent by which it warms or cools influence these perceptions? For example, a moderately warm change may be associated with a positive, personal communication, but stronger warmth may be associated with a negative event, either a fault within the device, or an unhappy friend/co-worker. Research has started to investigate what inherent responses, or semantic associations, users have when presented with different thermal changes [151, 230] but there are many such potential associations, which should be explored.

8.8.5 Pressure Input and Thermal Feedback Combined

Pressure input and thermal feedback are both inherent aspects of touch, and could be combined to create more engaging interfaces. Lederman & Klatzky [148] identified that the exploratory procedure of enclosure (wrapping the hand around an object) was sufficient for the investigation of both hardness (along with pressure) and temperature. The act of enclosing an object, and pressing on it a) applies pressure to the object to get feedback from it (regarding hardness) and b) provides thermal feedback about the object. These could be combined in an interface: Sue's phone is lying on the table beside her as she works. It is set on silent with no vibration so she is not disturbed. As the screen can flash on when a message arrives it is also faced down. In her own time she wants to know if she has any messages but, rather than picking it up, pressing the power button and looking at on-screen icons, thereby interrupting her workflow, she simply wraps her hand around the device, squeezing it. On one occasion the phone is ice cold, and so she knows she has no new messages. Later, she holds her phone again but this time it is moderately warm, indicating a small number of messages. Later still it is hot indicating a large number of messages, or the arrival of an important message.

An alternative combination of pressure and thermal feedback could be to give information about on-screen virtual buttons, using a pressure-sensitive touchscreen. Pressing lightly on a button provides a glance at the button's function, pressing hard enacts the button. An example use could be pressing on a suspicious link in a web-browser, or on a permanent deletion UI dialog button, which causes the phone to heat up, indicating danger, or cool down to indicate safety.

8.9 Conclusions

This thesis has investigated the feasibility of pressure-based input and thermal feedback for interaction with mobile devices. It followed research methodologies akin to those from perceptual and psychophysical science, to determine the fundamental characteristics of applied pressure and thermal perception using mobile devices, and then design interfaces based on these characteristics. For pressure input, the thesis focused on measuring how precisely individuals can control the amount of pressure they apply with one or more digits to mobile devices, when they are sitting, walking and provided with non-visual feedback. For thermal feedback, the research identified which forms of thermal stimulation would be

salient and comfortable when produced from limited hardware when the individual was walking, in order to develop ‘thermal icons’: structured, multi-dimensional thermal stimuli capable of conveying multiple pieces of information thermally.

Pressure input has been shown to provide accurate input in stable desktop interactions scenarios, but, while research has also studied pressure input on mobile devices, only one study has tested the influence of walking on applied pressure, which is essential to know for the development of appropriate interfaces. The research in this thesis has been the first to systematically test control of a wide range of pressure input when the user is walking. This research has shown that, while control of applied pressure degrades when walking, 10 distinct levels of pressure spread over 3.5 N can still be targeted with high accuracy. Both the number of levels and accuracy rates are above any previously found for mobile interaction and equal those found in static desktop scenarios. Because non-visual interaction is an important consideration for mobile devices, this thesis has also contributed an audio feedback design that can facilitate mobile pressure-based interaction: a second novel contribution. Using a rate-based control method, accuracy of pressure-based targeting can be as high as when sitting and being provided with visual feedback. Finally, this thesis research has shown that one-handed, multi-digit pressure-based input can perform as equally well as traditional two-handed multitouch input, while providing the benefits of requiring only one hand and leaving the screen free of obstructing fingers. This could potentially aid interaction with touchscreen devices when walking and encumbered, such as carrying bags or small children, as only one hand is required for input, and the content on screen is fully visible.

Traditional forms of non-visual feedback, namely audio or vibrotactile, are not always suitable for use, depending on the environment, and user preference for which modality is desired also varies. Thermal stimulation is an inherent means of receiving information in the natural world, and has the benefit of being entirely silent and personal. While thermal feedback has been used to convey information in complex Virtual Reality setups, its use for mobile interaction is much more limited. This thesis has included the first detailed study of designing thermal feedback specifically for mobile interaction. It identified which forms of thermal stimulation are reliably and comfortably detectable when both sitting and walking indoors, and so produced a set of guidelines for how best to design thermal feedback for mobile interaction. This thesis research has also shown the first examples of thermal icons: two-dimensional structured thermal feedback that can convey two pieces of information simultaneously, in a similar manner to Earcons and Tactons, in the audio and vibrotactile domains, respectively.

Pressure input and thermal feedback are two under-researched aspects of touch and the research in this thesis has successfully shown that both can be utilised to broaden haptic interaction in mobile human-computer interfaces. Pressure input could augment mobile interfaces with several accurate and continuous input channels, one for each of multiple digits. Eyes-free pressure input is also highly feasible. Thermal feedback is a promising means of conveying information, especially for indoor locations, where identification rates remained high. Thermal feedback can be salient, even when outdoors, and can provide a unique form of tactile feedback.

Appendices

To save on paper, the Appendices are available at:

http://www.dcs.gla.ac.uk/~gawilson/appendices.html

A. Experimental files for Experiment 1 and Preliminary study (Chapter 3)
B. Experimental files for Experiments 2 and 3 (Chapter 4)
C. Experimental files for Experiments 4 and 5 (Chapter 5)
D. Design of Hardware for Mobile Thermal Feedback
E. Experimental files for Experiments 6 and 7 (Chapter 6)
F. Experimental files for Experiment 8 and 9 (Chapter 7)

