DESIGN OF HAPTIC SIGNALS FOR INFORMATION COMMUNICATION IN EVERYDAY ENVIRONMENTS

by

MARIO JAVIER ENRIQUEZ

B.Sc., Instituto Tecnológico de Monterrey, 1997
M.Sc., University of British Columbia, 2002

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE STUDIES

(Computer Science)

THE UNIVERSITY OF BRITISH COLUMBIA
(Vancouver)

OCTOBER 2008

Abstract

Multi-function interfaces have become increasingly pervasive and are frequently used in contexts which pose multiple demands on a single sensory modality. Assuming some degree of modularity in attentional processing and that using a different sensory channel for communication can reduce interference with critical visual tasks, one possibility is to divert some information through the touch sense.

The goal of this Thesis is to advance our knowledge of relevant human capabilities and embed this knowledge into haptic communication design tools and procedures, in the interest of creating haptically supported interfaces that decrease rather than add to their users' sensory and cognitive load. In short, we wanted to create tools and methods that would allow the creation of haptic signals (accomplished via display of either forces or vibrations) extending beyond the one bit of communication offered by current pagers and cellular phone buzzers.

In our quest to create information-rich haptic signals we need to learn how to create signals that are differentiable. We also need to study ways to assign meanings to these signals and make sure that they can be perceived clearly when presented one after another even in environments where their recipient might be involved with other tasks. These needs frame the specific research goals of this thesis.

Most of the results described here were obtained through the study of tactile (in the skin) rather than proprioceptive (force feedback) stimuli. We begin by presenting several methods to create, validate and contrast tactile stimulus dissimilarity data and investigate the design of a waveform intended to be a tactile perceptual intermediate between a square waveform and a triangle waveform. Next, we explore methods to create and test tactile signal-meaning associations and document a surprising ability of participants to exhibit high recall of quickly learned associations at two weeks in a first examination of longitudinal recall of tactile stimuli. We then present methods to measure tactile stimulus masking and identify crucial perceptual thresholds relating to stimulus temporal spacing in an exploration into the masking effects of common-onset vibrotactile stimuli. Finally, we present methods to test haptic and multimodal perception in simulated scenarios including a method to simulate and control cognitive workload; and provide evidence that the commonly-used device of multimodal signal reinforcement can adversely impact performance in an ongoing primary task.

The research presented in this Thesis has implications for the design of signals to be used in displays that are emerging in embedded computing environments such as cars, games, cellular phones, and medical devices.

Table of Contents

List of Tables

List of Figures

Acknowledgements

This thesis is dedicated to my wife, Vanesa. Thank you for all your love, encouragement and support. I could not have done this without you.

I wish to thank my parents, Mario Sr. and Lupita for all their help and love and Manije joon for all her love and good wishes.

I am very grateful for the advice and encouragement from my supervisor, Karon MacLean. I would also like to thank Vince DiLollo, Alan Mackworth and Ron Rensink for their mentoring and support. Thanks are also due to the time and interest of all my labmates and friends. Thank you all.

1 Introduction

Picture yourself attending a conference in an unfamiliar city where you have a talk to give to a large audience. You listen to a recording of your talk on your headphones as you walk down a crowded street walking towards the convention centre. You do not know how to get to your destination, but your mobile telephone / GPS unit is providing continuous guidance cues through a tactile display. As you walk through a major intersection, you recognize a vibration pattern on your mobile telephone; you just received an urgent e-mail from your secretary. Throughout this time, your eyes have been free to attend to what should be your primary concern: keeping yourself safe from traffic and other road hazards.

In the future, we envision devices that will generate haptic signals (whether displaying forces, perceived through proprioception; or vibrations, perceived through the skin) that will not only inform us of isolated events (e.g., a mobile phone buzzing when you get a call), but also provide us with continuous notifications derived from sources such as GPS signals, time-and-safety-critical devices (e.g., operating room monitoring devices notifying a doctor of a patient's current condition) or interpersonal communication applications; all this without creating a disruption in our daily activities.

Understanding the principles and mechanisms of human haptic perception is an essential first step toward the design of useful haptic-based devices.

The first step towards this kind of abstracted haptic communication is to understand how synthesized haptic signals are perceived: which parameters are most salient, how these signals interact with other signals perceptually, how cognitively disruptive they are and what kinds of signals are easily distinguished. In addition to this, we need to investigate how we can associate meanings to these signals. Understanding these factors is a necessary prerequisite to the specification of higher-level signal characteristics which would eventually allow assignment of semantic meaning, and further allow the creation of more elaborate communications schemas.

The main focus of the work presented in this Thesis was to identify and provide the foundational building blocks that will facilitate the eventual design and development of devices that can use the sense of touch to communicate complex abstract information. In short, we wanted to create tools and methods that would allow the creation of haptic signals extending beyond the one bit of communication offered by current pagers and cellular phone buzzers. Because of convenience in setup and experimentation, most of our study examples are centered in tactile rather than proprioceptive display (Section 5.3 is a notable exception), but some of the results may be applicable to both haptic sub-modalities. Herewith, we use the term “haptic” in the cases when we mean both tactile and proprioceptive sensation or display; and either tactile or force feedback to refer more specifically to one type of display.

We approached this by first developing methods to create, validate and contrast haptic stimulus dissimilarity data. We then explored methods to create and test tactile signal-meaning associations and examined longitudinal recall of tactile stimuli meanings. With tactile communication as our goal, we set out to develop methods to measure tactile stimulus masking and identified crucial perceptual thresholds relating to stimulus temporal spacing in an exploration into the masking effects of common-onset tactile stimuli. Finally, we devised methods to test proprioceptive, tactile and multimodal perception in simulated scenarios including a method to simulate and control cognitive workload; and found evidence that the commonly-used device of multimodal signal reinforcement can adversely impact performance in an ongoing primary task.

In this chapter, we introduce a series of challenges that we believe to be important to consider in the design of haptic signals intended to communicate information. This is followed by a brief section on how the work presented in this Thesis relates to prior work in the area of haptic communication research. Our research objectives and contributions for the different sections of this document are summarized and the Thesis structure explained in Section 1.3. Finally, we conclude this chapter by presenting a list of publications resulting from the work reported in this Thesis.

1.1 Key Challenges for Haptic Communication Performance

Based on prior work by the author and others (Enriquez 2002; MacLean 2003), we hypothesize that before we can create useful informative haptic signals, we must investigate their design and usage considering not only the signals themselves, but also the contexts in which they might be used. In this section, we introduce four research challenges which have arisen out of our group's experience over the last decade, and which frame the present work.

1.1.1 Perceptually Guided Haptic Stimulus Design Methodologies

Prerequisites to usable haptic signals include *perceptual distinctiveness* and *signal richness* in the stimulus set. Stimuli intended to be used together in a set must not feel similar to each other or vary along too few perceptual dimensions, otherwise it might be difficult for users to create long-lasting associations to them (imagine a set of graphical computer icons consisting of 20 equally sized blocks, each a different shade of gray). Given that today's haptic displays can rarely be controlled in more than three perceivable dimensions if that (for example, frequency, amplitude and waveform of a vibrotactile stimulus), these dimensions must be exploited with care, ensuring that the stimuli created will maximize the perceptual differences amongst them.

1.1.2 Assigning Meanings to Haptic Signals

At the lowest level, devices and objects need to inform users of events, their identity or their current state or contents. You can know if a box is empty or not by shaking it. You can adjust the volume of your car stereo by reaching for the volume control and identifying it by shape and feel. However, this important basic affordance of objects is challenged by the advent of ubiquitous computing;

multi-purpose objects or controls often share a single physical control (e.g. a single knob on a car radio is used to adjust volume, balance, bass, treble, fader, etc. depending on the device's current state). The use of multi-function physical controls whose current active functionality cannot be identified by touch make it harder to determine what specific state a device is in or what function their controls are accessing at any given time without requiring visual attention.

Active haptic feedback could provide the missing identifying tactile cues for these types of devices. Devices that produce simple synthetic tactile signals, such as pager vibrators, have existed for years. However, we argue that this binary or amplitude-graded signal contains smaller amounts of intelligible information than may be possible with systematic, perceptually guided design and that in the future, it may support more expressive and nuanced communication through the sense of touch.

1.1.3 Tactile Signal Masking

A stimulus is said to be masked when interference from another stimulus prevents the recipient from explicitly detecting, identifying or localizing it. Modern computers are capable of displaying images on their screens at a rate of 60 frames per second. However, if we were to display completely different images containing different pieces of information in each of these frames, we could end up communicating nothing. The rapid presentation of certain types of information can prevent us from being able to perceive anything. This is a basic limitation of human perceptual abilities and an example of temporal stimulus masking. Tactile perceptual masking can arise from the presentation of other stimuli either through touch (uni-modal masking) or other senses (cross-modal masking).

We believe that effective haptic communication will not only depend on having signals that are distinguishable and identifiable, but also presented in a manner that prevents them from perceptually interfering with each other or with other important sensory stimulation that their recipient might be perceiving.

1.1.4 Disruptiveness and Perception under Cognitive Workload

Sensory overload is a common problem with contemporary user interfaces, particularly for those that connect users to computation embedded in portable devices and non-desktop environments (Manzey 1998). These interfaces are increasingly pervasive, often have complex functionality, and are frequently used in contexts which pose multiple demands on a single sensory modality. It is imperative that devices such as the mobile phone in the scenario at the start of this chapter are designed to communicate with us effectively as well as in a manner that does not cause unnecessary distractions from our everyday tasks. For example, in driver/pilot navigation aid systems it is critical that the information exchange remains reliable (signals are not missed) even in high-demand situations, but this information exchange must not interfere with safe driving/flying.

There is evidence that increasing a person's cognitive workload in conjunction with a primarily visual task makes it more difficult for them to notice additional visual signals (Patten 2004); however, it is not currently clear how signals in a *different* sensory modality such as touch will impact overall cognitive load. In order to be able to design effective haptic signals for use in multitasking environments, we must first understand how perception of these signals is affected by the presence of cognitive workload as well as environmental noise.

1.2 Relation to Prior Work in Haptic Communication Research

The bulk of current haptics human-factors research focuses on mapping basic human perceptual limits. These perceptual experiments usually fall into one of two categories: those where participants are asked to react to stimuli in some direct quantitative manner, and those where the requested response is intended to capture more subtly perceived attributes of the stimuli itself. In research focused on determining basic perceptual limits, the former category is more common.

A number of studies have investigated haptic perception and communication (as well as auditory and visual iconic communication) from different perspectives and with different goals (Craig 1985; Gaver 1988; Yazdani 1990). However, little research attention has been directed to the question of how to design information-bearing haptic signals and how to investigate haptic communication in more realistic simulated scenarios and contexts. We believe that testing the effectiveness of these haptic signals in more realistic situations will allow us to have a better understanding of how their effectiveness is affected by cognitive workload and the recipient’s changing priorities. These scenarios will also allow us to study how these signals will in turn affect cognitive workload and other tasks with which the recipient might be involved.

Given that the work presented in this Thesis covers a broad range of topics in different disciplines, a literature review of related background work is presented separately in every chapter.

1.3 Research Objectives, Approach and Contributions

Our current focus is an investigation of the necessary pre-requisites to designing a *set* of haptic signals that can effectively communicate abstract information. Our typical research vehicle is a simulated scenario in which the user is overloaded both perceptually and cognitively; we have employed this scenario at varying levels of abstraction, as appropriate. We are interested in creating and testing information-bearing haptic signals, and more broadly in developing methods and techniques that can be used to develop this type of signals for a variety of devices and contexts.

Throughout this Thesis, we introduce experiments that utilize different forms of haptic signals including vibrotactile signals and kinesthetic force feedback signals. Current vibrotactile and force feedback displays are limited in the variety and quality of the haptic sensations they can reproduce but even these displays can generate a wide array of sensations. Over the course of the studies presented here, we investigated several possible parameters, specifically perception of waveshape, frequency and rhythm patterns. These are characteristics that can be easily manipulated and applied to current haptic displays.

Our research is organized around a suite of objectives (presented as Thesis chapters) that encompass the research challenges identified above. In addition, this section provides an overview of the research contributions presented over the following five chapters. Figure 1-1 provides a diagrammatic overview of the Thesis structure.

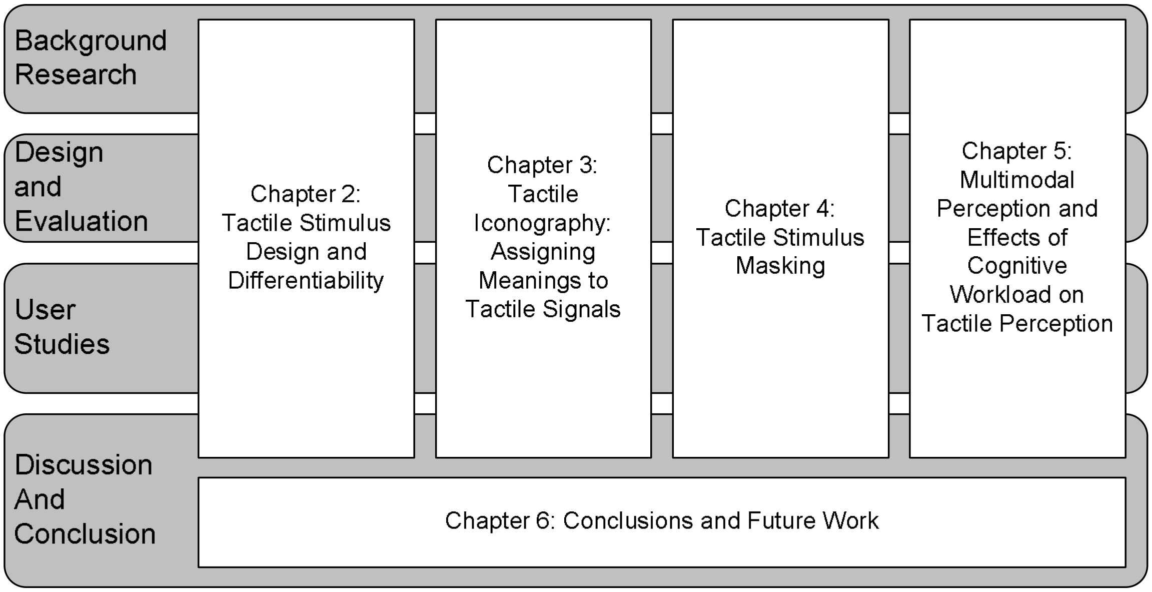


Figure 1-1 Overview of thesis structure and chapter contents

1.3.1 Haptic Stimulus Design and Differentiability (Chapter 2)

Our objective with this research thread is to develop tools and methods to help us attain an increased understanding of how to design haptic signals for communicating abstract information. In particular, we set out to investigate different methods and techniques to design and test haptic signals that are *clearly perceptible* and *perceptually distinct* (differentiable), and thus provide a firm methodological foundation for design work in this area.

A secondary goal of this thread was to verify the effectiveness of a stimuli dissimilarity data-gathering method previously developed by the author and others, through analytical comparison with another more traditional and accepted data gathering method.

Contributions

- Developed methods to compare and contrast perceptual stimulus dissimilarity data.
- Verified that the results obtained from an MDS analysis of dissimilarity data gathered by using a cluster sorting method are comparable to those obtained by using direct comparison.
- Created a waveform that can be systematically varied to render perceptually intermediate haptic waveforms from square to triangle.
- Discovered that the magnitude of the vertical components of the created haptic waveform is the perceptually salient characteristic and determines where this waveform lies in perceptual space.
- Developed software tools to allow rapid prototyping, editing and sharing of haptic sensations.

1.3.2 Haptic Iconography: Assigning Meaning to Haptic Signals (Chapter 3)

We foresee that in the future, everyday devices will communicate with us using meaningful haptic signals. However, these signals must be easily identifiable and convey meanings clearly. The objective of this research thread is to develop methods to investigate the recall of haptic concept-to-stimulus associations. Specifically, we created and adapted existing methods to assign meanings to sets of perceptually distinct tactile signals. Results show that participants are able to quickly learn the meanings associated to a set of tactile signals and (unexpectedly) recall these associations at least two weeks later. This ability has important implications for the design of haptic interfaces and is an important step towards our goal of more advanced communication through the sense of touch.

Contributions

- Developed and tested methods to assign meanings to tactile signals.
- Documented a surprising ability of participants to exhibit high recall of quickly learned associations at two weeks in a first examination of longitudinal recall of tactile stimuli.
- Found no difference in recall performance between arbitrary and user-selected associations between meanings and tactile signals.

1.3.3 Tactile Stimulus Masking (chapter 4)

If we are to build devices that communicate with us effectively using haptic signals, we must ensure that these signals are still perceptible when temporally or spatially contiguous. As demonstrated both in vision and touch, the most effective means of assessing such requirements is through *masking studies*, wherein perception is studied as a function of stimulus spacing.

Our objective with this research thread is to test and contrast two forms of tactile stimulus masking: backward and common-onset. With the purpose of designing information-bearing signals, we build on previous work in the area of tactile masking and focus on measuring stimulus *identification* performance rather than stimulus detection thresholds (Craig 1997). Our experiments test whether participants can identify a masked tactile stimulus, rather than just report its presence or absence. Results show a significant level of identification masking when tactile signals are both spatially and temporally close.

Contributions

- Designed and adapted methods to measure tactile identification masking for common-onset vibrotactile stimuli.
- Identified crucial but previously untested perceptual thresholds relating to common-onset stimulus temporal spacing.
- Discovered a significantly stronger masking effect for common-onset stimuli than for backward masking.

1.3.4 Realistic Scenarios: Haptic Perception in Multimodal and Cognitively Demanding Environments (Chapter 5)

Haptic signals taken beyond a controlled laboratory scenario must not only be perceptible and effective by themselves. These signals must be designed taking into account other activities their intended recipient may be involved with, as well as environmental noise. The objective of this research thread is to develop and test methods and techniques to measure the effects of cognitive workload on the perception of proprioceptive and multimodal (visual + tactile) signals. We created and adapted existing methods to synthesize controlled levels of cognitive workload, and devised a paradigm that allowed the evaluation of haptic and visual signals in a simulated high-cognitive-workload scenario. One key finding suggests that multimodal reinforcement of an incoming signal – commonly thought to improve response performance – actually exhibits a degrading effect on *overall* performance when workload is high.

With this body of work, we hope to demonstrate here that properly designed haptic signals can be a robust mode for communicating information even in the presence of cognitive workload, with implications for the design of interfaces that are less intrusive yet more effective.

Contributions

- Designed and developed experiments intended to ask participants to divert their attention into several and sometimes competing tasks.

- Found negative effects on warning signal performance in the presence of false alarms in a study of haptic warning signal reliability in a simulated driving scenario.
- Implemented methods to create controlled levels of cognitive workload.
- Found a possible negative effect of multimodal reinforced cues when used in high cognitive workload scenarios.

1.4 Publications Relating to this Thesis

The author has to date published a number of papers documenting the research reported in this Thesis.

- Enriquez, M., MacLean, K. (2003). The Hapticon Editor: A Tool in Support of Haptic Communication Research. 11th Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems (HAPTICS'03), Los Angeles.

In this paper we report the design of a tool to aid in the rapid prototyping and sharing of haptic signals. This tool has been used to aid in our ongoing research on haptic icons for low degree of freedom haptic displays. Some of the principles and goals illustrated in this software tool have since been extended in a different direction by others including the work by Swindells, Maksakov, et al. (Swindells 2006).

- Enriquez, M. J., MacLean, K. E. (2004). Impact of Haptic Warning Signal Reliability in a Time-and-Safety-Critical Task. 12th Annual Symposium on

Haptic Interfaces for Virtual Environments and Teleoperator Systems, IEEE-VR2004, Chicago, USA.

In this paper we report an initial investigation into multimodal perception and proprioceptive signals as a means to communicate continuous information in a simulated automotive driving scenario. The experiments described in this paper address the problem of driver behavior when information delivered through this new channel is not completely reliable.

- Enriquez, M. J., MacLean, K. E. (2008). "Backward and Common-Onset Masking of Vibrotactile Stimuli." Brain Research Bulletin, Special Issue on Robotics and Neuroscience 75(6): 761-769.

In this paper we report a series of experiments investigating vibrotactile stimulus masking. It is important to investigate the effects of stimulus masking when designing interfaces that may potentially be used to sequentially display different tactile sensations in close temporal proximity. The clear perception of these stimuli will depend not only on individual stimulus design, but also on the effects of interference from other stimuli preceding or following them.

- Enriquez, M. J., MacLean, K. E., Chita, C. (2006). Haptic Phonemes: Basic Building Blocks of Haptic Communication. Eighth International Conference on Multimodal Interfaces (ICMI’06), Banff, Alberta, Canada.

In this paper we report the first of two experiments designed to test the assignment of meanings to tactile signals to create haptic icons. The experiment described in this paper set out to investigate if people could learn arbitrary associations between basic tactile signal parameters and a set of meanings, and successfully remember these associations.

- Enriquez, M. J., MacLean, K. E. (2008). The Role of Choice in Longitudinal Recall of Meaningful Tactile Signals. Symposium on Haptic Interfaces for Virtual Environments and Teleoperator Systems, Reno, Nevada (Best haptics science paper award).

In this paper we report an investigation into the effect of giving participants a choice of tactile stimuli to associate to a set of concepts. To the best of our knowledge, this paper is the first to document longer term (two week) recall of meanings associated to vibrotactile signals.

2 Stimulus Design and Differentiability

It is our belief that the underutilized haptic sense can be used to absorb some of the demand that modern interfaces place on vision and audition. However, in order to be able to reach the point where we can utilize the haptic sense to aid in our everyday tasks, we must first understand how we perceive the kinds of synthetic haptic signals that today's display technologies are capable of reproducing. Current haptic displays can generate only a limited number of relatively simple different sensations. It is thus imperative that we investigate how we can design these signals in a manner that maximizes the effectiveness of current limited displays.

One of the major hurdles faced when designing synthetic haptic signals is a lack of understanding of how these signals will be perceived and categorized by the intended recipients. Proper design of these signals will be possible only when we understand what specific characteristics of these haptic signals are perceptually important.

The work described in this chapter builds on prior work by the author on haptic iconography (Enriquez 2002; MacLean 2003) where a set of 36 haptic stimuli were created using 3 different waveforms displayed at 4 different frequencies and 3 varying amplitudes. Experiments presented in (Enriquez 2002) addressed the question of how these haptic signals are perceived by using Multidimensional Scaling (MDS) (Cox 1988) as a tool to visualize perceptual axes for complex haptic stimuli. These experiments helped us identify a possible limitation of our MDS data gathering method and posed further questions regarding haptic stimulus design: specifically, we required a way to verify the method used to gather dissimilarity data in the experiment. Although MDS has been used in different disciplines for many years, it is a visualization tool and does not provide statistical comparisons; there are few methods to validate the data gathered to be used in an MDS analysis. In Section 2.4 we present an experiment designed to compare the results obtained by using a direct comparison method to gather MDS dissimilarity data to the cluster sorting method used by the author in prior work (Enriquez 2002).

The same perceptual maps derived from the experiments presented in (Enriquez 2002) suggested that haptic stimulus frequency and (to a lesser degree) amplitude map to perceptual continuums. Our choice of waveforms (square, sawtooth and sine), however, did not. It was these results that raised an interesting question: How can we create a perceptual waveform continuum? In particular, we were interested in determining which specific characteristics of these haptic waveforms are perceptually important. We try to answer these questions in the second part of this chapter (Section 2.5) where we present the design and evaluation of a haptic waveform intended to be perceptually intermediate between a square and sine waveforms. Finally, our desire to investigate perceptually intermediate haptic waveforms led to the design of a tool that allows rapid prototyping and editing of haptic sensations (Enriquez 2003).

This type of haptic waveforms need to be defined in a quick modify-user test cycle, requiring on-the-fly edit facilities Traditionally, the creation of complex haptic sensations requires time and software programming skills that make quick haptic signal prototyping and testing technically challenging. Our approach was to create a tool that grants its users various methods for creating new sensations including direct recording of manual trajectories and creation from a choice of basis waveforms, novel direct-manipulation icon editing mechanisms, integrated playback and convenient storage of icons to file (Section 2.6).

2.1 Objectives

In this chapter we present experimental techniques and tools that can be repeatedly re-applied to the needs of any given haptic device and application context. The validation of these tools and techniques and a quest to understand perceptual organization of created stimuli drove the work presented here.

Our specific objectives in this area were:

- *Develop methods to validate a cluster-sorting method used to gather haptic stimulus perceptual similarity data.* In previous work, we adapted a cluster sorting method to assess perceived similarity between components of a set of haptic stimuli (Enriquez 2002), MacLean 2003). To better understand the numerical impact of this data collection approach on the algorithm’s output, here we used a more commonly accepted data collection method (direct comparison) for the same stimulus set, treating the latter as a "gold standard". The MDS results obtained using these two data collection methods are compared and contrasted. We did obtain similar MDS results from both data gathering methods; the cluster sorting method allows for faster data gathering. This comparison thus provided a degree of empirical validation for the accuracy of the more efficient cluster-sorting method.

- *Investigate perception of haptic waveform characteristics as we design a haptic wave profile intended to be a perceptual intermediate between a square and triangle waveforms.* This work was part of our quest to understand haptic perceptual organization. It was our intent to create a set of haptic signals that would span perceptual space in two dimensions: frequency and waveform. In order to accomplish a spread in waveform, we required a set of perceptually intermediate waveforms from square to triangle. The work carried out in this project allowed us to understand key factors in how humans perceive and process simple synthetic haptic signals.

- *Design a haptic sensation editor that allows rapid prototyping, editing and sharing of synthetic haptic signals.* The goal of this project was to create a tool that would facilitate the creation of haptic sensations and allow non-expert users to build haptic sensations that could be easily edited and shared amongst users of the tool.

2.2 Contributions

- Verified that the results obtained from an MDS analysis of dissimilarity data gathered by using a *cluster sorting method* are comparable to those obtained by using *direct comparison*.
- Created a waveform that can be systematically varied to render perceptually intermediate tactile waveforms from square to triangle.
- Discovered that the magnitude of the vertical components of the created tactile waveform is the perceptually salient characteristic and determines where this waveform lies in perceptual space.
- Developed software tools to allow rapid prototyping, editing and sharing of haptic sensations

2.3 Related Work

The work presented in this section is divided into three subsections. Each of these subsections reports work related with each of the three major objectives of the work presented in this chapter: MDS validation techniques, perceptual waveform transformations and rapid prototyping of haptic sensations.

2.3.1 Understanding how we Perceive Haptic Sensations Using Multidimensional Scaling

Before a researcher can understand why an organism reacts to a stimulus in a specific manner, the researcher must first understand what aspects of the stimulus are attended by such organism. Identifying such aspects directly for complex stimuli can be hard to do. In order to simplify this identification, we use an exploratory statistical method known as Multidimensional Scaling (MDS).

MDS is a set of mathematical techniques that enable a researcher to uncover the “hidden structure” behind data. MDS is similar to principal component analysis (PCA), a mathematical procedure that transforms a number of (possibly) correlated variables into a (smaller) number of uncorrelated variables called principal components. However, PCA cannot take into account nonlinear structures, structures consisting of arbitrarily shaped clusters or curved manifolds, since it describes the data in terms of a linear subspace. PCA is a linear projection data reduction method while MDS is a non-linear projection method. Given that we do not yet know if the haptic perceptual space is linear, we chose to use MDS for our studies.

Multidimensional Scaling allows you to analyze N objects (in our case, haptic signals) according to their measured dissimilarity. A dissimilarity matrix is a set of values representing the perceived distances between each object in a set. MDS takes as input a dissimilarity matrix and generates a multidimensional configuration of the objects in an N dimensional space such that the distances in the Euclidean space approximate the dissimilarities specified by the matrix. MDS produces results in several representations ranging from one to N dimensions (Cox 1988). Dissimilarity data are always mapped with a variable degree of error to the geometrical space, and several measures are used to evaluate the goodness of fit. For instance, one popular method introduced by Kruskal (Kruskal 1964) is commonly used to report goodness-of-fit factors, consists in minimizing a stress function.

MDS is a powerful tool for analyzing complex scenarios. It simplifies the understanding of complex preference data by uncovering hidden structure. Many studies have been carried out in different disciplines utilizing this technique.

Mark Hollins et al. analyzed the perception of real surface textures using MDS (Hollins 1993; Hollins 2000). These textures were presented by moving them across the index finger of the participants who sorted them into categories on the basis of perceived similarity. Their test set consisted of 17 textures such as wood, sandpaper, and velvet. They obtained results mapped into a 3-dimensional space. Two axes were roughly associated with hard/soft and rough/smooth; the third was difficult to interpret. Their work shows an interpretation of the results based on the groupings in the MDS solution space that we adopted for our interpretations.

Of particular relevance is the work by Lawrence Ward who used MDS to study the perception of a set of pictures containing images of different natural and artificial (human generated) environments (Ward 1977; Ward L. 1981). For his work, Ward utilizes an innovative approach to obtain the dissimilarity data for the MDS analysis. Participants in these experiments were asked to rank the images five times using a different number of categories for each sort. The perceived dissimilarity for the picture set is calculated based on these sortings. The results obtained mapped to a space where one axis represented “naturalness” and another “scale”. The sorting methodology used in this work increases the efficiency of evaluating large sample sets, and improves repeatability and accuracy by avoiding the need to judge each item pair individually.

While Ward’s method was originally developed for visual stimuli, our group has adapted this technique to sets of haptic sensations. Brief computer-generated haptic signals were constructed by varying parameters such as frequency, magnitude and waveform and were presented to participants through a force feedback knob (Enriquez 2002; MacLean 2003) or a vibrotactile mouse (Chan 2005). Users were asked to classify the haptic sensations into different clusters; the sorting task was repeated five times varying the number of clusters. Results from an MDS analysis of this type of data have demonstrated a classification of the haptic sensations that follows intuition while providing an extra level of structural detail. In (MacLean 2003), for example, while frequency seemed to be the salient dimension overall, other perceptual dimensions such as waveform and (to a lesser degree) amplitude also emerged from the MDS plots.

In work which post-dates the experiment reported here, Pasquero, Luk, et al. (Pasquero 2006) examined possible deleterious effects inherent to the cluster sorting method without comparing it to direct comparison. Their goal was to determine if this data gathering method provided valid MDS results. They observed that the cluster sorting method’s global nature may be a weakness, because it creates a complex pattern of correlations among the elements of the dissimilarity matrix. However, their work confirms that results obtained with cluster sorting are nonetheless valid.

In this chapter, we compare the MDS results for two different methods to gather haptic signal dissimilarity data for MDS analysis: direct (or paired) comparison and an implementation of the cluster sorting method previously used by Ward et al. (Ward 1977) and further adapted to test haptic stimuli by the author in (Enriquez 2002).

2.3.2 Creating Perceptual Intermediate Precepts

To the best of our knowledge the work presented in Section 2.5 (Perceptual Transformation from Square to Triangle) is the first report of an attempt to create a perceptual morphing between two tactile waveforms. Relevant past work cited here comes from the auditory domain.

The work presented by Slaney et al. (Slaney 1996), describes techniques to automatically morph from one sound to another. In this paper, audio morphing is accomplished by representing the sound in a multi-dimensional space representing pitch and voicing information as a spectrogram. This space can then be warped or modified to produce a desired result. After matching components of the sound, a morph smoothly interpolates the sound amplitudes to describe a new sound in the same perceptual space. Finally, the representation is inverted to produce a sound. An important contribution of this work is the realization that audio morphing can effectively be separated into multiple, independent dimensions.

The work outlined by Bouvrie, et al. (Bouvrie 2006) seeks to develop a framework, which they call “inter-voice morphing”, for morphing between samples of speech that are identical in content, yet produced by different speakers. Given two spoken phrases, they attempt to smoothly morph between the characteristics that define the speakers in order to produce intermediate sequences that lie along the perceptual continuum connecting the two individuals. This project looks at the problem mainly from a signal processing perspective, in contrast to the learning or statistical approach.

2.3.3 Haptic Trajectory Acquisition and Haptic Sensation Editing

There has also been prior work in the area of recording haptic trajectories, a key method of input for the tool described in Section 2.6. The purpose of MacLean’s Haptic Camera (MacLean 1996), later expanded by Swindells and Maclean (Swindells 2007), was to systematically obtain input haptic trajectories for later reproduction. Their systems could obtain an approximate model for a real object’s haptic response, and play it back. However, the Haptic Camera collected input from passive devices rather than from a human hand, and the force model obtained for the device could be edited only parametrically.

Frei (Frei 2000) designed a mechanical device that could record trajectories and play them back. His goal was to create an entertaining and pleasing motion when combined and repeated, but not to facilitate editing of the created trajectories. However, the means of input provided some of the inspiration for our work.

Waveforms representing haptic behaviors can be edited using similar methods to those used for audio waveforms. An example of this is the audio icon work pioneered by Gaver (Gaver 1986), and further documented by Buxton et al. (Buxton 1990). These and other studies provide a starting point for haptic sensation development. People have successfully used audio tools to create haptic effects. For example, Chang and O’Sullivan (Chang 2005) used audio waveform tools to create recorded haptic icons to be played back through vibrotactile actuators in cell phones.

Swindells et al. (Swindells 2006) describe a more recent custom haptic sensation prototyping tool designed primarily for creating fast prototypes for 1 DOF haptic actuators such as knobs, sliders, pressure actuators, or temperature actuators. This tool introduced the concept of ‘haptic tiles’ to arrange and organize collections of haptic icon primitives and provided support for dynamic haptic properties and interactions. This work was carried out some time after the work presented in Section 2.6.

2.4 Perceptual Dissimilarity Data Validation

The purpose for the experiment described in this section was to compare the Multidimensional Scaling (MDS) results for two different experiment data gathering methods: Direct Comparison (Cox 1988) and a Cluster Sorting Method (Enriquez 2002).

Prior work by the author used a Cluster Sorting Method (CSM) to obtain MDS perceptual maps for a set of haptic sensations. In our implementation of the CSM to test synthetic haptic sensations, every participant is presented with graphical representations (graphical buttons or tiles that can be dragged to different locations of the screen using a computer mouse) of these sensations and asked to sort these into 3, 6, 9, 12, and finally 15 groups giving a total of 5 sorts (Enriquez 2002; MacLean 2003). The participants are asked to perform the sorting into groups based on perceived similarity between the icons. If any two icons felt alike, they were to be put in the same group, if different, then in different groups. As a result, if any two icons were placed in separate groups for all 5 different sorts, then the two icons got a dissimilarity value of 1000 (the maximum dissimilarity value possible). If for all five sorts, the two icons were in the same groups, they got a dissimilarity value of 0 (identical).

One of the advantages of using a Cluster Sorting Method is the shorter experiment sessions required to gather the stimulus dissimilarity data. When using Direct Comparison, every single pair combination of the items to be rated has to be presented to the participants for similarity rating.

The experiment described in this section utilizes the same 36 haptic sensations (2 sec. duration haptic sensations) described in (Enriquez 2002) using Direct Comparison to gather perceived dissimilarity data. The experiment was designed to have participants performing paired comparisons on this set of 36 different haptic sensations. The dissimilarity data gathered was then contrasted with prior results using a standard Multidimensional Scaling (MDS) algorithm.

The purpose of this contrast was to validate whether using Direct Comparison or Cluster Sorting as dissimilarity data gathering methods would yield comparable MDS results. The results of this comparison seem to suggest that indeed both direct comparison and cluster sorting produce dissimilarity data that yields similar MDS perceptual maps.

2.4.1 Approach

The experiment described here utilized a paired direct comparison methodology to obtain a rating of similarity for the set of 36 haptic sensations introduced by the author in (Enriquez 2002). The direct comparison method requires every subject to rate every possible pairing of the sensations in a scale from different to similar. All possible pair combinations have to be presented once to obtain a full dissimilarity matrix. These pairings are presented in random order and the participant is asked to give a measure of similarity between them. The experiment software written for this is interactive and self paced. It allows participants to rate the perceived similarity between pairs of haptic sensations by using a scale ranging from “same” to “different”.

The direct comparison stimulus similarity data is then fed to a MDS algorithm and its results compared and to those obtained using a cluster sorting method (Enriquez 2002) for the same stimulus set.

Formal experiments were carried out, utilizing the same 36 element haptic sensation set and the same hardware setup used in the author’s prior research (Enriquez 2002). The pre-existing software was modified to accommodate the change in methodology. 30 participants, each performing 210 comparisons of the 36 sensations in the set were recruited. The resulting dissimilarity values were analyzed using MDS.

2.4.2 Setup

The hardware setup used for the experiment is exactly the same that used for (Enriquez 2002; MacLean 2003). A generic 1.2 GHz Pentium III computer running Windows 2000 in real-time mode was used for all data collected in this experiment. An Immersion Impulse Drive I/O Board v. 1.0 provided I/O and amplification for the haptic interface.

The haptic interface was a direct drive actuated knob. The knob is rubber-covered brass with an outer diameter of 10.5 mm and a length of 16.5 mm. The 1mm thick rubber coating prevented slipping and allowed a better grip of the knob while minimizing compliance. The knob was mounted directly on the shaft of a 20-W Maxon DC motor model 118752. This 24-volt motor has a stall torque of 240 mNm and a mechanical time constant of 5 ms allowing a maximum frequency output of 200 Hz.

A Hewlett Packard model HEDS-5500 optical encoder with 4000 post-quadrature counts per revolution provides positional feedback. The motor/knob assembly was held horizontally with an adjustable vise on a table. A padded arm support was built to hold the participants' forearm comfortably while performing the experiments. The purpose of this armrest was twofold; the participants remained comfortable even for the one-hour long tests and they were forced to grasp the knob in a specific manner. This setup was designed to provide the participants with a single comfortable grasping position of the knob (Figure 2-1).

All participant sessions took place in a sound-proof experiment room in the UBC Computer Science Department's Imager Laboratory.

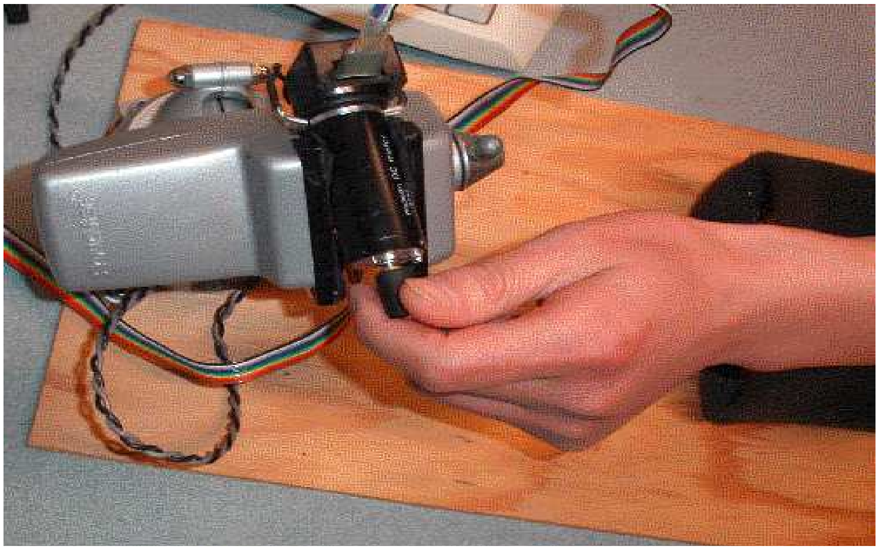


Figure 2-1 Proper grasp of the knob during experiments

2.4.3 Stimulus Set

Each haptic sensation was created by combining different frequencies, amplitudes and waveforms. Frequency values used were 0.5Hz, 5Hz, 20Hz and 100Hz. Torque amplitude values used were 12.3mNm, 19.6mNm, and 29.4mNm. Three different waveforms were used: triangle, square and sawtooth. The combinations of these parameters give a total of 36 different haptic sensations (Table 2-1).

Table 2-1 Haptic sensation parameters

Stimulus Number	Waveform	Amplitude (mNm)	Frequency (Hz)
1	Sine	12.3	0.5
2	Sine	12.3	5
3	Sine	12.3	20
4	Sine	12.3	100
5	Sine	19.6	0.5
6	Sine	19.6	5
7	Sine	19.6	20
8	Sine	19.6	100
9	Sine	29.4	0.5
10	Sine	29.4	5
11	Sine	29.4	20
12	Sine	29.4	100
13	Square	12.3	0.5
14	Square	12.3	5
15	Square	12.3	20
16	Square	12.3	100
17	Square	19.6	0.5
18	Square	19.6	5
19	Square	19.6	20
20	Square	19.6	100
21	Square	29.4	0.5
22	Square	29.4	5
23	Square	29.4	20
24	Square	29.4	100
25	Triangle	12.3	0.5
26	Triangle	12.3	5
27	Triangle	12.3	20
28	Triangle	12.3	100
29	Triangle	19.6	0.5
30	Triangle	19.6	5
31	Triangle	19.6	20
32	Triangle	19.6	100
33	Triangle	29.4	0.5
34	Triangle	29.4	5
35	Triangle	29.4	20
36	Triangle	29.4	100

2.4.4 Allocation of Comparisons Among Participants

One of the advantages of using a Cluster Sorting Method is the shorter experiment sessions required to gather the stimulus dissimilarity data. When using Direct Comparison, every single pair combination of the 36 haptic sensations has to be presented to the participants for similarity rating.

To calculate the required number of total comparisons for an experiment session, we use the following formula:

$$Comparisons = \frac{n(n-1)}{2} \qquad (1)$$

where n is the number of icons in the test set. In our case, having 36 icons would imply having a total of 630 comparisons. Pilot runs of the experiment showed that a participant performing these 630 comparisons would require about 3 hours for a single experiment run. We must note that when using direct comparison to gather dissimilarity data, it is not always necessary to have difference ratings for all pairs of stimuli from all subjects. Spence and Domoney (Spence 1974) investigated how incomplete dissimilarity matrices can be dealt with in perceptual MDS methods, when the data comes from a standard pair-wise comparison task.

With this in mind, the 630 comparisons were broken up into three subsets, where each participant would only perform one of the subsets in a 1 hour session. Therefore, each participant compared 210 pairs of haptic sensations. Ten full trials (repetitions) were carried out, thus requiring 30 participants for the experiment. A single repetition consisted of 630 comparisons and involved three users to complete. For each repetition, the 630 comparisons were randomly ordered. One trial consisted of one participant carrying out 210 comparisons (1/3 of a single repetition involving 630 comparisons).

Since for every repetition, the 630 paired comparisons were presented in random order, every participant got a different subset of comparisons to make. That is, every repetition consisting of 630 comparisons was randomly split into three sets of 210 comparisons, each of which was completed by three separate participants. Therefore, none of the participants performed the exact same 210 paired comparisons. We performed 10 repetitions (30 trials) for this experiment.

2.4.5 Procedure

A user interface written in Microsoft Visual Basic 6.0 led participants through the actual experiment, beginning with the collection of user data. For every repetition, three participants were given a block of 210 comparisons to make.

Prior to beginning each trial, participants signed a consent form as required by the university ethics review board (certificate number B01-0470). The form gave the participants some basic information about the type of experiment they were to perform, as well as an outline of the reimbursement and withdrawal conditions. A copy of this form can be found in Appendix A. After receiving instructions about the task to be performed, participants were required to wear noise-canceling headphones to block any audible artifacts generated by the haptic display.