References

1. Ahmaniemi, T., Marila, J. and Lantz, V. Design of dynamic vibrotactile textures. *IEEE Transactions on Haptics*, 2010. 3(4), pp 245-256.
2. Baba, T., Kushiyama, K. and Doi, K. ThermoGame: video game interaction system that offers dynamic temperature sensation to users. *Proceedings of SIGGRAPH 2010*, pp Article 51.
3. Baglioni, M., Malacria, S., Lecolinet, E. and Guiard, Y. Flick-and-Brake: Finger Control over Inertial/Sustained Scroll Motion. *Proceedings of CHI 2011 Extended Abstracts*, pp 2281-2286.
4. Barnard, L., Yi, J., Jacko, J. and Sears, A. An empirical comparison of use-in-motion evaluation scenarios for mobile computing devices. *International Journal of Human-Computer Studies*, 2005. 62, pp 486-520.
5. Benali-Khoudja, M., Hafez, M., Alexandre, J. and Kheddar, A. Thermal feedback interface requirements for virtual reality. *Proceedings of Eurohaptics 2003*, pp 438-443.
6. Bergmann Tiest, W.M. and Kappers, A.M.L. Cues for Haptic Perception of Compliance. *IEEE Transactions of Haptics*, 2009. 2, pp 189-199.
7. Bergstrom-Lehtovirta, J., Oulasvirta, A. and Brewster, S. The Effects of Walking Speed on Target Acquisition on a Touchscreen Interface. *Proceedings of MobileHCI 2011*, pp 143-146.
8. Bessou, P. and Perl, E.R. Response of cutaneous Sensory Units with Unmyelinated Fibers to Noxious Stimuli. *Journal of Neurophysiology*, 1969. 32, pp 1025-1043.
9. Birznieks, I., Jenmalm, P., Goodwin, A. and Johansson, R. Encoding of Direction of Fingertip Forces by Human Tactile Afferents. *Journal of Neuroscience*, 2001. 21, pp 8222-8237.
10. Blattner, M., Sumikawa, D. and Greenberg, R. Earcons and Icons: Their Structure and Common Design Principles. *Human-Computer Interaction*, 1989. 4(1), pp 11-44.
11. Brewster, S. Overcoming the lack of screen space on mobile computers. *Personal and Ubiquitous Computing*, 2002. 6, pp 188-205.
12. Brewster, S. and Brown, L. Tactons - Structured Tactile Messages for Non-Visual Information Display *Australasian User Interface Conference 2004*, pp 15-23.
13. Brewster, S., Chohan, F. and Brown, L. Tactile feedback for mobile interactions. *Proceedings of CHI 2007*, pp 159-162.
14. Brewster, S. and Hughes, M. Pressure-Based Text Entry for Mobile Devices. *Proceedings of MobileHCI 2009*, pp Article 9.
15. Brewster, S., Lumsden, J., Bell, M., Hall, M. and Tasker, S. Multimodal 'Eyes-Free' Interaction Techniques for Wearable Devices. *Proceedings of CHI 2003*, pp 473-480.
16. Brewster, S., Wright, P. and Edwards, A. A detailed investigation into the effectiveness of earcons. *Proceedings of ICAD 1994*, pp 471-498.
17. Brewster, S., Wright, P. and Edwards, A. An evaluation of earcons for use in auditory human-computer interfaces. *Proceedings of InterCHI 1993*, pp 222-227.
18. Brewster, S., Wright, P. and Edwards, A. Experimentally derived guidelines for the creation of earcons. *Adjunct Proceedings of HCI*, 1995.
19. Brown, L. and Brewster, S. Tactile crescendos and sforzandos: Applying musical techniques to tactile icon design. *Proceedings of CHI 2006 Extended Abstracts*, pp 610-615.
20. Brown, L., Brewster, S. and Purchase, H. A First Investigation into the Effectiveness of Tactons. *Proceedings of WorldHaptics 2005*, pp 167-176.
21. Brown, L., Brewster, S. and Purchase, H. Multidimensional Tactons for Non-Visual

Information Presentation in Mobile Devices. *Proceedings of MobileHCI 2006*, pp 231-238.
22. Brown, L. and Kaaresoja, T. Feel who's talking: using tactons for mobile phone alerts. *Proceedings of CHI 2006 Extended Abstracts*, pp 604-609.
23. Büring, T., Gerken, J. and Reiterer, H. Zoom interaction design for pen-operated portable devices. *International Journal of Human-Computer Studies*, 2008. 66, pp 605-627.
24. Butler, A., Izadi, S. and Hodges, S. SideSight: multi-touch interaction around small devices. *Proceedings of UIST 2008*, pp 201-204.
25. Cain, W. Spatial discrimination of cutaneous warmth. *The American Journal of Psychology*, 1973. 86(1), pp 169-181.
26. Caldwell, D.G., Lawther, S. and Wardle, A. Tactile perception and its applications to the design of multimodal cutaneous feedback systems. *Proceedings of IEEE Robotics and Automation 1996*, pp 3215-3321.
27. Campbell, C., Zhai, S., May, K. and Maglio, P. What You Feel Must Be What You See: Adding Tactile Feedback to the Trackpoint. *Proceedings of INTERACT 1999*, pp 383-390.
28. Casiez, G. and Vogel, D. The effect of spring stiffness and control gain with an elastic rate control pointing device. *Proceedings of CHI 2008*, pp 1709-1718.
29. Cechanowicz, J. and Gutwin, C. Augmented Interactions: A Framework for Adding Expressive Power to GUI Widgets. *Proceedings of INTERACT 2009*, pp 878-891.
30. Cechanowicz, J., Irani, P. and Subramanian, S. Augmenting the mouse with pressure sensitive input. *Proceedings of CHI 2007*, pp 1385-1394.
31. Chang, W., Kim, K.E., Lee, H., Cho, J.K., Soh, B.S., Shim, J.H., Yang, G.H., Cho, S.J. and Park, J. Recognition of Grip-Patterns by Using Capacitive Touch Sensors. *Proceedings of IEEE ISIE*, pp 2936-2941.
32. Cherycroze, S. Painful Sensation Induced by a Thermal Cutaneous Stimulus. *Pain*, 1983. 17(2), pp 109-137.
33. Christou, E., Grossman, M. and Carlton, L. Modeling variability of force during isometric contractions of the quadriceps femoris. *Journal of Motor Behavior*, 2002. 34, pp 67-81.
34. Chu, G., Moscovich, T. and Balakrishnan, R. Haptic Conviction Widgets. *Graphics Interface 2009*, pp 207-210.
35. Clarkson, E.C., Patel, S.N., Jeffrey, S.P. and Abowd, G.D., *Exploring Continuous Pressure Input for Mobile Phones*, in *GVU Technical Report; GIT-GVU-06-20*. 2006, Georgia Institute of Technology: Seattle, WA.
36. Claus, D., Hilz, M., Hummer, I. and Neundorfer, B. Methods of measurement of thermal thresholds. *Acta Neurologica Scandinavica*, 1987. 76(1), pp 288-296.
37. Crossan, A., Brewster, S. and Ng, A. Foot Tapping for Mobile Interaction. *Proceedings of BCS HCI 2010*, pp 418-422.
38. Crossan, A., McGill, M., Brewster, S. and Murray-Smith, R. Head Tilting for Interaction in Mobile Contexts. *Proceedings of MobileHCI 2009*, pp Article 6.
39. Crossan, A., Murray-Smith, R. and Brewster, S., *Instrumented usability analysis for mobile devices*, in *Handbook of Research on User Interface Design and Evaluation for Mobile Technology*, J. Lumsden, Editor. 2007, Information Science Reference: PA, USA. p. 927-944.
40. Crossan, A., Murray-Smith, R., Brewster, S., Kelly, J. and Musizza, B. Gait phase effects in mobile interaction. *Proceedings of CHI 2005*, pp 1312-1315.
41. Crossan, A., Williamson, J. and Brewster, S. Artex: Artifical Textures from Everyday Surfaces for Touchscreens. *Proceedings of CHI 2010*, pp 4081-4086.
42. Crossan, A., Williamson, J. and Brewster, S. Wrist rotation for interaction in mobile contexts. *Proceedings of MobileHCI 2008*, pp 435-438.
43. Darian-Smith, I. and Johnson, K. Thermal sensibility and thermoreceptors. *Journal of Investigative Dermatology*, 1977. 69, pp 146-153.
44. Darian-Smith, I., Johnson, K., LaMotte, C., Shinegana, Y., Kenins, P. and Champness, P. Warm fibers innervating palmar and digital skin of the monkey:

responses to thermal stimuli. *Journal of Neurophysiology*, 1979. 42(5), pp 1297-1315.
45. Edin, B. and Johansson, R. Skin strain patterns provide kinaesthetic information to the human central nervous system. *Journal of Physiology*, 1995. 487, pp 243-251.
46. Eijkman, E. Black-Box Analysis of the Skin Senses as a Multiple Communication Channel. *Biological Cybernetics*, 1989. 61(3), pp 163-170.
47. Eslambolchilar, P., Crossan, A., Murray-Smith, R., Dalzel-Job, S. and Pollick, F. Model-Based Target Sonification in Small Screen Devices: Perception and Action. *Handbook of Research on User Interface Design and Evaluation for Mobile Technology*, 2008, pp 478-506.
48. Feige, S. Can You Feel It?–Using Vibration Rhythms to Communicate Information in Mobile Contexts. *Proceedings of INTERACT 2009*, pp 800-803.
49. Fitts, P. The information capacity of the human motor system in controlling the amplitude of movement. *Journal of experimental psychology*, 1954. 47(6), pp 381-391.
50. Forlines, C., Shen, C. and Buxton, W. Glimpse: a novel input model for multi-level devices. *Proceedings of CHI 2005*, pp 1375-1378.
51. Fowler, S. and Notterman, J. Observed Short-Term Motor Memory Effect for Isometric Force Emission. *Perception & Psychophysics*, 1975. 17(4), pp 393-397.
52. Fujita, H. and Nishimoto, K. Lovelet: a heartwarming communication tool for intimate people by constantly conveying situation data. *Proceedings of CHI 2004*, pp 1553.
53. Gagge, A., Stolwijk, J. and Hardy, J. Comfort and Thermal Sensations and Associated Physiological Responses at Various Ambient Temperatures. *Environmental Research*, 1967. 1(1), pp 1-20.
54. Gallant, D., Seniuk, A. and Vertegaal, R. Towards more paper-like input: flexible input devices for foldable interaction styles. *Proceedings of UIST 2008*, pp 283-286.
55. Garcia, A. and Faus, L. Statistical analysis of noise levels in urban areas. *Applied Acoustics*, 1991. 34(1991), pp 227-247.
56. Givoni, B., Khedari, J., Wong, N.H., Feriadi, H. and Noguchi, M. Thermal sensation responses in hot, humid climates: effects of humidity. *Building Research & Information*, 2006. 34(5), pp 496-506.
57. Goel, M., Wobbrock, J. and Patel, S.N. GripSense: Using Built-In Sensors to Detect Hand Posture and Pressure on Commodity Mobile Phones. *Proceedings of UIST 2012*, pp 545-554.
58. Gooch, D. An Investigation into Communicating Social Presence With Thermal Devices. *MSc Dissertation*, 2009, University of Bath.
59. Goodwin, A. and Wheat, H. Sensory Signals in Neural Populations Underlying Tactile Perception and Manipulation. *Annual Review of Neuroscience*, 2004. 27, pp 53-77.
60. Gracely, R., Farrell, M. and Grant, M., *Temperature and Pain Perception*, in *Sensation & Perception*, S. Yantis, Editor. 2002, Wiley & Sons: New York. p. 619-651.
61. Gray, L., Stevens, J. and Marks, L. Thermal Stimulus Thresholds - Sources of Variability. *Physiology & Behavior*, 1982. 29(2), pp 355-360.
62. Green, B. The Effect of Skin Temperature on Vibrotactile Sensitivity. *Perception & Psychophysics*, 1977. 21(3), pp 243-248.
63. Green, B. Localization of Thermal Sensation - Illusion and Synthetic Heat. *Perception & Psychophysics*, 1977. 22(4), pp 331-337.
64. Green, B. Synthetic heat at mild temperatures. *Somatosensory Motor Research*, 2002. 19(2), pp 130-138.
65. Green, B. Temperature perception and nociception. *Journal of Neurobiology*, 2004. 61(1), pp 13-29.
66. Green, B. Temperature perception on the hand during static versus dynamic contact with a surface. *Attention, Perception & Psychophysics*, 2009. 71(5), pp 1185-1196.
67. Green, B. and Akirav, C. Individual differences in temperature perception: Evidence

of common processing of sensation intensity of warmth and cold. *Somatosensory Motor Research*, 2007. 24(1-2), pp 71-84.
68. Green, B. and Cruz, A. "Warmth-insensitive fields": evidence of sparse and irregular innervation of human skin by the warmth sense. *Somatosensory Motor Research*, 1998. 15(4), pp 269-275.
69. Green, B., Lederman, S. and Stevens, J.C. The effect of skin temperature on the perception of roughness. *Sensory Processes*, 1979. 3, pp 327-333.
70. Green, B., Roman, C., Schoen, K. and Collins, H. Nociceptive sensations evoked from 'spots' in the skin by mild cooling and heating. *Pain*, 2008. 135, pp 196-208.
71. Hagander, L., Midani, H. and Kuskowski, M. Quantitative sensory testing: effect of site and skin temperature on thermal thresholds. *Clinical Neurophysiology*, 2000. 111(1), pp 17-21.
72. Häger-Ross, C. and Johansson, R. Nondigital afferent input in reactive control of fingertip forces during precision grip. *Experimental Brain Research*, 1996. 110, pp 131-141.
73. Hallin, R., Torebjork, E. and Wiesenfeld, Z. Nociceptors and warm receptors innervated by C fibres in human skin. *Journal of Neurology*, 1981. 44, pp 313-319.
74. Halvey, M., Wilson, G., Brewster, S. and Hughes, S. "Baby It's Cold Outside": The Influence of Ambient Temperature and Humidity on Thermal Feedback. *Proceedings of CHI 2011*, pp 715-724.
75. Halvey, M., Wilson, G., Vazquez-Alvarez, Y., Brewster, S. and Hughes, S. The Effect of Clothing on Thermal Feedback Perception. *Proceedings of ICMI 2011*, pp 217-220.
76. Han, T., Alexander, J., Karnik, A., Irani, P. and Subrimanian, S. Kick: investigating the use of kick gestures for mobile interactions. *Proceedings of MobileHCI 2011*, pp 29-32.
77. Harding, L.M. and Loescher, A.R. Adaptation to warming but not cooling at slow rates of stimulus change in thermal threshold measurements. *Somatosensory & Motor Research*, 2005. 22(1/2), pp 45-48.
78. Harrison, B., Fishkin, K., Gujar, A., Mochon, C. and Want, R. Squeeze me, hold me, tilt me! An exploration of manipulative user interfaces. *Proceedings of CHI 1998*, pp 17-24.
79. Harrison, C. and Hudson, S. Using Shear as a Supplemental Two-Dimensional Input Channel for Rich Touchscreen Interaction. *Proceedings of CHI 2012*, pp 3149-3152.
80. Harrison, C. and Hudson, S.E. Abracadabra: wireless, high-precision, and unpowered finger input for very small mobile devices. *Proceedings of UIST 2009*, pp 121-124.
81. Harrison, J. and Davis, K. Cold-evoked pain varies with skin type and cooling rate: a psychophysical study in humans. *Pain*, 1999. 83(2), pp 123-135.
82. Heikkinen, J., Rantala, J., Olsson, T., Raisamo, R., Lylykangas, J., Raisamo, J., Surakka, V. and Ahmaniemi, T. Enhancing personal communication with spatial haptics: Two scenario-based experiments on gestural interaction. *Journal of Visual Languages and Computing*, 2009. 20, pp 287-304.
83. Heikkinen, J., Rantala, J., Olsson, T., Raisamo, R. and Surakka, V. Exploring the effects of cumulative contextual cues on interpreting vibrotactile messages. *Proceedings of MobileHCI 2011*, pp 1-10.
84. Henderson, N., White, N. and Hartel, P. iButton enrolment and verification times for the pressure sequence smartcard biometric. *Proceedings of E-smart 2001*, pp 124-134.
85. Henderson, N., White, N. and Veldhuis, R. Sensing pressure for authentication. *Proceedings of SPS 2002*, pp 1-4.
86. Hennecke, F., Berwein, F. and Butz, A. Optical Pressure Sensing for Tangible User Interfaces. *Proceedings of ITS 2011*, pp 45-48.
87. Henningsen, H., Ende-Henningsen, B. and Gordon, A. Contribution of tactile afferent information to the control of isometric finger forces. *Experimental Brain Research*, 1995. 105, pp 312-317.