Every trial began by having the participant feel 12 different haptic sensations, those whose parameters were the most extreme (Figure 2-2) (i.e. icons with highest and lowest frequencies). These 12 sensations were associated randomly with generically marked graphic buttons and displayed as a randomly ordered group to the participant. The participant could then in turn access the haptic rendering for each member of this subset of the complete set to familiarize him or herself with the range of haptic sensations to expect throughout the experiment and to pre-construct a mental model of expected similarity between the sensations.

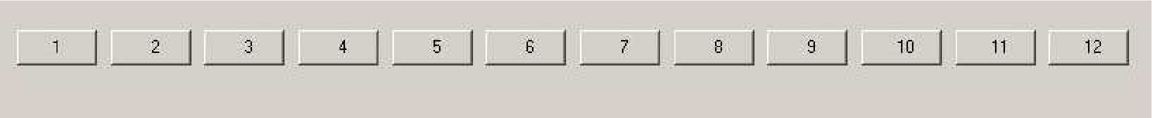


Figure 2-2 12 Graphical icons represent haptic sensations. These sensations were presented to each participant at the beginning of each experiment session.

Participants could display and feel the 12 different sensations as many times as they required. Following this, they were asked to complete a demonstration trial of the comparison test. At this time, the participant successively felt a pair of sensations, each of which being displayed for two seconds and separated by a two second break. The participant was then asked to rate how similar or different the two sensations felt (Figure 2-3) on a scale from “Same” to “Different”.

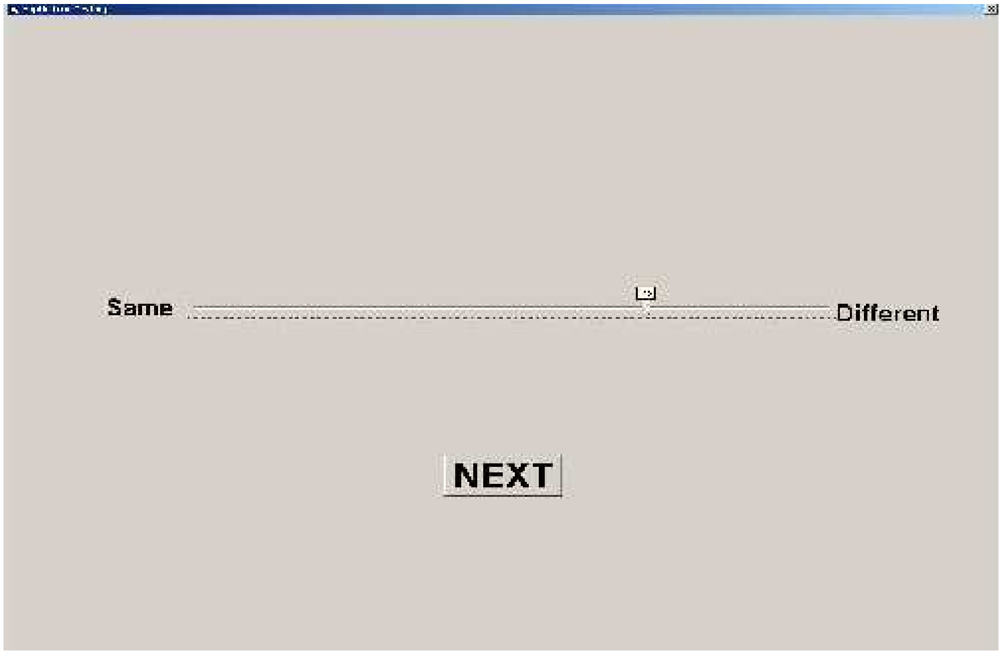


Figure 2-3 Comparison of a pair of haptic sensations

After completing this demonstration, participants completed the actual experiment, consisting of 210 comparisons. At a half way point, after 105 comparisons, the participant was prompted by a message to take a break.

The experiment results were stored as a partial dissimilarity half matrix calculated by the software for the session.

At the end of every session, the participants' questions regarding the purpose and application of the experiment were answered. Participants were asked to fill out a questionnaire and they were also provided a debriefing form with information about the experiment. Participants were paid $10 per hour spent in the experiment.

2.4.6 Participants

Recruitment for this experiment was done by advertising on the computer science undergraduate newsgroup, emailing all computer science graduate students, and putting up sign-up posters in both computer science buildings on campus.

30 participants were recruited for this experiment, 16 male and 14 female with ages ranging from 18-37 years of age, five of which reported themselves as left handed and the rest as right handed. None of the participants reported any disabilities or limitations in either their sight or touch senses.

2.4.7 Results

Each participant compared 210 pairs of haptic sensations, where a complete repetition contains 630 pairs. The results for each full (630 direct comparison) repetition were compiled from the three participants completing a randomly selected subset of 210 comparisons. The experiment run consisted of 10 full repetitions. Each of the possible 630 paired comparisons was performed 10 times. The 10 values for each paired comparison were averaged to produce a dissimilarity matrix which was fed to an MDS algorithm to obtain an MDS plot (Figure 2-4).

The MDS graph obtained from the direct comparison experiment data is presented in Figure 2-4. Figure 2-5 is the MDS graph resulting from the experiment using a cluster sorting method to collect dissimilarity data (obtained from (MacLean 2003)).

It must be noted that these two-dimensional graphs created using MDS are rotationally independent. That is, the results can be rotated pivoting on the center to any orientation. This means that the average MDS plots results for both experiments (one using direct comparison and the other a cluster sorting method) appear to be very similar. In both cases, frequency, which is represented by color in Figure 2-4 and Figure 2-5, is the predominant perceptually differentiable factor. Secondary groupings for waveform are more apparent in the graph resulting from the cluster sorting experiment.

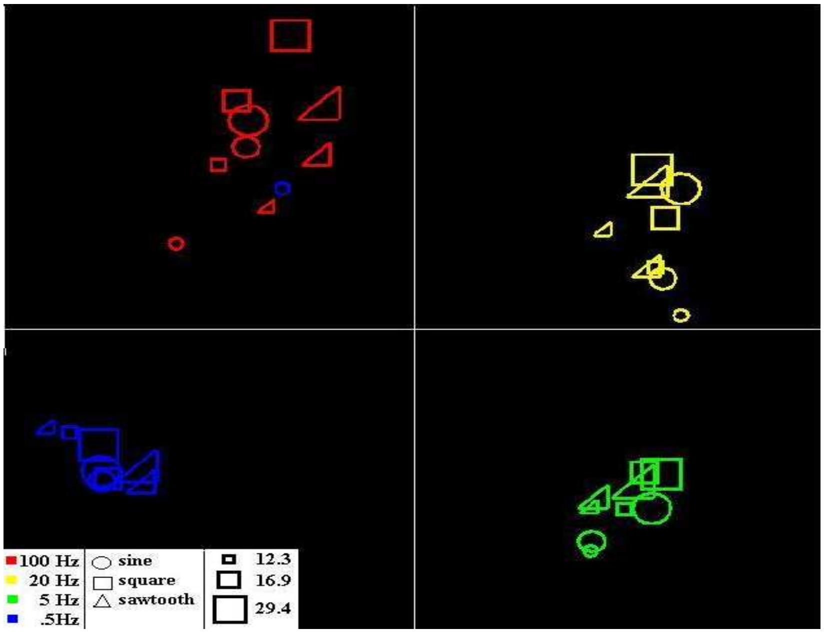


Figure 2-4 MDS experiment results for similarity data collected from 30 participants (10 trials) using direct comparison

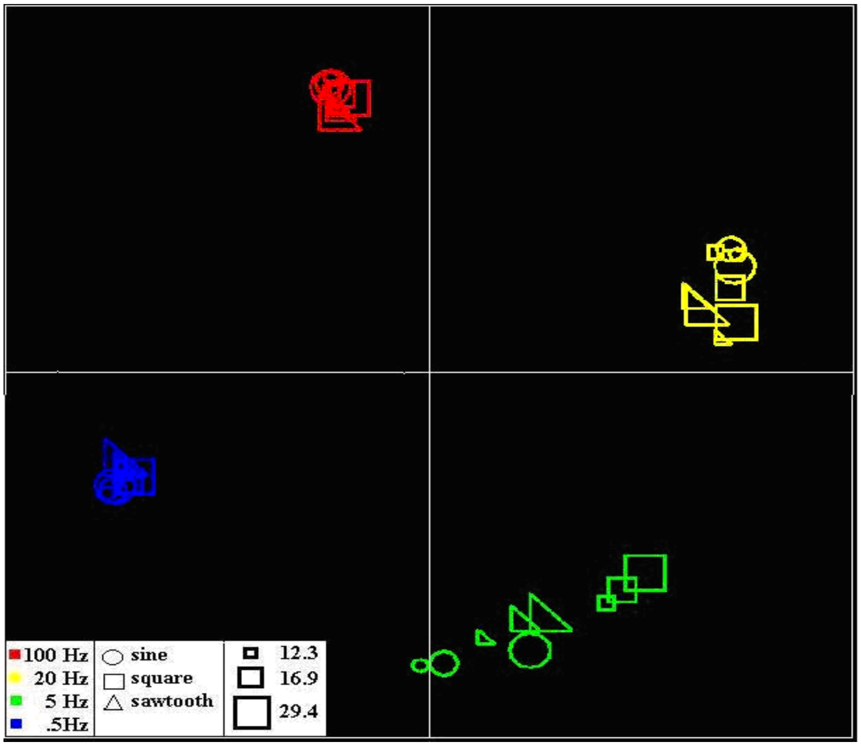


Figure 2-5 MDS experiment results for similarity data collected from 8 participants using a cluster sorting method

2.4.8 Discussion

Our main purpose with this experiment was to determine if a cluster sorting method would produce a MDS plot similar to those obtained by using a direct comparison method to gather dissimilarity data.

It must be noted that, MDS results are interpretative and thus our results and conclusions are interpretative and based on the observed MDS results obtained and on further lightweight qualitative analysis which was not inconsistent with the general conclusion (Enriquez 2008).

Both methods produced MDS graphs that show frequency having a most important role in differentiating sensations and waveform as the second most important characteristic. Both plots (Figure 2-4 and Figure 2-5) show primary groupings based on frequency and secondary groups based on waveform.

Looking at the MDS plots for the direct comparison method, we believe that there are two conclusions which could be made. First, each frequency is grouped together, thus all sensations with same frequency appear together. This implies that participants are very sensitive to frequency differences, and they perceive sensations with different frequency to be very different (as is the case in the results from the cluster sorting method). Second, for the yellow and red shapes, that is, the higher frequencies of 20 and 100Hz, the results are much sparser. This could lead us to believe that at higher frequencies, participants were more uncertain as to how to rate differences between the other parameters such as waveform and amplitude. However, this differs from the MDS plots observed for the cluster sorting method. In this case, participants were able to clearly distinguish and group the different waveforms and to a lesser-degree, the different amplitude values. We believe that the spread seen in the highest frequency in the plot derived from the direct comparison experiment is due to noise in the data gathered.

In summary, we found that the data gathered using the cluster sorting method resulted in MDS plots very similar to those obtained with direct comparison.

2.4.9 Conclusions

This project was designed to compare two different differentiability data gathering methods: Direct Comparison and a Cluster Sorting Method. Our goal was to validate MDS results obtained in the author's prior work (Enriquez 2002; MacLean 2003) and to determine the validity of the cluster sorting method used by the author and other members of our laboratory as a tool to gather dissimilarity data.

There are procedural advantages to the cluster sorting method which motivate us to use it. It provides a faster means to collect dissimilarity data and allows us to perform MDS analyses on stimuli without the complications arising from stimuli set sub-division. A full trial generating a dissimilarity matrix for a 36 element test set can be completed in about 1 hour. By comparison, direct comparison would require 3 one-hour sessions (and perhaps 3 participants).

Both data gathering methods yield similar MDS plots. Given the aforementioned benefits of the cluster sorting method, and based on our observations, we would recommend using a cluster sorting method when gathering dissimilarity data for MDS analysis for this type of stimuli. It must be noted that the cluster sorting method is not perfect. Pasquero et al. (Pasquero 2006) examines possible deleterious effects inherent to the cluster sorting method, without comparing it to direct comparison. However, their work confirms that results obtained with cluster sorting are nonetheless valid.

2.5 Perceptual Transform Wave from Square to Triangle

In video, morphing is a process of generating a range of images that smoothly move from one image to another. In a good morph, the in-between images all show one object smoothly changing its shape and texture until it turns into another object.

In the work just described (Enriquez 2002; MacLean 2003), we investigated the perceptual mapping of 36 haptic stimuli created by combining three discrete wave shapes (sine, square and sawtooth) with several periodic frequencies and amplitudes. We were able to treat frequency and amplitude as continuous variables. We could not do this for waveform. We therefore took a closer look at wave shape alone, and sought to identify the “morph” function and as well as the spacing of the perceptual intermediate states that would produce a continuously and linearly varying percept.

In the investigation described here, we looked at a single transformation between two waveform endpoints chosen for their known distinctiveness: triangle and square (the time rather than frequency domain was chosen for ease of manipulation). More broadly – beyond our current scope – it will be of interest to explore the perceptual dimensionality and salience of spectral variation in general: this study provides an initial foray into this space.

In this section, we describe an investigation in which we produced a waveform that is perceptually intermediate between a square and triangle waveforms. Using an iterative approach involving user testing, we converged on a waveform transform operated by a single parameter that generated linearly spaced, perceptually intermediate steps between a triangle and a square.

2.5.1 Waveform Design

For this experiment, we decided to use simple waveforms varying from a square to a triangle wave (Figure 2-6); prior work had revealed that these haptic signals were the most perceptually different of all that we tried. Conversely, a triangle and a smooth sine wave are nearly indistinguishable (Enriquez 2002).

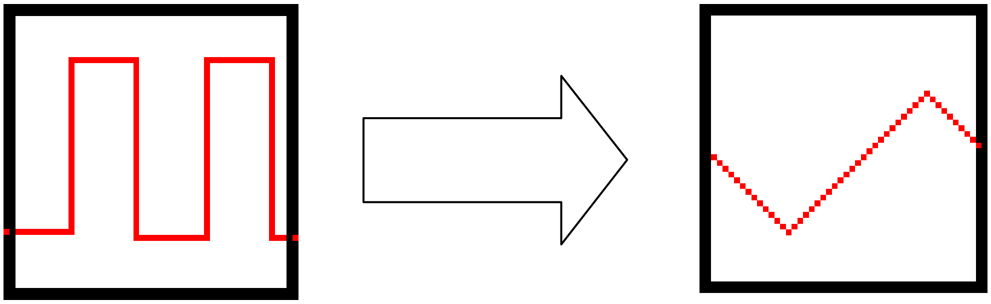


Figure 2-6 Finding a perceptually continuous and linear transform from square to triangle

The design of a perceptual waveform transform was an iterative process that involved continually testing the different designed waveforms in pilot tests with the help of other members of our laboratory. The first approach was to vary a square waveform by smoothly changing the slope angle of its vertical components (Figure 2-7). This transformation, however, produced an abrupt rather than gradual perceptual transition between the square and triangle waveforms. Pilot testers rated these waveforms as all feeling like a triangle waveform as soon as the slope was decreased beyond a few degrees (anything < 85 degree slope) and strictly as a square before this.

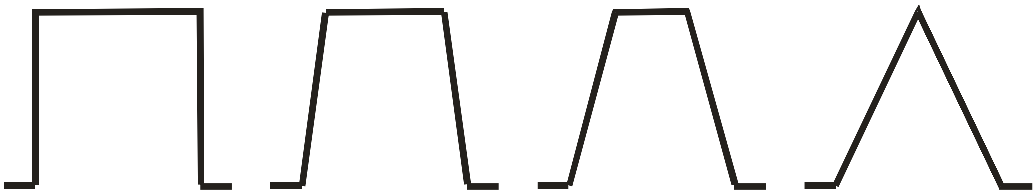


Figure 2-7 Unsuccessful waveform transform: varying slope from square to triangle.

Based on these results we then hypothesized that it was perhaps the vertical component of the waveform that influenced its perception as a square or triangle. We next tried creating a waveform that contained a vertical component. We could then vary the amplitude of this vertical component in relation to the waveform peak amplitude (outlined in red in figure 2-8). The resulting waveform can be recreated by the following weighted average formula:

$$amplitude = (1 - x) * square + x * triangle \qquad (2)$$

where both square and triangle waveforms are in-phase and at the same frequency and x determines the ratio of triangle vs. square. This waveform produced a near linear change in perception (between square and triangle) directly proportional to the value of x.

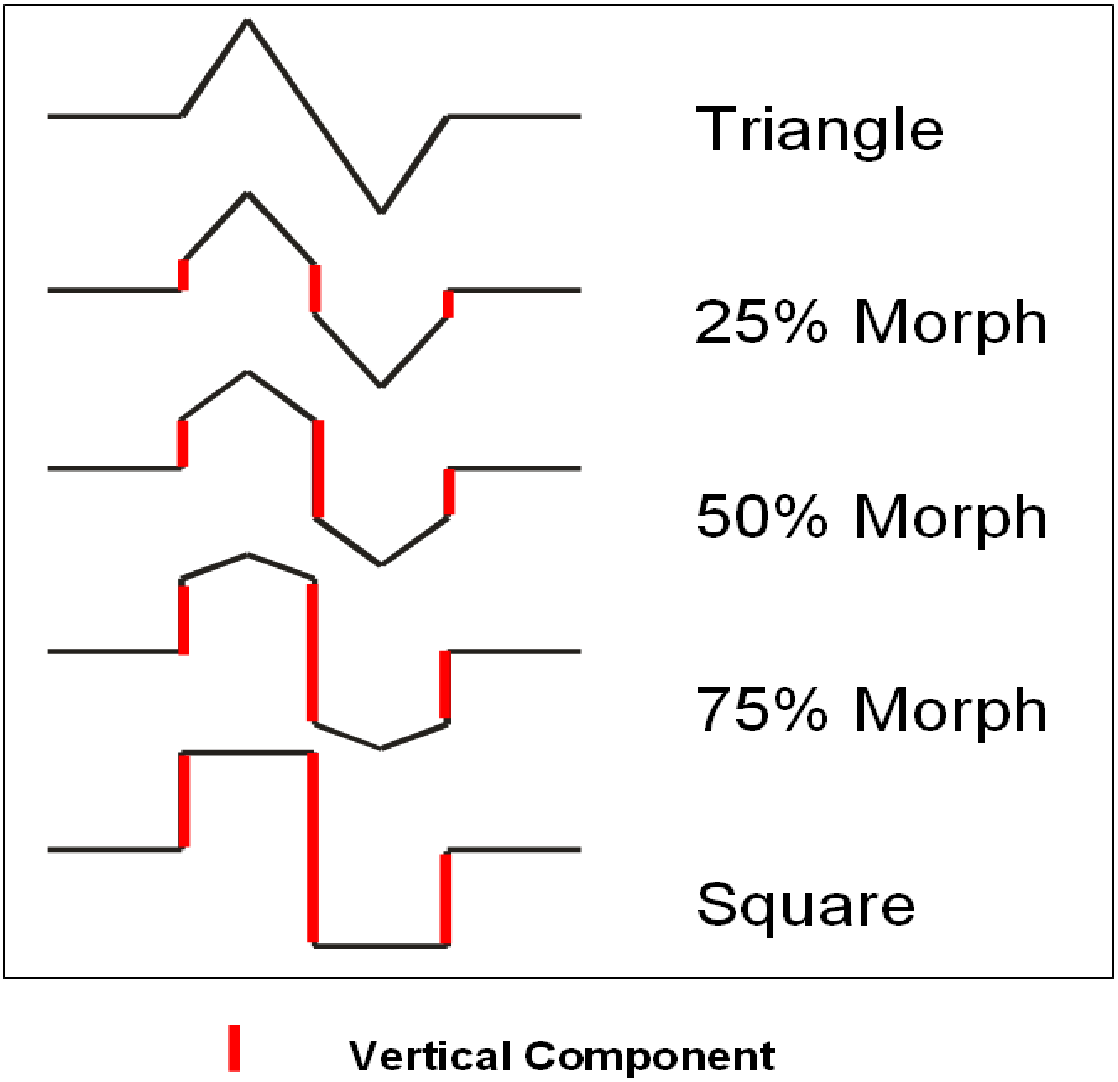
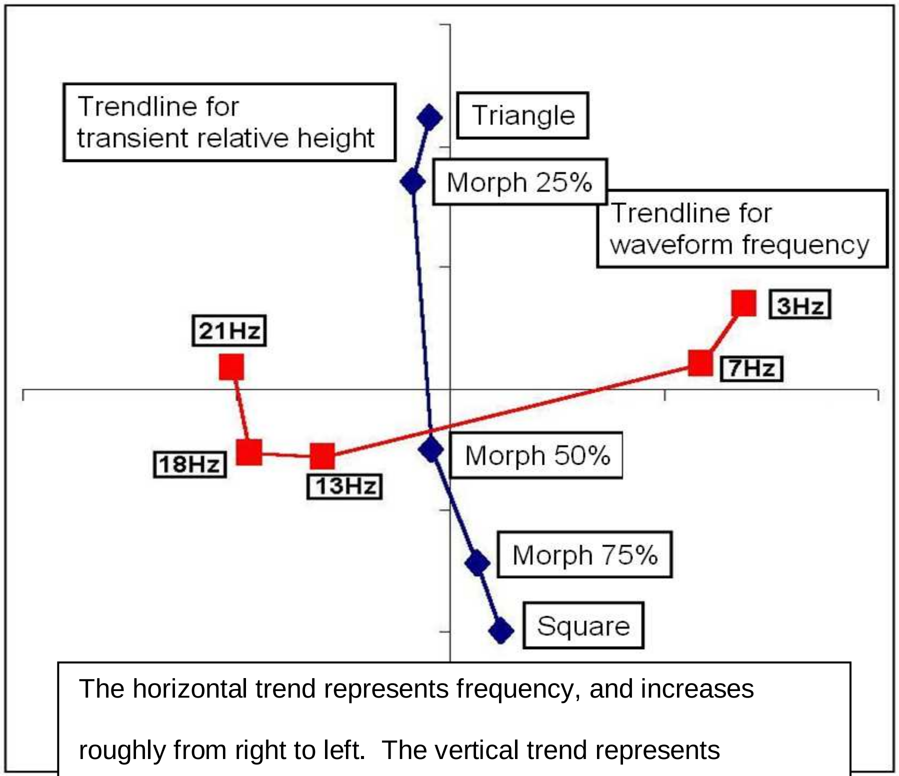


Figure 2-8 Successful waveform transformation: the perceptually important parameter is the amplitude of the vertical component of the waveform in relation to the waveform peak amplitude (outlined in red in the figure).This produces a near linear change in perceived difference between successive instances. Only one cycle of the waveforms is represented.

2.5.2 Validation

The morphing method represented in Figure 2-8 was validated with a multidimensional scaling analysis of perceived dissimilarity values following the procedures described above in Section 2.4 (Enriquez 2002). Figure 2-9 shows a graphical representation of the MDS perceptual map of a set of 25 haptic sensations composed of 5 waveforms (triangle, square and 3 intermediates) displayed at 5 different frequencies (3, 7, 13, 18 and 21 Hz) all presented at the *same fixed amplitude* (Enriquez 2006). 11 participants (6 male, 5 female; age range from 20-35 with median 27) took part in this study. The solution is presented in two dimensions for ease of visualization. Further dimensional MDS solutions resulted in minimal reduction of stress values.



The horizontal trend represents frequency, and increases roughly from right to left. The vertical trend represents waveform relative transient height: 1 = triangle; 2 = 0.25_morph; 3 = 0.5_morph; 4 = 0.75_morph; 5 = square

Figure 2-9 MDS perceptual distribution of stimulus test set including 5 different levels of the waveform transform (transient relative height).

These results show that the morphing derived from a linear change of a single parameter (vertical transient relative amplitude) indeed produced a series of waveforms that map to a near linear spacing between stimuli in the MDS space. Figure 2-9 shows a nearly straight line in the vertical axis formed by the different levels of morphed waveform varying from square to triangle (bottom going up). There is a slight curvature present near the two edges of the perceptual transform, both at the triangle and square extremes. We believe that a truly linear transform must include compensation for this non-linearity.

It is also worth mentioning that the different levels of transformation had an effect on the perception of frequency (peak amplitude for all waveforms used was fixed at a single clearly-perceptible level). This held particularly true at the highest (21Hz) and lowest (3Hz) frequencies used. Prior work showed similar interactions in perception between waveform and frequency (Enriquez 2002).

2.5.3 Conclusions and Next Steps

The perceptual transform we developed to create perceptually intermediate waveforms from square to triangle was quite successful, as confirmed through MDS validation. This analysis shows that those participants tested were indeed rating different levels of these morphed waveforms as being perceptually intermediates between a square and a triangle; furthermore, the degree of variability registered was comparable to the range expressed for a parameter already known to be highly salient (periodic frequency). Having the ability to create haptic stimuli by varying spectral features smoothly in perceptual space will aid in the design of better synthetic haptic signals, and will help interface designers to take better advantage of the current display technologies.

Another important contribution of this work was discovered in our attempts at creating this waveform. When using our haptic display setup, the most salient feature of a haptically rendered waveform is the amplitude of its vertical components in relation to waveform peak amplitude. It was the relative amplitude of these sudden changes that people seemed to focus on when rating these waveforms. This is an important contribution in helping us better understand how it is that we perceive and process these synthetic haptic waveforms.

It is important to note that the final method used to create this perceptual transform was devised through trial and error and through analysis of observations during a series of pilot tests. It may have been possible to achieve this with a more systematic and perhaps more automated method. We are aware that if we were to set out to find more dimensions in perceptual spectral space, we would require a more systematic approach.

2.6 Haptic Icon Editor: Fast Prototyping of Haptic Sensations

The work presented in this section describes the basic functions of a tool developed to aid in the prototyping of haptic sensations. This work has been published in the 11th Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems (HAPTICS'03), Los Angeles. (Enriquez 2003)

2.6.1 Introduction

Visual and auditory icons have long been integral to computer interfaces, as a means of indicating functionality, location and other low-dimensional information more efficiently than can displayed text (Huggins 1974; Yazdani 1990). Graphic icons, for example, are small and concise graphic representations of real or abstract objects. These icons should be easily identifiable by the user and can represent a spectrum of information, ranging from specific functions to abstract controls.

In everyday interaction with manual controls such as those found in a car, on a workbench or throughout a building, we use parameters such as shape, texture and muscle memory to identify and locate different functions and states of handles ranging from doorknobs to pencils and radio controls. With the introduction of “active” haptic interfaces, a single handle - e.g., a knob or a joystick - can control several different and perhaps unrelated functions. These multi-function controllers can no longer be differentiated from one another by position, shape or texture differences, and it becomes a design challenge to make both the existence of available functions and their identity apparent to the user. Active haptic icons may be able to solve this problem by rendering haptically distinct and meaningful sensations for the different functions. Michelitsch, Osen, et. al. took a different approach to solving this by making knobs that can change shape based on context (Michelitsch 2004).

A systematic approach to haptic sensation design requires tools that allow people without engineering background closer participation in the creative process, thus broadening and enriching the area. The Haptic Icon Editor, with its simple, efficient approach, is such a tool.

2.6.2 Hardware Setup

For our ongoing study of haptic icons, we are using a single degree of freedom (DOF) haptic display, configured as a knob (Figure 2-10). The low-DOF interface is appropriate since we anticipate that haptic icons will be most useful in simple, embedded interfaces rather than in high-end desktop systems. The haptic forces are displayed on the knob by a direct-drive DC motor with an optical encoder for feedback. This research setup employs a closed loop controller situated on a PC and communicating with the hardware through an I/O board.

Figure 2-10 Direct drive haptic display: A knob mounted on the shaft of a DC coreless motor equipped with an optical encoder allowed recording and reproduction of haptic sensations.

2.6.3 Haptic Sensation Amplitude

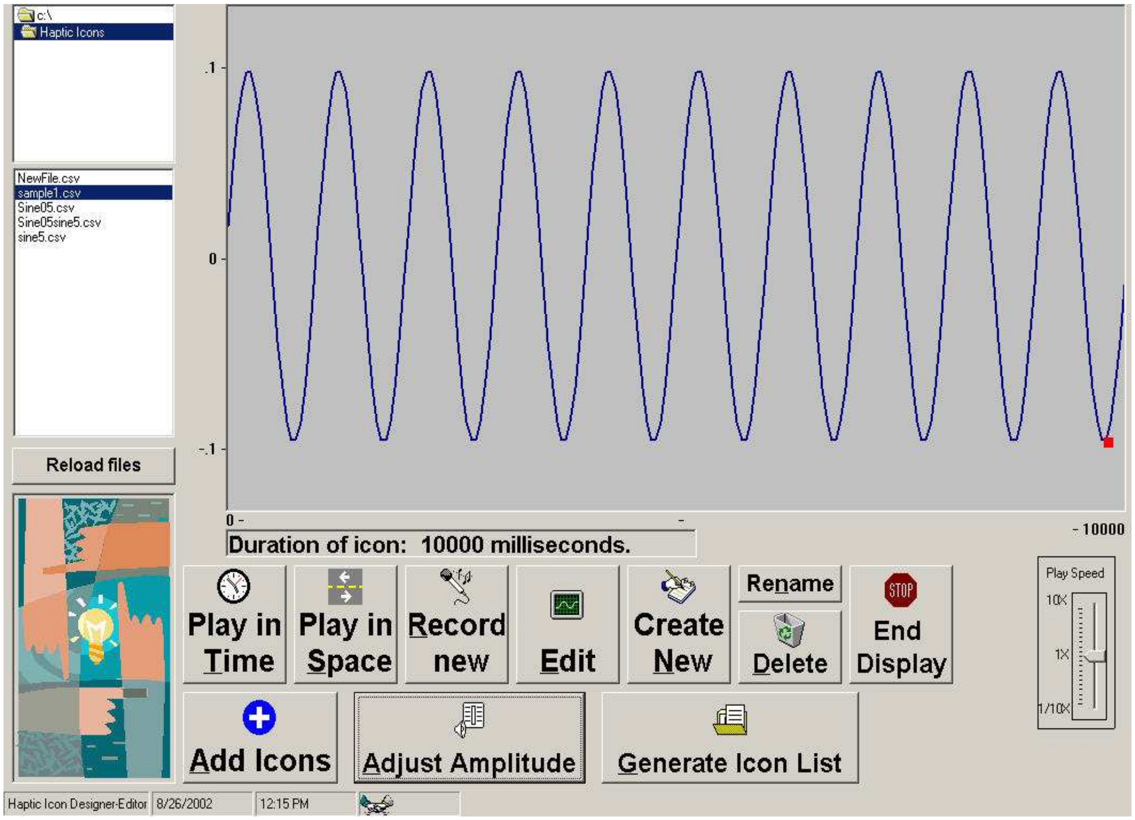
The default unit used to measure and display haptic waveform amplitude is knob revolution, i.e. 1.0 corresponds to one full revolution, or 2pi radians (Figure 2-11, Figure 2-12).

2.6.4 Basics of Operation

The Haptic Icon Editor works by managing and storing a representation of the haptic forces into files. This allows us to treat haptic sensations as any other digital media.

All operations performed with the Haptic Icon Editor affect the haptic waveform file being currently edited. Each waveform is stored in a separate file; and the waveform is activated by selecting it from the file list on the top left part of the main screen (Figure 2-11). When a new file is created by either direct recording or superposition of basic waveforms, it will be added to the file list. The user can perform any of the functions represented in the bottom of the main screen on the active sensation, including Play in Time, Play in Space, Record New, Edit, Create New, and Add Icons.

The area of the screen showing the sine waveform in Figure 2-11 presents a graphical representation of the haptic sensation being edited. The buttons (graphical icons) on the bottom of the screen represent the available functions for creating and editing an “opened” haptic waveform file and will be explained in detail in the following sections.



Duration of icon: 10000 milliseconds.

Figure 2-11 The haptic icon editor main screen.

2.6.5 Haptic Icon Editor Functions

2.6.5.1 Creating a New Icon

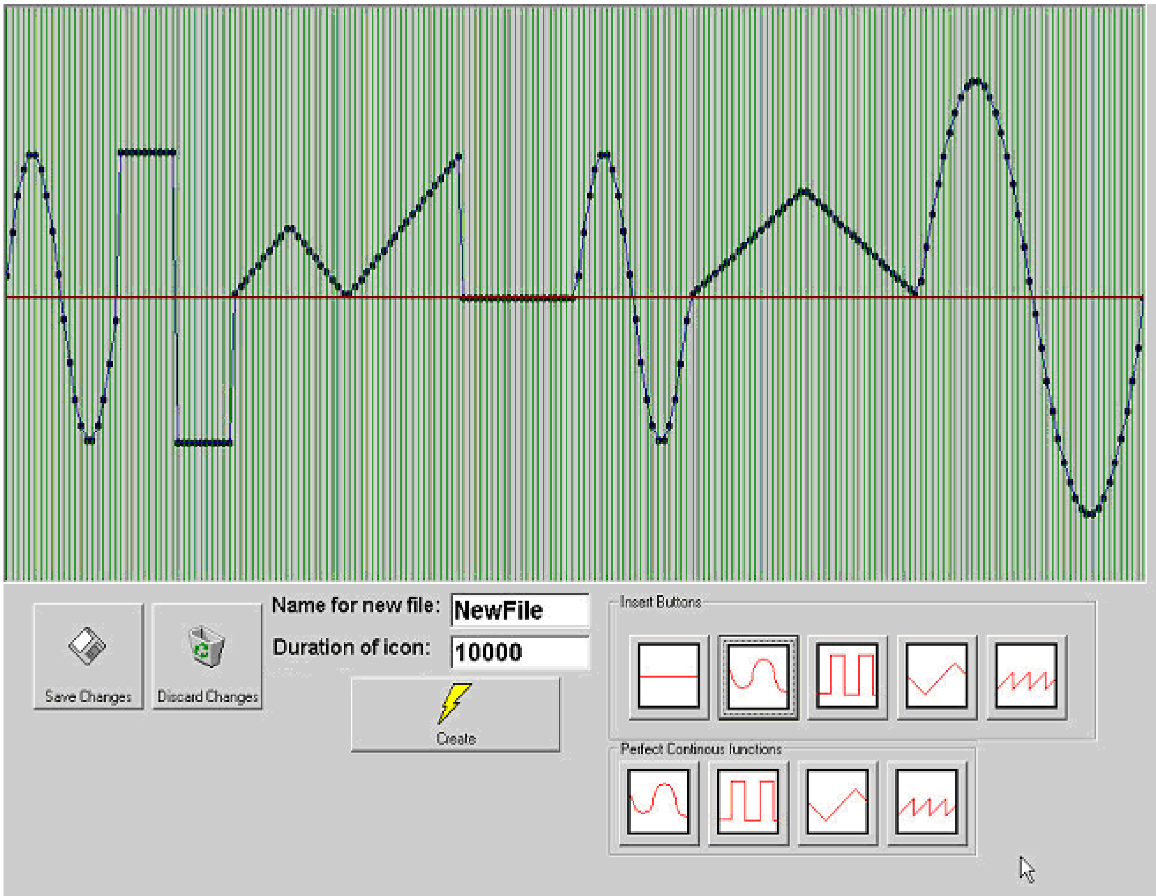
The Haptic Icon Editor allows you to create a new haptic waveform in two ways: direct recording of the user’s motions of the haptic knob and creation by addition of simple waveforms.

- Direct Motion Recording (Record New button)

This function allows the user to directly store the knob motions. The function records the movements for a specified duration and stores this as positional information in a file for later reproduction or editing.

- Creation of Sensations From Simple Waveforms (Create New button)

This function allows the user to create a new haptic waveform from scratch. The process begins by choosing and appending simple waveforms to create a haptic waveform file that can later be displayed through the haptic knob. The waveform data is stored in a new file when complete. When activated, the New Icon Screen is displayed (Figure 2-12).



Name for new file: NewFile

Duration of icon: 10000

Figure 2-12 Haptic icon creator screen This interface allows the creation of haptic sensations from basic waveforms

Figure 2-12 shows the haptic waveform creator screen. The functions in this screen allow the user to generate a new haptic waveform by building it from simple waveforms. The total duration for the haptic waveform can be specified in milliseconds in the space provided. The user can create the haptic waveform using one or more of the given basic functions. Each function is appended one after another. The length, frequency and amplitude for each waveform to be appended can be specified. The graph shown in Figure 2-12 has been created by concatenating seven simple waveforms of varying amplitudes, periods and durations. This file has a total duration of 10 seconds and will be stored with the name <NewFile>.

2.6.5.2 Editing Functions

The Haptic Icon Editor allows you to edit haptic waveforms in several different ways:

-Adjust Amplitude Function

-Graphic Editing of the Haptic Sensation

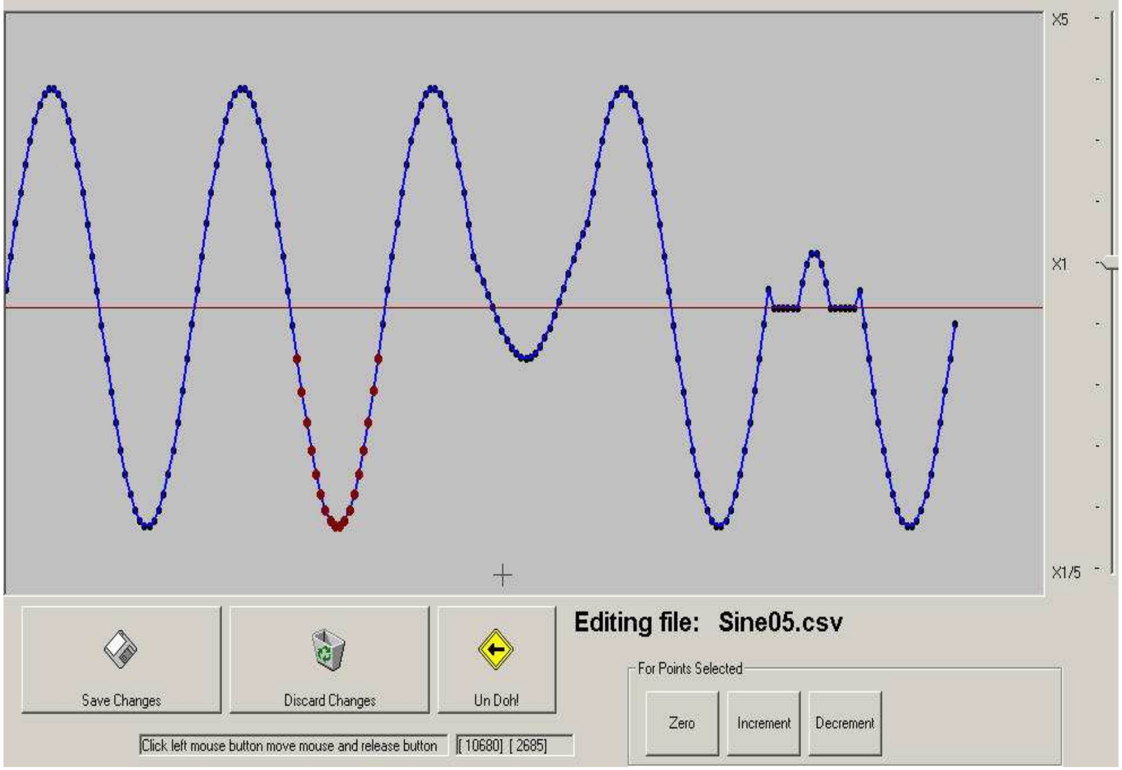
-Haptic waveform superposing

o Adjust Amplitude Function

This utility allows the user to either increase or decrease the overall amplitude of a haptic waveform by specifying a multiplicative scale factor. This is useful when the overall feel of the haptic waveform is either too weak or too strong but you wish to maintain its overall sensation. The default unit to use to measure and display waveform amplitude is a knob revolution.

- Haptic Icon Graphic Editor Function (Edit button)

Once a haptic information file has been created, the user can graphically edit the haptic waveform using simple mouse commands. Figure 2-13 shows the graphic editor screen.



Editing file: Sine05.csv

Click left mouse button move mouse and release button

Figure 2-13 Haptic icon editor screen This interface allows direct editing of a haptic waveform profile using the mouse

This screen shows the haptic function as a series of connected dots. Using the mouse, the user can select one or more of these dots and then by moving them, modify the shape of the haptic waveform. The selected dots can be moved up or down or set to center using the mouse and the editing functions on the lower right of the editor screen. When more than two dots are selected, moving the mouse up or down moves the selected dots in a parabolic shape with the center dot being moved the most.

- o Add Icons Function

This utility allows the user to generate new waveforms by superposing existing haptic waveforms. Figure 2-14 shows a capture of the haptic icon adder screen. Combining several simple waveforms can generate a more complex waveshape.

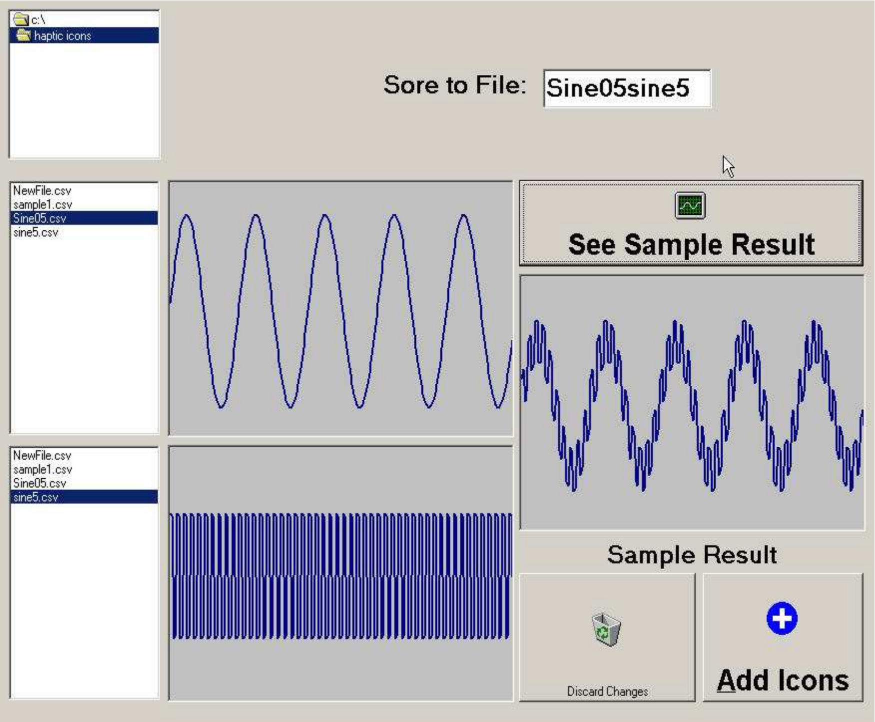


Figure 2-14 Icon adder screen This interface allows the superposition of existing haptic waveforms.