88. Hensel, H., *Thermal Sensations and Thermoreceptors in Man*. 1982, Springfield, IL: Charles C. Thomas.
89. Hensel, H. and Iggo, A. Analysis of Cutaneous Warm and Cold Fibres in Primates. *Pflugers Archiv European Journal of Physiology*, 1971. 329, pp 1-8.
90. Heo, S. and Lee, G. Force Gestures: Augmenting Touch Screen Gestures with Normal and Tangential Forces. *Proceedings of UIST 2011*, pp 621-626.
91. Hinckley, K., Pierce, J., Sinclair, M. and Horvitz, E. Sensing Techniques for Mobile Interaction. *Proceedings of UIST 2000*, pp 91-100.
92. Hirosawa, I., Dodo, H., Hosokawa, M., Watanabe, S., Nishiyama, K. and Fukuichi, Y. Physiological Variations of Warm and Cool Sense with Shift of Environmental-Temperature. *International Journal of Neuroscience*, 1984. 24(3-4), pp 281-288.
93. Ho, H. and Jones, L. Contribution of thermal cues to material discrimination and localization. *Perception & Psychophysics*, 2006. 68(1), pp 118-128.
94. Ho, H. and Jones, L. Development and evaluation of a thermal display for material identification and discrimination. *Transactions on Applied Perception*, 2007. 4(2), pp Artice 13.
95. Hoggan, E. Crossmodal interaction: using audio or tactile displays in mobile devices. *Proceedings of INTERACT 2007*, pp 577-579.
96. Hoggan, E. New parameters for tacton design. *Proceedings of CHI 2007 Extended Abstracts*, pp 2417-2422.
97. Hoggan, E. and Brewster, S. Crossmodal icons for information display. *Proceedings of CHI 2006 Extended Abstracts*, pp 857-862.
98. Hoggan, E. and Brewster, S. Designing audio and tactile crossmodal icons for mobile devices. *Proceedings of ICMI 2007*, pp 162-169.
99. Hoggan, E. and Brewster, S. Mobile Crossmodal Auditory and Tactile Displays. *Proceedings of HAID 2006*, pp 9-12.
100. Hoggan, E., Brewster, S. and Johnston, J. Investigating the Effectiveness of Tactile Feedback for Mobile Touchscreens. *Proceedings of CHI 2008*, pp 1573-1582.
101. Hoggan, E., Crossan, A. and Brewster, S. Audio or tactile feedback: which modality when? *Proceedings of CHI 2009*, pp 2253-2256.
102. Hoggan, E., Raisamo, R. and Brewster, S. Mapping Information to Audio and Tactile Icons. *Proceedings of ICMI 2009*, pp 327-334.
103. Hoggan, E., Stewart, C., Haverinen, L., Jacucci, G. and Lantz, V. Pressages: Augmenting Phone Calls with Non-Verbal Messages. *Proceedings of UIST 2012*, pp 555-562.
104. Hoggan, E., Trendafilov, D., Ahmaniemi, T. and Raisamo, R. Squeeze vs. Tilt: A Comparative Study Using Continuous Tactile Feedback. *Proceedings of CHI 2011*, pp 1309-1314.
105. Holland, S., Morse, D. and Gedenryd, H. AudioGPS: Spatial audio navigation with a minimal attention interface. *Personal and Ubiquitous Computing*, 2002. 6, pp 253-259.
106. Holman, D. and Vertegaal, R. Organic User Interfaces: Designing Computers in any Way, Shape or Form. *Communications of the ACM*, 2008. 51(6), pp 48-55.
107. Holman, D., Vertegaal, R., Altosaar, M., Troje, N. and Johns, D. Paper windows: interaction techniques for digital paper. *Proceedings of CHI 2005*, pp 591-599.
108. Hong, L., Brown, A. and Newell, K. Compensatory properties of visual information in the control of isometric force. *Perception & Psychophysics*, 2008. 70(2), pp 306-313.
109. Househam, E., McAuley, J., Charles, T., Lightfoot, T. and Swash, M. Analysis of force profile during a maximum voluntary isometric contraction task. *Muscle Nerve*, 2004. 29(3), pp 401-408.
110. Höver, R., Luca, M. and Harders, M. User-based evaluation of data-driven haptic rendering. *Transactions on Applied Perception*, 2010. 8(1), pp Article 7.
111. Ino, S., Shimizu, S., Odagawa, T., Sato, M., Takahashi, M., Izumi, T. and Ifukube, T. A Tactile Display for Presenting Quality of Materials by Changing the Temperature of Skin Surface. *Proceedings of IEEE International Workshop on*