Figure 2-14 shows a haptic waveform being generated by superimposing a low frequency sine waveform with a high frequency sine waveform. The resulting waveform can be stored with a name specified by the user. This added functionality allows the user a better palette for creating more complex functions to be used as haptic sensations.

2.6.5.3 Playback Functions

The creation of haptic icons is a highly iterative process, so it was critical for our tool to have an integrated and very easy to use playback functionality. Once a haptic information file has been created, there are two modes for displaying the file:

-Playback of the haptic waveform as a function of time

-Playback of the haptic waveform as a function of knob position

o Play in Time Function

This utility displays the previously created/edited waveform through the haptic display as a function of time. The icon is displayed by moving the knob to follow the positions indicated by the stored function for a specific time. Play in Time displays the data in the file as forces that vary through time, generating motions on the knob following the graph displayed from left to right.

When a haptic sensation was created through direct motion recording, the knob will mimic those motions previously stored. When the function was created from simple waveforms, the knob will follow the motions specified by the contours of the waveform as a function of time using a proportional-derivative control algorithm [PD] applied to the error between the waveform's position specification and the knob's measured position.

This playback method produces what we call passive haptic sensations. The user merely holds the knob and feels the forces expressed through it. There is no need for any exploratory motion from the user to perceive this type of haptic function.

As the haptic function is being displayed, a small red dot is superposed on the graph showing what part of it is being displayed on the knob at the time.

The Playback Speed slider control, located on the main screen of the Haptic Icon Editor (Figure 2-11), allows adjustment of the playback speed for the haptic sensation. This allows the user the possibility to record a haptic function at a slow pace, and then playing it back at a rate faster than it could be manually input. It can also be used to slow down the reproduction of haptic waveform to obtain a different feeling than the original recording provided.

- Play in Space Function

This utility presents the previously created/edited haptic waveform through the haptic display as a function of knob position. The user can actively explore the haptic function, receiving feedback forces that are dependent on the function being displayed and the position of the knob.

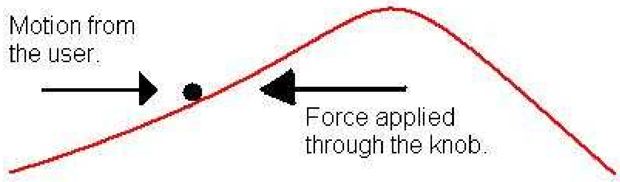


Figure 2-15 The play in space function operation

The haptic waveform being presented can be explored by rotating the haptic knob, producing force feedback proportional to the inclination of the function at the position specified by the haptic display (Figure 2-15). This gives the user the sensation of exploring a one-dimensional topographic map.

The knob reproduces the forces that would be felt when pushing a rolling object over the “terrain” of the function displayed. This allows testing of simple haptic textures that can be easily generated with this program.

The main screen displays a superimposed red dot on the graphic representation of the haptic waveform to allow the user to see what part of the function is being displayed at that specific knob position.

2.6.6 Using the Haptic Icon Editor

The Haptic Icon Editor has been an invaluable aid in designing and testing a collection of abstract haptic waveforms that are currently being used to conduct several psychophysical experiments in our lab.

2.6.7 Conclusions

We have presented a simple tool for creating, editing storing and displaying haptic waveforms. This tool has been used to aid in our ongoing research on haptic icons for low degree of freedom haptic displays and has been modified for use in a project aiming to promote physically embodied math learning for secondary students (Gerofsky 2008). This software also proved to be an invaluable aid in the design and testing of the perceptually intermediate waveforms described in Section 2.5.

Some of the principles and goals illustrated in this software tool have since been extended in a different direction by others including the work by Swindells, Maksakov, et al. (Swindells 2006).

3 Assigning Meanings to Haptic Signals

3.1 Introduction and Summary

A recurrent theme for our research group is to create ways in which the underutilized haptic sense can absorb more of the demand that modern interfaces place on vision and audition. Our group has used *haptic icons* (brief haptic stimuli associated with meanings) to provide users with information from a given device (Enriquez 2002). In Chapter 2, we introduced a series of experiments and tools developed with the intent of creating a set of haptic signals specifically engineered to be perceptually distinct. Given such a set of signals, we then set out to investigate the feasibility of assigning meanings to them with the intent of eventually using them in an interface. The obvious concerns about using touch to communicate information relate to (a) maximum recognizable size of stimulus set; (b) learnability of stimulus-meaning associations; and (c) longevity of those associations. All of these factors are exacerbated by the currently limited expressivity of haptic displays, particularly those suitable for embedded, mobile and wearable applications.

Some examples of information-rich haptic signaling are already appearing in commercial products, primarily in the realm of system-person communication. Cell phone manufacturers like Nokia are experimenting with distinctive vibration patterns (Immersion 2005; Brown 2006) hypothesizing that just as a ring-tone can be associated with a certain caller, a customized vibration can transmit the same meaning with less intrusiveness.

In the auditory domain, there have been some attempts at creating information-rich audio signals. These are in many cases representations of the objects or notions that embody a literal, direct meaning: for example, using the sound of a paper being crushed to indicate deleting a computer file. Most of us are familiar with both the sound of crumpling paper and the action of deleting a file, and can easily make the association.

While an intuitive approach might seem reasonable for the auditory sense, where we can reproduce most common sounds with current display technologies, it does not work as well with the sense of touch. Limitations of current force feedback and tactile display technologies and a lack of understanding of how it is that we perceive and process synthetic haptic signals make using these “intuitive” signals more difficult.

In this chapter, we introduce two experiments designed to test the feasibility of making arbitrary associations between meanings and synthetic tactile signals as well as methods to verify the recall of these associations.

These two experiments were designed to address the following questions:

- Can people learn arbitrary associations between basic tactile signal parameters and a set of meanings and successfully remember these associations?

- Can people learn the meanings associated to a set of arbitrarily created tactile signals?

- Will we observe a recall performance difference if people are given the chance to create their own associations between tactile signals and meanings?

- Will the associations be remembered two weeks after being learned without further reinforcement?

For the first of these experiments (Section 3.4), we tested whether participants could learn a set of meanings associated to different characteristics of a tactile waveform. We employed a set of 9 tactile signals created by combining 3 different waveforms presented at 3 different frequencies and assigned a meaningful concept to the different stimulus parameters (3 concepts for frequency and 3 concepts for waveform). In this manner, participants learned two meanings for each of the signals. After a brief (5 + 20 minute) training period, participants were able to learn and remember arbitrary signal parameter- concept associations with an average performance close to 80% for a period of 45 minutes after learning them.

A second experiment investigated the effects of giving participants a choice in selecting associations between *vibrotactile* signals and a set of meanings (Section 3.5). This experiment was designed to measure recall of the concept-meaning associations in two different cases: in the first, participants were given arbitrary signal-concept associations; in the second case, participants were allowed to choose which signal to use to represent each concept. We measured recall of the associations immediately after being learned and also measured the persistence of these associations two weeks after. We utilized two different sets of 10 vibrotactile signals and investigated the participants' ability to recall concepts associated with them. Two sets of 10 concepts were created to simulate what you would expect to find in an everyday interface such as a car radio or a GPS navigation unit. Recall of arbitrarily assigned signal – concept associations were compared with the associations chosen by the participants. Results show that participants could learn the concepts assigned to a set of 10 signals in less than 20 minutes and recall 86% of those associations learned, at least two weeks after the learning period. Furthermore, we found no difference in recall performance between arbitrarily assigned and participant selected associations.

To the best of our knowledge, the latter experiment is the first looking at longer-term recall of tactile signal meanings. Results from both experiments provided some initial data on the degree to which humans can learn and retain haptic stimulus-meaning associations, in contexts that varied in degree of realism. In general, our results are very encouraging. They suggest that everyday use of haptic icons with current vibrotactile display technology is feasible.

3.2 Contributions

- Developed and tested methods to assign meanings to haptic signals.
- Documented a surprising ability of participants to exhibit high recall of quickly learned associations at two weeks in a first examination of longitudinal recall of tactile stimuli.
- Found no difference in recall performance between arbitrary and user-selected associations between meanings and tactile signals.

3.3 Related Work

Past work which relates to the design of meaningful haptic signals, particularly for use in high-workload environments, include foundations for haptic perception and attentional processing, design and discrimination of “raw” (unassociated) haptic stimulus sets, embedded haptic feedback in high workload environments, and a catalog of specific projects employing haptics for abstract information display using either abstract or semantic approaches. We are not aware of any past longitudinal studies, or comparisons of different icon construction methods.

3.3.1 Haptic Perception and Attentional Processing

A comprehensive understanding of our haptic psychophysical capabilities is emerging through the work of researchers such as Klatzky and Lederman, who have documented our exquisite sensitivity to texture felt through a probe (Klatzky 2003). Tan et al. (Tan 1999) measured information transfer rates of 2-3 bits/second for tactile stimuli independent of duration: appreciable content can be conveyed through this channel. Further, there has been some work in using tactile stimuli to orient attention in another sensory modality; for example, by using taps on the back to direct gaze (Tan 2003).

Currently, the psychophysical research of greatest immediate relevance to haptic signaling relates to thresholds for resolving different excitation parameters (e.g., (Klatzky 2003; MacLean 2003)) and both temporal and spatial masking effects (e.g., (Gescheider 1995; Tan 2003; Enriquez 2007)). The values thus determined are heuristically useful for avoiding conflicts in first-pass stimulus prototyping. However, it is difficult to predict how parameters will be perceived when used together and further, how users will organize multidimensional stimuli within a group. The testing mechanisms described in chapter 2 and outlined here therefore remain essential until our psychophysical and cognitive sophistication greatly improves.

As discussed earlier, some currently recognized attentional theories support the approach of “offloading” information display onto the haptic channel (Wickens 2002). Other attentional research demonstrates linkages between vision and haptics (e.g., showing that tactile stimuli can be used to orient a user’s attention in another sensory modality by using taps on the back to direct gaze (Tan 2003). The latter suggests ways that haptics and vision can be used synergistically in high-load environments, but could also undermine the idea that different modalities will not interfere with one another's processing. Further work in this area is needed to better understand the perceptual processing "pipeline".

3.3.2 Designing and Validating Stimulus Sets

Prerequisite to usable haptic signals are perceived distinctiveness and structural richness in the stimulus set. We must ensure that every stimulus to be used in a set is distinguishable from one another. The only method we are aware of which provides a measure of the *relative* differentiability between two stimuli is based on perceptual Multidimensional Scaling (MDS); this method is used to extract perceptual axes for complex synthetic haptic icons (Enriquez 2002), and also used for synthetic stimuli in (van Erp 2003). The hardest part about this use of MDS is efficiently collecting high-quality difference data from users for relatively large stimulus sets. The author (MacLean 2003) employs an efficient cluster-sorting technique for this purpose, showing that a 36-item stimulus set constructed by varying frequency, magnitude and shape of 2-sec, time-invariant haptic wave shapes map to two perceptual axes. It suggests that expressive capability is maximized in one frequency subspace (7-25 Hz) for that particular force-feedback knob.

By comparison, MDS applied to a wide range of real stimuli suggest up to four perceived dimensions, some highly complex (Bergmann Tiest 2006). Synthetic approaches might enable designers to create new dimensions not present in natural sensations, and exceed this.

Brown et al. (Brown 2006) created 27 abstract tactile signals by varying rhythm, roughness and spatial location, based on prior indications of 3 differentiable levels for each of the 3 parameters individually. The design of this set did not consider parameter interactions, potentially explaining low recall performance.

Tang et al. (Tang 2005) tackle the problem of visual information overload by exploring how haptic feedback can be used as another means for information transmission. Their experiment shows that people can perceive and accurately process haptically rendered ordinal data while under cognitive workload.

Allen et al. present an exploratory study addressed at understanding the feasibility, with respect to ease of learning and usability, of efficient, eyes-free music play list navigation based on symbolic haptic renderings of key song parameters. In this study, users were able to learn haptic mappings for music parameters to usable accuracy with 4 minutes of training. These results indicate promise for the approach and support for continued effort in both improving the rendering scheme and implementing a haptic play list system (Allen 2005).

Most recently, Ternes and MacLean have devised a variant of the MDS methodology to handle larger sets, demonstrating its use on a set of 84 stimuli (Ternes 2007). This set was created through a careful analysis of rhythm perception, used with frequency and amplitude; new perceived sub-dimensions of rhythm were revealed.

3.3.3 Haptic Cues in Mobile Environments

Examples of simple tactile signaling can be found in commercial products. Some cell phones use distinctive vibration patterns (Immersion 2005): a customized vibration can transmit e.g., caller identity with less intrusiveness than a ring tone.

Also in the handheld domain, tactile feedback has similarly added context and cues for application navigation (Poupyrev 2004; Leung 2007), building our knowledge how mobile activities can benefit from this modality.

In 2001, BMW was the first to introduce haptic feedback on the automobile market with the iDrive™ in its 7-series vehicles (now found in other models as well). The iDrive is a force-feedback knob designed to help users access secondary vehicle functions such as audio and climate-control systems (Haller 2003). It varies the knob's feel (via programmed compliance and damping) to create a range of detent sensations (haptic bumps), with different sensations mapped to different control functions. It was initially greeted with skepticism because of its strangeness, poor usability, and the significant learning required. However, some design iterations combined with BMW's tenacity in holding the radical concept on the market long enough for user experience to develop has met with appreciation of its functionality (Day 2004). However, iDrive uses haptics to provide contextual cues (detents, stops, etc.) and not to communicate concepts or meanings. Perhaps the design of more information-rich cues would help make the device be more user-friendly.

3.3.4 Semantic vs. Abstract Icons, and Stress Testing

Semantic icons represent objects or notions through a literal, direct symbol. Gaver et al. (Gaver 1988; Gaver 1989) defined “Auditory Icons” as auditory representations of real objects and actions. The proposed advantage of using a semantic presentation is intuitiveness. Conversely, abstract approaches are similar to the auditory model used by Brewster et al. (Brewster 1992): “Earcons” are sounds and rhythms with no intrinsic or cultural meaning; their target or meaning has to be learned to be effective.

In a first instance of applying a semantic approach to tactile information display, Chan et al. (Chan 2005) developed 7 haptic icons to facilitate application sharing among distributed members of a group, by indicating request urgency in a custom turn-taking protocol. In both abstracted and situated environments, they found that the haptic icons (designed to be intuitive) could be learned to a high degree of accuracy in under three minutes and remained identifiable even under significant cognitive workload. The associations used were carefully explained to the users prior to the test.

Tang et al. (Tang 2005) prototyped a representational ordinal data display and tested it under visual overload. Their experiment showed that people can perceive and accurately process haptically rendered ordinal data while under cognitive workload, with accuracy ranging from 75-93% depending on representational model.

In (Brown 2006), Brown et al. used a magnitude representation for 9 signals composed from 3 levels each of “roughness” and pager motor intensity, to indicate respectively 3 cell phone message types and 3 priority levels. The idea was that the different intensity levels would intuitively represent different urgency levels. They found a recognition rate of 52% for roughness and 70% for intensity level.

Several examples lie mid-way along the abstraction spectrum. Van Erp et al. propose that familiarity with tactile rhythms drawn from popular music will aid in recalling concepts arbitrarily assigned to these tactile patterns; i.e. the abstract association would benefit from more memorable stimuli (van Erp 2003).

In Brown et al.’s 27-item rhythm-based set, participants were asked to identify concepts associated to the 3 different levels of the 3 different parameters used for each presentation of the stimuli (Brown 2006). Results showed an overall identification rate of 47.8% (chance performance would be 30% correct).

The idea of creating sets of intuitive representational haptic icons which do not require their users to learn their meaning is very attractive. Intuitive visual icons can be found everywhere, from computer desktops to signs in nature parks. However, creating these intuitive representational haptic icons is not an easy task. We are limited by both current display technologies as well as a poor understanding of how haptic signals are perceived and processed. It is for this reason that we decided to investigate whether users could build enduring mnemonic representations for abstract synthetic tactile signals (Section 3.5).

3.4 Haptic Signals with Meanings Study

The work described in this section has been published in the Eighth International Conference on Multimodal Interfaces (ICMI’06), Banff, Alberta. (Enriquez 2006)

3.4.1 Introduction

Sensory overload is a common problem with contemporary user interfaces, particularly for those that connect users to computation embedded in portable devices and non-desktop environments. These are increasingly pervasive, often have complex functionality, and are frequently used in contexts which pose multiple demands on a single sensory modality: e.g., accessing a cell phone while driving or sitting in a theatre, or using a remote control that requires visual inspection while watching TV in a darkened room.

As the intelligent aids we currently use in our daily tasks become more complex, they often entail a proportional increase in sensory overload. For example, the display of most cellular telephones conveys information far beyond basic “caller ID”, which means that its operation demands the visual sense for longer periods of time. Interface designers have an increasingly difficult task as they try to simultaneously address the market-driven need for more embedded functionality, while the number and form of the controls are limited by the physical dimensions of the device.

One possible approach to this complex set of design constraints is to divert some of the information flow through the touch sense. A viable implementation would allow for an increased information flow to be conveyed without overloading the visual or auditory senses, and, at the same time, not require additional interface controls: the haptic force feedback can be embedded into the existing set.

But haptic feedback is also well suited for a radically different contribution, by rendering abstract models or concepts as a new modality for communication. At the lowest level, devices and objects notify users of an event, their identity or their current state or contents. Simplistic versions, such as pager vibrators, have existed for years. However, we argue that this binary or amplitude-graded signal contains far less intelligible information than may be possible with systematic, perceptually guided design; and that in the future, it may support expressive and nuanced communication that qualifies as a new haptic language.

Following the work in (Enriquez 2002), we set out to build haptic icons to communicate simple concepts. In order to construct haptic icons, which can also be viewed as haptic words, we need to first understand the communication utility of their smallest distinguishable building blocks. We term these stimuli “haptic phonemes;” they can be constructed of simple waveforms with a fixed frequency and amplitude presented through a haptic display. These phonemes can be assigned meanings which, when combined to create haptic words, can represent increasingly elaborate families of concepts that are related both semantically and haptically.

3.4.1.1 Overview

The main goal of the experiment described here is to explore the degree to which a set of haptic stimuli can convey information alone, without requiring a reinforcing visual image. It is desirable to maximize the perceivable information density of a distinct haptic signal, ideally to rival the level of information content of a computer graphical icon. Our approach is to train a typical non-expert user to associate an arbitrary meaning with each of a set of haptic phonemes, and then test whether these associations can be consistently remembered.

The results obtained here suggest that participants are able to learn and later remember meaning–phoneme associations after a brief learning period. Furthermore, the associations are persistent for at least 45 minutes after the learning period.

3.4.2 Building Haptic Words

In our definition, haptic phonemes represent the smallest recombinant module of a physical haptic stimulus. While a haptic phoneme cannot be broken into smaller recombinant units, it does have multiple minimal dimensions which can be uniquely parameterized. For example, a single phoneme must have both a waveform (a specification of the temporal path its signal takes) and a frequency (the rate at which that path is traversed). We can leverage this by associating sub-meanings with the different dimensions – e.g., frequency might indicate urgency of a cell phone call, whereas waveform could specify identity. In this manner, phonemes form information units that, when put together, create more complex information units (words or phrases). We thus intend to assign meanings to haptic signal parameters (phonemes) which, when interpreted as a whole (words), result in more complex meanings.

Haptic words (or icons) can be built from haptic phonemes using two approaches (Figure 3-2):

1. Concatenation: Phonemes are combined serially to create a word (number and complexity of required haptic syllables must be determined via user experiments), loosely following an analogy with English word construction.

2. Superposition: Phonemes are combined in parallel to create a word of the same length as the longest original phoneme, following a musical chord analogy.

In the work described in this section, we consider phonemes in isolation.

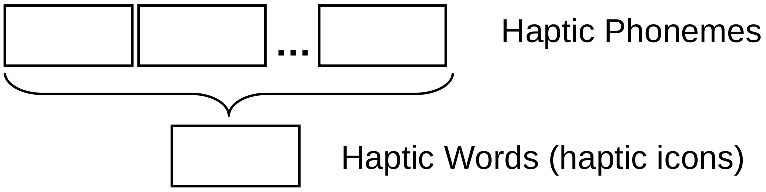


Figure 3-1 Building haptic words (haptic icons) through concatenation

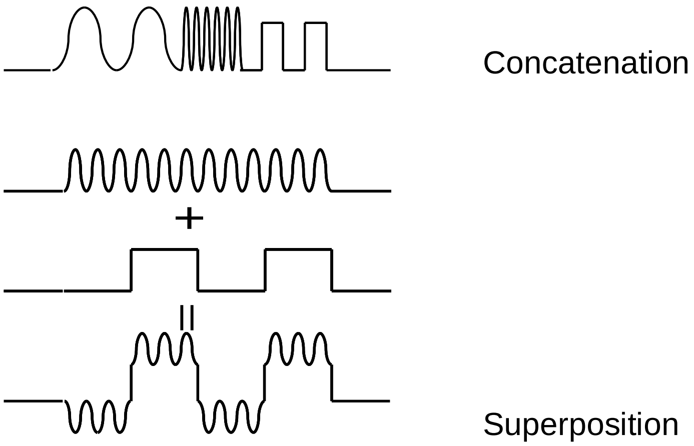


Figure 3-2 Using concatenation and superposition to create haptic words from haptic phonemes

3.4.3 Approach

We have previously tested participants ability to learn and remember semantically driven, associations intended to be intuitive and found encouraging results (Chan 2005). What kind of performance is possible with the harder case of deliberately arbitrary stimulus-meaning associations?

To test this proposition, we needed a mechanism that would facilitate learning the associations as well as a method to measure performance indicating how well the user is able to perceive and process the signals and relate them to their respective meaning. We chose to use a three-stage experimental approach which had participants performing two learning stages and a test stage. In the first stage, the self-guided learning stage, participants were first presented with a graphic interface that showcased a set of haptic sensations and allowed them to learn associations between the sensations and a set of previously determined arbitrary meanings with no semantic relation to the haptic stimuli. This was followed by a computer-guided enforced learning stage, where participants were asked to recognize and categorize a series of randomly presented haptic sensations into boxes labeled with their respective meanings, while receiving reinforcement feedback about errors. Finally in the test stage, participants repeated the enforced learning task but without reinforcement.

3.4.4 Phoneme Creation

In order to ensure that phoneme stimuli met our design specifications for discriminability, we followed a series of steps described here. The tools and methods used to develop these stimuli are described in more detail in Sections 2.4 and 2.5.

Prior work by the author (MacLean 2003) shows that a person can clearly distinguish a set of 10 or more haptic signals when delivered through a haptic knob. Here we used a similar haptic knob to deliver a set of 9 haptic stimuli by combining 3 levels of waveform and frequency. For simplicity, we used periodic waveforms with fixed frequencies; all signals were two seconds long.

We required a set of haptic stimuli which was (a) perceptually well-distributed in frequency and waveform and (b) perceptually equalized for amplitude (all signals “feel” the same amplitude). Good perceptual spread means that users perceive key stimulus attributes as varying linearly (MacLean 2003). This entailed determining a perceptual transformation that simultaneously modulated both dimensions (Chapter 2); for the waveform, we also had to ascertain a transformation path.

Finally, phonemes were created by assigning arbitrary meanings to these stimuli.

3.4.4.1 Perceptual Transform of Waveform

For this experiment, we decided to use simple waveforms varying from a triangle to a square wave; prior work had revealed that these haptic signals are perceptually distinctive; whereas a triangle and a smooth sine wave, for example, are indistinguishable (Enriquez 2002). As described in Section 2.4, we explored several waveform “morph” functions in search of one which varied linearly between these endpoints in a user’s perception, rather than feeling similar along most of its range. The most obvious path (a linear interpolation) failed to elicit a range of intermediate sensations.

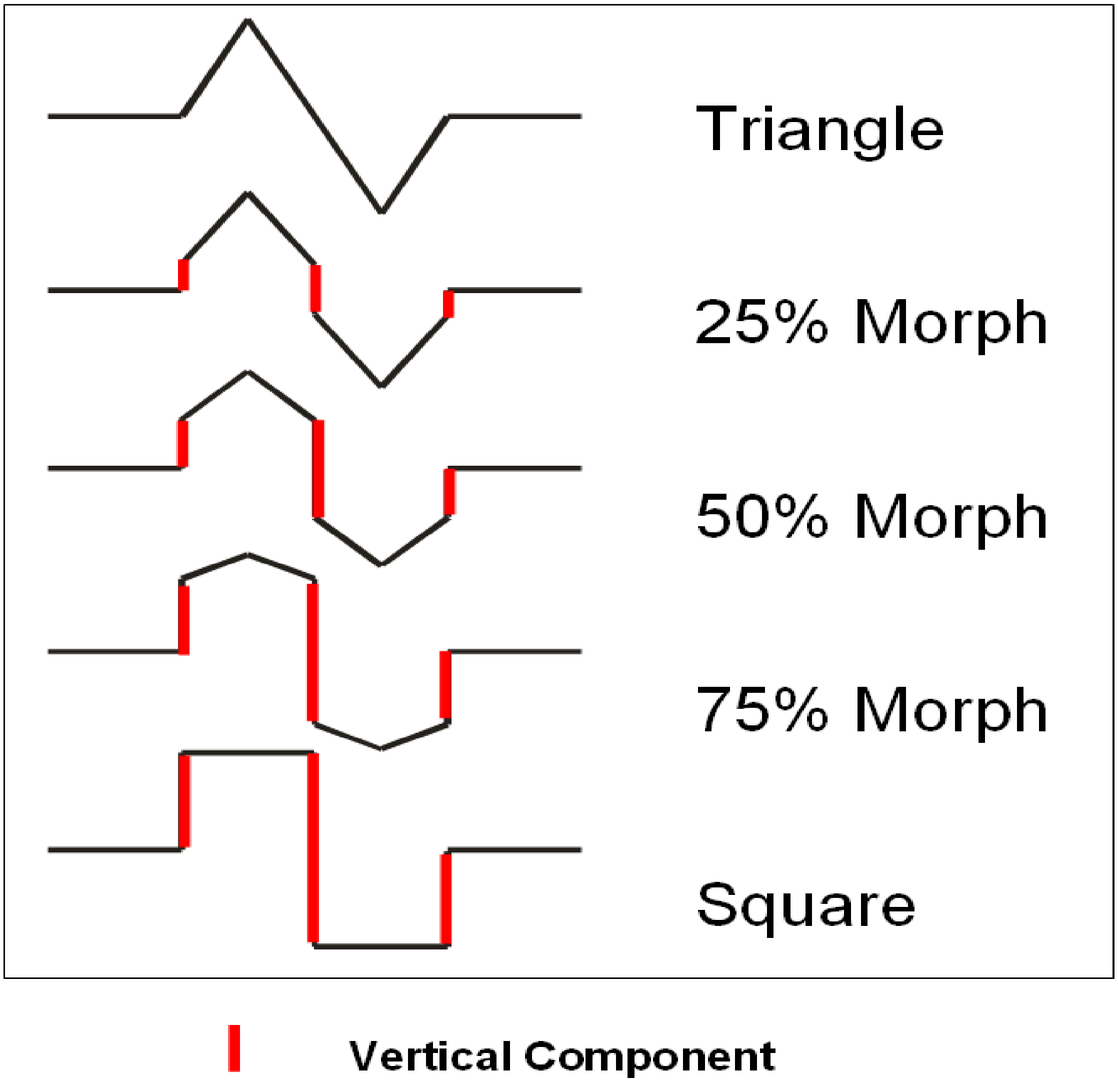


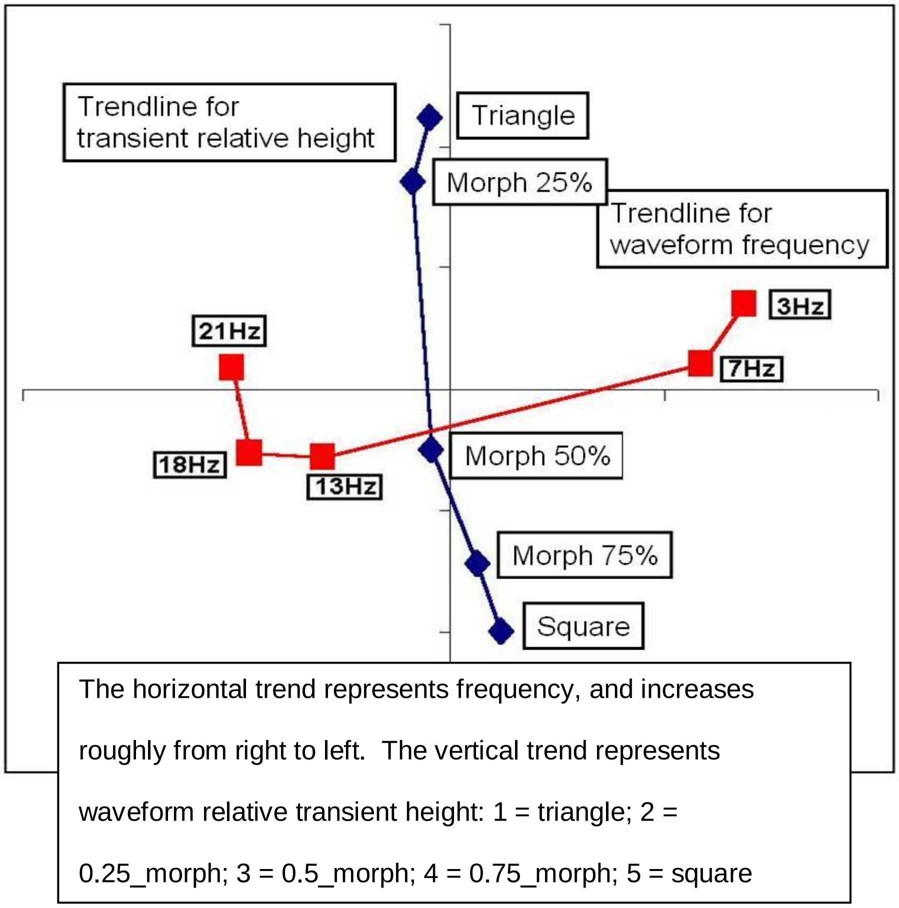
Figure 3-3 Perceptually intermediate waveforms from triangle to square (same as Figure 2-8)

The transform described in Figure 3-3 satisfied our criteria. A short user study revealed that the main factor that influenced the perceived sensation for this intermediate waveform is the amplitude of the vertical component *in relation* to the peak amplitude of the waveform. Relatively smaller vertical components feeling more like a triangle and larger ones more similar to a square waveform.

3.4.4.2 Multidimensional Distribution

To achieve uniform spacing of stimuli composed of simultaneously varying parameters, we chose a large initial set and then selected and adjusted a final set using a previously developed Multidimensional Scaling Analysis (MDS) tool (Enriquez 2002).

Our initial set contained 25 stimuli composed of 5 waveforms (triangle, square and 3 intermediates) displayed at 5 different frequencies (3, 7, 13, 18 and 21 Hz). Following MDS analysis (Figure 3-4), we selected a subset of 9 stimuli to maximize perceptual spread and differentiability. The selected stimuli were composed of triangle, Morph*0.4 (vertical line 40% of total amplitude) and Square waves presented at 7, 10 and 18 Hz.



The horizontal trend represents frequency, and increases roughly from right to left. The vertical trend represents waveform relative transient height: 1 = triangle; 2 = 0.25_morph; 3 = 0.5_morph; 4 = 0.75_morph; 5 = square

Figure 3-4 MDS perceptual distribution of initial stimulus set Family-Based Meaning Association (Same as Figure 2-9)

Table 3-1 Relative amplitudes for perceptual equalization of stimuli (2 sec duration), with assigned arbitrary meanings

	7 Hz *Grass*	10Hz *Flower*	18Hz *Tree*
Triangle *Blueberry*	0.62	0.55	0.45
Morph (0.4) *Strawberry*	0.62	0.55	0.45
Square *Orange*	0.62	0.55	0.45

Perceptual Equalization of Final Stimulus Set

Finally, we ran a short user study using Parameter Estimation by Sequential Testing (PEST) (Taylor 1967) on our chosen set, to adjust amplitudes for perceptual equality for all frequencies and waveforms. Table 3-1 shows the resulting relative amplitudes used in the study.

The main purpose of this experiment was to test the effectiveness of truly arbitrary associations between haptic phonemes and meanings. In order to ensure that the haptic phoneme – concept associations remained abstract, we chose nonsense meanings that were carefully screened to have no semantic connection to the stimuli themselves. However, we exploited the 2-dimensional property of the set and grouped meanings by what we will call families (associated with particular frequencies) and functions (associated with waveforms), allowing users memorize groups of items and thus assemble larger sets with less learning effort (Miller 1956).

The selected family (frequency-related) meanings were types of plants, while functions (waveform-related) were fruits; all 6 items were selected to suggest unique colors, shapes and textures (Table 3-1). Thus, a given stimulus was associated a dual meaning: an 18 Hz triangle wave could be regarded as either a Tree or as a Blueberry.

3.4.5 Methods

Experiment methods consist of setup, a 3-phase experiment protocol and recorded measures.

3.4.5.1 Physical Setup and Instructions

12 participants (6 male, 6 female; age range from 20-45 with median 27) took part in this study; most were university graduate students in the Department of Computer Science. Most had no experience with haptic displays, and the rest had moderate experience. Each was paid $10 for an approximately 1.5-hour session.

Phonemes were displayed on a direct-drive actuated knob, shown in Figure 3-5. The knob was rubber-covered brass with an outer diameter of 10.5

mm and a length of 16.5 mm. The 1mm thick rubber coating prevented slipping and allowed a better grip of the knob while minimizing compliance. The knob was mounted directly on the shaft of a 20W Maxon DC motor (RE025). This 24-volt motor has a stall torque of 240 mNm with a position frequency roll-off at 200 Hz. No position feedback was required since all signals were delivered open-loop.

Participants were seated at a table so that they could comfortably rest their hand and hold an actuated knob (Figure 3-5) which was placed on a raised platform matched by a padded armrest. To mask auditory noise from the haptic display, participants wore headphones and listened to white noise throughout the test session. They received graphical feedback from a Dell 17” LCD monitor positioned approximately 60 cm away, and made responses with a standard mouse and by typing on a keyboard in front of the monitor.

At the beginning of each experiment run, participants read instructions presented on the computer screen and were queried for questions. Following the three experiment stages, at the end of the experiment, participants were debriefed about their experience, and solicited for subjective reactions.

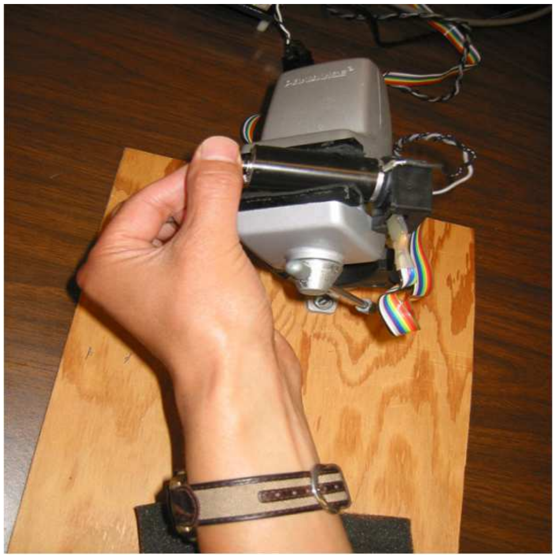


Figure 3-5 Participants held a haptically enabled knob through which haptic phonemes were displayed (Same as Figure 2-1)

3.4.5.2 Protocol

Following instruction, the experiment had three phases: self-guided learning, enforced learning and testing.

3.4.5.3 Self-Guided Learning Phase

Using the GUI shown in Figure 3-6, participants repeatedly selected different Family-Function combinations and felt the corresponding haptic stimulus (consisting of the corresponding waveform at the specified frequency displayed by the haptic knob). Participants could spend up to 5 minutes in this phase and for uniformity, they were not allowed to return to the self-guided learning interface after this point.

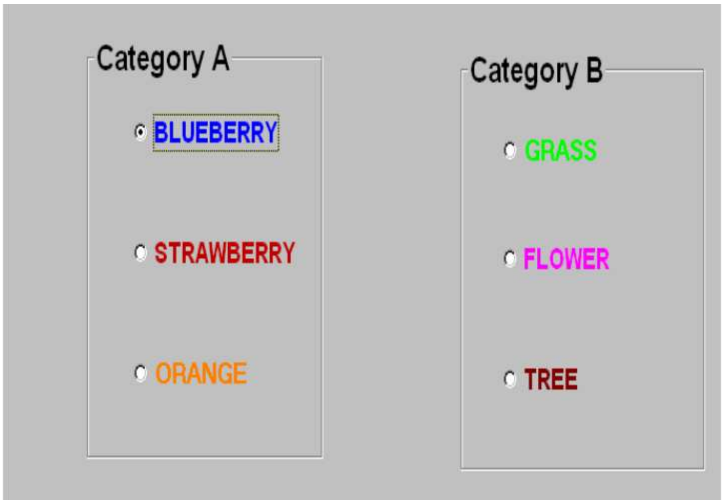
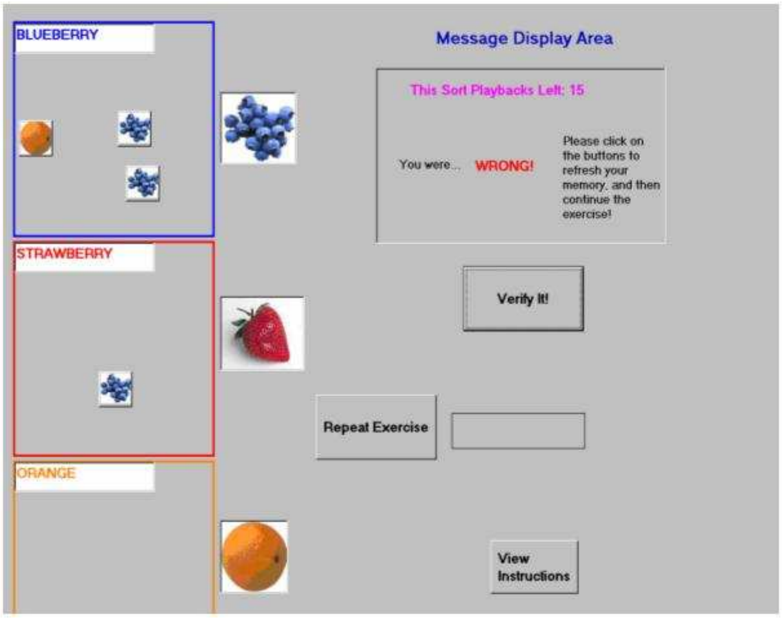


Figure 3-6 Self guided learning GUI (phase 1 of experiment)

3.4.5.4 Enforced Learning Phase

Throughout this stage, participants interacted with Identification and Reinforcement views of the Enforced Learning GUI (Figure 3-7). Participants were presented with 4 randomly selected phonemes (presented as draggable icon tiles) and three destination boxes labeled with either Families (Grass, Flower and Tree) or Functions (Blueberry, Strawberry, and Orange). A left mouse-click on a phoneme tile triggered playback of the corresponding haptic stimulus. With a right mouse-click, the tile could be dragged into the chosen box.

This phase had 20 trials presented in random order: 10 had participants sorting phonemes into boxes labeled by Family, and 10 had participants sorting the phonemes by Function. |A trial consisted of sorting 4 phonemes. This means that participants sorted a total of 80 phonemes into either Family or Function boxes during this stage. Participants were allowed to feel a given tile's stimulus up to 15 times before placing it. To discourage participants from grouping tiles by compared similarity (as opposed to recall of associations) a tile could not be moved or felt again once it was placed into a Family or Function box. At the end of each trial, participants were given visual error feedback and required to re-sort incorrectly identified tiles (Reinforcement view).



This Sort Playbacks Left: 15

You were... WRONG!

Please click on the buttons to refresh your memory, and then continue the exercise!

Figure 3-7 Enforced learning GUI (Reinforcement - second phase of experiment)

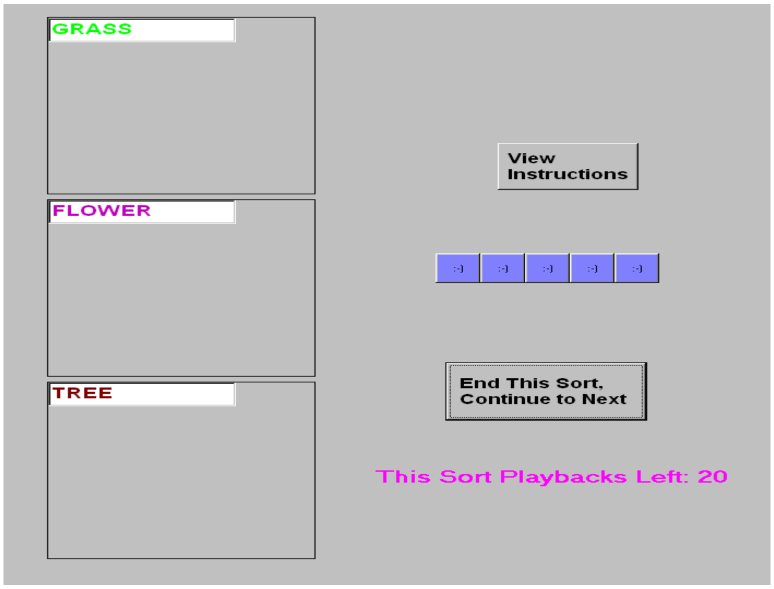
The vertical layout of the destination boxes matched the order in which the buttons were labeled in the self-guided learning phase (Figure 3-6). Our goal was to benchmark learnability of associations for the case of arbitrary associations. It was acceptable for this learning to be demonstrated with either a spatial or semantic mapping, both of which are arbitrary.

3.4.5.5 Test Phase

All participants proceeded to the test phase regardless of performance in the 20-trial enforced learning phase. We chose to measure performance following a uniform learning period, rather than the learning time required to achieve a specified performance level, in complement to the approach taken by (Chan 2005).

The test phase consisted of 20 randomized repetitions of the 9 phonemes tiles, with 10 repetitions sorted into Family boxes and 10 into Function boxes. 180 sort items were randomly allocated to 36 trials with 5 tiles each. To minimize fatigue, a 5-minute break was enforced after Trial 18.

A test trial was identical to the identification step of an enforced learning trial, with the exception that participants sorted 5 tiles instead of 4 (Figure 3-8) and were allowed to display a trial 20 rather than 15 times. We used 5-tile trials here (rather than 4 as during learning) to avoid participant re-use of systematic strategies (other than association knowledge) developed during learning. Following the sort, participants did not receive feedback on their performance.



This Sort Playbacks Left: 20

Figure 3-8 Test GUI (third phase of experiment)

3.4.5.6 Measures

We collected several measures including number and identity of phonemes placed in each destination box by experiment condition, and enforced learning and test phase durations.

3.4.6 Results

12 participants (6 male, 6 female; age range from 20-45 with median 27) took part in this study; most were university graduate students in the Department of Computer Science. Most had no experience with haptic displays, and the rest had moderate experience. Each was paid $10 for an approximately 1.5-hour session.