Robot Human Communication, 1993. 2, pp 220-223.
112. Iwasaki, K., Miyaki, T. and Rekimoto, J. AffectPhone: A Handset Device to Present User's Emotional State with Warmth/Coolness. *Proceedings of BIOSTEC 2010*, pp.
113. Jagacinski, R., Repperger, D., Moran, M., Ward, S. and Glass, B. Fitts' Law and the Microstructure of Rapid Discrete Movements. *Journal of Experimental Psychology: Human Perception and Performance*, 1980. 6(2), pp 309-320.
114. Jami, L. Golgi tendon organs in mammalian skeletal muscle: Functional properties and central actions. *Physiological Reviews*, 1992. 72, pp 623-666.
115. Johansson, R. Sensory Input and Control of Grip. *Novartis Foundation Symposium*, 1998. 218, pp 45-59.
116. Johansson, R. and Westling, G. Roles of glabrous skin receptors and sensorimotor memory in automatic control of precision grip when lifting rougher or more slippery objects. *Experimental Brain Research*, 1984. 56, pp 550-564.
117. Johnson, K., Darian-Smith, I. and LaMotte, C. Coding of incremental changes in skin temperature by a population of warm fibers in the monkey: correlation with intensity discrimination in man. *Journal of Neurophysiology*, 1979. 42(5), pp 1332-1353.
118. Johnson, K., Darian-Smith, I. and LaMotte, C. Peripheral neural determinants of temperature discrimination in man: a correlative study of responses to cooling skin. *Journal of Neurophysiology*, 1973. 36, pp 347-370.
119. Jones, L. Matching forces: constant errors and differential thresholds. *Perception*, 1989. 18(5), pp 681-687.
120. Jones, L. Thermal Touch. *Scholarpedia*, 2009. 4(5), pp 7955.
121. Jones, L. Visual and Haptic Feedback in the Control of Force. *Experimental Brain Research*, 2000. 130, pp 269-272.
122. Jones, L. and Berris, M. Material Discrimination and Thermal Perception. *Proceedings of HAPTICS 2003*, pp 171-178.
123. Jones, L. and Hunter, I. Effect of Fatigue on Force Sensation. *Experimental Neurology*, 1983. 81(3), pp 640-650.
124. Jones, L. and Hunter, I. Force Sensation in Isometric Contractions - A Relative Force Effect. *Brain Research*, 1982. 244(1), pp 186-189.
125. Jones, L. and Hunter, I. Perceived Force in Fatiguing Isometric Contractions. *Perception & Psychophysics*, 1983. 33(4), pp 369-374.
126. Jones, L. and Lederman, S., *Human Hand Function*. 2006, new York: Oxford University Press.
127. Jones, L. and Piateski, E. Contribution of tactile feedback from the hand to the perception of force. *Experimental Brain Research*, 2006. 168, pp 298-302.
128. Jones, L.A. and Berris, M. The Psychophysics of Temperature Perception and Thermal-Interface Design. *Proceedings of HAPTICS 2002*, pp 137-142.
129. Kabbash, P., Buxton, W. and Sellen, A. Two-handed input in a compound task. *Proceedings of CHI 1994*, pp 417-423.
130. Kane, S. and Wobbrock, J. Getting off the treadmill: evaluating walking user interfaces for mobile devices in public spaces. *Proceedings of MobileHCI 2008*, pp 109-118.
131. Kanosue, K., Sadato, N., Okada, T., Yoda, T., Nakai, S., Yoshida, K., Hosono, T., Nagashima, K., Yagashita, T., Inouse, O., Kobayashi, K. and Yonekua, Y. Brain activation during whole body cooling in humans studied with functional magnetic resonance imaging. *Neuroscience Letters*, 2002. 329, pp 157-160.
132. Kenshalo, D. Somesthetic Sensitivity in Young and Elderly Humans. *Journal of Gerontology*, 1986. 41(1), pp 732-747.
133. Kenshalo, D., Holmes, C.E. and Wood, P.B. Warm and Cool Thresholds as a Function of Temperature Change. *Perception & Psychophysics*, 1968. 3(2A), pp 81-84.
134. Kenshalo, D. and Scott, H. Temporal course of thermal adaptation. *Science*, 1966. 151(1), pp 1095-1096.
135. Kenshalo, D.R., Decker, T. and Hamilton, A. Spatial Summation on the Forehead,

Forearm and Back produced by Radiant and Conducted Heat. *Journal of Comparative and Physiological Psychology*, 1967. 63(3), pp 510-515.
136. Kenshalo, D.R., Nafe, J.P. and Brooks, B. Variations in Thermal Sensitivity. *Science*, 1961. 134, pp 104-105.
137. Kilbreath, S., Refshauge, K. and Gandevia, S. Differential control of the digits of the human hand: evidence from digital anaesthesia and weight matching. *Experimental Brain Research*, 1997. 117, pp 507-511.
138. Kilbreath, S.L. and Gandevia, S.S. Neural and Biomechanical Specializations of Human Thumb Muscles Revealed by Matching Weights and Grasping Objects. *Journal of Physiology*, 1993. 472, pp 537-556.
139. Kildal, J. 3D-Press: Haptic Illusion of Compliance when Pressing on a Rigid Surface. *Proceedings of ICMI-MLMI 2010*, pp Article 21.
140. Kildal, J. Evoked Friction on a Smooth Touch Device. *Proceedings of CHI 2011 Extended Abstracts*, 2011, pp 2233-2238.
141. Kildal, J. Tangible 3D Haptics on Touch Surfaces: Virtual Compliance. *Proceedings of CHI 2011 Extended Abstracts*, 2011, pp 1123-1128.
142. Kushiyama, K., Baba, T. and Doi, K. Thermal design display device to use the thermal tactile illusions: Thermo-Paradox. *Proceedings of SIGGRAPH 2010*, pp Article 99.
143. Kushiyama, K., Baba, T., Doi, K. and Sasada, S. Thermo-Pict Neo. *Proceedings of SIGGRAPH 2010*, pp Article 42.
144. Kushiyama, K., Inose, M., Yokomatsu, R., Fujita, K., Kitazawa, T., Tamura, M. and Sasada, S. Thermoesthesia: About collaboration of an artist and a scientist. *Proceedings of SIGGRAPH 2006*, pp Article 23.
145. Lai, C., Niinimaki, M., Tahiroglu, K., Kildal, J. and Ahmaniemi, T. Perceived Physicality in Audio-Enhanced Force Input. *Proceedings of ICMI 2011*, pp 287-294.
146. Lautenbacher, S. and Strian, F. Sex-Differences in Pain and Thermal Sensitivity - The Role of Body Size. *Perception & Psychophysics*, 1991. 50(2), pp 179-183.
147. Lederman, S. The 'Callus-thenics' of Touching. *Candian Journal of Psychology*, 1976. 30, pp 82-89.
148. Lederman, S. and Klatzky, R. Hand movements: A window into haptic object recognition. *Cognitive psychology*, 1987. 19, pp 342-368.
149. Lee, B., Lee, H., Lim, S., Lee, H., Han, S. and J, P. Evaluation of Human Tengential Force Input Performance. *Proceedings of CHI 2012*, pp 3121-3129.
150. Lee, D., McGillis, S. and Greenspan, J. Somatotopic localization of thermal stimuli .1. A comparison of within- versus across-dermatomal separation of innocuous thermal stimuli. *Somatosensory Motor Research*, 1996. 13(1), pp 67-71.
151. Lee, W. and Lim, Y.K. Thermo-Message: Exploring the Potential of Heat as a Modality of peripheral Expression. *Proceedigns of CHI 2010 Extended Abstracts*, pp 4231-4237.
152. Levesque, V., Oram, L., MacLean, K., Cockburn, A., Marchuk, N., Johnson, D., Colgate, J. and Peshkin, M. Enhancing Physicality in Touch Interaction with Programmable Friction. *Proceedings of CHI 2011*, pp 2481-2490.
153. Li, Y., Hinckley, K., Guan, Z. and Landay, J. Experimental analysis of mode switching techniques in pen-based user interfaces. *Proceedings of CHI 2005*, pp 461-470.
154. Lin, M., Cheng, Y., Yu, W. and Sandnes, F. Investigation into the feasibility of using tactons to provide navigation cues in pedestrian situations. *Proceedings of OZCHI 2008*, pp 299-302.
155. Lin, M., Goldman, R., Price, K., Sears, A. and Jacko, J. How do people tap when walking? An empirical investigation of nomadic data entry. *International Journal of Human-Computer Studies*, 2007. 65, pp 759-769.
156. MacLean, K. Using haptics for mobile information display. *Proceedings of Pervasive*, 2008, pp 175-179.
157. Mai, N., Avarello, M. and Bolsinger, P. Maintenance of low isometric forces during prehensile grasping. *Neuropsychologia*, 1985. 23(6), pp 805-812.