3.4.6.1 Enforced Learning

The enforced learning phase had an average duration of 19.3 minutes (StDev. 4.8, min 10.6 and max 28.2 minutes). On average, participants made 19.4 out of a possible 80 errors (76% correct, where chance would be 33%; StDev 6.17, min 12, max 33 errors). During this phase, participants made on average 76% correct responses. Individual performance during this phase is shown in Figure 3-9.

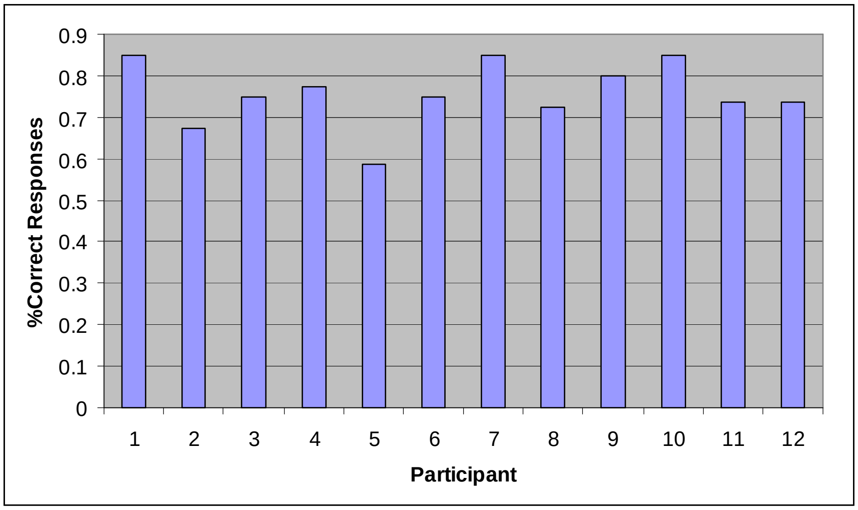


Figure 3-9 Individual performance during enforced learning: % correct responses

3.4.6.2 Average Identification Performance

Figure 3-10 summarizes the correct and erroneous phoneme identifications made in the test phase, for waveform (Function) and frequency (Family). Correct identification averaged across all three waveforms was 22.0/30 (73%) where chance performance would be 33%. Identification of meanings assigned by frequency was slightly better, with, an average for all frequencies of

24.2/30 (81% correct, chance=33%). Unfortunately the experiment was not designed in a way that allows for a simple statistical test. However, the graph presented in figure 3-10 suggests that participants were able to learn the associations between meanings and haptic phonemes.

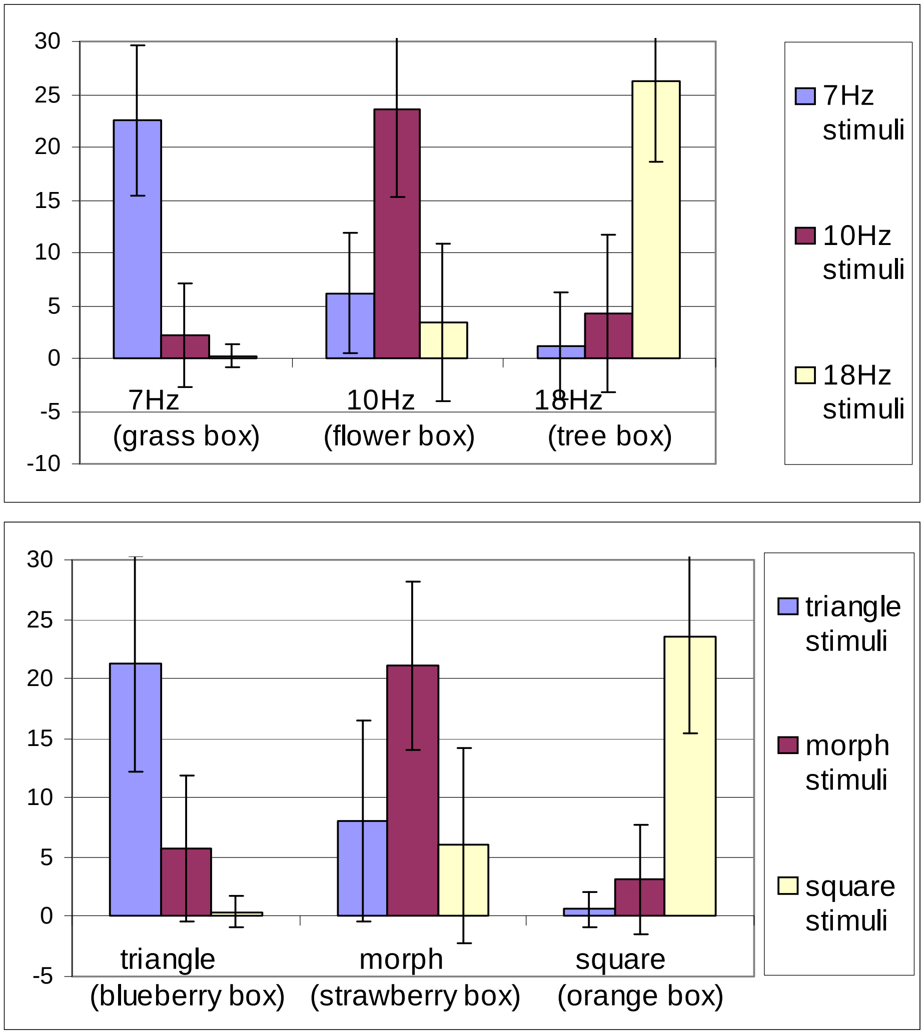


Figure 3-10 Average number of placements made in test phase. Correct placements are enumerated by the bars marked '*'. Each figure represents average behavior across participants of 90 item sorts per subject, where 30 correct placements into each of the three destination boxes would be a perfect performance.

Figure 3-10 also illustrates confusion patterns. For example, participants on average confused the triangle waveform 8.1 /30 times (27%) with the morphed waveform, but only 0.66 (2%) with the square; 7Hz stimuli were confused 6.2 (21%) times with 10Hz, but only 1.3 (4.3%) times with 18Hz. While the ordinal structure of the stimuli tends to drive miss-associations towards the center categories (morph and 10 Hz), it is useful to consider asymmetry within each plot and the two plots with each other: Triangle-to-Morph errors are made more often than Square-to-Morph, and there are more confusions overall for waveform than frequency.

3.4.6.3 Individuals Analysis

Figure 3-11 shows the aggregate identification performance for each subject alone. Most participants apparently formed strong, correct haptic stimulus-meaning associations. A small number appeared to form consistent incorrect associations. For example, Participant 1 consistently identified the morph stimulus at 7Hz for morph at 10Hz (70% of the time) and the square stimulus at 7Hz for square at 10Hz (70%). Participant 7 consistently identified the triangle stimuli for morph (80% of the time). From this we conclude that both Participants 1 and 7 are a representative example of the category of participants who consistently learned a certain association, albeit the wrong one.

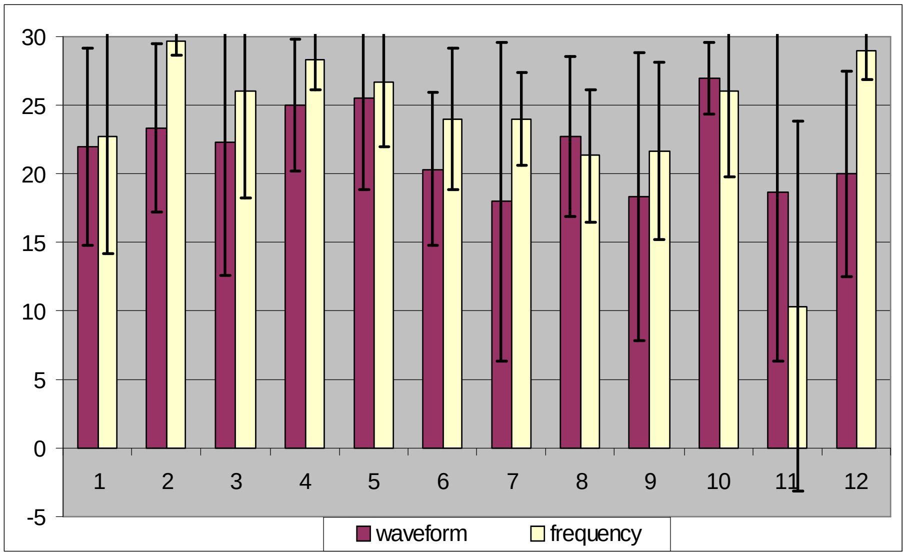


Figure 3-11 Individual identification performance (average of correct responses for each destination box)

Other participants learned some mappings well, but had difficulties with others. For example, Participant 4 achieved near-perfect scores for the triangle (90% accuracy) and morph (87% accuracy) stimuli but exhibited poor results for the square waveform stimuli (56% accuracy).

The best-performing participants (4, 5 and 10) made mistakes in only one of the frequency or waveform associations.

3.4.7 Discussion

Overall, the results of this experiment show that with training of about 25 minutes, participants demonstrated learning of correct associations for the difficult case of arbitrary associations with a performance rate of 73 or 81% for waveform and frequency respectively (chance 33%). An individuals analysis showed that some of the participants were consistently making the same mistakes, suggesting that incorrect but consistent associations were learned. We view these results as promising, particularly in consideration of the intentionally non-intuitive associations and brief training period. We already know that mappings intended to be more intuitive can lead to even shorter learning periods (Chan (Chan 2005), observed a 3-minute average for performance-gated learning of a set of similar size and structure but metaphorically mapped).

3.4.7.1 Validation of Methodology

3.4.7.2 Adequacy of Learning and Learning Protocol

During enforced learning, participants made on average 76% correct responses, which in turn resulted in comparable and reasonable test phase performance (average 77%). Test phase performance indicates that learning worked overall, and its similarity to enforced learning phase performance implies that learning had reached a plateau.

Nevertheless, the design of the learning phases might be improved. It is possible that some of the wrong associations made by the participants could have been avoided with a ‘smarter’ training method. In particular, we believe performance could be improved by allowing participants to return to the self-guided learning interface when consistent mistakes in the enforced learning phase have been detected (Figure 3-11). In addition, a slightly longer self-guided learning phase (here it was limited to 5 minutes) might have allowed a shorter and more effective enforced learning phase.

3.4.7.3 Conservatism of Results

The reported results could be considered to be conservative (for the abstract association case) in that they do not take into consideration these incorrectly associated stimulus-meaning pairs.

The experiment was designed to test for properly associated stimulus characteristic-meaning pairs. However, we did not test for identification of both meanings associated to any stimulus simultaneously. We believe that if we had done this, it might have made the associations easier to remember. We believe that requiring participants to dissociate frequency and waveform (the two signal attributes) made the task harder. Having participants learn separate meanings for the different attributes probably complicated the process of creating mnemonics to remember the signal parameter meanings.

It must be mentioned that allowing participants to feel the signals multiple times probably helped in their performance.

3.4.7.4 Individuals Performance

During the enforced learning phase, participants made on average 76% correct responses, which in turn resulted in reasonable test phase performance, producing an average of 23.1 (77%) correct responses (StDev = 8.0; max =29.7; min =18).

Participants seemed to experience increased difficulty in isolating the mid-range stimuli, but did significantly better for the extremes of the set. This comes as no surprise since we expected the anchoring effect to facilitate the identification of stimuli whose characteristics fall into either ends of their respective scale (Paulhus 1991).

There were no participants who performed generally poorly throughout the experiment. Rather, we distinguish three categories of performance:

1. Did well on all categories (6/12)
2. Did well overall, but consistently learned the wrong association for some categories (4/12)
3. Did well on some categories, but performed poorly on others (2/12)

Most participants (Types 1 and 2) were able to construct a strong and consistent mental binding, despite the fact that the associations presented to them were deliberately designed to hold no intuitive meaning. The predominance of this ability leads us to believe that the majority of users will be able to learn the desired stimuli-arbitrary meaning associations.

With regards to Type 3 participants, we hypothesize that some participants do not have the haptic equivalent of “perfect pitch”, although they are not necessarily “tone deaf”. Such a deficit could stem from either an inability to make distinctions, or from cognitive difficulty in making and remembering associations. To distinguish these, we would need to pair an association test with a stimulus distinctiveness test for every individual.

At the other end of the scale, some participants might have a natural talent in memorizing haptic signals and represent the higher proficiency we observed in this study. These participants probably have a natural ability both to make haptic distinctions and form cognitive associations.

3.4.7.5 Extending the Haptic language: Family and Function

Giving meanings to different characteristics of a haptic phoneme allows us to increase the information density of a haptic language. In this experiment, we utilized haptic phonemes created by varying two characteristics (frequency and waveform) amongst three possible values. We hypothesized that this approach would allow more information-rich haptic signals to be created.

To actually test this premise, a direct comparison of flat and family-based sets must be made. In either scenario, the experiment paradigm used here was not designed to ascertain maximal learnable set size, but instead measured performance for a given set size under a given learning procedure. Further work will be required to determine how large a number of distinct haptic phonemes can be used in a set while still remaining easy to learn and effective.

3.4.7.6 Enduring Associations

We measured the number of times that participants correctly sorted the family + function haptic representations into their corresponding family or function. This was taken as an indication of how difficult it is for a person to learn these arbitrary associations. We were also interested in measuring the change in ability to respond correctly to the signals that occurs during 45 minutes of testing.

We found that overall, participants were able to correctly remember the associations throughout a 45 minute test phase after the enforced learning phase (p=0.028). This suggests that the associations were stored in a longer-lived form than short term memory. It is reasonable to expect that without further demand, these associations would soon degrade following the end of the experiment; further work will be required to establish their longevity given regular use and reinforcement.

3.4.7.7 Design Implications

Test results show that after a brief (5 + 20 minute) training period, participants are able to learn and remember arbitrary haptic phoneme - meaning associations with an average performance close to 80%; a benchmark which we view to be conservative with respect to a real application which would probably benefit from more intuitive mappings and ongoing learning reinforcement. This suggests that a similar set of haptic phonemes could be successfully utilized as a means for a device to communicate simple information to its user.

Confusion patterns suggest that some people might need more time to learn correct haptic phoneme-meaning associations. Results suggest that most participants formed strong mental bindings for all the phonemes with the correctly (or incorrectly) learned associations; and the remainder could make good associations in most cases. Furthermore, it should be noted that most mistakes in identifying the haptic phonemes occurred when participants mistook either end of a characteristic's scale with that of the middle value. Perhaps selecting phoneme sets with only two values per dimension (avoiding mid-values), would result in sufficiently improved identification performance to justify additional dimensions to increase set size. For example, would people do better with three dimensions with two values on each (23=8) rather than two dimensions with three values (32=9)?

3.4.8 Conclusions and Future Work

Inspired in part by human ability to parse non-intuitive graphical icons, the findings presented here suggest that haptic signals can be a robust way to communicate meaningful information to a user: arbitrary associations between haptic phonemes and meanings can be learned to a usable performance level after a 25 minute training period and remembered consistently for a relatively long period of time (45 minutes).

This opens the possibility for the design of interfaces that take advantage of the sense of touch as a communications channel. If arbitrary associations can be learned, this simplifies the task of designing haptic interfaces intended to communicate meaningful information to their users. More intuitive associations, when available, can probably be learned more easily still. Further, we hypothesize that with regular and pervasive reinforced exposure, larger set sizes could be learned to better accuracy. More work is needed to establish this.

In addition to the small experiments proposed in Discussion, most of which are targeted at increasing usable phoneme set sizes, we need to explore ways of increasing the amount of information that can be encoded in a single haptic information module. For example, the set of 9 haptic phonemes tested in this project could theoretically be used (through concatenation or superposition) to create a larger set of haptic words (or icons) which could convey more complex meanings and perhaps open the way for the development of a far-reaching haptic language. How much complexity can be perceptually and cognitively decoded from haptic messages?

.Another important consideration, given the likelihood of multitasking / time-and-safety-critical working environments, is the robustness of haptic icons or phonemes to workload. For example, could users still appropriately utilize the haptic phonemes from this study in a real world situation, such as in an automobile navigation aid or a cell phone identify feature used while walking down a busy urban street? Methodologies for exploring these questions is being developed (e.g.,(Chan 2005; Tang 2005)) but the general concept of designing interfaces for high workload is one with an open future.

3.5 User vs. Arbitrary Icon-Meaning Associations Study

The work described in this section has been published in the Symposium on Haptic Interfaces for Virtual Environments and Teleoperator Systems, Reno, Nevada. (Enriquez 2008)

3.5.1 Introduction

The primary concerns when designing information bearing haptic signals (haptic icons) relate to (a) maximum recognizable size of stimulus set; (b) learnability of stimulus-meaning associations; and (c) longevity of those associations. All of these factors are exacerbated by the currently limited expressivity of haptic displays, particularly those suitable for embedded, mobile and wearable applications where we envision these types of signals to be most beneficial.

Information density of stimulus sets: The experiment reported here uses a modest set size of 10 vibrotactile stimuli, and as such does not directly address the challenge of large sets. However, other work has shown the feasibility of creating a distinguishable set which meets the limits of foreseeable human associative limits, with novice users, a moderately expressive display and close attention to stimulus design (Ternes 2007; Ternes 2008). This promising situation will improve with better devices, an experienced user base and a better understanding of how we perceive and process these types of signals.

We also point out that icons can carry varying amounts of information. Information transfer can be increased by increasing set size (while ensuring the set remains identifiable), or alternatively by increasing the perceivable information content of individual signals; or possibly both. The best approach depends on the need of the application. For both, the threshold of perceivability is modulated by characteristics of the anticipated operating environment, and complex icons might be harder to recognize under workload. Design heuristics and techniques for this aspect of icon construction is emerging in other work (e.g., (Chan 2005)), but is not the goal of the project described in this (Section 3.5).

Models for creating and learning associations: As further discussed in Section 3.3, common approaches to creating stimulus-meaning associations are abstract or semantic, with the first offering the ability to control, optimize and thus maximize the size of a usable stimulus set, while the latter seems as if it would be easier to learn and remember. These two approaches have not, however, been compared for efficacy side-by-side. To do so, the difference between them must be considered more broadly: we suggest that at issue is not whether the designer perceives a semantic association, but whether the learner does. Through our past experience with designing and deploying haptic icons, we have noticed that users often have personal opinions about appropriate associations which the designer cannot possibly predict. We therefore hypothesized that the safest way of supporting semantic associations is to let the user “roll their own.”

Longevity of associations: The final and likely the most critical prerequisite for a successful implementation of information bearing haptic signals is the potential for an enduring association. Once learned in an initial session, will they persist without reinforcement after an interval of time? How long will they persist? Do different associative and learning mechanisms influence success?

The experiment reported here is a first effort to shed light on the last two questions.

3.5.2 Overview

The experiment presented here was designed to test the hypothesis that allowing participants to choose their own stimulus-meaning associations would, by permitting leverage of their own implicit mental models, improve various subjective and objective metrics relating to learning and retention of those associations. Furthermore, this experiment examines the degree to which users retain learned meanings after a two-week interval, without the benefit of any interim reinforcement.

Our typical practice for supporting users' learning of stimulus-association meanings, regardless of the mechanism used by the designer to build the set, has been an iterated reinforced learn-test cycle (e.g., (Chan 2005; Enriquez 2006)). In the present experiment, we used this approach in a first condition termed arbitrary associations, using pre-assigned, randomly chosen stimulus-meaning matches for a set of 10 meanings. We also tested a second condition termed user-chosen, wherein users selected associations for another 10 meanings from a set of 20 perceptually differentiable tactile stimuli. In both conditions, we otherwise used the same methodology for reinforcement learning, subsequent recall testing and eliciting subjective responses. To broaden the test's external validity, we evaluated these two conditions in two simulated interface scenarios: a hand-held navigation unit and an automobile radio control.

The 20-element stimulus set we used was constructed by varying rhythm patterns of two-second duration, presented at different amplitudes through a vibrotactile display.

Our results suggest that participants are able to learn and later remember stimulus-meaning associations after a brief learning period at 80% recall. Furthermore, association persistence at two weeks after the learning period was 86% of the originally learned associations (with no further training). We saw no significant difference in average performance between arbitrary and user-chosen associations. Interestingly, many participants reported that they believed the arbitrarily-chosen associations had been designed with metaphorical intent; i.e. they discovered their own mnemonic associations. Perhaps this could be an explanation for the undifferentiated result. Post-experiment interviews also revealed that participant’s expectations and confidence levels for their actual performance sharply lagged their actual performance. None believed that they could recall more than a few of the previously learned associations.

These findings have important implications for the design of interfaces intended to communicate information through touch: they underscore the eminent feasibility of using haptic icons in everyday interface design, suggesting high learnability, independent of designer assignment mechanism. They also improve our understanding of the cognitive steps employed by users in their learning process. A better grasp of this process is essential if we are to maximize the haptic channel's potential, by designing icons optimally and supporting users in learning them.

3.5.3 Approach

Our group and others have previously tested people's ability to learn and remember both semantic and deliberately arbitrary (Section 3.3) stimulus-meaning associations and found encouraging results for both. Here, we sought to compare these two cases directly, and to use longitudinal recall as one important new metric.

To test this proposition, we needed a mechanism for learning the associations as well as a performance measure indicating how well the user is able to perceive and process the signals and relate them to their respective meaning

Thus, we ran this experiment in two sequential sessions separated by two weeks, using the same participants in both. In one of his/her sessions, the participant chose which signals would represent the set of 10 meanings for that session. In the other, he or she were given arbitrary signal-meaning associations for a different 10-meaning set. The order of these sessions was randomized among participants. The two sets of meanings were drawn from two embedded contexts and counter-balanced with individual and condition order.

Each session had three stages (Figure 3-12). The first was self-guided learning, where users used a graphic interface that showcased a set of tactile stimuli and allowed them to learn associations between them and their meanings (Chan 2005; Enriquez 2006). This was followed by a computer-guided reinforced learning stage, where users were asked to identify a series of randomly presented tactile signals and drag them into boxes labeled with their respective meanings, while receiving feedback about errors. In the final test stage, users again performed this learning task but without error reinforcement.

In this experiment, we use objective measures of recall performance for both arbitrary and user selected associations (both immediately after training and two weeks after) as well as subjective measures of participant opinion regarding task difficulty and confidence levels for the learned associations.

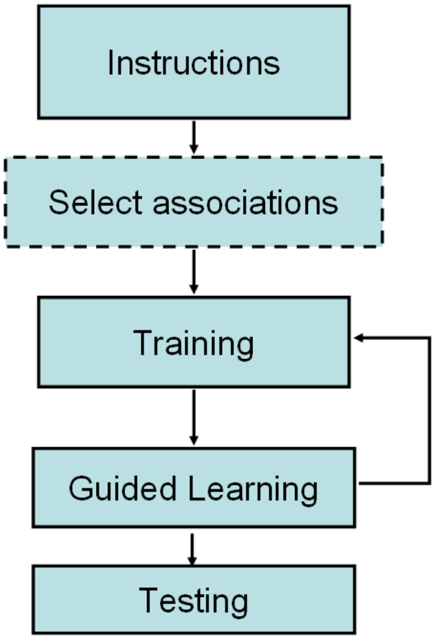


Figure 3-12 Experimental procedure

3.5.4 Methods

Experiment methods consist of display and setup, stimulus design, a 3-phase experiment protocol and recorded measures.

3.5.4.1 Tactile Display

Our experiments were carried out using a custom display integrating one Audiological Engineering (AE) VBW32 vibrotactile display (www.tactaid.com, visible in Figure 3-13). These voice-coil-based transducers, which are used commercially in hearing aids, are capable of producing precisely timed (on/off within 2 ms) waveforms at a useful range of frequencies and amplitudes, with maximum efficiency at 250 Hz; and can be driven directly by a computer's sound card. Tactile displays using similar technology can be found in commercially available mobile phones, PDA's and GPS navigation units.

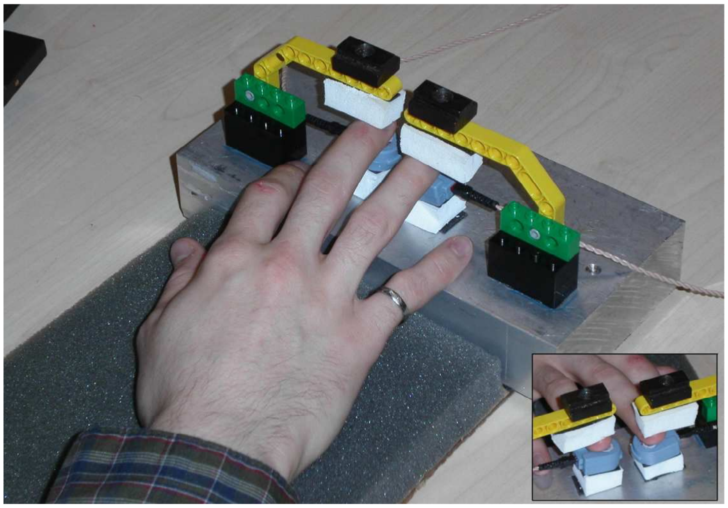


Figure 3-13 Participants placed the index finger of their non-dominant hand on a tactile display

3.5.4.2 Stimulus and Meaning Set Designs

The design for the vibrotactile stimulus set used for this study took as a starting point the 84-element set of rhythm, frequency and amplitude mentioned above (Ternes 2007). Although we only required 20 stimuli here, we felt it would be interesting to add additional textural diversity through the addition of some more naturalistic stimulus, i.e. more broad-spectrum than the single-frequency tones of the source set; prior work has shown that having richer timbre in auditory signals aids in recalling meanings associated with them (Brewster 1994). We hypothesize that similarly, richer tactile signals would lead to higher recognition rates as well. Therefore, we chose 11 disparate stimuli from the 84-element rhythm set (drawn from various distinctive areas of that set's MDS map). These were complemented with 9 additional signals created manually by auditory recording of sounds through a microphone, such as taps on the microphone and scratching over a rough surface, to reach our target set size of 20.

Table 3-2 Simulated interface scenarios used for functions

GPS Navigation	Left, Right, Forward, Back, Up, Down, Faster, Slower, Stop, Go
Automotive Audio System	Volume, Balance, Bass, Treble, Mid Range, Fader, Mute, Tuner, CD, AM

Signal meanings were drawn from two scenarios representative of the type of multi-tasking, attentionally demanding contexts where we anticipate haptic icons will be most useful: a hand-held navigation unit, and an automobile radio control. 10 were used, a group size felt to represent a reasonably broad utility; Table 3-2 lists their specific values. We took this approach to avoid unattractive alternatives of meanings that were completely unrelated (unfeasible) or which could not be semantically organized by participants in some unexpected way – that is, we asserted the connections among the group, leaving it unambiguous.

3.5.4.3 Physical Setup and Instructions

The design of the apparatus (Figure 3-13) was driven by needs for consistent hand position and finger pressure, as well as vibration isolation to prevent crosstalk between the stimulus sites. It utilized one AE display mounted on a 3 cm thick aluminum plate and insulated with 1 cm thick latex foam rubber commonly used to mechanically isolate sensitive electronic equipment from vibration. The participants placed their non-dominant hand on another foam pad which was attached to the aluminum plate; weights mounted on articulated plastic arms held his/her index finger against the transducers with a constant pressure of 30 grams.

The tactile display was interfaced through the sound card in a 2.5 GHz Pentium 4 computer running Windows XP. To mask auditory noise from the haptic display, participants wore noise-canceling headphones and listened to white noise throughout the test session. They received graphical feedback from a Dell 22” LCD monitor positioned approximately 60 cm away, and made responses with a standard mouse and by typing on a keyboard in front of the monitor.

At the beginning of each experiment run, participants read instructions presented on the computer screen and were queried for questions. Following the three experiment stages, at the end of the experiment, participants were debriefed about their experience, and solicited for subjective reactions.

3.5.4.4 Protocol

Every participant carried out two sessions of the experiment (U: User-Chosen associations and A: Arbitrary associations), conducted two weeks apart. For each session, participants had to learn associations between 10 signals and 10 meanings. In the user-chosen session, participants were given 10 meanings drawn from one of the two contextual scenarios used, and asked to choose their favored signals from the full set of 20 signals described above. For their arbitrary session, participants were presented with arbitrarily chosen associations between the set of 10 meanings and 10 of the 20 signals (the same associations for all participants). The order of the two different sessions was randomized among participants.

Our two conditions (arbitrary or user-chosen associations) and two meaning scenarios (navigation or radio) thus resulted in 4 participant types (Table 3-3).

Table 3-3 Allocation of experiment run types. Participants were assigned to one of the four experiment run types.

Session type	1st Session	2nd Session (+2 wks)
1	User-Chosen+Radio	Arbitrary + Nav
2	User-Chosen + Nav	Arbitrary + Radio
3	Arbitrary + Radio	User-Chosen + Nav
4	Arbitrary + Nav	Chosen + Audio

For the session run second (regardless of type), participants began the session with a brief recall test to measure how well they could recall the tactile signals associated to meanings learned two weeks prior. Participants had no opportunity to re-learn the associations before this recall test.

Each session had three phases: *training, guided learning* and *testing* (Figure 3-12). Participants were allowed to switch between the self-guided and reinforced learning stages as many times as they required until they decided to proceed to the testing phase. Once in the testing phase, they could not return to either the self guided or reinforced learning phases.

Training Phase: Using the GUI shown in Figure 3-14, participants could repeatedly click on each of 10 different buttons labeled with meanings and feel the corresponding vibrotactile stimulus. Participants were allowed to return to the self-guided learning interface from the reinforced learning phase if desired.

Guided Learning Phase: In the Identification and Reinforcement views of the Guided Learning GUI (Figure 3-15), participants were presented with 10 labeled meaning boxes, purposefully ordered along a different dimension from

Figure 3-14, along with 10 draggable stimulus tiles. A left mouse-click on a tile triggered playback of the corresponding stimulus. With a right mouse-click, the tile could be dragged into a meaning box.

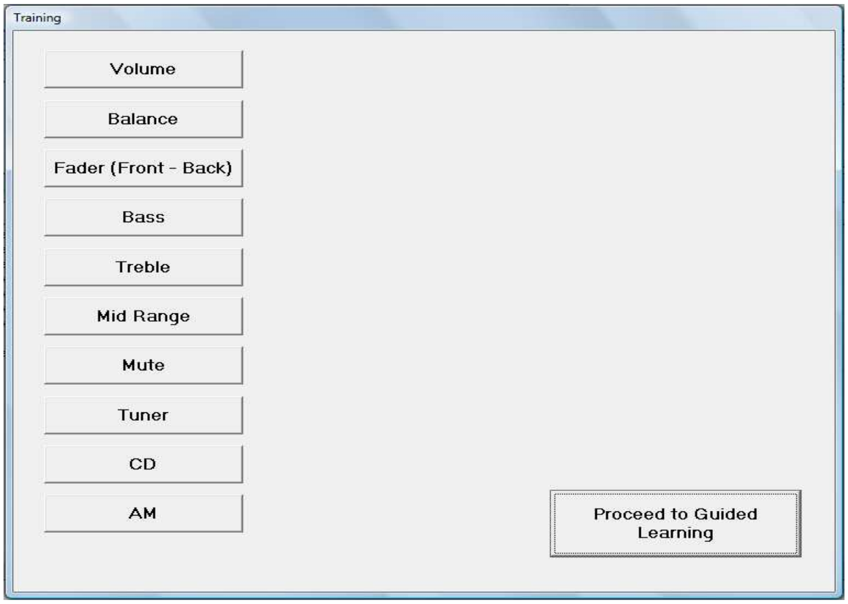


Figure 3-14 Participants learned pre-defined meanings for 10 haptic icons, by feeling and matching to targets

Participants could feel a given tile's stimulus any number of times before placing it. To discourage participants from grouping tiles based on relative comparisons (as opposed to absolute recall of associations) a tile could not be moved or played once it had been placed. At the end of each trial, participants were given visual feedback regarding any errors (correct tiles turned green and incorrectly placed tiles turned red) and could return to the self-guided learning phase if they wished to do so. Once a participant had sorted all 10 tiles into their meaning boxes three times, they could proceed to the testing phase.

Figure 3-15 Guided learning and testing phase GUI. In training, participants tested their knowledge of the signals' meanings and received feedback as to their placement of the tactile signal tiles (green for correct placements, red for incorrect). Placement feedback was not given during the Test Phase.

All participants proceeded to the test phase regardless of performance in the reinforced learning phase. Given that we intended to measure recall performance as a function of condition, we chose this approach rather than having participants train to a preset performance level.

Test Phase: A test trial was identical to the identification step of a Guided Learning trial, with the exception that participants did not receive feedback on their performance, which was recorded. The test phase consisted of 10 randomized trials. To minimize fatigue, a 5-minute break was enforced after Trial five.

3.5.4.5 Measures

Our objective metrics were number and identity of tactile signals placed in each meaning box by experiment condition, and reinforced learning and test phase durations. We also measured the level of association recall from the first session after two weeks (before the second experiment run). That is, for n participants, we obtained 2n observations of immediate recall (n for each condition), and n observations of 2-week recall (n/2 for each condition).

To obtain subjective responses, we conducted an open-ended interview with each participant after each experiment run, in which we inquired about their thoughts about the tasks they had performed and the level of difficulty for the two different experimental conditions. Just before the short recall test at the start of the second session we asked participants how many of the associations they thought they could remember after two weeks. For the interview after the second and last run of the experiment, we also inquired whether choosing associations helped in learning them.

3.5.5 Results

3.5.5.1 Participants

10 male and 2 female science graduate students were paid for their participation in the study (22-34 years, all right handed).

3.5.5.2 Identification Performance

Figure 3-16 shows the results for identification performance for the first session (regardless of session type) for each of the 12 participants, paired with recall of those same associations two weeks later (tested before the 2nd session with a new set). On average, participants correctly recalled 80.1% of the stimulus-concept associations immediately after this first learning period (left bar in each pair). Two weeks later, participants correctly recalled 70.1% of the same associations (right bar), or 86% of those associations recalled in the first test.

Figure 3-17 presents average recall performance immediately after learning for all sessions, grouped by simulated interface scenario type and user choice of associations. Figure 3-18 shows the same data broken apart more specifically. A single-factor within-participants ANOVA shows a statistically significant difference for scenario ($p<0.002$, $F=18.857$) and session order ($p<0.002$, $F=20.056$). The difference between User-Chosen and Arbitrary associations was not significant at $p\leq 0.05$.

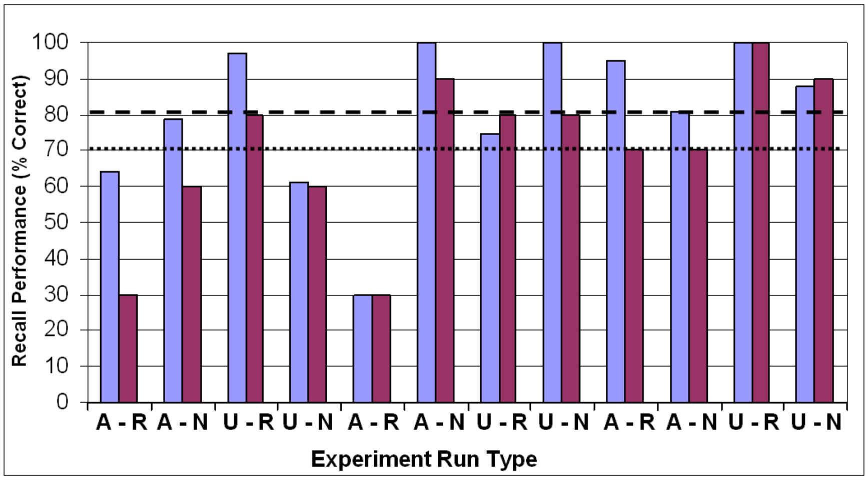


Figure 3-16 Longitudinal recall. Identification performance for first set of associations learned, by participant: immediately after learning (left blue-shaded bar) and as recalled two weeks later (right red-shaded bar). Bar labels indicate session type (U: User Chosen or A: Arbitrary associations; R: Radio or N: Navigation scenario). The dashed and dotted lines represent average performance immediately after learning and recall two weeks later respectively.

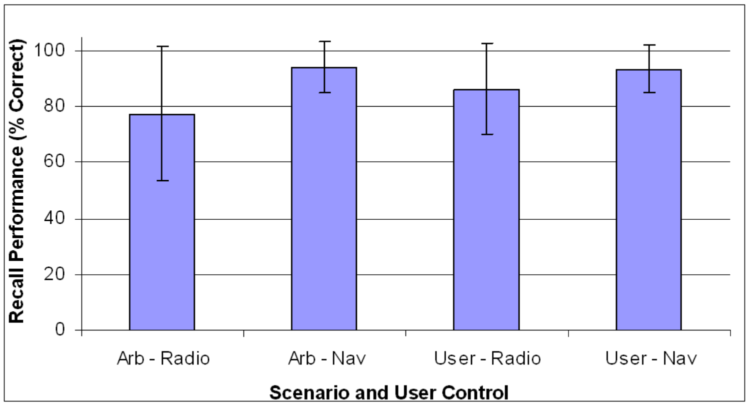


Figure 3-17 : Identification performance immediately after learning, grouped by interface scenario and user control over associations. 24 sessions are represented (two per participant; eight values per bar).

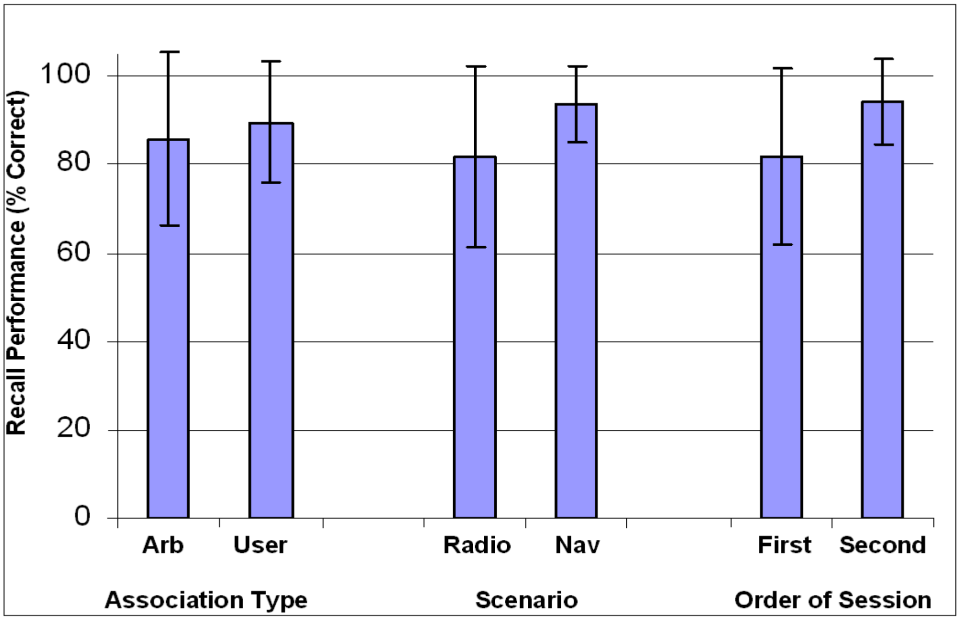


Figure 3-18 Performance grouped separately by: association type, scenario, and order of session (0 or 2 weeks). Performance improved significantly after a two week gap, and differed for the interface scenarios.

3.5.5.3 Subjective Opinions

Following their first session, participants expressed that learning the associations was easier than they had expected. Some commented that the arbitrarily chosen associations were a good match to the functions in the interface (they did not learn that the associations were completely arbitrary until after completing the second session). One individual had trouble learning his first session's associations (arbitrary), stating that they "did not make sense". This participant had the lowest score for the first session's associations and their 2-week recall (both 30% correct). However, in the second session's associations, where this individual chose his own associations, his recall performance was 100%.

At the start of the 2nd session, none of the participants believed that they could recall all the associations learned two weeks prior. Most participants predicted recall of zero associations; only two expected they could recall three or more.

After the second session, participants were asked to comment on whether having a choice for the stimulus-function associations made a difference. Most indicated a preference for choice. One (the one who had trouble learning the arbitrary associations) commented that choice was a necessity, given that the associations should make sense if you are to learn them. Only one participant believed that having a choice did not matter.

All participants were astonished to learn of their recall performance and commented that they did not believe their performance could be so good after such a short learning period. Furthermore, most participants reported that learning the associations was easier than they expected and that choosing the signals themselves simplified learning even further.

3.5.6 Discussion, Conclusions and Future Work

Our primary research questions were:

1. Can a user learn and later recall a set of 10 concepts associated with tactile stimuli?

2. Do these associations endure after two weeks?

3. Are associations easier to recall when users choose the tactile stimuli-concept associations themselves?

3.5.6.1 Immediate and Longitudinal Performance

We found that both arbitrary and user-selected associations between tactile signals and meanings can be learned to a usable performance level (80%, where chance would be 9%)) after a 10-15 minute training period. This corresponds to 8/10, or a single switched assignment. Associations were remembered consistently over time (86% recall after two weeks of the initially learned associations) even without further reinforcement. To the best of our knowledge, this is the first report of long term recall of synthetic haptic signal associations.

These are promising results, which are consistent with positive answers to our first two research questions. While the evaluation described here was small, it was designed to be representative of realistic use contexts. If users can indeed learn the meanings of a set of vibrotactile signals this quickly and persistently, haptic icons could be a practical method to present information such as device state or function identity in everyday contexts.

Further, we hypothesize that with regular and pervasive reinforcement, larger set sizes could be learned to better accuracy regardless of associative method. There are two basic stages of learning: acquisition (initial learning) and maintenance (repeated exposure to a single type of stimulus). The latter enables the organism to learn about stimulus properties and allows for longer term recall (Hall 1991). More work is needed to establish this here.

3.5.6.2 Learning and Performance by Association Method

Arbitrary associations were learned to the same level of recall as those with user-chosen associations, for the conditions present here. These results seem to show that the theoretical advantage of having the kind of metaphorical cue as presented by Chan et al. or more literal representation of the real world (Gaver 1989) was not necessary for the given haptic icons.

However, based on subjective responses it appears that participants often created their own metaphors for the arbitrarily composed sets; they were not dependant on a designer to build and explain the associations. The fact that they were able to do so might be one of the most interesting and unanticipated results of this experiment. It begs the immediate follow-up question of whether they would have been less able or willing to create their own mnemonics if they were explicitly told up front that the stimulus-meaning associations were ungrounded in any intentional meaning (at least some apparently believed that designer-created meanings existed).

Setting aside the tantalizing issue of how people are able to carry out what we thought would be a very difficult task: if borne out in more extensive evaluations on larger sets, the implication is greatly simplified icon design. The downside of a semantic approach to icon design is the difficulty of making it scale.

However, more intuitive associations, when available, can probably be learned more easily still; this seemed to be the view of our participants, regardless of their performance. Chan et al. found higher levels of success for haptic icons when users were given an explanation for their design and their related associations (Chan 2005), although this was not a controlled condition. An interesting direction to explore is whether what we observed was a similar learning performance but at differing levels of effort; and whether, with more difficult learning tasks, this effort differential might translate to a measurable performance deficit.

3.5.6.3 Actual versus Subjective Recall Performance

Participants did not believe they could recall stimulus-meaning associations at two weeks, and yet objective recall was only 10% (1/10 matches) less than immediately after training.