158. Marentakis, G. and Brewster, S. A Comparison of Feedback Cues for Enhancing Pointing Efficiency in Interaction with Spatial Audio Displays. *Proceedings of MobileHCI 2005*, pp 55-62.
159. Marentakis, G. and Brewster, S. Effects of Feedback, Mobility and Index of Difficulty on Deictic Spatial Audio Target Acquisition in the Horizontal Plane. *Proceedings of CHI 2006*, pp 359-386.
160. Marentakis, G. and Brewster, S. A Study on Gesture Interaction with a 3D Audio Display. *Proceedings of MobileHCI 2004*, pp 180-191.
161. Marks, L. and Gonzalez, R. Skin Temperature Modifies Pleasantness of Thermal Stimuli. *Nature*, 1974. 247(5441), pp 473-475.
162. McCallum, D., Mak, E., Irani, P. and Subramanian, S. PressureText: pressure input for mobile phone text entry. *Proceedings of CHI 2009*, pp 4519-4524.
163. McGee, M., Gray, P. and Brewster, S. The Effective Combination of Haptic and Auditory Textural Information. *Haptic Human-Computer Interaction*, 2001, pp 118-126.
164. McGookin, D. and Brewster, S. Advantages and issues with concurrent audio presentation as part of an auditory display. *Proceedings of ICAD 2006*, pp 44-50.
165. McGookin, D. and Brewster, S. An Investigation into the Identification of Concurrently Presented Earcons. *Proceedings of ICAD 2003*, pp 1-4.
166. Meh, D. and Denislic, M. Quantitative assessment of thermal and pain sensitivity. *Journal of the Neurological Sciences*, 1994. 127(1), pp 6.
167. Mithal, A. and Douglas, S. Differences in movement microstructure of the mouse and the finger-controlled isometric joystick. *Proceedings of CHI 1996*, pp 300-307.
168. Miyaki, T. and Rekimoto, J. GraspZoom: Zooming and scrolling control model for single-handed mobile interaction. *Proceedings of MobileHCI 2009*, pp 81-84.
169. Mizobuchi, S., Chignell, M. and Newton, D. Mobile text entry: relationship between walking speed and text input task difficulty. *Proceedings of MobileHCI 2005*, pp 122-128.
170. Mizobuchi, S., Terasaki, S., Keski-Jaskari, T., Nousiainen, J., Ryynanen, M. and Silfverberg, M. Making an impression: force-controlled pen input for handheld devices. *Proceedings of CHI 2005*, pp 1661-1664.
171. Mower, G. Perceived Intensity of Peripheral Thermal Stimuli is Independent of Internal Body Temperature. *Journal of Comparative and Physiological Psychology*, 1976. 90(12), pp 1152-1155.
172. Murakami, T. and Nakajima, N. DO-IT: deformable object as input tool for 3-D geometric operation. *Computer-Aided Design*, 2000. 32, pp 5-16.
173. Mustonen, T., Olkkonen, M. and Hakkinen, J. Examining mobile phone text legibility while walking. *Proceedings of CHI 2004 Extended Abstracts*, pp 1243-1246.
174. Nakashige, M., Kobayashi, M. and Suzuki, Y. "Hiya-Atsu" media: augmenting digital media with temperature. *Proceedings of CHI 2009*, pp 3181-3186.
175. Narumi, T., Akagawa, T., Seong, Y.A. and Hirose, M. Characterizing the Space by Thermal Feedback through a Wearable Device. *Virtual and Mixed Reality*, 2009. LNCS 5622, pp 355-364.
176. Narumi, T., Akagawa, T., Seong, Y.A. and Hirose, M. Thermotaxis. *Proceedings of SIGGRAPH 2009*, pp Article 18.
177. Newell, K. and McDonald, P. Information, coordination modes and control in a prehensile force task. *Human Movement Science*, 1994. 13(3-4), pp 375-391.
178. Oakley, I. and O'Modhrain, S. Tilt to scroll: evaluating a motion based vibrotactile mobile interface. *Proceedings of IEEE Eurohaptics 2005*, pp 40 - 49.
179. Oakley, I. and Park, J. Did you feel something? Distracter tasks and the recognition of vibrotactile cues. *Interacting with Computers*, 2008. 20(2008), pp 354-363.
180. Olausson, H., Wessber, J. and Kakuda, N. Tactile Directional Sensibility: Peripheral Neural Mechanisms in Man. *Brain Research*, 2000. 866, pp 178-187.
181. Oron-Gilad, T., Salzer, Y. and Ronen, A. Thermoelectric Tactile Display Based on the Thermal Grill Illusion. *Proceedings of EuroHaptics 2008*, pp 343-348.