Two areas of future investigation come to mind. On the cautionary side, it is conceivable that a lack of confidence could be detrimental in real usage; i.e. an un-trusted warning might be dismissed; at least one study by our group has suggested this with respect to warning signals (Enriquez 2004). Taking a more positive view, we wonder whether the disparity we observed is simply a matter of unfamiliarity; i.e. will trust in one’s “tactile intuition” come with regular, reinforced use?

3.5.6.4 Individual Differences

Even though results show that most participants could learn arbitrarily assigned associations, one individual struggled (30% recall, compared to 100% for 2nd-session user-chosen associations); stating that they “did not make sense to him”. Based on other similar anecdotal observations such as this, it would be unsurprising to find a wide and possibly bimodal distribution in human tactile acuity and/or higher level signal processing. And just as there are “visual” and “auditory” learners, perhaps some individuals will easily learn haptic associations in their own right, while some others will require a metaphorical reference to another modality to ground them.

3.5.6.5 Information Density and Larger Icon Sets

We also need to explore ways of increasing the amount of information that can be encoded in a single haptic information module. For example, the set of 20 tactile signals tested in this project could theoretically be used (through concatenation) to create a larger set of haptic phrases, each of which could convey more complex meanings and perhaps open the way for the development of a far-reaching haptic language. What is the limit to the complexity that can be perceptually and cognitively decoded from haptic messages?

To be broadly useful, haptic icon set size must be somewhat scalable, and certainly larger than 20 items; work by others in our group suggests discernable sets of 75-100, using today’s vibrotactile display hardware (Ternes 2007). These sets will perhaps suffice so we can concentrate on cognitive bottlenecks while awaiting hardware improvements. This raises two questions: Can participants remember associations for a “large” set of haptic icons, and how do we systematically design both the stimulus sets and the associations to them to optimize learnability?

Most participants in this study seemed to have little difficulty in creating metaphors to remember the haptic icons’ meanings; but how will this scale to larger sets? To increase scalability, more work is required to determine what underlies intuitive associations. Furthermore, if these haptic icons are to be used in different applications or interfaces, we must set standards to ensure that their meanings remain consistent throughout.

3.5.6.6 High Workload Environments

Another important consideration, given the likelihood that these haptic icons will be used in multitasking / time-and-safety-critical working environments, is their robustness to workload. Could users utilize the tactile signals used here in a real-world situation, while driving an automobile or using a handheld GPS while walking down a busy urban street? Methodologies for exploring these questions are being developed (Driver; Chan 2005; Tang 2005; Enriquez 2007) but the general concept of designing interfaces for high workload is one with an open future.

3.5.7 Summary

Results from this experiment have provided some initial data on the degree to which humans can learn and retain tactile stimulus-meaning associations, in a somewhat situated context, as well as some subjective observations on how they might be performing this learning feat. In general, our results are very encouraging; suggesting that everyday use of haptic icons with current vibrotactile display technology is feasible on the basis of learnability.

We have suggested many directions for future work. Of these, the most immediately essential have to do with scalability of set size, a more detailed look at longitudinal learning, and use in realistic, attentionally demanding contexts.

4 Tactile Stimulus Masking

The work described in this chapter (beginning with Section 4.3) has been published in the Brain Research Bulletin, Special Issue: Robotics and Neuroscience (Enriquez 2008).

4.1 Introduction and Summary

In the future we envision devices that will seamlessly communicate with us through the sense of touch. In order to achieve this kind of transparent communication, we should be able to design haptic signals in a manner that allows their recipient to perceive them clearly, even when presented in close temporal and perhaps spatial contiguity. In Chapter 2, we introduced two experiments designed to test information-bearing haptic signals. In these experiments, the signals were presented with enough temporal separation as to ensure that they would not interfere with each other. As we move towards more realistic scenarios, it is important to investigate possible adverse effects arising from close temporal and spatial presentation of these types of signals.

In this chapter, we present an experiment designed to test the effects of tactile stimulus masking. Stimulus masking refers to the interference of one stimulus with another causing a decrease or lessening in perceptual effectiveness. The clear perception of different tactile sensations in close temporal proximity will depend not only on individual stimulus design, but also on the effects of interference from other stimuli preceding or following them,

The experiment presented here was designed to investigate backward and common-onset masking of vibrotactile stimuli using a commodity display. Backward masking occurs when a stimulus follows the presentation of a second target stimulus preventing (masking) its perception. Common-onset masking occurs when both target and masking stimulus are presented simultaneously but the masking stimulus persists after the target stimulus has finished (masking the perception of the target).

We used a two-channel setup, presenting stimuli to the middle and ring finger pads of a participant’s right hand. The stimuli consisted of 250 Hz sinusoidal waveforms displayed at a fixed amplitude in various combinations of duration (0, 30 or 300 ms) and stimulus onset asynchrony (0 or 30 ms). Our results confirm the existence of a statistically significant masking effect for both forms of haptic masking explored, with common onset exhibiting a significantly larger masking effect than backwards. However, an analysis of confidence in response levels shows no difference between the two successful masking techniques. We discuss mechanisms that could be responsible for these results, which have implications for the design of user interfaces that rely on tactile transmission of information.

4.2 Contributions

- Designed and adapted methods to measure tactile identification masking for common-onset stimuli.

- Identified crucial but previously untested perceptual thresholds relating to common-onset stimulus temporal spacing.

- Discovered a significantly stronger masking effect for common-onset stimuli than for backward masking.

4.3 What is Masking?

A common definition for stimulus masking is “the interference of one perceptual stimulus with another causing a decrease or lessening in perceptual effectiveness” (Howard 1995). For our purposes, we will consider a stimulus to be masked when interference from another stimulus (differing either in time or location or in both) prevents the recipient from explicitly identifying it.

Our own motive for understanding tactile masking is to support perceptual design of user interfaces that convey information through touch. We predict that these interfaces will often be used in multitasking contexts that are filled with distractions (Chapter 5). Two perspectives pertain. Sometimes, a designer will wish to avoid inadvertent masking of signals: for example, temporal masking due to “packing” stimuli closely in time in an effort to maximize information transfer (Tan 1999; Enriquez 2002; MacLean 2003). At other times, the designer might wish to deliberately mask perceivable information-bearing tactile stimuli as a tool to isolate the factors that affect our ability to process tactile patterns sequentially, and their relation to attention and signal detection (Marcel 1983; Greenwald 1996; Merikle 2000), or to produce actionable signals that minimize attentional demands.

Our focus for this experiment was on the latter, and in the study described here we seek practical methods (usable in commodity applications) for masking information-bearing tactile signals.

4.4 Related Work

Our knowledge of haptic single-stimuli perception is exemplified by experiments of Srinivasan, Tan and others which use synthetic stimuli to determine various human capabilities, including pressure, stiffness, position resolution and force magnitude (Srinivasan 1989; Srinivasan 1990; Tan 1999); while Klatzky & Lederman have studied texture perception extensively, most recently touching through a stylus (Klatzky 2003). These and other studies lay the foundation upon which we can further explore haptic perception and begin to build a haptic language. However, because of the real-world environment in which this language will be used (full of distractions and competing demands on our attention) we also need to understand how haptic signals are masked.

We differentiate the haptic masking studies we will review here along two dimensions: characteristics of the stimulus being masked, and properties of the masking technique itself. These studies typically investigate either stimulus detection (a stimulus is perceivable as present or absent) or stimulus identification (where the stimulus incorporates some manner of variation in pattern, e.g., spatial layout or rhythm, and is thus capable of delivering information based on its identity). Masking techniques that have been commonly studied include forward (masking stimulus precedes target stimulus; attributed to temporal integration), backward (masking stimulus follows presentation of target stimulus), and sandwich (target stimulus is both preceded and followed by maskers) masking.

Numerous studies have investigated the masking effects of tactile stimuli. Many of these have focused on how masking affects the *detection* of simple vibrotactile stimuli (Weber 1978; Gescheider 1989; Gescheider 1995; Oxenham 2001). In these studies, different tracking methods are used to determine detection thresholds for stimuli in the presence of different forms of maskers. Some utilized collocated target and masker stimuli, with the masker being band-limited noise and the target a sinusoidal waveform (Gescheider 1989). Another paradigm utilizes targets and maskers presented at different frequencies – e.g., (Weber 1978). These results have provided a foundation for other investigations into masking effects of more complex, information-rich stimuli.

Researchers have also begun to study temporal and spatial masking effects on identification of different types of tactual stimulation patterns (intended to carry detectable information beyond presence/absence) delivered to various areas of the body, e.g., (Craig 1987; Tan 2003). These studies investigate the effects of stimulus masking on different vibration patterns presented through an array of vibrotactile displays used to convey meanings in a similar fashion to the raised dots used on an electronic Braille display.

Aligned with the goal of the experiment reported here, some recent studies using relatively complex stimuli, representing either temporal and spatial patterns have reported several different forms of masking which can occur for the sense of touch (Craig 1985; Rinker 1998; Rinker 1998; Klatzky 2003; Tan 2003; Srinivasan 2005). Of particular relevance is a series of experiments by Tan et al. which targeted temporal masking properties of complex patterns designed for information transfer (Tan 2003). In that study, stimuli were delivered to the left index finger of three participants who were asked to identify target signals masked by forward, backward, and sandwiched paradigms with Stimulus Onset Asynchronies (SOA) of up to ±640ms. The SOA is the temporal interval between the onsets of two stimuli. Seven perceptually distinct stimuli composed of one, two, or three spectral components (2-4, 30 and 300 Hz) were constructed at each of two signal durations (125 or 250 ms). The lower frequencies presented at shorter durations resulted in partial waveform representations. The masking stimuli were selected from the same stimulus set as the target stimuli. Results show a masking effect (average 70% of correct responses, with performance increasing with SOA) for the different types of masking. For these complex stimuli, participants often confused characteristics of the masker with those of the target; and there was considerable variation in individual performance.

Craig et al. performed a series of experiments investigating the ability of participants to localize a tactile pattern presented at one of several locations on their left index finger, in the presence of a second tactile masking pattern (Craig

1989). The target stimulus, generated on a 6 x 24 array of stimulators, was presented either by itself or in the presence of an extraneous stimulus (masker) that either preceded (200-0ms SOA) or followed (0-200ms SOA) the target. The masking stimuli were identical in form to the target stimuli. The localizability of the target was affected by the SOA between the target and masker with masking being strongest (68% correct responses) when the masker followed the stimulus at relatively short SOA's (0-30 ms). In another study (Craig 1997), Craig et al. found that the identification of a spatial target pattern presented to one finger may be interfered with by the presentation of a second pattern to either the same or a second finger in both forward and backwards masking paradigms.

Evans observed the strongest masking effects at target durations under 100ms (Evans 1987). Both Tan (Tan 2003) and Evans et al. (Craig 1987) found that degree of masking was influenced by the complexity of the stimuli employed; participants were able to identify simpler spatial patterns more accurately. Tan used long complex stimuli and longer SOA's (>125ms) in order to accommodate low-frequency spectral content, and observed lower and less consistent masking effects. However, Tan's study also showed that percent correct scores were highest with the simplest target patterns (those that contained one spectral component).

Di Lollo and Enns have shown an application of another form of masking for visual stimuli, called common-onset or object substitution masking (Enns 1997), where the masking stimulus is presented simultaneously with a clearly visible target stimulus but the surrounding masker remains after the target stimulus has been removed. In vision, this form of masking is attributed to interruption masking which occurs when the mask appears before the target has been fully processed and represents a competition for higher level processes involved in object recognition. The term “object substitution” is used to describe the latter category because the mask appears to do more than interrupt the perceptual process and instead seems to become the new focus of object recognition mechanisms.

Di Lollo and Enns offer a theory of how common onset masking works for vision (Di Lollo 2000; Enns 2000): they suggest that object substitution occurs whenever there is a mismatch between the re-entrant visual representation (in their experiments, the participant’s representation of the target) and the ongoing lower-level activity produced by current sensory input (the persistent masker). In the case of vibrotactile stimuli applied to two fingers, the re-entrant representation theory would play out as follows. Initially, two signals (one from each stimulus) are sent through the nervous system to the homunculus in the somatosensory cortex, where a representation of the skin and other senses is stored. The prefrontal cortex requests a re-entrant confirmation of one of the response hypotheses (finger 1, 2 or both) from the homunculus. By this time, the stimulation is present in only one finger and this mismatching information is transferred back to the prefrontal cortex. Using a similar form of common-onset masking in vision, researchers have been able to effectively mask otherwise clearly visible stimuli (Enns 1997; Di Lollo 2000; Enns 2000).

4.5 Objectives and Overview

The goal of our research was to investigate the masking characteristics (backwards and common-onset) of simple vibrotactile stimuli presented to the fingertips using commercially available, relatively inexpensive transducers and stimuli presented at standardized levels, with the longer-term goal of integrating this type of transducer and stimuli into existing and new interfaces for tactile communication, e.g., in mobile devices. Simple yet information-rich stimuli can be useful for communicating navigational cues and event notification signals in these devices (Chapter 3 and Chapter 5). However, it is imperative to first understand how these stimuli interact with one another when presented in tight spatial and temporal proximity. Furthermore, our larger goal of non-intrusive threshold or sub-threshold level communication requires optimizing methods for deliberately masking these stimuli effectively and consistently, yet without recourse to lengthy and sometimes complex individualization processes, . This experiment was designed with the latter purpose in mind.

To the best of our knowledge, we are the first to investigate common-onset masking for fixed amplitude vibrotactile signals presented to separate contra-lateral loci and the first to present a measure of participant certainty in recognizing the test stimuli. The results obtained here have immediate practical applications for the design of tactile interfaces, and also improve our understanding of the underlying perceptual processes involved when decoding these simple vibrotactile signals.

4.6 Method

4.6.1 Approach

In this experiment, we compared the degree of masking produced in backwards (BWM) and common-onset (COM) masking methods relative to unmasked signals (CTRL), with simple vibrotactile stimuli delivered to a participant’s finger pads. Two fingers were used (middle and ring), in order to support testing of the COM method using single-frequency targets and maskers. Masking method was manipulated with the presence or absence of a fixed SOA (duration determined in pilot tests), and we considered both performance and subjective response confidence in a three-alternative identification task (signal present on left, right or both fingers).

We designed our stimulus sets to balance the goals of (1) maximizing the effects of masking, (2) producing target stimuli capable of simple information transfer (such as navigation cues), and (3) producing maskers that are minimally intrusive. Given our goal to re-use these signals while we determine the most effective masking methods, we conservatively chose simple (single-frequency sinusoid) stimuli as being generally hardest to mask (Tan 2003), and then (to minimize intrusiveness) in pilot studies roughly identified minimum effective durations of both target and masker as well as appropriate SOA values. Finally, we chose the middle and index fingers after pilot studies identified them as having similar levels of sensitivity (Sherrick 1953) for the chosen stimuli. We employed repeated measures of a target identification task, with both target and masker presented at the same fixed amplitude to both fingers and for every participant. Our choice of method and our signal design reflects our intent to use these signals in later studies to investigate higher level perceptual processes: fixed-level standardized stimuli will allow us to investigate the effectiveness of these information-bearing stimuli as well as how perception of these stimuli relates to confidence levels. This differs from previous work in tactile masking where signal intensities are carefully adjusted for every participant and thresholds for stimulus detection are determined through an adaptive procedure (Gescheider 1989).

4.6.2 Apparatus

Our experiments were carried out using a custom display integrating two Audiological Engineering (www.tactaid.com) VBW32 vibrotactile displays (visible in Figure 4-1). These voice-coil-based transducers, which are used commercially in hearing aids, are capable of producing precisely timed (on/off within 2 ms) waveforms at a useful range of frequencies and amplitudes, with maximum efficiency at 250 Hz; and can be driven directly by a computer's sound card. Tactile displays using similar technology are finding their way into commercially available mobile phones, PDA's and GPS navigation units.

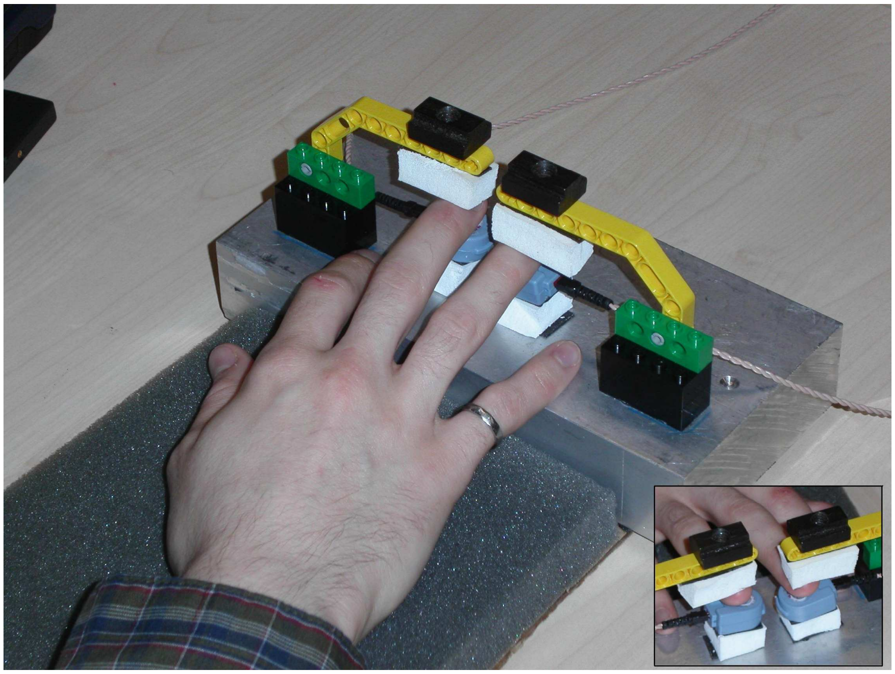


Figure 4-1 Tactile display hardware The participant's middle and ring finger are pressed against the vibrotactile displays using 30 gram weights to maintain a constant pressure. (Same as Figure 3-13)

The design of the apparatus (Figure 4-1) was driven by needs for consistent hand position and finger pressure, as well as vibration isolation to prevent crosstalk between the stimulus sites. It utilized two AE displays mounted on a 3 cm thick aluminum plate and insulated with 1 cm thick latex foam rubber commonly used to mechanically isolate sensitive electronic equipment from vibration. The participant's hand rested on another foam pad which was attached to the aluminum plate; weights mounted on articulated plastic arms held his/her fingers against the transducers with a constant pressure of 30 grams. User pilot tests confirmed that no crosstalk occurred with this arrangement.

The tactile display was interfaced through the sound card in a 2.5 GHz Pentium 4 computer running Windows XP. Participants wore noise-canceling headphones to block any audible artifacts that the device might produce.

4.6.3 Experiment Task and Instructions

In both the pilot and main experiments, we used a three-alternative forced-choice performance task followed by a two-alternative forced-choice subjective task. Participants were read instructions from a script before the beginning of the experiment. They were told that every trial would consist of a single stimulus presentation after which they would be asked to respond with one of three options: stimulus present on middle finger (answering “left”), stimulus present on ring finger (answering “right”) or stimulus felt on both fingers (answering “both”). The participants responded by using the left hand to press a key on the computer keyboard with overlays showing “left”, “right” or “both”. After responding by identifying the stimulus presented, participants were asked to rate the level of confidence in their response by answering “certain” or “uncertain” (again using a keyboard overlay).

4.6.4 Stimuli

We used in-phase sinusoidal stimuli with identical amplitudes presented at 250 Hz. Human tactile sensitivity is highest around 250 Hz (Sherrick 1986). Throughout the experiment, we presented stimulus pairs consisting of various combinations of three durations and two Stimulus Onset Asynchrony (SOA) levels (including zero for COM) to the participant’s middle and/or ring finger.

In a series of pilots, we determined an appropriate range of stimulus duration and levels. In the first of these pilots, we adjusted stimulus amplitude using Parameter Estimation Through Sequential Testing (PEST) (Taylor 1967) so that, when presented randomly to either the middle (referred to hereafter and to participants as “left”) or ring (“right”) finger of the right hand, participants could accurately identify the target finger for stimuli with durations of 10-500ms 95% of the time. This resulted in stimulus amplitude of 10 dB above threshold. We used this fixed-amplitude for both target and masker for every participant in the experiment.

For all our comparisons, we used three stimulus durations: long (masking), short (target) or none (for use in control trials). For this study, we first chose a target (short signal) duration of 30 ms (the shortest reliably perceived when unmasked; 10-50 ms were tested). A masker (long signal) of 300 ms was then chosen for minimum length in effective masking of the chosen target signal (150-500 ms tested). We note that these thresholds are to some extent specific to the apparatus as well as the experimental setup used.

SOA differentiates BWM from COM (zero SOA), which are otherwise identical. We used an SOA of 30ms for BWM; because it demonstrated the most effective masking of the 10-50 ms range explored in pilot studies (masking level began to drop as 50 ms was approached). These SOA values are consistent with prior work on tactile pattern masking (Craig 1989).

Figure 4-2 is a graphical representation of the stimuli pairs used, grouped as control (CTRL), in which only one stimulus was applied, common-onset masking (COM) and backward masking (BWM).

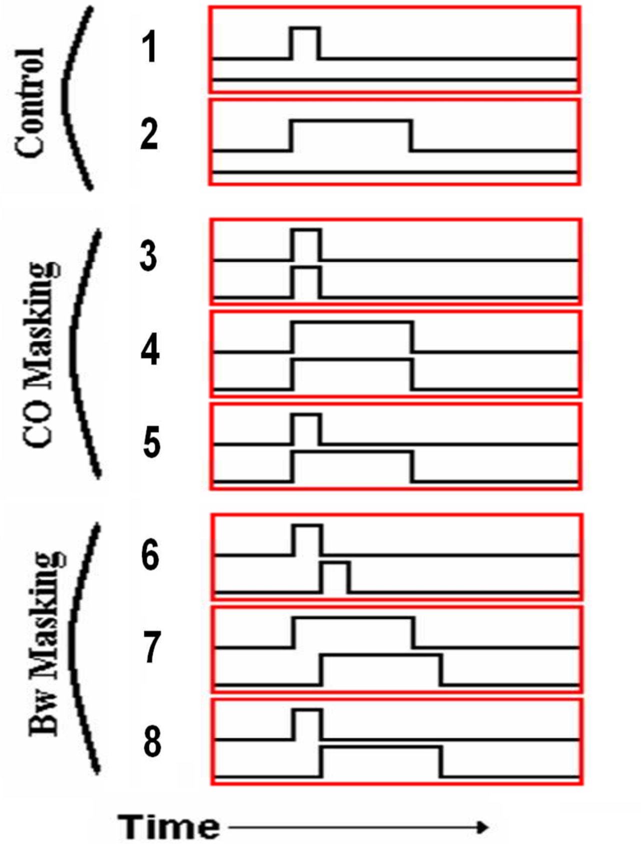


Figure 4-2 Eight stimuli types employed Each box represents the stimuli presented to the middle or "left" (upper) and ring "right" (lower line) finger for a different kind of trial; mirrors of all except symmetric stimulus pairs 3-4 were also used. Short pulses are 30 ms, long pulses are 300 ms, and a flat line indicates no stimulus delivered to that finger. SOA's are either 0 or 30 ms.

4.6.5 Experiment Design

Masking presence and type (the latter dictated by SOA) were the independent variables for this study. As illustrated in Figure 4-2, single-finger trials (CTRL) were used as controls. COM (common onset) trials were those where two stimuli of any length were initiated with zero SOA. BWM (backwards masking) trials were those where two stimuli of any length were initiated with 30 ms SOA.

Participants were instructed to respond “left” or “right” if any stimulus (short or long) was noted on only the left/middle or right/ring finger, respectively; and “both” if stimuli were detected on both fingers. Thus, “both” indicates successful identification of a target stimulus despite presence of a masker.

A total of 18 trials were delivered in a single repetition: 8 types with 6 mirrored (stimuli 1, 2, 5, 6, 7 and 8), and two extra balancing applications of the control pairs one and two and their mirrors. The latter ensured that every repetition had an equal number of “left”, “right” and “both” correct responses and was intended to minimize response bias. For each participant, 10 full repetitions were conducted, and trial order was randomized within repetitions (a different random order for each repetition and participant).

4.7 Results

4.7.1 Participants

Eleven university students, 5 female and 6 male, participated in the experiment. All were 22-27 years of age, right handed and were paid $10 in cash for a 35-minute session. All participants reported normal tactual function.

4.7.2 Identification Performance

We obtained a measure of masking performance in the form of overall rate of correct responses for each of the 8 stimuli configurations tested (from 14 distinct pairs, including mirrors). Each bar in Figure 4-3 was obtained by counting the correct responses out of 10 repetitions for all 11 participants and dividing this number by the total number of presentations of that particular stimulus. The error bars represent the standard error of this average. The graphic pairs below each bar specify the stimulus parameters for the middle (left graphic) and ring (right graphic) fingers: absent, short and long black regions indicate stimulus durations of 0, 30 or 300 ms respectively, and a black region atop a short white region indicates a 30 ms delay. For example, in Figure 4-3, stimulus 7 represents the presentation of a long stimuli applied to the middle finger and a long (300 ms) delayed (30 ms) stimuli applied to the ring finger (backwards masking of the middle finger) along with its mirror, i.e. backwards masking of the ring finger. A correct response to these stimuli would be “both”.

For statistical analysis, control stimuli 1 and 2 were grouped together giving a single factor, one-way 7 level ANOVA. We observed a statistically significant main effect on identification performance of masking type ($F(1, 11)=32.856$, $p < 0.001$). Visual inspection of the mean for each stimulus (see Figure 4-3) indicated seven groupings (control stimuli 1&2 lumped together).

These were tested using pairwise ANOVA comparisons, adjusted for the family of significant differences at the .05 level. The stimuli pairs for which we observed a statistically significant difference are listed in Table 4-1.

Table 4-1 Significant differences in paired comparison for masking type

Stimulus 1	Stimulus 2	P value
1,2 (lumped)	5	p < 0.000
1,2 (lumped)	8	p < 0.001
3	5	p < 0.004
4	5	p < 0.000
4	8	p < 0.007
5	6	p < 0.041
5	7	p < 0.000
5	8	p < 0.000
6	8	p < 0.000
7	8	p < 0.002

Common onset masking (stimulus 5) produced the greatest rate of erroneous responses, i.e. the most effective masking, with the ANOVA comparisons showing significant differences between COM (stimulus 5) and all forms of BWM ($p < 0.041$).

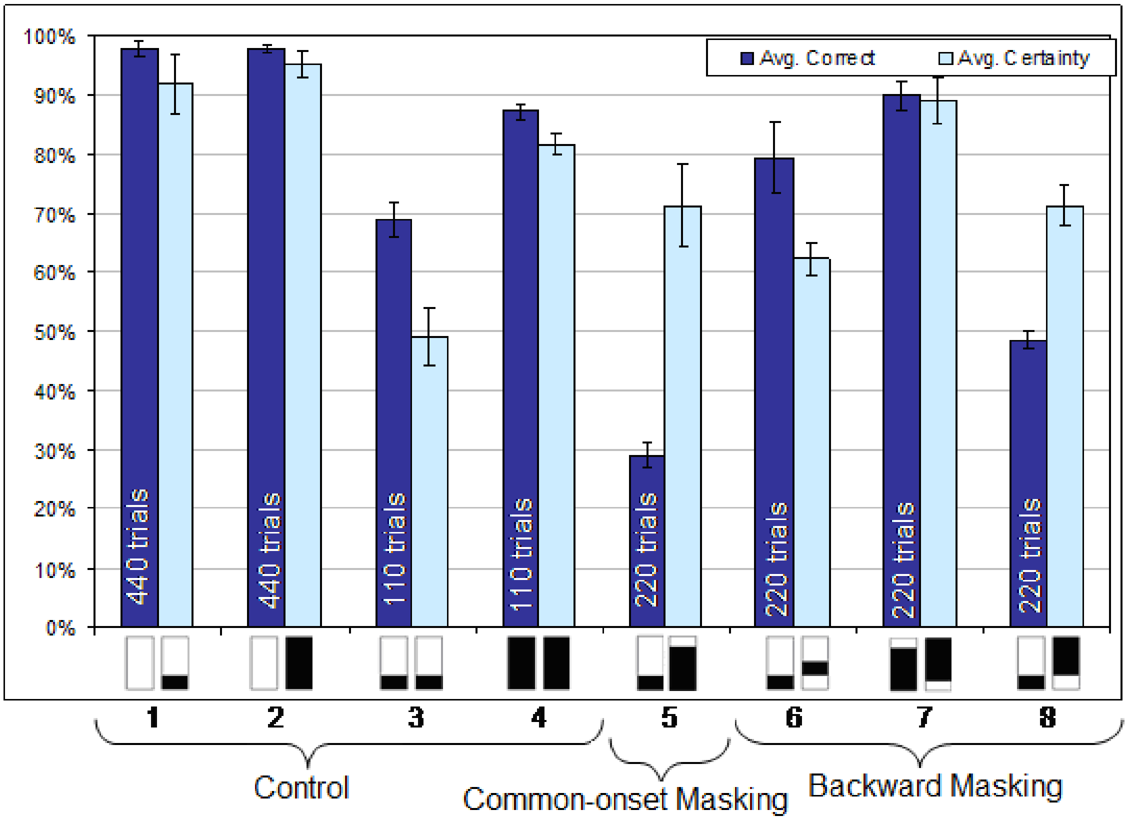


Figure 4-3 Overall percentage identification performance and confidence levels, by stimulus type. The stimulus number matches Figure 4-2. The number of trials represented by each pair of bars is shown on each bar. Lower performance values indicate a stronger masking effect. Stimuli 1, 2, 5, 6, 7 and 8 were mirrored; stimuli 1, 2 and their mirrors were applied twice for overall left/right balance. See Section 4.6.4 for further description of this figure.

We analyzed incorrect responses to determine whether errors were due to overlooking a masked target. We found that nearly 100% of incorrect responses indeed involved missing the short (target) stimulus: in these cases, the response to a masked trial was the longer stimulus, rather than either ‘Both’ (the correct response) or the short stimulus (a different possible incorrect response). Thus together with the observation of near-perfect responses for both long and short un-masked single-finger stimuli, we can conclude that an error is equivalent to a successfully masked target. The types of error and error rates for each of the stimuli are presented in Table 4-2. Visual inspection of individual identification performance data indicated that although there are differences in overall task performance amongst participants as evidenced in the standard deviation bars for Figure 4-3, trends across stimuli pairs are consistent for all participants.

Table 4-2 Error rates and stimulus confusion matrix. This table shows how the errors were distributed for each of the stimuli types (as described in Figure 2) used in the study. Response distribution values indicate the % of participants who provided each of the possible responses; % Correct simply repeats the distribution value that was in fact the correct answer. Most bars in Figure 2 are the average of the two rows of this table which represent mirrors of the same stimuli.

Stimulus Type	Stim #	Middle Finger		Ring Finger		% Correct	Response distribution			Conf Level
		Delay	Duration	Delay	Duration		Middle	Ring	Both	
Control	1	0	0	0	30	98.6%	0%	**99%**	1%	91.8%
	2	0	0	0	300	97.7%	0%	**98%**	2%	94.5%
	1	0	30	0	0	97.3%	**97%**	0%	3%	91.8%
	2	0	300	0	0	98.2%	**98%**	0%	1%	95.9%
Common Onset Masking	3	0	30	0	30	69.1%	25%	5%	**69%**	49.1%
	5	0	30	0	300	21.8%	0%	78%	**22%**	71.8%
	5	0	300	0	30	36.4%	64%	0%	**36%**	70.9%
	4	0	300	0	300	87.3%	10%	3%	**87%**	81.8%
Backward Masking	6	0	30	30	30	75.5%	2%	23%	**75%**	59.1%
	8	0	30	30	300	33.6%	4%	63%	**34%**	70.9%
	7	0	300	30	300	89.1%	5%	6%	**89%**	88.2%
	6	30	30	0	30	83.6%	10%	6%	**84%**	65.5%
	8	30	300	0	30	63.6%	35%	1%	**64%**	71.8%
	7	30	300	0	300	90.9%	5%	5%	**91%**	90.0%

We note that there was a difference in performance within the mirrored stimuli for both COM and BWM, observable in Table 4-2 and most evident for stimuli 5 and 8. However, this difference was not statistically significant ($p<0.199$ for COM and $p<0.307$ for BWM) and we thus continued to group these stimuli in the larger analysis. We believe this was due to a difference in tactile acuity between the middle and ring fingers. Our stimuli were presented at a standardized level and were not balanced for possible differences in perceptual sensitivity between middle and ring fingers, because of our goal of testing worst-case “plug and play” use. From Table 4-2, we observe that while the target stimulus was in general masked more effectively for COM than BWM for the same mirrors of stimuli 5 and 8 (lower percent correct rates), the ring finger was more sensitive (the short target stimuli noted more often with a “both” response for both the BWM as well as the COM stimuli.

4.7.3 Response Confidence

The confidence participants reported in their responses, regardless of actual performance, is shown in Figure 4-3: responses of ‘1’ (confident) are counted and normalized to the total number of trials. Confidence is similar for both types of masking (67.4% and 74.2% for COM and BWM, respectively), and lower than for the unmasked control trials (93.5%); a single-factor, 7-level ANOVA suggests a statistically significant effect of masking type on confidence ($p<0.005$, $F=8.764$). Post-hoc comparisons indicate a difference between both masking methods and the control stimuli (p<0.001), but not between COM and BWM.

There are two notable observations to be made of the confidence results. First, the lowest level of confidence (49.1%) was accorded to the simultaneous presentation of two short signals to both fingers (COM stimulus 3) and substantially lagged actual performance (69.1%). Secondly, confidence levels for COM stimulus 5 and BWM stimulus 8 (in both cases, masking of a short target) were identical (71.4%). Actual performance levels for those stimuli (21.8%, 36.4% for the two mirrors of COM 5 and 33.6%, 63.6% for BWM 8) were much lower than confidence levels in these instances. Taken together, these two observations suggest that while the masker's length (30 vs. 300 ms) did not substantially change performance, it did substantially change confidence in performance. That is, participants had high confidence that only one signal was present, when in fact both were.

4.8 Discussion

4.8.1 Masking Performance

The results obtained indicate that some form of masking is possible under the selected conditions and with the hardware tested. Both of our base stimuli (30 and 300ms) could be accurately (98.0 %) identified when presented in isolation, but identification performance dropped to 31-87% when combined with a masker in some form. A review of individual participants results show that while differences exist, individuals exhibit the same general pattern of performance across stimuli (i.e. a graph like Figure 4-3 has roughly the same shape for every participant, but at slightly different amplitudes).

In their general trend, our results are consistent with previous work in the areas of tactile backward masking and visual common-onset masking, where a decreased level of target identification accuracy is induced by the introduction of a masking stimulus presented contiguously (temporally and spatially) to the target stimulus. However, differences in our methodology and type of stimuli offer new and more fine-grained insights with respect to overall masking results obtained. We will first develop these by comparing our BWM results with previous studies that employed a similar methodology to ours, and then proceed to look closely at new comparisons possible within our own data.

4.8.2 Type of Errors and Stimulus Complexity

Evans (Evans 1987) reported 20-65% overall error rates under backward masking at SOAs in the range of 26–106 ms; of these, 20-30% were attributed to the use of the masker as response (the remainder of errors were random). In contrast, our results for both BWM and COM stimuli show nearly 100% of errors (out of overall error rates of 49-90% for our three BWM stimulus variants) being made by using the masker as the response (Table 4-2). We believe that the increased specificity in type of error which we found (as opposed to differences in overall error rate, which are harder to compare given differences in setup) is due to the simplicity of the stimuli used here (only three possible cases: left, right or both) as well as the masking paradigm employed. In our design, a successfully masked stimulus was one where the participants respond with the masker.

In the same study, Evans reported that for backward masking, percent-correct scores indicated that when stimuli were masked, participants were able to identify simpler spatial patterns more accurately. Similarly, Tan’s (Tan 2003) percent-correct scores were highest with target patterns that contained one spectral component, and lowest with those that were more complex (containing three spectral components). In the experiment reported here, every stimulus was composed of a single spectral component and carried 1.5 bits of information (participants could answer left, right or both). It thus varied both spatially, and temporally in duration and alignment of stimuli between fingers. We therefore hypothesize that the masking effect would be stronger if we were to employ more complex stimuli.

4.8.3 Other Potential Mechanisms: Temporal Integration

Temporal integration is often cited as an explanation for decreased levels of accuracy in stimulus recognition when stimuli are presented closely in time (Craig 1982; Evans 1986). These studies suggest that target identification may be disrupted because the target and non-target form a composite pattern through temporal integration. For example, in vision, if two semicircles (one left and one right) are target and masker respectively (presented one after another), then integration would result in perception of a full circle composite. In this case, the two circle halves are presented in both close temporal and spatial proximity, with each half circle presented to a different location at a different time.

We believe it is unlikely that the incidence of masking observed in our study is a result of temporal integration: this would imply that the target and masker form a composite percept that is the temporal and/or spatial sum of both signals. Our hypothesis is that with our paradigm, temporal integration would work against any of the masking techniques used and would in fact improve stimulus identification accuracy: temporal integration of a short stimulus presented to one finger and a longer stimulus presented to another finger would form a composite percept of two fingers being stimulated (which would be the correct response). Based on findings in vision (where, for example, "_" and "|" are integrated and perceived as "+"), the composite percept is most likely to be that of both fingers being stimulated, i.e. the expected correct response (Parker 1992).

4.8.4 Common Onset Masking as Compared to Backwards Masking

To the best of our knowledge, this is the first study to investigate common-onset masking of vibrotactile signals (COM) using standardized stimulus levels presented to separate but contiguous loci as represented by stimulus pair 5; and one of a few studies which assess any form of vibrotactile COM (Gescheider 1989; Weisenberger 1994). Di Lollo et al. (Di Lollo 2000), investigated this form of masking for visual stimuli (which are not typically standardized by individual); our study was designed to mimic their setup while using vibrotactile stimuli. In vision, the masker takes the form of 4 dots presented simultaneously and surrounding but not touching a target shape. The target can appear in one of 8 possible locations on a screen. The dots remain for a period of time after a brief presentation of the target shape. Participants are unable to identify a target shape within the dots and report only the presence of the dots. In our study, a short vibrotactile stimulus is presented simultaneously with a second stimulus which remains present after the short stimulus has ended. The target can be presented to the middle finger, the ring finger or both loci simultaneously. Participants are unable to perceive the short stimulus and report only the long one.

In the present study, we found that COM provided the most effective masking overall. In particular, COM stimulus 5 produced a participant response accuracy of 29.1% correct as compared to 48.6% for BWM stimulus 8, whereas confidence levels were similar. Our data are consistent with the view that this difference in performance is the result of a combination of two different masking mechanisms similar to those observed in vision. Both COM and BWM can in theory be subject to the backwards (interruption) mechanism of masking (which Enns found to be strongest at 0 < SOA <100ms for vision) (Enns 1997); but COM additionally may be affected by camouflage (noise) masking, which is strongest at SOA's ~ 0 (Turvey 1973; Ganz 1975; Breitmeyer 1984; Gescheider 1989;

Oxenham 2001). These time estimates could plausibly be used as a first crude approximation of durations required for tactile signal processing if we assume that they involve a substantial cognitive component, as is believed to be the case for visual pattern processing.

This theory is substantiated by the observation that COM stimulus pairs 3 and 4, which unlike the other COM stimuli should not be subject to backwards masking as the two stimuli were of the same length, resulted in 69 and 88% correct responses for the 30 and 300ms duration stimuli respectively. This represents about 40 and 12% masking relative to unmasked signals, and is comparable to the masking performance difference of 19.5% observed between stimulus pairs 5 (camouflage plus backwards masking) and 8 (backwards masking alone). Thus, we can posit an additive effect of these two mechanisms.

Another possible contributor to the observed difference between COM and BWM is that our backward-masked stimuli might have generated a salient sensation of motion on the fingers of the participants, due to the (30ms) delay between the onset of the target and masking signals. The presence of this sensation of motion was reported by two of the participants for some of the trials. Previous work has shown an increased sensitivity to perception of motion for vision (Exner 1888) but to the best of our knowledge, this increased sensitivity has not been investigated for the sense of touch. Further research is required to follow up on this possibility.

From a high-level theoretical standpoint, our results might be explained by the proposed existence of the same type of higher level perceptual processes that have been recently explored in vision, i.e. re-entrant processing of the target stimulus (Di Lollo 2000). As described above in Section 4.4, this theory states that perception of an object (or stimulus) is the result of a series of hypothesis-confirmation stages: when a stimulus is first detected, a hypothesis is built as to what the stimulus is. This hypothesis is later confirmed or modified based on subsequent gathering of information. The work by Sillito and Bullier (Sillito 1994; Hupe 1998) provides evidence for the theory of re-entrant processing. This theory provides an attractive explanation of what we have observed. In the case of stimulus 5 (COM, short target and long masker), the initial hypothesis is that there is one stimulus being presented to each finger. This hypothesis is later rejected when the data available at a subsequent time (after the short target has terminated) points to a single stimulus being present. In this way, after an initial stimulation ascends through the perceptual system, an iterative-loop system acts to reduce noise to establish the most plausible perceptual interpretation.

4.8.5 Confidence and its Relation to Masking

To the best of our knowledge, this is the first study to report confidence in responses to masked vibrotactile stimuli. It is important to look at confidence levels in relation to performance levels for the different stimuli utilized. A high level of confidence indicates that the participants were certain about their interpretation of the stimuli being presented. When this is combined with a high error rate (low % correct responses), it implies that masking that went beyond "confusing" participants to convincing them of the error case. Conversely, a low confidence level indicates that the participants felt they were unable to clearly perceive the stimuli, and might be paired with either a high but uncertain actual performance or genuinely confused, poor performance (chance or below chance levels).

Confidence levels for stimuli 1, 2 (no masking) and 4, 7 (2 long stimuli) closely match their performance levels (difference <5%). These stimuli were correctly identified on 80-95% of the trials. Confidence is always slightly lower than the high identification performance for these stimuli, but it is perhaps unjustified to make such precise comparisons of subjective and objective parameters such as these. Instead, we will use these confidence levels and their relation to respective performance levels as baselines for relative comparisons below.

Stimuli 3 and 6 (COM and BWM versions respectively of short stimuli on both fingers) both exhibit a performance / confidence disparity of about 20%, as compared to <5% above. This means that participants were correctly identifying the stimuli (70-80% accuracy) significantly more often than they believed (49-62% confidence levels, with lowest confidence for COM stimulus 3).

Confidence levels are equal (71%) amongst the two stimuli that show the strongest masking effects: 5 (COM) and 8 (BWM), both involving a short target and long masker. For both of these stimuli, confidence levels were considerably higher than performance levels. The COM stimulus was correctly identified only 29% of the time and the BWM stimulus was identified 49% of the time (chance = 33%). This suggests that participants were confident that they had perceived and interpreted the stimuli correctly more often than they did, with disparities of 42% (COM) and 22% (BWM). The higher confidence / performance values as compared to stimulus pairs 3 and 6 is very likely due to the difference in length of masker; 5 and 8 have a long masker which dominates the target more effectively than that of 3 and 6. However, the target in both cases is clearly detectable when alone.

From the standpoint of designing information-bearing signals that will be intentionally masked from conscious perception at periods of high cognitive workload, we can speculate that high confidence-performance disparities are positive: the goal here would be to achieve high correct identification performance, without a conscious awareness of the target information having been received. Conversely, however, uncertainty about perception might contribute to cognitive load at the same time that attentive processing of the unmasked signal has been averted. However, further work is required to substantiate such theories.

4.9 Conclusions

To the best of our knowledge, this is the first report of an investigation of the effects of common onset masking of vibrotactile stimuli presented to separate but contiguous locations using fixed amplitude stimuli, and the first report of participant confidence levels when identifying the masked vibrotactile stimuli. Most of the existing work on tactile masking has focused on testing different forms of forward and backward masking techniques and to the best of our knowledge, none have reported participant confidence.

Specific contributions of the present study include observations that for synthetic vibrotactile signals, (a) common-onset masking (exemplified as simultaneous presentation of a short target with a longer masker) shows the strongest masking effect among the set of masking techniques tested, (b) backward masking presents lower, yet significant, masking levels and (c) confidence levels for participant responses are affected equally by backward and common-onset masking. We propose that the performance differences we have observed between COM and BWM may be explained by an additive effect of two complementary masking mechanisms similar to those which have been observed in visual signal processing. Meanwhile, the pattern of confidence levels we observed suggest that if we intend to deliberately mask stimuli yet maintain confidence levels as high as possible we should consider using common-onset masking.

The stimulus design and masking paradigms presented comprise an innovative method for investigating the masking characteristics of simple vibrotactile patterns. The use of simple potentially information-bearing stimuli has allowed a better understanding of the underlying processes that occur when identifying masked stimuli. Although common-onset masking has been explored for the sense of vision, our work has not only applied similar techniques to the sense of touch but has also opened up the possibility of further investigating possible commonalities between high-level visual and haptic perceptual processes when stimuli are masked.

The experimental method and analysis techniques developed can be used both to determine communication bandwidth when using multiple vibrotactile displays for interactive devices (to maximize communication ratios), and to further study tactile perceptual processes.

5 Haptic Communication under Cognitive Workload

5.1 Introduction and Summary

Modern user interfaces – complex and time-critical – must support users who multi-task. However, we have a poor understanding of how computer-user communication degrades with attentional demand and tradeoffs inherent in introducing new display modalities into high-demand environments. Touch is an under-examined candidate for offloading visual and/or auditory channels for information exchange tasks.

In Chapter 3, we introduced two experiments designed to test the association of meanings to a set of tactile signals. The results obtained in these experiments were promising. They showed that touch could be a good candidate to unload some of the information presented to us by everyday devices. However, a real world implementation of these signals will require investigating perception of these signals in cases where a recipient's other senses are being used at the same time. Furthermore, the design of experiments designed to test the effectiveness of different signal modalities in multitasking environments can be complicated. It is hard to attain both experimental control and realistic scenarios.

In this chapter we introduce two experiments designed to test key aspects of multimodal perception in simulated multi-tasking and high-cognitive-workload scenarios, in an attempt to better understand how it is that touch-based communication will work in everyday situations. Our goals with these experiments were:

(1) Determine the effects of faulty force feedback warning signals as we would expect to find in real world interfaces where signals are derived from sensor information (Experiment 1, Section 5.3). We wanted to know how users would react to force feedback warning signals that appear to be faulty. Results confirm a deleterious effect of false alarms on overall signal efficacy but we could not find evidence of a negative influence of missed warning signals.

(2) Compare the effectiveness of tactile and visual guidance cues in a simulated high workload scenario (Experiment 2, Section 5.4). Our purpose with this experiment was to test tactile signals when used in environments where their recipients are engaged with several things at the same time. Results suggest tactile signals might be a more robust communication channel than visual signals when the recipient is involved in a primarily visual task.

The first of these experiments (Section 5.3) was designed to test for perception of a continuously-varying, model-based haptic signal in a simulated driving environment. Simulating realistic applications requires a good understanding of how to construct more life-like but often less controllable experiment scenarios. In this experiment, we study this problem in the context of advanced automobile interfaces. We employ a throttle pedal with programmable force feedback to indicate potentially undesirable situations in the external environment and to gently but steadily guide the driver away from them. We have found evidence that within this scenario, errors in such a warning signal can have a negative effect on the behavior of the driver within the conditions studied. These experiments required a complex protocol and necessarily permitted a variety of participant tactics. Results show that the presence of false alarms (system warns user of a nonexistent condition) within a set of trials had a deleterious effect on overall haptic signal effectiveness, eliminating the warning signal improvement. On the other hand, misses (system does not report an existent condition), had no such influence. Post-experiment analysis revealed that very subtle variations in participant instruction produced large differences in tactics and consequent experiment outcome.

The second experiment (Section 5.4) investigates multimodal perception in a simulated high-workload scenario. In this study we compared salience-calibrated tactile, visual and multimodal navigation cues during a driving-like task, and examined the effectiveness and intrusiveness of the navigation signals while varying cognitive workload and masking of task cues. We found that participants continued to utilize tactile navigation signals under high workload, but their usage of visual and reinforced multimodal navigation cues degraded; further, the reinforced cues under high cognitive workload disrupted the visual primary task. While multimodal cue reinforcement is generally considered a positive interface design practice, these results demonstrate a different view: dual-modality cues can cross a distraction threshold in high-workload environments and lead to overall performance degradation. Conversely, our findings also indicate that tactile signals can be a robust, intuitive and non-intrusive way to communicate information to a user performing a visual primary task.

For clarity, the related work section in this chapter is divided and presented along with each of the two experiments reported here.

5.2 Contributions

- Designed and developed experiments intended to ask participants to divert their attention into several and sometimes competing tasks.
- Found negative effects on warning signal performance in the presence of false alarms in a study of force feedback warning signal reliability in a simulated driving scenario.
- Implemented methods to create controlled levels of cognitive workload.
- Found a possible negative effect of multimodal reinforced cues when used in high cognitive workload scenarios.

5.3 Impact of Haptic Warning Signal Reliability in a Time-and-Safety-Critical Task

The work presented in this section has been published in the 12th Annual Symposium on Haptic Interfaces for Virtual Environments and Teleoperator Systems, IEEE-VR2004, Chicago, USA (Enriquez 2004).

5.3.1 Introduction

In Chapter 3, we presented two experiments intended to test tactile signals as a means to communicate abstract information. These experiments were conducted in a laboratory scenario, where participants were given clear instructions on their task and were not required to perform any other activities at the time of the experiment.

As we integrate haptic feedback into sophisticated real applications, we need to better understand how to conduct more life-like – and often less controllable – experimental scenarios. These experiment scenarios generally entail a realistic context and/or relatively complex tasks; and in the attempt to generate context, may invite involved or deliberately imprecise instructions. In such a situation, how can we get the participants to focus on the desired aspects of the experiment without revealing to them critical experiment information? How should we instruct participants so as to produce a desired performance tactic when tasks are complex and often cannot be clearly explained for experimental purposes? Finally, how can we draw strong conclusions from performance and response data collected in a deliberately uncontrolled environment?

With this experiment, we examine these questions in the context of advanced automobile interfaces. Our paradigm employs programmable force feedback in the primary driving controls (here, the throttle pedal) to indicate potentially undesirable situations in the external environment and to gently but steadily guide the driver away from them.

For our experiments we consider a scenario that presupposes the existence of “drive by wire” automotive throttle control systems, whereby a pedal position sensor and electronic signal replace the traditional all-mechanical linkage from pedal to engine control module. These systems have begun to appear in the last several years for their virtue of improving fuel efficiency and throttle response. However, their existence incidentally affords a redefinition of how the primary controls feel to the driver and further allow the use of a newly-bidirectional channel to deliver new kinds of information in a new format. Given the critical nature of the driving task and in particular of the role played by the throttle and its feel, it is essential that such new interfaces be well designed. The experiments described here address one aspect of this larger problem: driver behavior when information delivered through this new channel is not completely reliable (e.g. an automotive, forward-looking radar sensor might fail to detect the presence of an obstacle in the path of the vehicle.)

5.3.2 Related Work

5.3.2.1 Haptic Constraints

The concept of using programmed force feedback to subtly inform and/or modify user behavior in a real time manipulation task is not new. Rosenberg first proposed using haptic virtual fixtures to constrain user motion through a space in ways analogous to the use of a ruler or compass in mechanical drafting (Rosenberg 1994). More recently, others have employed dynamically and sometimes automatically generated fixtures in applications such as surgical teleoperation; for example Payandeh & Stanisic demonstrated improvement in terms of performance, workload and task training time (Payandeh 2002), and Okamura's group has been optimizing characteristics of the haptic signal itself (Nolin 2003).

Most relevant to the work presented here are the following studies using force feedback to guide or constrain using haptic steering wheels. Steele & Gillespie looked at shared control of steering in a car and noted improved tracking performance and reduced visual demand in a visual tracking task (Steele 2001). More recently, Griffiths and Gillespie (Griffiths 2004) showed that a fixtures-based assistance improved lane keeping and reduced visual demand in a driving task. Forsyth and MacLean (Forsyth 2006) used fixtures based on a system-known path to be followed in a steering task, and addressed problems of instability in high-bandwidth following by constituting the control signal from a look-ahead prediction algorithm.

These and other studies consistently document the potential for appropriately displayed haptic feedback to provide information that enhances performance and reduces user effort in demanding real time tasks. However, we have yet to consider characteristics of the information used to generate the informative signals; in particular its reliability, and how this may play out in the ability of the user to utilize that information.

5.3.2.2 Virtual Models of Physical Systems

Haptically portrayed models of familiar physical systems can make a haptic aid more intuitive (Snibbe 2001). We hypothesize that use of this approach in a driving situation can influence driver behavior towards a more conservative driving pattern in a subtle and non-irritating way, and potentially without the driver’s explicit attention or awareness. However, we do not know how a user might respond to a haptic signal based on a virtual physical model when the signal cannot be guaranteed to be reliable.

5.3.2.3 Warning Signal Effect & Signal Reliability

Tipper (Tipper 2003) found a classic and robust warning signal effect (Mowrer 1940; Bertelson 1967) in response time when participants were given a haptic warning (a buzz on the hand) 100-1000 msec before receiving a visual stimulus to which they were to respond by pushing a computer key: response time improved in proportion to the advance warning given. Signal reliability has been shown to play a role in the way people process information contained within the signal (Tanner 1954; Green 1966). Based on this, Tipper proceeded to manipulate the reliability of the warning signal by corrupting it successively with 25% false negatives (“misses” or MI), 25% false positives (“false alarms” or FA) or a mix of these two types of errors. She found that the presence of FAs within a set of trials eliminated the warning signal improvement in response time even for those trials where the signal was present (“valid trials”); MI trials, on the other hand, had no such influence on the valid trials. Mixed errors produced the same negative effect as purely FA errors.

We argue that the reason for this “bleeding” of a deleterious effect on subject behavior when a warning signal is subject to false positives is due to the subject’s destroyed trust (whether conscious or not) in the reliability of the warning signal. This data suggests that false negatives do not similarly destroy trust. However, it was collected in a highly abstract context.

5.3.2.4 Sensor Reliability and Potential Impact on User Trust

It is generally very difficult to guarantee a technical system’s perfect performance. In our situation of an intelligent system that warns a user of a critical situation, imperfect performance might occur when the system finds a critical situation when one does not exist (FA’s), or fails to find one when it does exist (MI’s). Further, a class of “perceptual” errors can occur through no fault of the technology: if the system finds and signals a warning for any situation that truly exists but which the user never perceives, the user may erroneously believe that system has delivered a false positive. In terms of impact on the user's trust of the system, this "perceptual" false alarm is indistinguishable from a "technical" false alarm. A user interface that takes input from sensors must therefore accommodate potential imperfections in the source input by understanding how the user will react to various amounts and types of sensor inconsistency or unreliability.

5.3.3 Driving Simulator

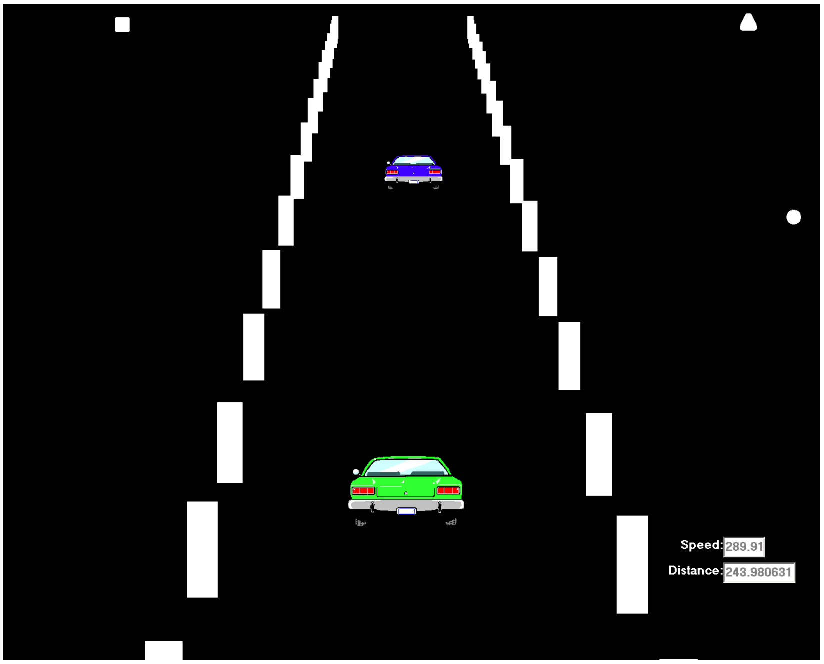
We wished to (a) establish whether use of a warning signal displayed as a haptic model of a familiar physical system can modify driving behavior in its perfect (reliable) form, and (b) explore how the same signal when unreliable might impact the driver's ability or willingness to make effective use of this information. We therefore developed a graphically simple driving simulator that reproduced several key aspects of a complex driving environment. A visual tracking task was executed via a force feedback pedal that superimposed an Active Pedal representation on the usual pedal spring force (Figure 5-1).

Figure 5-1 Setup. “Driver” at simulator with force feedback pedal.

For this analysis, the physical system we modeled is that of a spring attached to the front of the car with a rest length equal to a nominal following gap behind the car ahead. When the driven car approaches the leading car, the driver feels the “compression” of this spring as an additional resistance through the throttle pedal: he must push a little harder to maintain the same gap. The smaller the gap between cars, the greater this extra push. This is, of course, only one of the possibilities for augmenting the information presented by the driving interface.

5.3.3.1 Graphical Interface

The graphical interface (Figure 5-2) portrayed two cars on a road. The participant controlled the speed of the following car - which is stationary in the reference frame of the screen - using the pedal. The motion of the participant's car was conveyed by the rate at which road posts move toward the bottom of the screen. The speed of the (upper) lead car and ultimately its distance from the bottom car varied according to a pseudo-randomized control algorithm outlined below.



Speed: 289.91
Distance: 243.980631

Figure 5-2 The main screen. Participant controls the speed of the car on the bottom of the screen using the FF pedal in tandem with a "brake".

5.3.3.2 Workload Task: “Road Signs”

People often perform more than one task while driving; adjusting the radio, talking on the phone and using navigation systems absorb driver attention. In order to test the pedal force feedback in a multitasking environment, we provided an additional workload task: shapes (Figure 5-3) appeared at random locations and time intervals on the road margins and slowly faded away. Participants were asked to press the <ENTER> key when a particular shape (the triangle) was presented.

Figure 5-3 Workload shapes which appear at random locations on the “road” margins Their size relative to other graphical features can be seen in the previous figure.

The effort required for this task was adjusted during pilot experiments by varying the size, number, frequency and distinctiveness of shapes until pilot participants felt the workload task was “reasonably challenging” and we felt it was competing substantially for attention with their primary driving task. In these experiments, the workload task appeared about 2-16 times per minute.

5.3.3.3 Speed Control: FF Pedal and Brake

The simulator included a force feedback pedal with a position sensor interfaced through an IO board for force display and throttle input; the participants also used the keyboard space bar as a “brake”. The pedal position input determined the acceleration of the participant-controlled car in the simulation. In the following, "lead" refers to leading vehicle, "car" refers to the participant’s vehicle, "TTC" is Time to Contact (time until a collision should current relative velocities be maintained) and "THW" is Time Headway between the two. We employ a desired THW of two seconds, i.e. the car crosses a point on the road two seconds after the lead. *Xcar*, *Xlead* and *Xrel* refer to the position of the participant’s car, the lead car and the distance between them. In a similar manner, *Vcar*, *Vlead* and *Vrel* refer to the cars’ velocity.

$$X_{rel} = X_{lead} - X_{car} \qquad \text{An increasing } X_{rel} \text{ is good.} \tag{3}$$

$$V_{rel} = V_{lead} - V_{car} \qquad \text{An increasing } V_{rel} \text{ is good.} \tag{4}$$

$$THW = \frac{X_{rel}}{V_{car}} \qquad TTC_{control} = \frac{X_{rel}}{-V_{rel}} \tag{5}$$

$$THW_{desired} = 2 \text{ Desired THW.} \tag{6}$$

$$OutputForcetoPedal = \left[a\left(\frac{1}{THW} - \frac{1}{THW_{desired}} \right) + \frac{b}{TTC_{control}} \right] C \tag{7}$$

The variables *a* and *b* are constants that define the displayed force profile given the distance between the cars and their velocity. We used values that maintain the relation b=-a/15 based on simulated results. *C* is a constant gain used to keep the total pedal force within a comfortable and comparable to a mechanical pedal system range.

5.3.3.4 Following-Car Dynamics

The position of the throttle pedal is used to calculate the “force” applied to the participant’s car; the car’s actual acceleration profile also depends on a wind and road-drag model component as well as on the vehicle’s mass and internal friction. If the brake is not pressed:

$$Acceleration = \frac{F_{throttle} - F_{drag}}{M_{car}} \tag{8}$$

where

$F_{throttle}$ = Force generated by motor = f(pedal position)

F_{drag} = Drag force proportional to car’s speed

M_{car} = Mass of car

If $F_{throttle} < F_{drag}$, the car gradually slows down. If the brake is pressed:

$$Acceleration = \frac{-F_{brake}}{M_{car}} \tag{9}$$

where F_{brake} is a constant brake force.

5.3.3.5 Lead-Car Dynamics

In order to observe the participant’s response to Active Pedal activations in a finite amount of time, we needed the participant to interact with the AP fairly often. We aimed for three activations per minute as an acceptable facsimile of driving on a busy highway. This was achieved by adaptively adjusting the erraticness of the lead car velocity. Our algorithm randomly changed the lead car’s virtual accelerator pedal and brake positions, and its consequently computed velocity, at discrete, randomly determined intervals ranging from 5 to 18 seconds. At all other times, the lead car maintained a constant accelerator and brake setting. Erraticness could be set at four levels from “steady” to “abrupt”. The program evaluated the rate of Active Pedal activations every 30 seconds and adjusted erraticity level as needed.

5.3.4 Experiment Design

Experiment Units: All **trials** shared the same overall structure: the participant "drove" through 20 AP activations or “events”, where each event is delineated by an activation (triggered when THW dropped below two seconds). The trial ended after 20 activations with an approximate duration of 6.7 minutes (three events / minute). In post processing we segmented each trial into these 20 observations, and computed performance metrics independently on each segment. We used seven **trial types** representing 4 variables: workload task (W) present / absent; Active Pedal (force feedback) present or absent (P); and False Alarms (F) and Misses (M) committed by the AP at various frequencies. Table 5-1 identifies 4-letter trial labels.

A **Session** consisted of a practice trial followed by five experiment trials; each of a different type. Every participant completed trial types 0000, W000, 0P00 and WP00 (WL and reliable AP present/absent) in random order, followed by one of type WPF0, WP0M or WPFM (either False Alarms, Misses or both plaguing the AP signal in the presence of the workload task).

Table 5-1 Experiment trial types

Trial Description				**Label**
Work Load	Active **P**edal	**F**alse Alarms	**M**isses	
NO	NO	NO	NO	0000
YES	NO	NO	NO	W000
NO	YES	NO	NO	0P00
YES	YES	NO	NO	WP00
YES	YES	25%	NO	WPF0
YES	YES	NO	25%	WP0M
YES	YES	12.5%	12.5%	WPFM

Training: All participants performed a practice trial where the different combinations of parameters that would be presented in the following 5 trials (AP+- and WL+-) were experienced. However, False Alarms and Misses were not experienced, nor was their possibility mentioned.

Participants: We used 36 participants in three runs of the experiment (12 per run). The participants were between 18-40 years of age, 14 female and 22 male, all with valid driver's licenses and normal vision and motor capability.

Instructions: For all three runs of the experiment, participants were told they were competing in a virtual "driving rally" with scoring on race time, safety errors and performance in the workload task. In the first two runs, instructions were read from a script, by a different individual for each run. For the third run, instructions were conveyed by a video recording of an experimenter relating the same script. At the time, we considered the instructions for all sessions and runs to be effectively identical.

5.3.5 Analysis

Analysis was conducted via Matlab scripts and Visual Basic code created for data segmentation, computation of performance metrics, collation of segment results, statistical comparisons and graphical display. This section describes how several analysis issues were handled.

5.3.5.1 Data Segmentation

To delineate the 20 activation events in each session, we defined a segment to begin as the participant leaves the critical THW zone from the previous segment and continue through to the end of the next critical zone penetration.

5.3.5.2 Performance Metrics

We used three performance metrics to examine the impact of warning signal reliability on driving behavior. These metrics were computed for every segment of every non-practice trial for each participant. For all three, more positive values indicate worse performance (see Figure 5-4).

P_{crit} (Critical Zone Penalty): weighted integral of time spent inside the critical region (THW < its nominal 2-second value) where the AP signal is activated. The closer the driver is to the lead car, the higher the penalty:

$$mP_{crit} = \sum_{k=nEnterZone}^{nLeaveZone} \frac{1}{THW_k} \tag{10}$$

where *nEnterZone* and *nLeaveZone* refer to the time steps during which the critical zone was entered and departed respectively. THW_k is the Time-Headway at that time step; its inverse is larger when the driver is closer to the lead car.

Brake: # of samples in the segment where the “brake” was pressed, multiplied times sample period.

Crashes: # of crashes during a segment (THW = 0).

5.3.5.3 Statistical Comparisons

To determine relative driving performance among the different experiment conditions, we compared distributions of segment performance metrics rather than trial and/or segment mean values. Mean values of metrics like amount of braking or number of crashes exhibit large variance by their nature, and even statistically significant differences may not be very meaningful. Distributions, on the other hand, retain information related to frequency and likelihood of these kinds of events occurring under the different conditions studied.

A **Kolmogorov-Smirnov (KS) test** statistically evaluates the difference between data distributions. The response distributions include all observations of a particular metric for a given set of conditions: the KS test then provides the likelihood that two such distributions are different (Chakravarti 1967). Specifically, KS uses as a test statistic the maximum difference over all x values of the cumulative distributions of the two data sets X_1 and X_2. Mathematically, this can be written as:

$$\text{KS test statistic} = \max(|F_1(x) - F_2(x)|), \tag{11}$$

where $F_1(x)$ is the proportion of X_1 values <= x and $F_2(x)$ is the proportion of X_2 values <= x.

It should be noted that the KS test does not distinguish between differences due to distribution means, shapes or variances; this is acceptable for our purposes since all of these are relevant, and in general the differences we found appeared due to a combination of these factors.

5.3.5.4 Data Normalization

Because we observed substantial between-participant variation and we were most interested in the effect on individuals of varying experimental conditions, we normalized the data on participants by computing the mean of all observations in a given metric for each participant and then removing that mean from the participant observations before comparison. (This meant that negative values were possible for the metrics).

5.3.5.5 Statistical Comparisons

We performed four different statistical comparisons on the described performance metrics, each based on a KS test between two distributions. In order to utilize a 2-distribution test to compare three distributions, we therefore had to carry out three pair-wise comparisons.

1. Effect of Reliable AP and of Workload

Does AP feedback help when reliable? What effect does our workload model have? To measure the effect of AP, trial were lumped as (0000+W000) and (0P00+WP00), then compared in a 2-way test. For workload, the same trials were lumped as (0000+0P00) and (W000+WP00). Each test utilized 36 participants x 4 trials x 20 segments = 2880 observations.

2. No AP vs. Reliable AP vs. 25% Misses AP

What is the impact of Misses on performance? This 3-way test compared trial types W000, WP00 and WP0M for the 12 participants who performed **WP0M** (**4 from each experiment run**). Each of the three component 2-way tests utilized 12 x 2 trials x 20 = 480 observations; 12x3x20=720 observations in all were involved in the three comparisons.

3. No AP vs. Reliable AP vs. 25% False Alarms AP

What is the impact of False Alarms? This 3-way test also utilized 720 observations and compared trial types W000, WP00 and **WPF0** for a second subset of 12 participants.

4. No AP vs. Reliable AP vs. 12.5% M + 12.5% F

What is the impact of mixed False Alarms and Misses? This 3-way test also utilized 720 observations, and compared trial types W000, WP00 and **WPFM** for the final subset of 12 participants.

5.3.6 Results

In Figure 5-4, we see the baseline effects of the reliable AP signal (top half), and of workload (bottom). The Active Pedal signal (as implemented in our simulator) reduced the magnitude of all selected metrics, proving to be a significant aid over the no-AP case. However, our mechanism for imposing workload demonstrated mixed results, hurting performance for one of the metrics (braking), less significantly improving performance for another (P_{crit}) and having no significant effect on the third metric (Crashes) within the conditions studied.

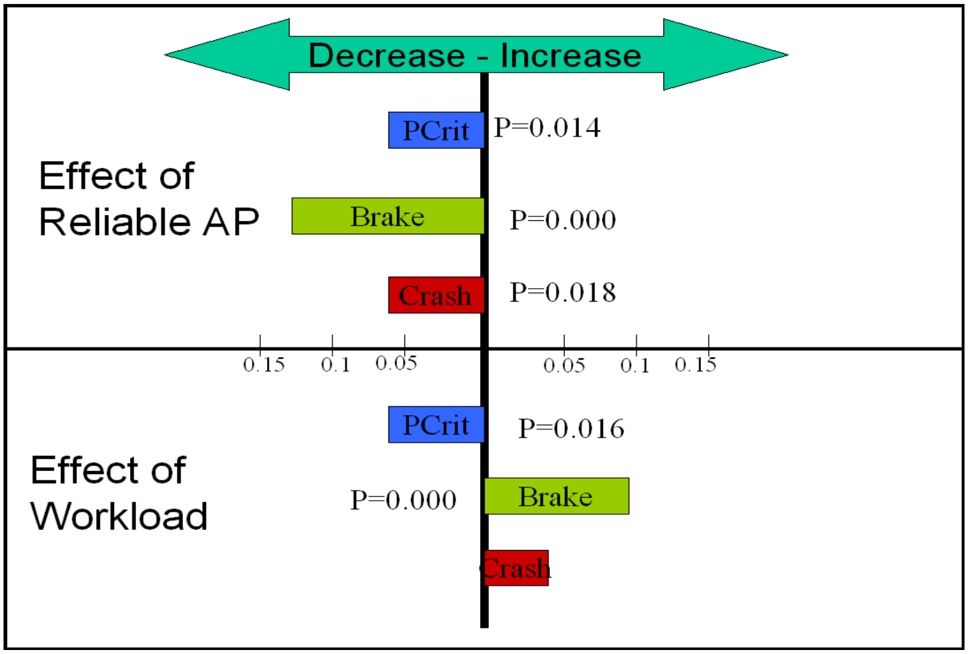


Figure 5-4 results for test 1 show the effect of reliable AP and of workload for the three metrics (all 36 participants). Data from the same 2880 observations have been compared after lumping by presence/absence of AP signal (top) or workload task (bottom). The P-value indicates the statistical significance of the noted difference according to the KS test, whenever P<0.050. The x-value is the dimensionless KS test statistic, i.e. the maximum difference between the two cumulative distribution functions. The x-direction of the arrows indicates whether the change in the performance metric denotes an increment or decrement in the related metric. *For these metrics, a more negative value indicates a more conservative driving pattern.*

Using a similar convention, Figure 5-5 shows the results of the 3-way KS comparisons (composed of three 2-way tests) of the trials that used no AP signal, a reliable AP signal, or a particular type of unreliable signal. In the top graph in Figure 5-5, the signal for this twelve-participant subset was corrupted by Misses. The Brake and Crash metrics do not show a significant alteration in driving behavior when a reliable AP signal was employed (No AP vs. Rel AP). However, P_{crit} shows a significant increase in time spent in the critical zone for reliable-AP trials (large positive blue bar), countering the 36-participant result shown in Figure 5-4 (negative blue bar), and the two other 12-participant results for this comparison for False Alarms and Mixed Error participant subsets below. In the comparison of reliable with Miss-prone AP trials (Rel AP vs. Misses), P_{crit} and Brakes indicate that for these 12 participants, driving style was ***more extreme when the warning signal was reliable than when it was subject to misses***. The final row in this graph (No AP vs. Misses) is consistent with the first two for its only significant metric (Brakes): a miss-prone signal is better than none.

Proceeding in this manner through the remaining two graphs of Figure 5-5, we see in summary that a reliable AP signal usually results in a performance improvement (I.E. a tendency towards more conservative driving) over no AP signal (P_{crit} for the first group is the only exception), and that this result is often significant. False Alarms results in a performance most similar to that of no signal at all, i.e. the presence of false alarms appears to “wipe out” the benefits of the reliable signal for our specific setup and experiment design (Brakes metric). The presence of Mixed Errors results in a behavior intermediate between a reliable signal and none, for all metrics. The twelve FA participants seem to have been less reactive than the other 24, showing little diversity in performance for any metric except Brake (which followed the results of Mixed Errors).

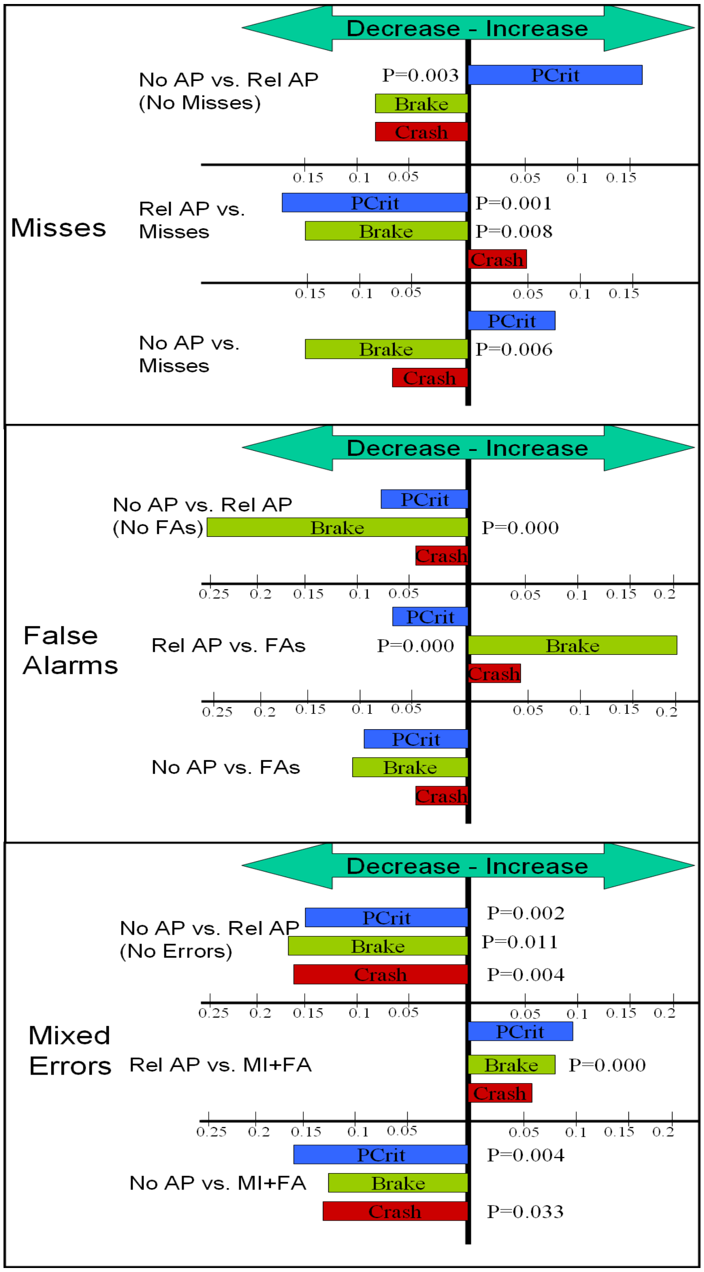


Figure 5-5 KS results are compared two at a time for (in each of the three graphs) three cases: No Active Pedal (AP) signal, a reliable AP signal, and an unreliable AP signal corrupted by one of three categories of errors (Tests 2-4). The KS test P-value is noted when significant at P=.05, and the x-value is the dimensionless KS test statistic. A negative value indicates that the rightmost condition in the pair produced a lower value for the measure specified by the bars. For example, in the first case, the Active Pedal condition resulted in less use of the brake and less crashes but more incursions into the critical zone than the no Active Pedal condition.

Unsurprisingly, Crashes shows the least consistent results among the three metrics. It represents the most extreme error, and the one most likely to be influenced by variations in the participant's accustomed driving style and the current driving mindset.

5.3.7 Impact of Instructions

Considering the complex nature of the experiment and some initial non-intuitive observations, we were prompted to examine our data in greater detail. A close examination of the individual results revealed that each participant's overall driving behavior correlated with that of others in the same run.

Figure 5-6 shows the average values obtained for the Critical Zone Penalty (P_{crit}) for the three separate runs of the experiment (12 participants each; and each run included participants tested with all types of signal error).

Here we can see that for all three runs, the presence of an Active Pedal improves performance overall (by reducing the average P_{crit}). Also evident from this figure is the difference in overall magnitude for the three separate runs of the experiment, regardless of AP signal presence or absence.

The only variable we have been able to identify that could explain this phenomenon is the delivery of participant instructions for each run. A different experimenter administered each of the first two runs of the experiment, reading the same script; becoming suspicious, we instructed 3rd-run participants with the aid of a video recording, using the voice of a third experimenter. We conjecture that intonation, expression and verbal emphasis might have varied enough between the different experimenters to encourage different degrees of driving conservatism among each set of participants.

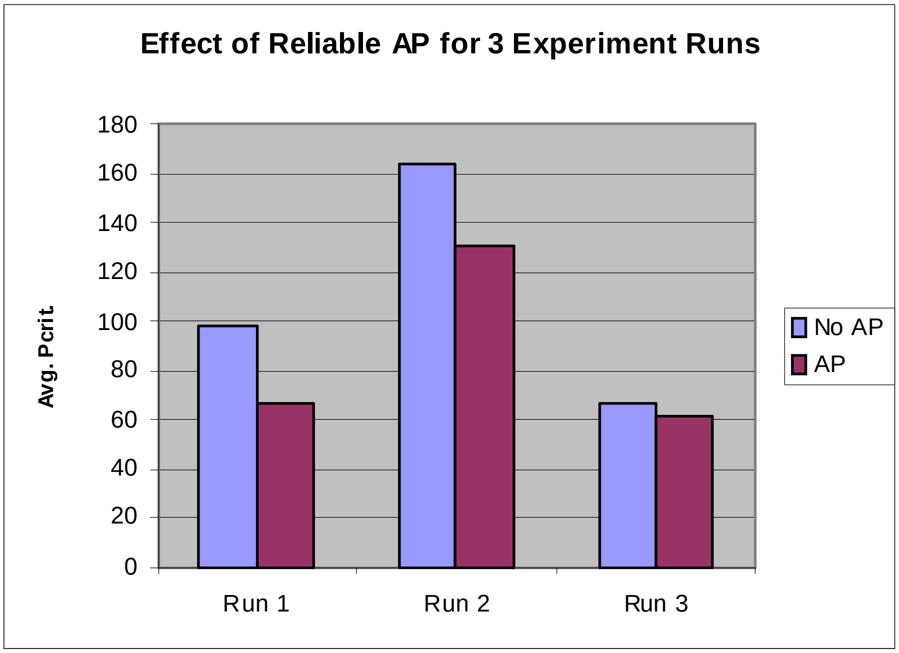


Figure 5-6 P_{crit} metric for the 3 experiment runs (12 participants each) after grouping by presence/absence of the AP signal. Data from all reliable trials (0000+W000 vs. 0P00+WP00) are compared. A more negative value indicates improved performance. Reliable AP always improves performance over no AP, but overall performance varies substantially between the three runs.

5.3.8 Discussion

In summary, we can observe that for this simulator and the tested combinations of experiment conditions:

(i) AP forces (always vs. never present) had a significant impact on all of the three metrics considered (p=0.014, 0.000, 0.018 respectively; 36

participants and 2880 observations). This is a strong result. Our workload task, on the other hand, did not appear to have a consistent effect on these metrics for the conditions tested. Further work will investigate the effect of workload more directly.

(ii) The improvements observed when using a reliable AP are lost when the signal is plagued with False Alarms (25%) in this particular experimental context. Performance with FA's is never significantly different from that with no signal (e.g., a Drive-By-Wire system with no additional force feedback). However, this pattern of degradation is significant only for the Brake metric for this participant subset. Mixed errors (FA+Misses), produced behavior similar to that of just FA's, for those results that are significant, although with a smaller magnitude. This effect of false alarms suggests that this type of error may undermine the improvements gained with the uncorrupted AP and are consistent with those found by Tipper (Tipper 2003), but to our knowledge this is the first time they have been documented in a semi-realistic driving context (i.e. continuous task subject to additional workload tasks).

(iii) Within the conditions studied, the presence of 25% Misses significantly ***improves*** performance over the cases of both a reliable and an absent AP signal, for some metrics (in Figure 5, compare the Rel AP vs. Error AP for Misses relative to those for FA and Mixed errors; the pattern is markedly different). This is perhaps the most surprising result among Tests 2-4.

Why might the presence of missed events (Misses) in the warning signal improve performance over the case of a perfectly reliable AP signal – under these conditions and measured by these metrics? It is believed that a warning signal of any type places individuals in a state of heightened alert and thus decreases reaction times (Bertelson 1967). However, it may also be the case that when individuals come to fully trust a warning signal, they may not feel the need to attend so closely to the task, particularly when a second task is competing for that attention – and this may result in decreased performance, despite the warning signal. Conversely, we theorize that our participants seem to make good use of a signal that is always trustworthy when it does trigger, but cannot be depended on to trigger for every valid target, without abdicating responsibility for finding those other events. This result may not appear in the case of false positive signals because the user may then feel that the signal is never trustworthy. If so, determination of the cognitive or perceptual level at which this distinction is made will require further investigation.

This change in driving behaviour from ‘trusting the system’ to ‘losing trust’ could be gradual or abrupt. However, given the experiment design and the data gathering method used, it is not possible to verify this. The data gathered in regards to the use of the pedal signal is not precise enough. The errors were introduced at random time intervals (within the trials that had errors) and the actual trials were too short (<4 minutes) to lend themselves to this type of analysis. Further work will be required to determine the rate of change in trust level.

(iv) There is a noticeable difference in participant behavior (and thus performance according to our measures) for the 3 separate runs of the experiment. This can be clearly seen in Figure 5-6, where there is an evident difference between the overall values for the P_{crit} metric for the three separate runs of the experiment. The same trend was observed in the two other metrics (not shown due to length restrictions).

We theorize that these differences are a result of variation in the participants' understanding of their assigned task. The three experiment runs were administered by different individuals. A careful postmortem suggested that these individuals inadvertently placed a subtly different emphasis on different aspects of the instructions for each run, thus creating three different driving mindsets that could explain the evidence seen in Figure 5-6: *Slow/Calm* (Experiment 3), *Fast/Aggressive* (Experiment 2) and *Somewhere In Between* (Experiment 1).

The instructions were designed to situate the participants in a "drive conservatively but quickly" mindset. This gave the participant the responsibility of enacting a compromise between two often-conflicting goals, as most of us do in real-life driving on a daily basis. However, a simulator is not a real car and brings no real consequences to aggressive driving. If the experimenter read the instructions with a greater emphasis on "conservative" as opposed to "quickly", the participant's behavior might be different for that particular run of the experiment. This is what seems to have occurred.

5.3.9 Conclusions

The experiment described here confirms previous evidence of a deleterious effect of interspersed false positives (in contrast to the neutral effect of false negatives) on the ability to use a binary haptic warning signal. This work extends these findings to a substantially more sophisticated scenario involving a semi-realistic driving simulation with a pedal-controlled tracking task in the presence of additional workload, with intuitively generated continuous force feedback delivered through the pedal, and for a set of metrics which evaluate “conservative driving”.

This experiment has also introduced new possibilities regarding the potentially positive performance impact of interspersed false negatives in a warning signal for our specific context.

We conclude that participant instruction can strongly influence their attitude when immersed in complex scenarios such as the Active Pedal driving simulator. A post-experiment analysis of the results leads us to conclude that our instructions inadvertently created 3 different kinds of driver mindsets (Slow, Moderate, and Fast). Specifically, we believe that the three experimenters tended to encourage the participants to drive more or less aggressively through both vocal emphasis in reading written instructions, and ad-hoc clarifications. The level of impact of Active Pedal force feedback (AP) varies given these different driver mindsets. At least within the conditions studied, the AP seems to have a stronger influence in moderating driving behavior for people who are driving aggressively.

In general, the strong sensitivity of this type of highly contextualized, stakes-based experiment to experimenter-influenced participant strategy underscores the need for care in experiment design and protocol as well as careful analysis of results to better understand the gathered data.

As implemented by us, the presence / absence of a workload task made no measurable difference in the impact of AP on driving performance. Possible causes for this are:

a) Our WL task was not hard enough to impact on the "automaticity" of the driver's mental state.

b) The principal response to the WL task was to drive less aggressively in general, a condition in which the AP had less effect. Thus WL (in this case) may have changed participant behavior, but independently of the AP.

We emphasize that the conclusions presented here apply only to our proposed haptic feedback model (Active Pedal) and the additional information it might provide to the driver and not to the general Drive-By Wire case.

In future work, we plan to further investigate the subtleties of warning signal reliability for complex scenarios such of that described here, and to innovate on mechanisms for reliable experimentation in these situations.

5.4 Tactile Vs. Visual Guidance in a Maze

User interfaces are becoming more complex and users increasingly rely on them to perform tasks in parallel, often in distracting environments. In driver navigation support systems, for example, the information exchange must be reliable (critical signals should not be missed, even when the driver is busy; whereas it is tolerable or even desirable for non-critical signals to be overlooked at busy times); but the exchange must not interfere with safe driving. This highlights a key design tradeoff in the design of multitasking systems: the need to balance a signal's detectability with its intrusiveness. An intrusive signal is one which diverts attention from other important tasks; we define this trait operationally as a measurable negative effect on the performance of an ongoing primary task, e.g., for visual detection. Clearly, the easiest way to create a signal which is more robust to workload is to also make it more intrusive, but this can have immediate and negative consequences on overall workload.