182. Oulasvirta, A., Tamminen, S., Roto, V. and Kuorelahti, J. Interaction in 4-second bursts: the fragmented nature of attentional resources in mobile HCI. *Proceedings of CHI 2005*, pp 919-928.
183. Pang, X., Tan, H. and Durlach, N. Manual discrimination of force using active finger motion. *Perception & Psychophysics*, 1991. 49(6), pp 531-540.
184. Pertovaara, A., Kauppila, T. and Hämäläinen, M. Influence of skin temperature on heat pain threshold in humans. *Experimental Brain Research*, 1996. 107, pp 497-503.
185. Pertovaara, A. and Kojo, I. Influence of the rate of temperature change on thermal thresholds in man. *Experimental Neurology*, 1985. 87(1), pp 439-445.
186. Poupyrev, I., Nashida, T. and Okabe, M. Actuation and Tangible User Interfaces: the Vaucanson Duck, Robots and Shape Displays. *Proceedings of TEI 2007*, pp 205-212.
187. Quinn, P. and Cockburn, A. Zoofing! Faster List Selections with Pressure-Zoom-Flick-Scrolling. *Proceedings of OZCHI 2009*, pp 185-192.
188. Raisamo, R., *Evaluating different touch-based interaction techniques in a public information kiosk*. 1999, University of Tampere: eprints.kfupm.edu.sa. p. 1-11.
189. Raisamo, R. A Multimodal User Interface for Public Information Kiosks. *Proceedings of PUI Workshop*.
190. Ramos, G. and Balakrishnan, R. Pressure marks. *Proceedings of CHI 2007*, pp 1375-1384.
191. Ramos, G. and Balakrishnan, R. Zliding: fluid zooming and sliding for high precision parameter manipulation. *Proceedings of UIST 2005*, pp 143-152.
192. Ramos, G., Boulos, M. and Balakrishnan, R. Pressure widgets. *Proceedings of CHI 2004*, pp 487-494.
193. Rantala, J., Raisamo, R., Lylykangas, J., Ahmaniemi, T., Raisamo, J., Rantala, J., Makela, K., Salminen, K. and Surakka, V. The role of gesture types and spatial feedback in haptic communication. *Transactions on Haptics*, 2011. 4(4), pp 295-306.
194. Rath, M. and Rohs, M. Explorations in sound for tilting-based interfaces. *Proceedings of ICMI 2006*, pp 295-301.
195. Rekimoto, J. Tilting operations for small screen interfaces. *Proceedings of UIST 1996*, pp 167-168.
196. Rekimoto, J. and Schwesig, C. PreSenseII: bi-directional touch and pressure sensing interactions with tactile feedback. *Proceedings of CHI 2006*, pp 1253-1258.
197. Ren, X., Yin, J., Zhao, S. and Li, Y. The Adaptive Hybrid Cursor: A Pressure-based Target Selection Technique for Pen-Based User Interfaces. *Proceedings of INTERACT 2007*, pp 310-323.
198. Reynolds, H., Smith, N. and Hunter, P. Construction of an Anatomically Accurate Geometric Model of the Forearm and Hand Musculo-Skeletal System. *Proceedings of IEEE EMBS 2004*, pp 1829-1832.
199. Rogers, S., Williamson, J., Stewart, C. and Murray-Smith, R. AnglePose: Robust, Precise Capacitive Touch Tracking Via 3D Orientation Estimation. *Proceedings of CHI 2011*, pp 2575-2584.
200. Roudaut, A., Lecolinet, E. and Guiard, Y. MicroRolls: expanding touch-screen input vocabulary by distinguishing rolls vs. slides of the thumb. *Proceedings of CHI 2009*, pp 927-936.
201. Salminen, K., Surakka, V., Raisamo, J., Lylykangas, Pystynen, J., Raisamo, R., Makela, K. and Ahmaniemi, T. Emotional Responses to Thermal Stimuli. *Proceedings of ICMI 2011*, pp 193-196.
202. Sato, K. and Maeno, T. Presentation of Sudden Temperature Change Using Spatially Divided Warm and Cool Stimuli. *Proceedings of EuroHaptics 2012*, pp 457-468.
203. Sawhney, N. and Schmandt, C. Nomadic radio: speech and audio interaction for contextual messaging in nomadic environments. *ACM Transactions on Computer-Human Interaction*, 2000. 7(3), pp 353-383.
204. Schildbach, B. and Rukzio, E. Investigating selection and reading performance on a

mobile phone while walking. *Proceedings of MobileHCI 2010*, pp 93-102.
205. Schwesig, C., Poupyrev, I. and Mori, E. Gummi: a bendable computer. *Proceedings of CHI 2004*, pp 263-270.
206. Schwesig, C., Poupyrev, I. and Mori, E. Gummi: user interface for deformable computers. *Proceedings of CHI 2003*, pp 954-955.
207. Scott, J., Brown, L.M. and Molloy, M. Mobile Device Interaction with Force Sensing. *Proceedings of Pervasive 2009*, pp 133-150.
208. Serina, E., Mockensturm, E., Mote, C. and Rempel, D. A structural model of the forced compression of the fingertip pulp. *Journal of Biomechanics*, 1998. 31, pp 639-646.
209. Shi, K., Irani, P., Gustafson, S. and Subramanian, S. PressureFish: a method to improve control of discrete pressure-based input. *Proceedings of CHI 2008*, pp 1295-1298.
210. Shi, K., Subramanian, S. and Irani, P. PressureMove: Pressure Input with Mouse Movement. *Proceedings of INTERACT 2009*, pp 25-39.
211. Siek, K., Rogers, Y. and Connelly, K. Fat finger worries: How older and younger users physically interact with PDAs. *Proceedings of INTERACT 2005*, pp 267-280.
212. Sines, P. and Das, B. Peltier Haptic Interface (PHI) for improved sensation of touch in virtual environments. *Virtual Reality*, 1999.
213. Slifkin, A. and Newell, K. Noise, information transmission, and force variability. *Journal of Experimental Psychology: Human Perception and Performance*, 1999. 25(3), pp 837-851.
214. Sosnoff, J. and Newell, K. Intermittent visual information and the multiple time scales of visual motor control of continuous isometric force production. *Perception & Psychophysics*, 2005. 67(2), pp 335-344.
215. Srinivasan, M., Beauregard, G. and Brock, D.L. The impact of visual information on the haptic perception of stiffness in virtual environments. *Proceedings of the ASME Dynamics Systems and Control Division*, 1996. 58, pp 555-559.
216. Srinivasan, M. and Chen, J. Human performance in controlling normal forces of contact with rigid objects. *Advances in Robotics*, 1993. 49, pp 119-125.
217. Srinivasan, M. and LaMotte, R. Tactual discrimination of softness. *Journal of Neurophysiology*, 1995. 73(1), pp 88-101.
218. Stephens, J. and Taylor, A. Effect of Visual Feedback on Physiological Muscle Tremor. *Electroencephalography and Clinical Neurophysiology*, 1974. 36(5), pp 457-464.
219. Stevens, J. and Hooper, J. How Skin and Object Temperature Influence Touch Sensation. *Perception & Psychophysics*, 1982. 32(3), pp 282-285.
220. Stevens, J. and Marks, L. Spatial summation and the dynamics of warmth sensation. *Perception & Psychophysics*, 1971. 9(5), pp 391-398.
221. Stevens, J.C. Temperature and the Two-Point Threshold. *Somatosensory & Motor Research*, 1989. 6(3), pp 275-284.
222. Stevens, J.C. Temperature Can Sharpen Tactile Acuity. *Perception & Psychophysics*, 1982. 31(6), pp 577-580.
223. Stevens, J.C., *Thermal Sensibility*, in *The Psychology of Touch*, M.A. Heller and W. Schiff, Editors. 1991, Lawrence Erlbaum: New Jersey.
224. Stevens, J.C. and Choo, K. Temperature sensitivity of the body surface over the life span. *Somatosensory & Motor Research*, 1998. 15(1), pp 13-28.
225. Stevens, J.C. and Marks, L. Spatial Summation of Cold. *Physiology and Behaviour*, 1979, 22, pp 541-547.
226. Stevens, J.C., Marks, L. and Simonson, D. Regional Sensitivity and Spatial Summation in Warmth Sense. *Physiology & Behavior*, 1974. 13(6), pp 825-836.
227. Stewart, C., Hoggan, E., Haverinen, L., Salamin, H. and Jacucci, G. An Exploration of Inadvertent Variations in Mobile Pressure Input. *Proceedings of MobileHCI 2012*, pp 35-38.
228. Stewart, C., Rohs, M., Kratz, S. and Essl, G. Characteristics of Pressure-Based Input for Mobile Devices. *Proceedings of CHI 2010*, pp 801-810.

229. Strigo, I., Carli, F. and Bushnell, M. Effect of ambient temperature on human pain and temperature perception. *Anesthesiology*, 2000. 92(3), pp 699-707.
230. Suhonen, K., Muller, S., Rantala, J., Vaananen-Vainio-Mattila, K., Rasiamo, R. and Lantz, V. Haptically Augmented Remote Speech Communication: A Study of User Practices and Experiences. *Proceedings of NordiCHI 2012*, pp 361-369.
231. Tan, H., Pang, X. and Durlach, N. Manual resolution of length, force, and compliance. *Advances in Robotics*, 1992. 42, pp 13-18.
232. Tang, H., Beebe, D. and Kramer, A. Comparison of tactile and visual feedback for a multi-state input mechanism. *Proceedings of IEEE EMBS*, pp 1697-1700.
233. Tang, H., Beebe, D.J. and Kramer, A.F. A multilevel input system with force-sensitive elements. *International Journal of Human-Computer Studies*, 2001. 54, pp 495-507.
234. Taus, R., Stevens, J.C. and Marks, L. Spatial Localization of Warmth. *Perception & Psychophysics*, 1975. 17(2), pp 194-196.
235. van Erp, J. Guidelines for the use of vibro-tactile displays in human-computer interaction. *Proceedings of Eurohaptics 2002*, pp 18-22.
236. Vazquez-Alvarez, Y. and Brewster, S. Eyes-Free Multitasking: The Effect of Cognitive Load on Mobile Spatial Audio Interfaces. *Proceedings of CHI 2011*, pp 2173-2176.
237. Wahren, L., Torebjork, E. and Jorum, E. Central suppression of cold-induced C fibre pain by myelinated fibre input. *Pain*, 1989. 38, pp 313-319.
238. Westling, G. and Johansson, R. Responses in glabrous skin mechanoreceptors during precision grip in humans. *Experimental Brain Research*, 1987. 66, pp 128-140.
239. Wettach, R., Behrens, C., Danielsson, A. and Ness, T. A thermal information display for mobile applications. *Proceedings of MobileHCI 2007*, pp 182-185.
240. Whitton, J. and Everall, J. The thickness of the epidermis. *British Journal of Dermatology*, 1973. 89, pp 467-476.
241. Wigdor, D. and Balakrishnan, R. TiltText: using tilt for text input to mobile phones. *Proceedings of UIST 2003*, pp 81-90.
242. Wigdor, D., Forlines, C., Baudisch, P., Barnwell, J. and Shen, C. Lucid touch: a see-through mobile device. *Proceedings of UIST 2007*, pp 269-278.
243. Williams, L. and Bargh, J. Experiencing physical warmth promotes interpersonal warmth. *Science*, 2008. 322, pp 606-607.
244. Wilson, G., Brewster, S. and Halvey, M. The Effects of Walking and Control Method on Pressure-Based Interaction. *Proceedings of CHI 2011 Extended Abstracts*, pp 2275-2280.
245. Wilson, G., Brewster, S. and Halvey, M. Towards Utilising One-Handed Multi-Digit Pressure Input. *Proceedings of CHI 2013 Extended Abstracts*.
246. Wilson, G., Brewster, S., Halvey, M. and Hughes, S. Thermal Feedback Identification in a Mobile Environment. *Proceedings of HAID 2013*, Article 2.
247. Wilson, G., Brewster, S., Halvey, M. and Hughes, S. Thermal Icons: Evaluating Structured Thermal Feedback for Mobile Interaction. *Proceedings of MobileHCI 2012*, pp 309-312.
248. Wilson, G., Brewster, S.A., Halvey, M., Crossan, A. and Stewart, C. The Effects of Walking, Feedback and Control Method on Pressure-Based Interaction. *Proceedings of MobileHCI 2011*, pp 147-156.
249. Wilson, G., Halvey, M., Brewster, S. and Hughes, S. Some Like it Hot? Thermal Feedback for Mobile Devices. *Proceedings of CHI 2011*, pp 2555-2564.
250. Wilson, G., Hannah, D., Brewster, S. and Halvey, M. Investigating One-Handed Multi-Digit Pressure Input for Mobile Devices. *Proceedings of CHI 2012 Extended Abstracts*, pp 1727-1732.
251. Wilson, G., Stewart, C. and Brewster, S. Pressure-Based Menu Selection for Mobile Devices. *Proceedings of MobileHCI 2010*, pp 181-190.
252. Wimmer, R. and Boring, S. HandSense: discriminating different ways of grasping and holding a tangible user interface. *Proceedings of TEI 2009*, pp 359-362.
253. Wobbrock, J., Findlater, L., Gergle, D. and Higgins, J. The Aligned Rank Transform

for nonparametric Factorial Analyses Using Only AVOVA Procedures. *Proceedings of CHI 2011*, pp 143-146.
254. Wobbrock, J., Myers, B. and Aung, H. The performance of hand postures in front- and back-of-device interaction for mobile computing. *International Journal of Human-Computer Studies*, 2008. 66(2008), pp 857-875.
255. Yang, G. and Kwon, D. KAT II: Tactile Display Mouse for Providing Tactile and Thermal Feedback. *Advanced Robotics*, 2008. 22, pp 851-865.
256. Yang, G. and Kwon, D. Thermo-Tactile Interaction Using Tactile Display Device. *Proceedings of IFAC 2008*, pp 851-865.
257. Yang, G.H., Jones, L. and Kwon, D.S. Use of Simulated Thermal Cues for Material Discrimination and Identification with a Multi-Fingered Display. *Presence*, 2008. 17(1), pp 29-42.
258. Yang, G.H., Jones, L.A. and Kwon, D.S. Use of Simulated Thermal Cues for Material Discrimination and Identification with a Multi-Fingered Display. *Presence*, 2008. 17(1), pp 29-42.
259. Yarnitsky, D. and Ochoa, J. Studies of heat pain sensation in man: perception thresholds, rate of stimulus rise and reaction time. *Pain*, 1990. 40, pp 85-91.
260. Zhai, S., *Human performance in six degree of freedom input control*, in *Graduate Department of Industrial Engineering*. 1995, University of Toronto: Toronto. p. 166.