The haptic (touch) sense may provide a solution for effective communication during driving and other situations typified by high cognitive demand where visual and auditory channels are overloaded or unavailable. Given that the primary tasks of both driving a car and piloting an aircraft currently rely predominantly on visual perception, we believe it may be useful for secondary displays such as navigation aids to communicate some information through haptic signals rather than the visual and audio information displays that are prevalent in modern systems. Indeed, previous research has shown that haptic signals can be successfully used to communicate directional information, for example by giving cues to pilots to help them control the attitude of their aircrafts (Rupert 2000; Van Veen 2003), to orient spatial attention (Tan 2003) and to control music parameters (Verillo 1992).

However, these and other studies did not examine the possible impact of this signaling on other cognitive and user processes. There is evidence that increasing drivers’ cognitive workload in conjunction with a primarily visual task makes it more difficult for them to notice additional visual signals (Patten 2004); but it is not currently clear how signals in a different sensory modality will impact overall cognitive load. Multiple resource theory suggests that at least from a perceptual standpoint, a different modality could offer diminished interference (Wickens 1992). To our knowledge, there have been no studies which examined haptic cue intrusiveness and detectability at the same time, or compare these parameters with impact of cues in other modalities.

In the experiment presented here, we explored the premise that during a predominantly visual task, people’s capacity to detect and respond to haptic navigation signals is less susceptible to the negative effects of visually-derived workload than their ability to detect and respond to visual navigation signals, because of diminished interference due the use of separate channels. In our experiment, we employed an ongoing visual detection task to simulate key aspects of the visual and attentional demands of driving, allowing us to investigate the effectiveness (detectability versus intrusiveness) of communicating navigation information visually, haptically and through both modalities at the same time.

Our results show evidence supporting our basic premise that touch is a robust and usable channel for communicating navigational information during a primarily visual task, and more so than the visual channel. We furthermore found that multimodally reinforced cues (combining both haptic and visual cues) were more intrusive, in highly loaded situations, than either unimodal cue. These results have implications for the design of navigational interfaces that are more informative for a given level of intrusiveness than is the status quo of visual signaling.

In the remainder of Section 5.4, we will discuss related research; our setup, including methodology for simulating workload, perceptually calibrating the visual and haptic signals, and simulation of environmental masking; and our experiment design. Finally, we present and discuss our findings, their implications and proposed future work.

5.4.1 Related Work

When a person is engaged in an ongoing visual task, it may be more effective to communicate supplementary information via a different sensory modality. Here, we briefly highlight the most relevant findings from several areas related to this proposition, including those that shed light on conditions for cross modal signal reinforcement versus enhancement; on the roles of workload and signal reliability in signal detection; and in the use of tactile signals to orient spatial attention as we propose to do with this experiment.

It is beyond the scope of this experiment to delve into the topic of human vibrotactile sensation and signal detectability. Relevant overviews include (Heller 1991; Verillo 1992; Klatzky 2003; Jones 2008).

5.4.1.1 Cross-modal Interference

When simultaneous signals perceived through the same sensory modality carry different types of information or require different responses, they are more likely to interfere (i.e. the recipient could have difficulty in distinguishing and/or processing them) than if these signals are perceived through separate modalities (Wicken's Multiple Resource Theory (Wickens 1984; Wickens 2002). This makes it at least hypothetically more efficient and/or robust to communicate through a combination of modalities. Alais et al (Alais 2004) further found that each modality has its own attentional resources – there is less interference when two concurrent tasks involve separate modalities (in their case, vision & audition) rather than when both involve the same modality, as long as the two tasks do not direct attention to different spatial locations and thus recreate visual interference.

This question is, however, not completely clear-cut. For example, another study using fMRI observations of two concurrent but unrelated tasks, one cognitive and one sensory, did indicate neural area overlap (Just 2001). That is, there are other places in the processing pipeline where interference can occur.

Since sensory tasks are often closely coupled with their cognitive components, this cannot be ignored.

Correspondingly, the change in communication efficiency that may occur when the two signals are presented through different modalities will likely only become apparent when resources somewhere in the pipeline are heavily used – that is, when workload is high. When cognitive resources are readily available, the results cited above suggest that response to any combination of signals may be limited by physical perception capacity alone.

For the reminder of this section, we will use Wicken’s (Wickens 1984) definition of workload: the demand that is placed on mental resources. In our own experiment design, it will be seen that we manipulated workload by use of a primarily cognitive task, which also engaged a perceptual channel (audition) not used in the other tasks.

5.4.1.2 Cross-modal Enhancement

In the instance of simultaneous multimodal signals containing redundant or reinforcing information, behavioral and neural imaging evidence suggest that cross-modal integration (additive or even multiplicative effects) may be expected. In their classic study 15 years ago, Stein and Meredith trained cats to respond to visual and auditory stimuli, alone and in combination (Stein 1993). They found that low-salience, simultaneous stimuli at the same spatial location increased response rates well beyond what would be predicted by combining the low response rates for the individual stimuli; whereas for higher salience signals, the improvement in response rates was not as dramatic. More recently, Macaluso and Driver (Macaluso 2001) used fMRI to demonstrate that the intraparietal sulcus encodes spatial information for both visual and tactile stimuli, suggesting that similar perceptual processes might be involved for both modalities.

Functionally, Gray and Tan (Gray 2002) have successfully directed visual attention with clearly perceptible dynamic tactile cues in a low-workload setting.

5.4.1.3 Impact of Workload on Signal Detection in All Modalities

Performing multiple and/or challenging tasks simultaneously will have an impact on single or multimodal perception and processing. We are most concerned with the severe attentional fragmentation that occurs in mobile, multitasking environments, when the pressure to constantly switch contexts based on interruptions can cause an additional load on attentional and cognitive resources and lead eventually, to a breakdown of fluent interaction (Oulasvirta 2005; Iqbal 2007). These studies found task switching occurring every few seconds in such situations.

5.4.1.3.1 Intrusiveness of Performing Additional Perceptual Tasks during a Visual Task

Previous work has shown that increasing visual or auditory cognitive workload makes it harder to notice visual signals. For example, Patten et al. (Patten 2004) found that talking on a cell phone while driving slows reaction time for identification of visual targets, and that more complex conversations involving the driver have a larger effect than simpler conversations. Miura (Miura 1990) found that in a real driving situation, increasing cognitive demand (measured by increased traffic volume) increases time to notice a small light projected onto the windshield, and decreases the maximum distance between the light and the gaze fixation point for any notice at all. Rantanen and Goldberg (Rantanen 1999) have also shown that increasing cognitive workload reduces the size of the visual field.

5.4.1.3.2 Haptic versus Visual Signaling during a Primary Visual Task

One study has looked for and found differences in detection ability between visual and tactile signals in the presence of workload, when used in conjunction with a visually-dominated primary task. Engstrom et al. found that hit rates in a peripheral cue detection task were adversely affected by difficult driving environments for both visual and tactile cues (Engstrom 2005). This study was designed to determine whether peripheral tactile signals (vibrations applied to the wrist) showed sufficient degradation in response to overall workload to replace peripheral visual signals as a workload metric, during a real driving scenario in which the signals carried simple presence/absence information. Their results show both tactile and visual signal detection rates were sensitive (dropping an average of 12%) to visual and cognitive secondary tasks (counting backwards by sevens or dialing a cellular phone) in a range of driving environments (actual highway, rural or city routes), with the two types of cues subject to slightly different degradation patterns. However, a univariate ANOVA showed no significant effect for stimulus modality (tactile vs. visual) for either hit rate or response time.

5.4.1.4 Signal Detection Theory and Role of Signal Reliability

Behind the larger proposed scheme of providing informative cues to users who are being subjected to taxing workload is an assumption that the cues are indeed informative. In fact, there are two major sources of signal unreliability in the envisioned circumstances: errors on the part of the imperfect intelligent system in generating the signals, and errors on the part of the user in detecting them amongst a complex background of masking sensory and cognitively demanding tasks. Both of these sources can take the form of either “misses” or “false alarms”; and in the widely accepted model proposed by Signal Detection Theory, the relative statistical likelihood of either form is determined by the user’s or intelligent system’s “criterion”, i.e. conservatism in making a choice as to whether a given stimulus is signal or noise (Green 1966; Wickens 2002).

There has been considerable study and theory development of the impact of signal reliability on user’s utilization of those signals, of which we mention only a few interesting examples. Looking at unimodal cue reliability and task complexity, (Maltz 2003) found that cues have a net benefit in complex but not simple tasks; tending to exacerbate unhelpful overload in simple cases. The work presented in Section 5.3 of this Thesis and further published in

(Enriquez 2004) studied the impact on driver acceptance of cues subjected to misses or false alarms in a simulated, multitasking driving scenario. Results suggested that while false alarms caused the driver to reject the cues, signals subject to misses tended to increase driver performance in the cued aspect of behavior over levels found for perfectly reliable signals. Our explanation for this is that false alarms were considered a “waste of time” whereas misses were trustworthy when present but not sufficient alone; therefore, in the latter cases driver performance benefited from both the cues and full driver vigilance.

Most recently, Dixon and Wickens described the last of a series of studies which substantiate a new model of operator responses to misses and false alarms (Dixon 2007). This model is consistent with the hypothesis put forth by (Enriquez 2004), but requires accounting for the extra benefits accruing to “miss”-plagued signals in cost of the user’s “scanning” for that signal instead of performing some other task.

5.4.1.5 Spatial Orienting

There is considerable evidence that haptic feedback is well suited to indicating directional information. Driver and Spence (Driver 2004) summarize psychological support for strong cross-modal links for spatial attention between vision and touch. In more applied contexts, Tan, Lim and Traylor (Tan 2000) created a 3-by-3 tactor array (9 vibrotactile displays) which was used to indicate directions on a user’s back, and van Veen and van Erp (Van Veen 2003) created a vest with 60 embedded tactors, which has been used to help helicopter pilots maintain stability in simulated flights. Rupert (Rupert 2000) showed that tactor arrays on the torso can help pilots maintain an accurate sense of attitude (pitch and roll) in the absence of any visual attitude information. Ho et al. (Ho 2005) found that tactile stimulators speeded attention shifts to the location of critical visual events, regardless of their spatial predictiveness.

5.4.1.6 Summary: Relation of This Experiment to Past Work

For our experiment, we followed (Engstrom 2005) by investigating the effects of cognitive workload on haptic and visual signals. However, our study exhibits several key differences from Engstrom's. As with other more recent (unimodal and low-workload) directional studies, we used signals (salience-equalized in the baseline case) that carried extra information in the form of navigational cues on the premise that additional cognitive processing may be required to utilize such signals than for a pure signal detection task. We employed a simulated environment for greater control, while carefully reproducing key aspects of the real context. Our study specifically focuses on how signals in a different sensory modality might impact overall cognitive load, by examining performance in the primary task as well as detection of peripheral signals.

Finally, we studied the effect of multimodal signals (both reinforced and alternating). The clear importance of signal reliability (actual or perceived) on operator behavior led us to include a “mixed” cue type, where the user does not know which modality to expect the next cue. Reinforced signals are generally regarded as more salient, and we were curious how this would play out in a demanding situation; indeed this is where we found our most interesting results.

5.4.2 Approach

In this experiment, we wished to address three specific research questions. Will sensory and cognitive workload due to a visual search-respond task equally degrade visual and haptic cues? Secondly, are visual and haptic cues equally intrusive on such a task? And third, how do multimodal signals (reinforced and alternating) compare to unimodal signals in detectability and intrusiveness?

5.4.2.1 Experiment Paradigm

To examine these premises, we chose a spatial navigation paradigm of negotiating a graphically-rendered 3-dimensional maze (pictured in Figure 5-7), where the user’s only source of information about “correct” turn choice comes from visual or haptic navigation cues. Computer-aided navigation is a common task increasingly supported by in-vehicle aids which communicate a variety of information to human users; furthermore, navigation is often performed while drivers or pilots are under heavy cognitive workload from traffic and other factors (Baldwin 2003).

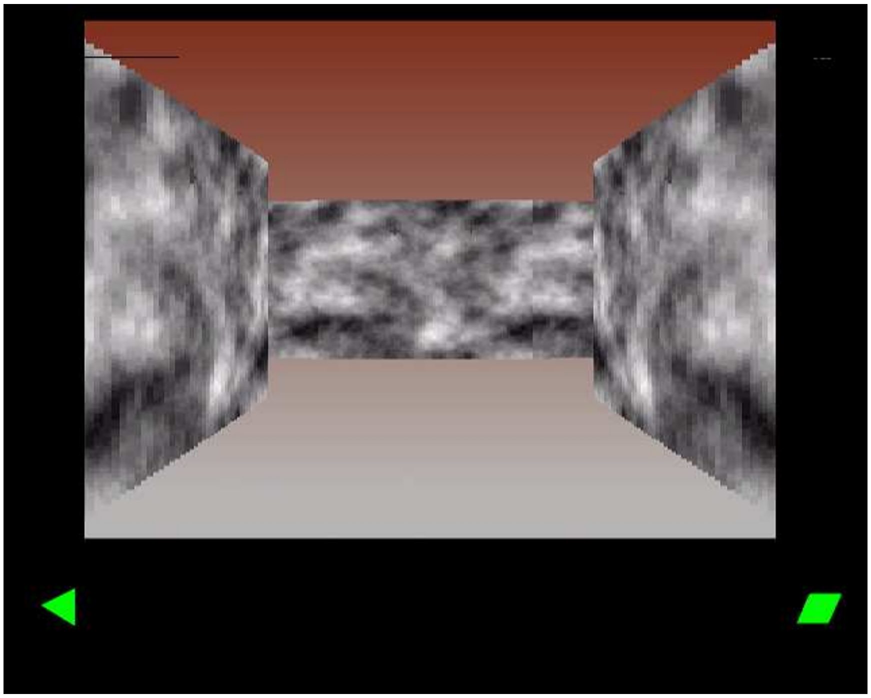


Figure 5-7 Screen shot of the maze, with a visual left-turn cue displayed. The visual direction cues were presented on their respective side of the screen while a different shape of similar salience was presented simultaneously on the opposite side of the screen, such that users had to actually process the two images to identify and interpret them.

5.4.2.2 Tasks: Primary Visual Search with Navigation and Workload Manipulation

The user was given three tasks to be carried out simultaneously. These are described here in overview and with further detail in Section 5.4.3, and depicted in Figure 5-8.

(1) A primary task of continuous visual search. Spatial navigation is generally carried out while one's primary visual attention is focused on the road,

path or the instrument panel, with time and safety-critical implications. We simulated this primary activity with a foundational visual search and response task.

- Regardless of condition, participants were instructed to watch for targets (crosshairs on the maze walls) and respond by immediately stepping on a pedal.
- Decreases in rate of target acquisition were used as an indicator of intrusion of other tasks on this one.

(2) A secondary navigation task, guided by visual and/or haptic cues. We were particularly interested in a change in the user’s ability to utilize these cues when cognitive workload was increased.

- We presented tactile and/or visual directional cues as participants approached each intersection in the maze. Participants responded by pressing left or right buttons to command a turn in the respective direction at the intersection.
- Navigational accuracy was measured by rate of correct turn choices, and participants provided subjective estimates of their percent correct turns following each condition.

(3) An additional cognitive workload task, which was either present or absent in a given trial. We used an auditory/cognitive task to manipulate

cognitive workload over and above that of the primary and secondary tasks: audition was chosen to avoid biasing the other haptic/visual tasks.

- In all conditions, participants listened to spoken passages through a set of headphones. In trials with the added cognitive workload condition, they were asked to count the number of sentences they heard and report the number at the end of the trial block.
- It was not feasible to measure performance due to limitations of our experimental design that did not take into consideration the fact that most participants lost count of the sentences being heard.

5.4.2.3 Simulating Environmental Noise in Navigation Cues

In real scenarios such as everyday driving, visual and haptic cues would be perceived against a background of environmentally derived sensory noise: a driver processes a continuous, visual stream loaded with both relevant and irrelevant stimuli, while feeling road vibration on the steering wheel. We therefore strove to simulate a realistic level of environmental masking, aiming for signals that were noticeable most but not all of the time.

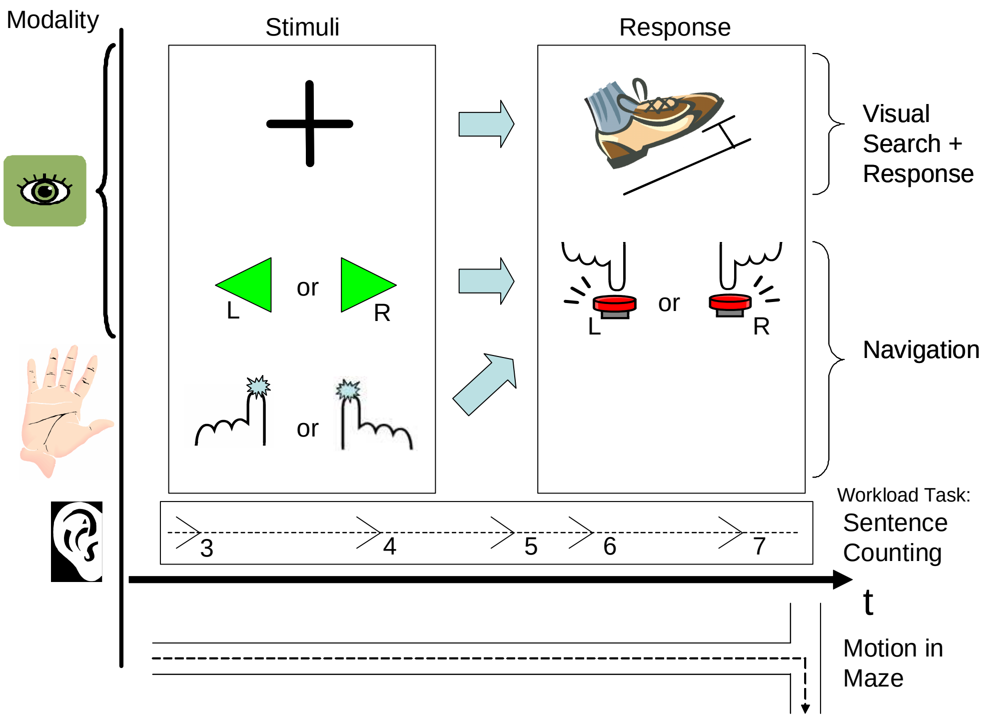


Figure 5-8 Schematic representation of one trial. The visual primary task was to notice the randomly timed appearance of a small crosshair embedded on the maze walls and react to it by pressing on a pedal. As the participant also listened continuously to an audio recording and counted the number of sentences (workload task, present in some conditions), s/he was presented with a tactile (buzz on either the left or right index finger) or visual (green directed triangle) navigation cue indicating the direction to turn at the next intersection, which the participant indicated by pressing the corresponding turn button. In this figure, the position of the icons along the time axis indicates relative times at which these events might occur.

Pilot studies showed large individual differences in visual and tactile stimulus detection ability. In order to make an objective comparison of the relative effects of workload on the two navigation signal types, we devised and carried out a custom calibration procedure to adjust stimulus-to-noise ratios so that each participant would have the same baseline performance level.

5.4.3 Methods & Materials

In this section, we describe experiment design, the setup and simulated environment task specifics and the dependent measures utilized.

5.4.3.1 Design and Factor Manipulations

We used a 4x2 factor, within-participants design, with four multimodal variants in navigation cue and two levels of cognitive workload (Table 5-2).

Factor 1: Navigation Cue Type. In conditions employing the haptic-only signal type (H), all of the navigation signals during that block of trials were haptic, and in conditions employing the visual-only signal type (V), all navigation signals were visual. We also included a reinforced navigation cue condition (H+V) to investigate the effects of cross-modal integration, where both a haptic and a visual signal were presented simultaneously for every trial in the block. Finally, we included a mixed navigation cue condition (H|V) where either a haptic signal or a visual signal was presented for every trial, to explore the impact of broadening attentional requirements.

Factor 2: Workload. Additional cognitive workload was either present or absent (denoted by the subscripts $_{WL+}$ and $_{WL-}$ respectively).

Table 5-2 Eight experimental conditions

Condition		Description
Haptic-Only (H)	Workload (WL+)	Vibration on left or right finger indicates turning direction
	No Workload (WL-)	
Visual-Only (V)	Workload	Triangle on screen indicates turning direction
	No Workload	
Reinforced (H+V)	Workload	Vibration and triangle both indicate turning direction
	No Workload	
Mixed (H|V)	Workload	Either a vibration OR a triangle indicates turning direction
	No Workload	

5.4.3.2 Apparatus

The experiment apparatus consisted of the physical setup and the simulated maze software.

5.4.3.2.1 Physical setup

Participants were seated at a table with each hand resting comfortably on its own haptic display box (Figure 5-9), with the index finger on the vibrotactile display and the middle finger on the button. The two boxes were fixed to the table approximately shoulder width apart (38 cm). Participants viewed the maze on a 17” LCD monitor placed ~60 cm away, and listened to recordings of spoken passages for the workload task through a set of noise-canceling headphones.

Figure 5-9 Physical setup. Participants placed one hand on each haptic display box and viewed the maze on the computer screen. The index fingers rested on the voice coil transducers (blue, mounted on blue boxes) to receive the haptic navigation cue; the middle fingers used the yellow response buttons. An oscillating motor providing background haptic noise is attached to the inside of each display box (not visible here). Not shown in this picture were two foam pads supporting wrists and forearms.

Two types of vibrotactile displays were used for the experiment (Figure 5-10). The target tactile signals were displayed directly to the two index fingertips using two Audiological Engineering VBW32 skin transducers (www.tactaid.com) driven through the computer's sound card. These voice-coil-based transducers are capable of producing precisely timed (on/off within 2 msec) waveforms at a useful range of frequencies and amplitudes, with maximum efficiency at 250 Hz; they were vibrationally isolated from their mounting box so that sensations would be felt only on the fingertip. Tactile noise was generated with a pair of T.P.C. model FM37E flat coreless vibration motors, attached to the boxes holding the voice coil actuators and felt through the whole hand when active. These pager motors oscillated at approximately 133Hz with a fixed amplitude.

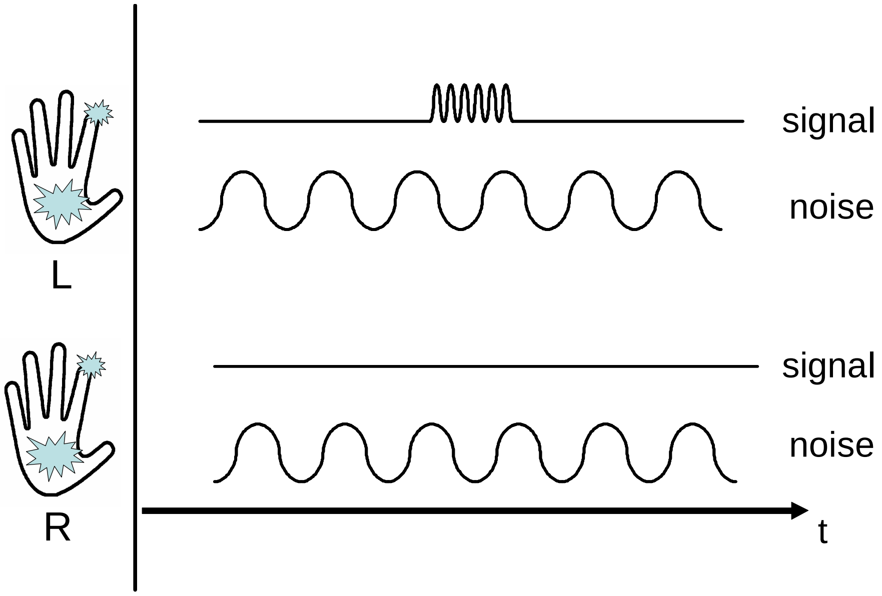


Figure 5-10 Haptic noise (ongoing vibration from pager motor, indicated on palm here but felt by whole hand) and a haptic left-turn signal (occasional burst signal from voice coil display, applied to index finger). Relative signal amplitudes were adjusted for each participant; their representation here is approximate but representative.

Each of the tactile display boxes had a button for indicating chosen turn direction. A foot pedal was used to collect responses to the visual search task (below).

5.4.3.2.2 Maze

Participants advanced at fixed and constant velocity through a virtual maze rendered three-dimensionally with a first-person point of view (Figure 5-7 and Figure 5-11). Every intersection was a “T” where only left or right turns were possible. If the participant reached the intersection wall and did not turn, he stopped. The maze intersections were generated randomly in real time. Participants were not allowed to back up and re-approach any intersection. The length of each corridor segment was varied randomly by 50%.

5.4.3.3 Task Details

The implementation details for each task are provided here.

5.4.3.3.1 Visual Detection Task (Primary Task)

Small crosshairs of a color present in the mottled maze walls appeared randomly on the maze wall at the end of the corridor in 40% of the trials (Figure 5-11). Participants were asked to watch for these targets as their highest priority task, and respond by pressing the foot pedal when they saw one. Two targets were always presented simultaneously in mirror locations on the left and right half of the wall at the end of the corridor, to prevent the turning direction from being influenced by target asymmetry. Pairs of targets could appear in one out of four possible locations. Participants responded to the primary visual detection task by activating a foot pedal with one foot (the same throughout the experiment, and chosen by the participant) upon sighting a crosshair.

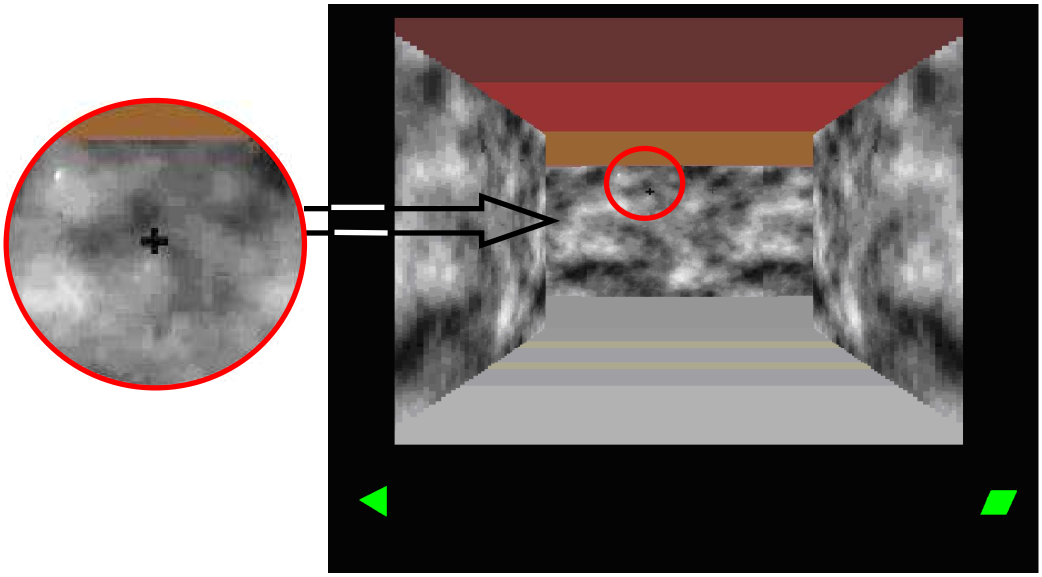


Figure 5-11 Close-up view of one of the crosshairs for the visual target detection task The crosshairs always appeared as a mirrored pair in one of 4 possible configurations on the wall at the end of the corridor.

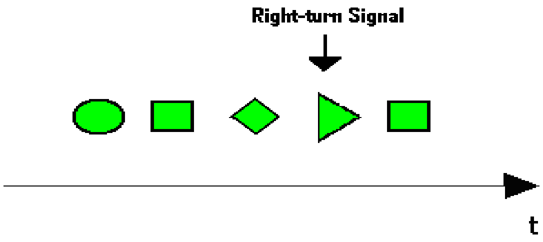


Figure 5-12 Rapid serial visual presentation of the visual navigation cue (fourth position) and its distractors, as they might appear for several moments on the right side of the screen A similar stream would be appearing simultaneously on the left side of the screen but without containing a turn signal.

5.4.3.3.2 Navigation Task (Secondary)

Haptic and/or visual navigation cues indicated the direction to turn at each intersection in the maze. A cue was presented in one or both modalities before each maze intersection; the signal was timed to appear randomly during an interval of 4.0 to 0.5 seconds prior to estimated arrival at the next intersection. Both types of cues were rendered against background noise, to simulate a real context and to require users to attend to the signal’s information content rather than its mere presence.

Visual turning cues consisted of triangles appearing in the left and right corners of the screen below the maze (Figure 5-11): a triangle pointing to the left in the left location indicated an upcoming left turn, and similarly for a right turn. To make the visual signals noisy, we also displayed a variety of randomly shaped but similar-salience distractor images in these locations in rapid succession (Figure 5-12), using the Rapid Serial Visual Presentation technique ((Potter 1976); interval described below under Salience Calibration). This limited the amount of time available to perceive and process each image and distinguish navigation cues from noise. By decreasing the presentation time for each visual image, we increased visual noise and thus made the targets harder to detect, with the aim of realism.

A haptic cue was a 200ms, 250Hz vibration presented to the index finger of the participant’s left or right hand using the high performance voice coil vibrotactile display. A vibration on the left index finger indicated a left turn at the next intersection, and a vibration on the right index finger indicated a right turn. To make the haptic signals noisy, we applied a uniform level of background vibration using the pager motors which were felt by the whole hand (Figure 5-9).

Haptic signal-to-noise ratio was varied by adjusting the more controllable voice coil based vibration amplitude.

Participants turned left or right by pressing a button on the haptic display box with the middle finger of the left or right hand respectively.

Signal Salience Calibration: Stimulus-to-noise ratios for the navigation cues were adjusted independently for each participant such that by the end of the calibration phase and without the workload task, a participant would respond correctly to each type of signal (haptic or visual) delivered alone approximately 80% of the time. Effects of adding the workload task could then be directly compared in terms of resulting performance in both the primary visual detection and navigation tasks.

This calibration (in both modalities) was accomplished using an adaptive method (Parameter Estimation by Sequential Testing, or PEST, developed by Taylor & Creelman (Taylor 1967)) where signals are adjusted successively during a number of trials until the desired level of performance is reached. This resulted in haptic cue amplitudes (sound card output to voice coils) varying from 0.2 to 2 across all participants with a median of 0.5 Volts. The visual shapes presentation interval had a median of 155 ms, ranging from 68 to 563 ms. and the same values were used for the left and right sides for both visual and haptic parameters for any given participant.

5.4.3.3.3 Workload Task

In all conditions, participants listened to spoken passages through a set of headphones. In the added cognitive workload conditions, they were asked to count the number of sentences they heard throughout the trial block. This task required participants to maintain a count in memory, thus creating a continuous workload. At the end of the trial set, participants reported that number and indicated whether or not they had lost count.

The spoken passages were taken from news articles, and read / recorded by the experimenter. The content of the chosen passages was mundane, and topics likely to arouse emotion or substantial interest were avoided. The average sentence length for each passage was between 19 and 22 words (median = 21.3), and no sentences contained less than 10 or more than 35 words. On average, 11-13 sentences were heard in a trial block (i.e. the value of an accurate count).

We did not consider the difference between the actual number of sentences heard and the number counted to be an indication of performance on the workload task: losing count would likely result in guessing, in which case the number reported would not be a meaningful measure of performance.

5.4.3.4 Procedures

A trial consisted of a user experiencing one navigational intersection, and blocks in most cases consisted of 30 such trials run continuously. The average trial duration was 3 seconds, and dependent on the (variable) corridor length since velocity was fixed. Post-training, a 30-trial block typically took about 90 seconds to execute; and at the end of each block of trials, participants typed answers to a series of subjective questions using a standard keyboard.

5.4.3.4.1 Instruction and Training

To learn the procedure, participants first listened to instructions from an audio recording. Participants were instructed to emphasize visual target detection (crosshairs); and further told that all other tasks were secondary and equally important amongst themselves. In light of the strategic nature of the effort required, participant instructions were carefully designed and presented via a recording to avoid confusion and/or experimenter influence on task emphasis (Enriquez 2004).

Participants then practiced navigating the maze. For the first two 30-trial practice blocks (one with visual navigation cues (V) and one with haptic cues (H), order counterbalanced), there was no spoken passage playing, no background vibration (noise) on the haptic navigation cue display boxes, and the shapes on the screen changed slowly. Participants then practiced listening to a passage and counting the number of sentences they heard (Workload Task) with no navigation or visual task. Finally, two inclusive practice blocks were done – one V, the other H – with the background vibrations on, the shapes changing quickly, and counting sentences of a passage playing through the headphones. Thus, one typical training sequence would be

$V \rightarrow H \rightarrow$ Workload task $\rightarrow V_{WL+}$ with noise $\rightarrow H_{WL+}$ with noise,

but for half of the participants, the order of the V and H training blocks was reversed.

5.4.3.4.2 Calibration

The salience levels of the haptic and visual signals were next separately calibrated, so that the participant was able to navigate 80% of the turns correctly. Participants listened to spoken passages during the calibration but were not required to count the sentences in them, as was the case in non-workload experimental blocks. This was done to ensure that any differences observed with the introduction of the sentence counting task were due to the added cognitive workload and not to auditory noise or other perceptual factors.

5.4.3.4.3 Experiment Blocks

Finally, eight blocks of the experiment trials described in Table 5-2 (one of each type) were carried out in random order for each participant, with a short break (~0.5 minute) between each. The (H), (V) and (H+V) blocks had 30 trials (30 intersections) each. The (H|V) blocks consisted of 30 haptic trials and 30 visual trials, for a total of 60 trials. In all cases half of the signals indicated left turns and half indicated right turns.

At the start of each block of trials, a pop-up message box told participants what type of signals they would receive in order to minimize the adaptation period at the beginning of each block. Following each block of trials, a different message box asked participants to estimate (as a percentage) how often they knew which way to turn, and in the cognitive workload conditions, asked them how many sentences they heard and whether they had lost count.

5.4.3.4.4 Debrief

At the end of the experiment, participants were asked to fill out a questionnaire containing multiple choice questions about the H and V navigation signals, with respect to how comfortable they were and the difficulty of attending to them, and open-ended questions where they described the strategies they used.

5.4.3.5 Dependent Measures

For the primary visual target detection task, the percentage of visual targets (crosshairs) detected out of all targets presented during the block was measured, and the mean time between the appearance of the crosshair target and the response was computed. A “successful” response was one in which the crosshair was responded to, by pressing a pedal, before the participant reached the end of the maze corridor on whose end wall it appeared. “False” responses (i.e. pedal presses when no crosshairs had appeared) were ignored (these occurred no more than 10% of the time).

For each experimental condition, performance in the secondary navigation task was measured as the % of correct turns in a given block. Reaction time is not a meaningful measure for this navigation task, since participants were asked not to specify turn direction after receiving a signal until they reached the next intersection. Confidence was indicated by the participant's report (as a percentage value) of how often he or she knew which way to turn for each condition.

5.4.4 Results

Thirteen individuals (seven male, aged 18-75 years with median 28.2, mostly university students) participated in this study. Profile details are reported in Table 5-3:Q6-10. In summary, most had no experience with tactile displays, and the remainder had moderate experience. Eight played video games less than once per month, and the remainder far more than this. Driving experience varied in a roughly normal distribution across the group. Each participant was paid $10 for a 1-hour session.

Table 5-3 Questionnaire responses These questions were asked once at the end of the experiment session.

Q1. Which type of turning signals did you find more comfortable?

(30%) The visual signals | (62%) The tactile signals | (8%) Both types were equally comfortable

Q2. When you were asked to count sentences, which type of turning signals did you find it more difficult to pay attention to?

(54%) The visual signals | (23%) The tactile signals | (23%) Both types were equally difficult to pay attention to

Q3. Were there any occasions where you stopped paying attention to a task altogether, because you were overwhelmed with trying to do too many things at once?

(77%) Yes | (23%) No

Q4. Were there any occasions where you were confused about what you were supposed to do, or you could not remember what you were supposed to do?

(23%) Yes | (77%) No

Q5. Did you use any strategies to try and improve your performance?

(77%) Yes | (23%) No

Q6. How much previous experience have you had with tactile displays?

(62%) None | (38%) Some | (0%) Extensive

Q7. How often do you play, or have you played video games?

(8%)	(54%)	(15%)	(15%)	(8%)	(0%)
Never	A few times / year	A few times / month	A few times / week	Almost every day	Every day

Q8. How often do you drive, or have you driven a car?

(8%)	(15%)	(15%)	(31%)	(15%)	(15%)
Never	A few times / year	A few times / month	A few times / week	Almost every day	Every day

Q9. How much musical training have you had?

(38%) None | (54%) Some | (8%) Extensive

Q10. Is English your first (native) language? (Yes/No)

(54%) Yes | (46%) No

Resulting mean values for all metrics and conditions are listed in Table 5-4. A repeated-measures, two-factor ANOVA was performed for all metrics; its significant results and relevant post hoc tests (using a threshold of p=0.05), are shown in Table 5-5. These results are elaborated on by task in the following sections.

Table 5-4 Means and standard deviations from all metrics The first two sections display high-level pooling of the two independent factors: workload (present or absent) and navigation cue type (generally four levels). The Correct Turns metric is analyzed with five navigation cue levels: the mixed cue condition (H|V) is subdivided into (H|V)H and (H|V)V – correct turns following H or V cues respectively; the other metrics could not be similarly subdivided within blocks. All values reflect the means of 13 blocks (one block of 30 trials (60 trials for un-split (H|V) blocks) per condition per participant).

Condition	**Visual Search Hits (%)** ± STD	**Visual Search Reaction Time (s)** ± STD	**Correct Turns (%)** ± STD	**Confidence (%)** ± STD
Pooled: by navigation cue type				
(H)	**55.4** ± 35.6	**1.6** ± 0.3	**80.5** ± 15.0	**61.4** ± 24.1
(V)	**53.0** ± 36.6	**1.6** ± 0.5	**78.3** ± 10.3	**51.1** ± 21.7
(H+V)	**48.9** ± 34.9	**1.7** ± 0.3	**84.5** ± 11.9	**60.0** ± 22.1
(H|V)	**53.2** ± 36.7	**1.8** ± 0.7	(H|V)$_H$ **73.0** ± 18.1 (H|V)$_V$ **72.4** ± 9.9	**44.1** ± 21.9
Pooled: by workload absent or present				
WL–	**59.2** ± 32.7	**1.8** ± 0.8	**79.7** ± 15.4	**54.2** ± 23.6
WL+	**55.3** ± 34.2	**1.5** ± 0.5	**75.7** ± 13.0	**49.3** ± 23.1
Unpooled: all conditions ordered by workload absent, then present				
(H), WL–	**59.9** ± 33.0	**1.7** ± 0.2	**81.3** ± 17.8	**69.3** ± 19.7
(V), WL–	**57.6** ± 35.5	**1.7** ± 0.4	**82.6** ± 09.0	**56.8** ± 21.0
(H+V), WL–	**62.6** ± 29.2	**1.7** ± 0.3	**90.5** ± 11.8	**68.8** ± 17.6
(H|V), WL–	**56.5** ± 32.8	**2.1** ± 1.1	(H|V)$_H$ **73.6** ± 17.4 (H|V)$_V$ **71.0** ± 11.8	**49.8** ± 15.8
(H), WL+	**60.2** ± 32.8	**1.6** ± 0.4	**79.7** ± 12.2	**62.1** ± 19.0
(V), WL+	**58.9** ± 31.5	**1.5** ± 0.5	**74.1** ± 11.2	**48.8** ± 22.6
(H+V), WL+	**43.3** ± 36.3	**1.7** ± 0.3	**78.5** ± 12.0	**59.7** ± 15.6
(H|V), WL+	**58.8** ± 36.2	**1.6** ± 0.4	(H|V)$_H$ **72.3** ± 19.5 (H|V)$_V$ **73.8** ± 7.7	**45.0** ± 22.6

Table 5-5 Significant ($p<0.05$) primary, interaction and relevant post-hoc effects identified by a repeated-measures ANOVA performed on workload (WL+/-) and navigation cue type (either 4 or 5 levels as described in Table 5-4 caption). The condition with a higher ("better") value is always listed to the left, and post-hoc p-values reflect a Bonferroni correction. There were no significant effects for Visual Search Response Time. The one marginal interaction and its post-hocs are listed as well (T3).

Row	**Type of effect**	**Condition 1 (higher value)**	**Condition 2 (lower value)**	***p***	**Partial η^2**
METRIC: Primary Task – Visual **Search** Hits					
S1	Interaction	Workload (2) × Navigation cue type (4 levels)		**0.015**	0.668
S1$_A$	Post-hoc	(H+V), WL–	(H+V), WL+	**0.007**	
S2	Primary	Between Subjects (n=13)		**0.000**	0.784
METRIC: Navigation Task – Rate of Correct **Turns**					
T1	Primary	WL–	WL+	**0.028**	0.341
T2	Primary	Navigation cue type (5 levels)		**0.039**	0.641
T2$_A$	Post-hoc	(H+V)	(H|V)$_V$	**0.034**	
T3	*Interaction*	*Workload (2) ×Nav cue type - MARGINAL*		*0.070*	*0.584*
T3$_A$	*Post-hoc*	*(V)$_{WL-}$*	*(V)$_{WL+}$*	*0.026*	
T3$_B$	*Post-hoc*	*(H+V)$_{WL-}$*	*(H+V)$_{WL+}$*	*0.022*	
T4	Primary	Between Subjects (n=13)		**0.000**	0.993
METRIC: Navigation Task – **Confidence**					
C1	Primary	WL–	WL+	**0.008**	0.482
C2	Primary	Navigation cue type (4 levels)		**0.008**	0.718
C2$_A$	Post-hoc	(H)	(H|V)	**0.005**	
C2$_B$	Post-hoc	(H+V)	(H|V)	**0.007**	
C3	Primary	Between Subjects (n=13)		**0.000**	0.934

5.4.4.1 Individual Differences

We observed substantial individual differences: in our within-subjects design, a significant between-subjects effect was confirmed for each metric

($p<0.000$ in all cases; see Table 5-5, rows S2, T4 and C3). This is borne out by the size of the standard deviations reported in Table 5-4 and Figures 5-13 to 5-15. They are especially pronounced in the primary task (Visual Search), which is perhaps explainable by the strategic freedom possible in allocating attention to this task (see below).

Considerable analysis of the individual performances did not reveal any unusual distributions or correlations with the profile items reported in Table 5-4 (i.e. video game experience, musical training, driving experience and native language). Our objective and subjective measures appeared to be normally distributed by participant. However, we did in some instances see some apparent variation in metric standard deviation as a function of cue type, primarily in rate of correct cued turns; this will be discussed more in Section 5.4.5.2.1.

5.4.4.2 Primary Task: Visual Search

Participants were instructed to give their primary attention to the visual search task, wherein a crosshair target appeared on the maze's back wall in 40% of trials, counterbalanced by other conditions. We measured success of detecting the crosshair at all, and the time taken to notice it. There was no significant effect of block order, suggesting that learning of the task was not a factor.

5.4.4.2.1 Target Detection Success

The results for visual target (crosshair) detection are illustrated in Figure 5-13 and Table 5-4. Overall average performance (success in noticing the crosshair before reaching end of the maze corridor) was 57.3% (59.2

and 55.2 without and with workload, respectively). This task was unforced, so there is no concept of “chance” performance against which to compare it; the detection was intentionally designed to be difficult to reveal differences, and the baseline performance level achieved is therefore not of interest.

The ANOVA identified a significant interaction between navigation cue type and presence of the workload task (p=0.015; Table 5-4, row S1), but no primary effect of either workload or cue type. As seen in Figure 5-13, the interaction was due largely to a dramatic effect of increased workload on the reinforced (H+V) condition (a 30.8% reduction in crosshair detection, p=0.007; Table 5-4, row S1A). This result is discussed in Section 5.4.5.1.3.

5.4.4.2.2 Target Acquisition Time

The ANOVA found no significant difference in visual target acquisition time (average time between the presentation of the visual targets and the participant’s pedal press response).

5.4.4.3 Secondary Navigation Task

In the navigation task, direction choice was cued by various combinations of haptic and visual signals. We reasoned that a condition-linked decrease in correct turn responses would indicate a corresponding increase in difficulty of noticing and responding to these cues. We also obtained participant self-reports of confidence in their turning performance following each cue-type block. Figure

5-14 shows performance in both of these metrics with workload levels pooled; Figure 5-15 shows the same data broken down by workload.

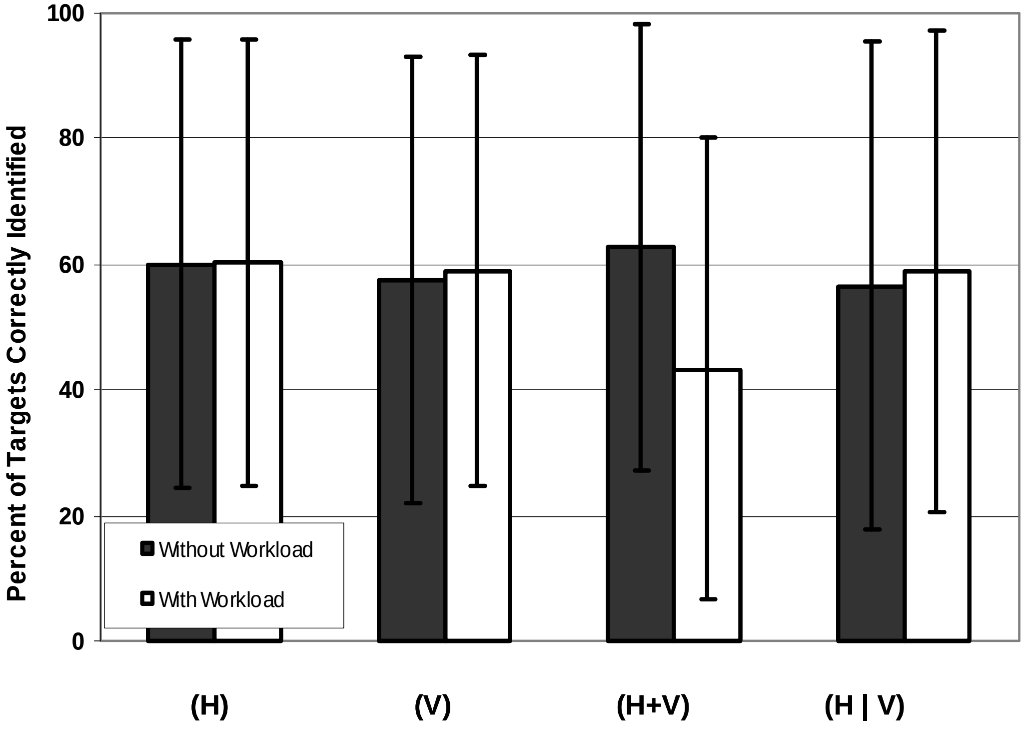


Figure 5-13 Visual target detection performance as a function of navigation signal type and presence of workload. In all figures, error bars represent standard deviations. For this unforced, single-alternative detection task there is no concept of “chance” performance.

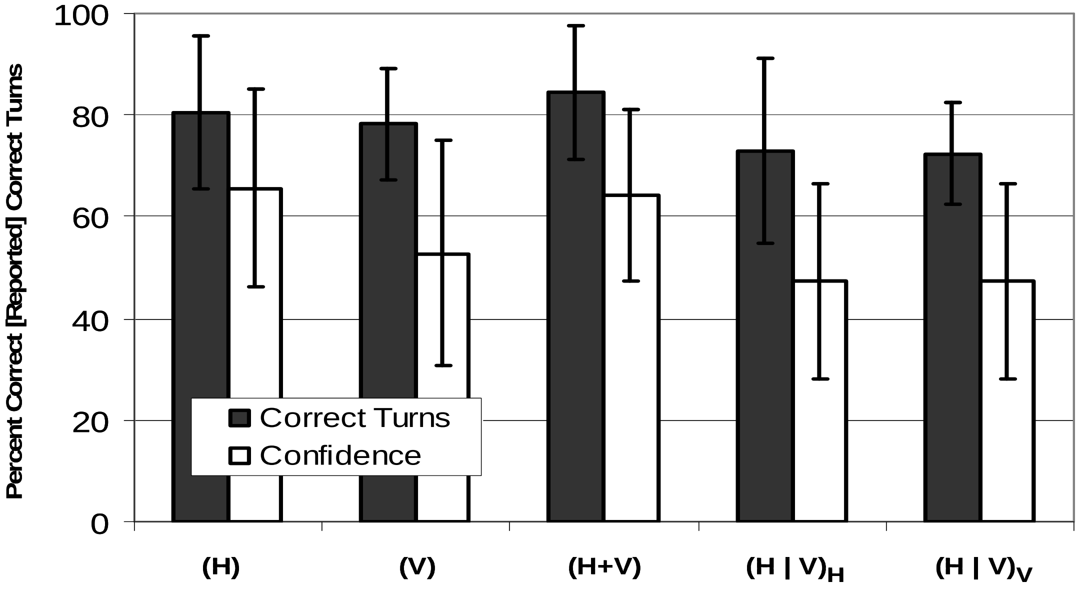


Figure 5-14 Navigation performance and subjective confidence across navigation signal type conditions (workload lumped, i.e. first section of Table 5-2, column 3). Task difficulty was pre-calibrated to deliver a baseline 80% correct turn response to (H)$_{WL-}$ and (V)$_{WL-}$ cues. The last two column pairs subdivide the Correct Turns results for the mixed (H|V) condition, showing the haptic and visual cue subset, respectively; (H|V) Confidence ratings are not subdividable, but are repeated for comparison purposes. A subjective confidence rating of 50% means that the participant believed he/she was making the correct choice 50% of the time in that block, and guessing the other 50%; a 0% rating would mean he/she was guessing all the time.

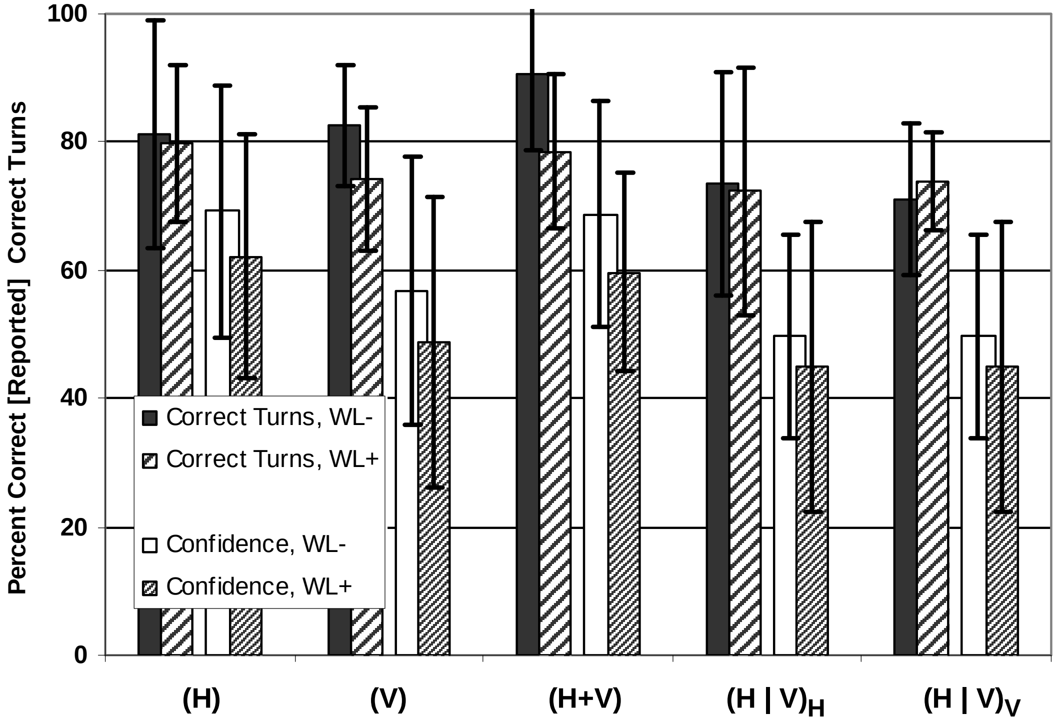


Figure 5-15 Navigation performance (percent correct turns) and confidence as a function of signal type and workload. Figure 5-14 is repeated with its two response variables subdivided by workload (hatched bars indicate high workload conditions).

5.4.4.3.1 Rate of Correct Turns

The signal-to-noise ratios for tactile and visual signals were previously calibrated to produce an 80%-correct workload-free navigation performance level when these navigation signals were delivered uni-modally. Actual performance in the corresponding experimental conditions was indeed close to this: turn responses were correct, on average, for 81.3% and 82.6 of trials in (H)WL– and (V)WL– blocks, respectively. Learning over the 8 experiment blocks likewise appeared discountable as a factor, as performance in these baseline HWL– and

VWL– blocks did not depend on their sequence number. Thus, it appears that our calibration held throughout the experiment trials.

For the metric of Rate of Correct Turns we used 5 rather than 4 cue type conditions, subdividing the mixed condition (H+V) into responses following visual cues (H|V)V and haptic cues (H|V)H ; (H|V) blocks included twice as many trials for this purpose (see Section 5.4.3.4).

ANOVA results confirmed significant main effects of both workload (p=0.028; Table 5-4, row T1) and navigation cue type (p=0.039; row T2); but their interaction was just above the significance threshold (p=0.07, row T3). Within post-hoc analysis of navigation cue type, only the (H+V) versus (H|V)V effect differences are significant at p=0.05 (row T2A). This is visible in Figure 5-15, where the left bars in each group show that reinforced (H+V) cues resulted in the best performance (84.5%) while mixed cues did the worst (72.7% overall). Subdivided, (H|V)V and (H|V)H are similar in mean value, but (H|V)V exhibits a tighter standard deviation.

Workload impact is less clear-cut. Figure 5-15 repeats Figure 5-14 with workload effects introduced (left-most pair in each cue-type grouping are for the Correct Turns metric). Workload does appear to impact the different cue types differentially, with (V) and (V+H) cues taking the heaviest hit (10.3% and 13.3% reductions, respectively). Both of these results are significant in post-hoc analysis (Table 5-5, rows T3A-B), but are left ambiguous by the marginality of the overall ANOVA interaction. In contrast, when purely haptic signals were given, the addition of the workload task did not appear to have any particular effect on people's ability to respond. The (H|V) conditions are also largely unaffected by added workload. These ambiguous results are explained by the large individual variances observed; a larger experiment will be required to resolve them.

Thus, overall we found evidence that while reinforced (H+V) turn cues resulted in the best navigation response performance in the absence of workload, these same reinforced signals also showed the greatest performance drop when workload was added. There were suggestions of a similar workload-linked drop for purely visual signals (V). There are no such suggestions when the signals were solely haptic (H). Mixed turn-cue conditions appeared universally difficult, with lower navigation performance overall (similar to V_{WL+}) which was however relatively insensitive to workload.

5.4.4.3.2 Confidence

After each block, participants reported the percentage of time they were confident they were choosing the correct turn direction (Figure 5-14 and Figure 5-15, right bars in each cue-type grouping; and Table 5-4).

A repeated-measures ANOVA indicated significant main effects of workload (p = 0.008; Table 5-5, row C1) and navigation cue type (p = 0.008; row C2) on participant's estimates of how often they knew which way to turn. Post-hoc analysis shows that the cue type effect is due to differences in turning confidence for (H) and (H+V) cues (highest) relative to mixed (H|V) cues (rows

C2A-B). Consistently with performance results, people reported that they knew which way to turn more often with haptic (H) or reinforced (H+V) signals than with mixed signals (H|V). This confidence differential was not evident (in ratings or verbal comments) with visually indicated turn cues.

The workload-cue type interaction was not significant at p=0.05, and visually this can be seen in Figure 5-15: workload has a similar deleterious effect on confidence ratings similarly for all turn cue types.

5.4.4.3.3 Performance Prediction Based on Confidence Reports

It is interesting to compare reported confidence with actual navigation performance. We can use the heuristic that if the participant's insight into his own accuracy is perfect, actual performance should equal the percentage of responses of which the participant is confident, plus half of the remainder; for which a chance guess should yield 50% correct. This is roughly what we do see – i.e. participants appeared to be fairly accurate in self-assessing their performance. The largest deviations to this were under-predictions in the absence of workload, for V and H+V cue types (by 4.2% and 6.1% respectively).

5.4.4.4 Participant Opinions

A majority of participants found the tactile signals to be the most comfortable; a majority also thought it was more difficult to pay attention to visual signals compared to tactile signals while counting sentences. Their responses confirmed that at least at times, the combined experiment tasks were overwhelming (an intended result); but most indicated that at all times, they understood what they needed to do and how to do it.

A detailed breakdown of questionnaire responses can be found in Table 5-3, and the various questions are discussed as relevant in Section 5.4.5.

5.4.5 Discussion

The goal of this experiment was to compare participants' use of visual, haptic and variously combined signals to communicate information to a user who is engaged in a primary visual task, and to investigate the impact of cognitive workload on performance in both tasks. Understanding impact on users of multimodal signaling is important in terms of providing design heuristics to support multitasking users: such interfaces need to support reliable communication, while avoiding unnecessary distraction in the common context of primary visual tasks.

Specifically, our experiment was designed to explore the following questions:

Q1. Do visual (V) and haptic (H) navigation signals (and their variants) impinge differently on the visual stimulus detection task?

Q2. Are (V) and (H) navigation signals equally susceptible to the negative effects of cognitive workload during a visually-based search and response task?

Q3. Will multimodal, reinforced signals (H+V) be more robust to workload than unimodal cues? Will they be more intrusive?

Q4. To obtain the full benefit of multimodality, is it necessary for cue modality to be reliable and consistent – given that in a real situation, some cues will inevitably be overlooked? This was tested with the alternating (H|V) condition.

We investigated these questions using a primarily visual spatial navigation paradigm, because it shares features with the common multitasking scenario of navigation during driving or piloting a plane. Our visual target detection task was intended to simulate some aspects of the visual attention needed to drive or fly.

In the following, we address each of these research questions in turn, finding in each of the questions supporting evidence of varying strength, along with an absence of countering evidence. We also comment on non-hypothesized observations found in our data, with the most interesting being a failure in the utility of reinforced signal cues when cognitive workload exceeds a threshold level.

5.4.5.1 Modality Intrusiveness

We proposed that visual navigation cues had the potential to be more attentionally intrusive than haptic cues on activities unrelated to navigation, because the primary task being interrupted is visual; and that this difference was likely to manifest itself when workload was increased (Q1 in our list above).

To explore this, we must start from a baseline performance approximated by the non-workload, unimodal conditions $(H)_{WL-}$ and $(V)_{WL-}$. Recall that the salience of the unimodal turn cues was effectively equalized to attain uniform performance in the turning task, without workload (Sections 5.4.3.4.2 and 5.4.4.3.1). Performance in the visual search task could not be simultaneously controlled; however, we did see similar visual search performance for $(H)_{WL-}$ and $(V)_{WL-}$ (Figure 5-13).

5.4.5.1.1 Overall Sensitivity of Primary Search Task to Workload and Cue Type

The primary visual detection task was designed to be difficult even without additional workload: over all conditions, the best performance in target detection ever achieved was about 60% success. Workload clearly exacerbated the stress which participants felt: post-experiment queries confirmed that most participants had been occasionally loaded to the point of task failure (Table 5-5:Q3). Despite this, there was no significant primary effect of either workload or navigation cue type on primary search task performance.

Up to a point, strategy was very likely the primary reason for this. Participants were told that the visual target detection task was the most important of the three they had to engage in; and after the experiment, 73% of participants confirmed that they employed strategies in their behavior (Table 5-3:Q5). The lack of any difference in measured intrusiveness of most kinds of navigation signals may simply reflect good ability to preserve performance in that task at the expense of performance elsewhere.

The fact that there is an exception to this pattern (i.e., one type of navigation cue where workload did have a considerable impact on primary task performance, as is evident in Figure 5-13) suggests that above a threshold, the visual search task was sensitive to workload. That is, participants were indeed operating close enough to their overall capacity to exhibit a decline even in the most important task under some conditions.

Further insight is available by reviewing factor sensitivity in the secondary navigation task, where performance did exhibit workload-linked degradation for some cue types (Figure 5-15, and discussed further in Section 5.4.5.2).

Thus, assuming a pool of cognitive resources with the primary task assuming strategic precedence, it does appear that both workload and cue type made differential demands on that pool, visible by considering performance broadly across all tasks.

5.4.5.1.2 Susceptibility of Visual Search Task to Visual versus Haptic Resource Competition

More specifically, our original thesis (Q1) regarding intrusiveness was that visual turn cues would compete for resources with the primary visual search task more effectively than haptic turn cues. This was not obvious on the basis of the visual search results alone: $(H)_{WL+}$ and $(V)_{WL+}$ resulted in comparable performance.

However as explained above, we must also consider the other tasks with which the user was simultaneously engaged. In Navigation performance, responses to turn cues with a visual component – (V) and (V+H) – do suggest greater sensitivity to workload than to (H) cues (Figure 9; marginally significant).

In addition, 54% of participants reported difficulty in attending to visual signals under added workload, compared to half that for haptic signals. Together, these data do indicate a strategic trading-off of processing resources, in which secondary tasks relying on vision are unduly affected. It must be noted that these results might be due in part to the natural alerting function of tactile stimuli. A basic function of the motor system of all animals is to protect the body from attack or collision. One type of defensive reaction, a fast, stereotyped response, is called the startle reflex which can be elicited through touch and audition (Casto 1989; Yeomans 2002; Cooke 2003).

5.4.5.1.3 Intrusiveness of Reinforced (H+V) Navigation Cue under Workload

Unexpectedly, we did find visual search performance impacted by a significant interaction between workload and cue type, attributable entirely to reinforced navigation cues (H+V). These were associated with a 13% drop in targets noticed when workload was added to the mix.

Multimodal reinforcement is generally thought to enhance signal detection, and indeed we see this gain for $(H+V)_{WL-}$ navigation performance (Figure 5-15). However, the cost of this higher salience appears to be increased intrusiveness into other, competing tasks. In this case, this manifests itself only when workload is increased, presumably pushing the user over a resource margin. Cognitive demands were then high enough that primary task performance suffered along with, as discussed below, performance in the navigation task for $(H+V)_{WL+}$.

5.4.5.2 Workload Robustness

In this section, we examine facets of cued navigation performance and confidence to explore our other three questions, which consider relations between cue modality and the cue's robustness to workload (Q2); as well as propositions about cue reinforcement (Q3) and reliability (Q4).

5.4.5.2.1 Large Individual Variance in Response to Haptic Cues

Before embarking on the discussion of users' ability to accurately and confidently utilize our various turn cue conditions, it is worth noting the relatively large between-subject variation in turn correctness when the cue has a haptic component. Referring to Table 5-4, 3rd section and 3rd column, we see that in absence of workload, standard deviation (StDev) is notably largest for $(H)_{WL-}$ and $(H|V)_{H,\,WL-}$; in contrast, turn performance in the $(V)_{WL-}$ condition has the smallest StDev of all the conditions.

This appearance of differential sensitivity changes when workload is added: $(H)_{WL+}$ StDev narrows to become consistent with $(V)_{WL+}$, but response to $(H|V)_{H, WL+}$ becomes even more scattered. Surprisingly, responses to visual cues in the mixed condition $(H|V)_{V, WL+}$ become very consistent among individuals, although average performance is nearly identical to $(H|V)_{H, WL+}$.

No such patterns were observed in participants' reported confidence; i.e. individuals were apparently consistent (in terms of small StDev's) across both turn cue type and workload level, in how well they think they responded to the turn cues – regardless of their actual performance or the average value of their ratings.

We were not able to test this individual-differences pattern for significance, but it is striking. If real, the most likely explanation has two parts:

(a) The unfamiliarity of the (H) turn cues could result in inconsistent utilization across individuals (some are likely more receptive to this modality than others, and none are accustomed to using it in this way).

(b) At the same time, the resource-competition effects we see elsewhere degrading responses to (V) cues under workload, could also be acting to increase the relative utility of the (H) cues under workload. In the mixed condition, the unreliability of the cue's source would tend to make this "the worst of both worlds", explaining the especially wild variance in response to haptic cues in the mixed condition, or $(H|V)_{H, WL+}$: haptically challenged or inexperienced individuals under workload who do not know where to look for the next cue will be especially disadvantaged.

5.4.5.2.2 Visual Turn Cues in Competition with a Visual Search Task?

In Section 5.4.5.1, we predicted (but did not find) increased intrusiveness of solely visual navigation cues over their haptic analogs on an unrelated, strategically primary and largely visual task, on the rationale of competition for a scarce perceptual resource (vision). The user, however, had volition over how to allocate that resource.

Our navigation performance (Correct Turns) data suggests but falls just short of statistically validating the accompanying idea that (as instructed) participants completed the primary visual task at the cost of responding to visual turn cues. We make this argument on the observation of exaggerated negative impact on visual cues of added workload (Q2), where resources are stretched to point of revealing weaknesses. Specifically, the first two column groupings in Figure 5-15 show that (V) turn cues suffered more, on average, from workload than to (H) turn cues. As explained in Section 5.4.4.3.1, this difference appears as significant in post-hocs but the multivariate ANOVA’s overall workload-cue type interaction is marginal, making its interpretation ambiguous.

We find more solid support for Q2 in the subjective results. Confidence in turn performance was similarly reduced by workload for all cue types (Figure

5-15). However, Figure 5-14 shows that independently of workload, users were considerably more confident about their response to (H) and (H+V) cues than to (V) or (H|V) cues regardless of workload, and this difference is confirmed by post-hoc significance.

Consistently, in their subjective responses at the end of the experiment (Table 2), 62% of participants reported they found haptic signals most “comfortable”, compared to 30% preferring visual signals – a factor of two. Only 8% had no preference. Likewise, a 2:1 ratio found visual, rather than haptic, signals “hardest to attend to” in the presence of the extra workload task, with a quarter seeing no difference. These reactions are not due to familiarity, since 62% of participants reported zero past experience with tactile displays, and the remainder only a little. These subjective reports underscore that additional effort was associated with the visual cues.

Relating our performance data to that of others, findings of the most similar past study are fairly consistent to ours. Engstrom et al. (Engstrom 2005) showed that hit rates for visual and tactile signal detection dropped 15% and 13% respectively from their baselines with the introduction of their worst-case workload task; both drops were significant, but whether the difference in these reductions is significant was not of direct interest to the authors and not tested (perhaps more interesting is the different patterns of sensitivity in the two signal types, among the various driving situations and tasks tested). Engstrom et al did find a nonzero workload-based reduction in response to tactile cues, whereas we did not. We do note that the experimental conditions and workload tasks were different for these two studies; the haptic cues differed, and Engstrom's was highly realistic (actually driving a car and talking on a cell phone) whereas ours was well controlled, including baseline equalization of hit rates for both tactile and visual signals by controlled (rather than environmental) haptic masking. Thus, many uncontrolled factors could explain the observed absolute differences; however, the overall trend is fairly consistent.

Summarizing our evidence for Q2, our results and other's suggest that actual utility (implying detectability) of haptic cues is more robust to workload than visual cue utility, when in competition with a visual primary task. Moreover, users are considerably more confident about and comfortable with cues containing a haptic component, despite the modality's unfamiliarity. Together, these strongly imply support for Q2, i.e. that haptic cues are more effective than visual cues in this situation.

More data would further clarify the situation. An extension of this study would benefit from considering whether long-term familiarity with spatially informative haptic signals might increase both overall performance and reduce individual variation in ability to utilize them; and from including a condition without any primary visual task – to see whether (H) and (V) performance and confidence then become more equal.

5.4.5.2.3 Importance of Reliability: Mixed Modality Cues

We will next examine our fourth question about the importance of signal reliability on its utilization (Q4): according to signal detection theory, we expected that when users are unable to predict the next cue's modality, performance would suffer. Navigation performance in response to cues of unpredictable modality (H|V) was indeed overall lower than both the unimodal and reinforced multimodal cues, with the latter difference significant (Figure 5-14). Meanwhile, workload had an apparently null effect on navigation performance for mixed cues (Figure 5-15), perhaps explained by fact that mixed-cue navigation was already at the lowest performance observed. The only difference in Correct Turn responses to (H) and (V) cues in the mixed condition – $(H|V)_H$ and $(H|V)_V$ respectively – was a greater consistency across individuals for the latter; explaining why only $(H|V)_V$ appears in the significant post hocs.

We can further argue that our mixed cue condition effectively reproduces the case of reinforcing and visual cues, when a user is visually distracted (Figure 5-15): these cue types both give good results in low workload situations, but with workload they both drop to the performance that mixed cues always give, regardless of workload. This could be explained as follows: when the user is distracted and missing cues (usually visual cues), then the signal becomes effectively as unreliable as the mixed-modality cues.

This theory is suggested by significant post-hocs showing workload-linked performance drops for (V) and (V+H), whose overall interaction ANOVA falls just short of significance (Section 5.4.4.3.1). Consistently, Confidence ratings for (V) are not significantly different than (V|H) levels. There are no such hints for purely (H) signals, for which responses are singularly unaffected by additional cognitive workload and confidence remains relatively high.

5.4.5.2.4 Reinforcement: Benefit Fails in Cognitive Overload?

We close our workload robustness discussion with our third research question about the utility of reinforced turn cues (Q3), since here we found our most intriguing results.

Without workload, reinforced multimodal signals $(H+V)_{WL-}$ was responsible for an improvement in navigation performance over the average for unimodal conditions H_{WL-} and V_{WL-} (an increase of 8.6% correct turns, Figure 5-15); this was unsurprising, given known cognitive effects of reinforcement.

However, an interesting effect occurred when workload was introduced $(H+V)_{WL+}$. Multimodally reinforced navigation signals were not robust to workload. The addition of cognitive workload caused a mean navigation performance decrease of 12.0% in the reinforced condition, compared to mean decrease of 5.0% in the unimodal conditions (1.5% for haptic and 8.5% for visual) in the unimodal conditions (Figure 5-15 and Table 5-4, third section). Stated another way, the benefit due to reinforcement disappeared when workload was added, with performance returning to unimodal levels (i.e. an average of 76.9% for (H) and (V), and 78.5% for (H+V)).

This workload-linked difference in (H+V) navigation performance is significant in a post-hoc associated with the marginal overall ANOVA interaction (Section 5.4.4.3.1); so more data is needed to clarify it specifically. However, it is further upheld by the unique and significant drop in workload-linked performance in the primary visual search task (Figure 5-14, and discussed earlier in 5.4.5.1.3). That is, in the larger picture of the user's effort, the $(H+V)_{WL-}$ condition was clearly implicated in higher overall workload and task degradation.

We suggest that this result can be explained by an "overload" effect, where more cognitive resources are needed to process information received from both modalities at the same time (even when that information reinforces the same percept), due to the "loudness" of the incoming information. This effect may only be apparent when cognitive workload is added because the workload increases the total demand for cognitive resources past a threshold, so that all tasks can no longer be carried out easily at the same time.

5.4.6 Conclusions and Future Work

In this experiment, we simulated common high-workload, visually-dominated scenarios (such as driving or walking in crowded, high-traffic areas while engaged in secondary tasks such as navigation and verbal communication) with a set of three simultaneous, highly controlled and monitored tasks, among which the user was required to strategically allocate inadequate sensory and cognitive resources. The experiment was devised to shed insight, in a controlled setting, on the role of modality (haptic, visual, reinforced and mixed) in intrusiveness and workload robustness of guiding cues in a secondary task, as evinced by performance and subjective confidence in the various tasks.

Visual search (detection of targets appearing at random screen locations) was the primary task; simultaneously, haptic and visual navigation cues directed participants through a maze, with salience adjusted to give uniform baseline performance. A cognitive workload task (counting heard sentences) was added in half of the conditions.

5.4.6.1 Primary Findings

The following findings were either suggested or supported (as specified below) through a combination of strong single observations, and weaker but consistent, triangulated observations from a number of sources with an absence of contradictory evidence.

5.4.6.1.1 Intrusiveness and Workload-Robustness of Haptic versus Visual Secondary Cues

Overall resources were clearly limited under full workload conditions; however users were mostly successful in managing their effort strategically so as to protect the task they'd been instructed to focus on (primary visual search) and maintain it at its baseline level, which was probably perceptually rather than cognitively limited. However, taking all tasks into consideration, visual cueing in the secondary task “cost” more than haptic cueing in terms of task performance, and was subjectively more distracting and less comfortable, even in the absence of workload. In this case of a visual primary task, visual cues were more intrusive than haptic cues.

Adding the cognitive workload task made it harder (objectively and subjectively) to respond correctly and directly to the visual navigation cues; the haptic signals proved immune, presumably because visual attention was also needed for the visual target detection task. Visual cues were more susceptible to workload-linked degradation in the immediate task than were haptic cues.

Both of these observations are consistent with the idea that there are separate pools of attention for each modality (Alais 2004); in this case, introducing the signals via the haptic sense left visual resources for the primary visual task.

5.4.6.1.2 “Overload” of Multimodally Reinforced Cues

Multimodal reinforcement followed typical patterns of improved signal response in the absence of workload. However, with workload an it appears that an important threshold seems to have been crossed; the multimodal signal made it more difficult to maintain performance across the board. Whereas the secondary navigation task took the brunt of workload-linked performance drops, in the reinforced cue condition the damage was also inflicted on the primary task despite the participant's demonstrated ability to strategically prioritize in other cases. We suggest that this implies a mechanism that went beyond resource exhaustion to active distraction.

We also noticed that a reinforced cue in the presence of workload may have "looked" the same as an unreliable "alternating" cue, since in a distracting environment some cues might be missed. On the surface, this would tend to argue for reinforcement – to make sure an important signal gets through, one way or another. But subjectively, the reality may be more confusing than if it is expected only on one channel; since the resulting "misses" will create a greater perception of signal unreliability. The criticality of the signal is also relevant.

5.4.6.1.3 Importance of Cue Reliability

Finally, our "alternating" cue condition, where a navigation cue might be either haptic or visual at random, gave unambiguously poor results. This result is consistent with signal detection theory: multiple potential channels to scan for a result means more overall effort and additionally, more likelihood of an overlooked (as opposed to mis-interpreted) cue.

From a design heuristic standpoint, this implies that in an attentionally demanding environment, unimodal cues may be more effective than cues from more than one modality when they are not temporally aligned (reinforcing). However, we have shown that reinforcement may be subject to unreliable perception when at least one of the paired cues is missed, as well as concerns about sensory overload when both are registered. Combined with its potential for sensory overload, this suggests that both multimodal cue variants (reinforced and alternating) we tested are better replaced with unimodal signals when high workload is anticipated.

In summary, countering their unambiguous value in calmer situations, the increased salience of multimodally reinforced cues in a stimulus-saturated environment is damaging to overall performance; and might have limited positive effect due to perceived unreliability. A more focused study will be required to clarify this interesting implication.

5.4.6.2 Design Implications

The findings presented here have several important implications for the design of effective and safe user interfaces. Together, they suggest that haptic cues can be an intuitive way to communicate navigation information to a user; in the context of a visual primary task, they are here shown to be more reliable, robust to workload and subjectively comfortable and clear to users, while intruding less on the primary task.

It is also important to be aware of problems that may arise due to cross-modal integration when even consistent, reinforcing signals are given multimodally. Our results indicate that multimodally reinforced cues increase cognitive demand more than unimodal cues: when dual-modality signals are in use, the addition of non-visual workload raises cognitive demands and impairs detection of visual targets. Furthermore, when signals are presented through randomly alternating modalities, detection performance drops significantly. Finally, we suspect that under workload, reinforced multimodal signals register as alternating (unreliable source) signals, and thus are subject to a double handicap. In short, unimodal signals may be more reliable and less effortful and intrusive than multimodal signals when high levels of cognitive workload is anticipated.

Potential applications for haptic cues of this type include most real-time, safety-critical environments with a continuous guidance interface component, including automotive and aircraft systems. A growing niche in the consumer world is for pedestrian navigation systems in handheld mobile devices. It is important to note that while we employed one type of directional haptic cue (spatial distributed vibrotactile stimuli); similar results might apply to a wide variety of other haptic stimuli and devices.

5.4.6.3 Future Work

Several issues which will further inform the design of optimal haptic guidance systems need to be further investigated.

In the short term, it would be useful to extend the current investigation as suggested throughout, e.g., to clarify our results here with additional data and to include a condition omitting the primary visual task for more fundamental base-lining. However, given triangulation from various sources, we have reasonable confidence in the present claims.

More interesting would inclusion or substitution of a haptic primary task for the visual one used here, to confirm that the asymmetry of effect we observed would then be reversed. This experiment will be harder to devise, given the weak role that haptic sensation plays in modern communications interfaces. We are also curious about how performance will change when haptic information transfer is more familiar, in either a primary or secondary role; longitudinal study is required to ascertain this.

Longer term, it will be important to create more realistic instantiations of the target environment, and eventually to actual driving with a real-time navigational aid. Each increment in realism will be accompanied by a loss of experimental control, underscoring the importance of the current step.

More generally, we believe that haptic communication can be employed for more complex information transfers than the binary directional cues examined here. We do not yet know how much information can be encoded in a single haptic message, although the two studies presented in Chapter 3 and work by others suggest that these signals will contain more than the 1 bit of information of the directional signals used here. The new evidence of haptic signals' relative robustness to workload presented here is encouraging, but how will more complex signals fare? These questions and others require further analysis and research.

6 Conclusion

As computing infrastructure becomes ubiquitous, computation is embedded ever more seamlessly into our surroundings (Bohn 2004). Computer-controlled, multi-function interfaces have become increasingly useful and desirable interaction tools. These interfaces are increasingly pervasive, often have complex functionality, and are frequently used in contexts which pose multiple demands on a single sensory modality: e.g., accessing a cell phone while driving or in a theatre, or using a vision-dependent remote control while watching TV in a darkened room. On the assumption that there is some degree of modularity in attentional processing and that using a different sensory channel for communication can reduce interference with critical visual tasks (Duncan 1997; Wickens 2002; Alais 2004), one possible approach to follow is to divert some information through the touch sense. Haptics has the potential to be useful as an attention-conserving communication channel, but only if designed with reference to human abilities. Understanding the principles and mechanisms of human haptic perception is an essential first step toward the design of useful haptic-based devices.

The overriding goal of this Thesis was to advance our knowledge of relevant human capabilities and embed this knowledge into haptic communication design procedures, in the interest of creating haptically supported interfaces that decrease rather than add to their users' sensory and cognitive load*.* The complexity of the intelligent aids we currently use in our daily tasks contributes to sensory overload. For example, the display of most cellular telephones conveys information far beyond basic “caller ID”, which means that its operation demands the visual sense for longer continuous periods of time exactly when visual attention is most fragmented (Oulasvirta 2005).

In this Thesis, we describe a series of experiments designed to further our understanding of haptic signal perception, design and usage in communicating simple meanings under different conditions. Our user studies identified a series of perceptual and cognitive limitations for processing simple vibrotactile stimuli and uncovered a surprising ability to identify and remember meaningful tactile signals over a long period of time. Our haptic icon editor was the first design tool for haptic icons and it explicitly recognized the need for and contributed to an iterative, visually supported design process. Finally, we have developed a series of software tools and methodologies intended to test different perceptual limits and to help create perceptual maps for different haptic stimuli. Throughout this work, we have made basic perceptual discoveries and created design tools that address various points and challenges in the meaningful-haptic-signal design process. These tools and methodologies have been used and extended further by researchers both within our group as well as in other universities (Chan 2005; Pasquero 2006; Swindells 2006; Ternes 2007).

6.1 Primary Contributions

Summaries of contributions for each of the major Thesis sections are described below.

Chapter 2:

Validating Perceptual Data for MDS Analysis — Although Multidimensional Scaling has been used in different disciplines for many years, there are few methods to validate the data gathered to be used in an MDS analysis. We presented a possible approach to try to solve this using a visualization tool to contrast and compare stimulus dissimilarity data (Section 2.5). The tool serves as an aid to try to obtain a measure of validity for the data gathered for later MDS analyses as well as to get a better understanding of how the tested stimuli were perceived by the recipients.

Haptic Perceptual Intermediate Waveform Design— In our effort to better understand perception (and cognitive processing) of haptic signals, we explored the concept of perceptual intermediate states for sensations that had not been previously used on a continuum. Through a systematic investigation, we established the presence of such a continuum as well as its perceptually linear spacing in the case of waveform perception. We created a waveform that could serve as an intermediate between a triangle and square waveforms. This allowed us to create a stimulus set varying perceptually in two dimensions:

frequency and waveform. This stimulus set was later used to test the feasibility of assigning meanings to properties of said stimuli.

Haptic Stimulus Editor — A systematic approach to haptic stimuli design requires tools that allow people without an engineering background closer participation in the creative process, thus broadening and enriching the area. We developed a tool to allow the rapid prototyping, editing and sharing of haptic stimuli. This tool was the first to exploit the affordances provided by an easy to edit visualization of a haptic signal. We developed this tool based on this idea of simplicity, demonstrated its potential and showed how important it was when used in conjunction with all the other tools we have developed in this process. Some of the principles and goals illustrated in this software tool have since been extended in a different direction by others including the work by Swindells, Maksakov, et al. (Swindells 2006).

Chapter 3:

Assigning Meanings to Haptic Signals — The tools and methods described in Chapter 2 paved the way to the creation of sets of perceptually distinct tactile sensations spread across a perceptual map. These signals were a necessary prerequisite to investigate the possibility of creating information-bearing tactile signals. Two experiments were designed to investigate assigning meanings to tactile signals.

In the first of these experiments, we assigned meanings to tactile signal parameters. In this manner, participants learned meanings associated with levels of frequency and waveform. Results show that participants were able to learn and remember arbitrary signal parameter- concept associations with an average performance close to 80% for a period of 45 minutes after learning them. The second experiment investigated the effects of giving participants a choice when assigning meanings to a set of tactile signals and tested long term recall (two weeks) of the learned associations. Results for this experiment show that participants had a surprising ability to learn the arbitrary associations as well as those chosen by them (within a 20 minute period) and remember an average of 86% of these learned associations two weeks later without reinforcement. These studies were the first to report the assignment of meanings to abstract tactile signals and the first to report long-term recall of these associations.

Chapter 4:

Tactile Stimulus Masking — Having developed sets of meaningful tactile signals that could be identified when presented alone, we set out to investigate possible (accidental as well as deliberate) perceptual interference of said tactile signals when presented in close temporal contiguity.

A series of pilots helped determine parameters to use in an experiment designed to test perceptual masking of tactile stimuli. This experiment was designed to investigate two mechanisms for temporal masking of vibrotactile stimuli (backwards and common-onset) using a commodity display. Results confirm the existence of a statistically significant masking effect for both forms of masking explored, with common onset exhibiting a significantly larger masking effect than backwards. To the best of our knowledge, this experiment is the first to analyze confidence in response levels to identification of masked stimuli. The knowledge gained from this experiment allows us to estimate proper temporal stimuli separation to prevent interference and will also allow us to further investigate perceptual and cognitive mechanisms by purposefully masking tactile stimuli from perception.

Chapter 5:

Haptic Signal Perception under Cognitive Workload— In Chapter 3, we learned that we could create meaningful tactile signals which could be identified when presented in isolation. However, it is our intent to have these signals eventually used in everyday interfaces. These are often used in situations where the recipients' other senses (and cognitive processes) are occupied with other tasks. With this in mind, we conducted two experiments designed to test haptic perception in simulated multi-tasking and high-cognitive-workload scenarios. In the first of these experiments we investigated the effects of force feedback warning signal reliability. We found that the introduction of false alarms reduced the effectiveness of a said warning signal while the introduction of misses did not appear to have this effect. The second experiment compared salience-calibrated tactile, visual and multimodal navigation cues during a driving-like task, and examined the effectiveness and intrusiveness of the navigation signals while varying cognitive workload and masking of task cues. We found that participants continued to utilize tactile navigation signals under high workload, but their usage of visual and reinforced multimodal navigation cues degraded; further, the reinforced cues under high cognitive workload disrupted the visual primary task. The findings reported in these experiments show that haptic communication in the real world and even under considerable levels of cognitive workload might indeed be a viable alternative.

6.2 Future Work

Continuation of the work described in this Thesis would benefit from efforts in the following areas:

- Design and analysis of a larger collection of haptic icons and comparisons with other modalities (visual icons) could further improve our understanding of haptic perception and perhaps allow us to determine shared cognitive limitations. How many of these icons can we learn? Will different contexts allow the re-use of the same icons with different meanings? How different must icons be to allow us to tell them apart consistently?
- Longer term reinforced learning and recall testing of haptic icons is required before these are truly usable. We need to determine a method that will allow a number of users to gradually learn and continually use a set of haptic icons during everyday activities. This would allow us to verify whether the use of haptic icons in everyday

devices is feasible. For example, will reinforced (persistent) learning allow us to grow familiar with haptic icons in the same way we familiarize ourselves with their graphical counterparts? How long does it take for people to get used to haptic icons?

- We require further investigation into the effects of cognitive workload on perception of haptic signals as well as the effects these signals might have on other tasks with which a user might be involved. It would be ideal to test these signals in more realistic scenarios. For example, can we still identify the meanings of haptic signals as we are driving a real (as opposed to simulated) car? Will the cognitive effort of identifying these signals have an adverse effect on other tasks we do at the same time? Are haptic signals any better than visual signals when conducting a more realistic primarily visual task?
- Additional studies could be carried out to investigate an increase in the information density of haptic icons. These could test different methods to make haptic icons shorter or to form part of sequences of haptic icons which give way to more complex meanings. How much information can we convey per unit of time using haptic icons? Can this increase with repeated exposure to the signals?

More generally, future work should be focused on trying to improve our understanding of human perceptual and cognitive limitations with regards to haptic signals. This will allow us to further investigate the possibility of creating a more complex haptic language.

References

Alais, D., Morrone, C., Burr, D. (2004). "Separate attentional resources for vision and audition." Journal of Neuroscience **1-12**.

Allen, M., Gluck, J., MacLean, K., Tang, E. (2005). An initial usability assessment for symbolic haptic rendering of music parameters. ICMI 2005, 7th International Conference on Multimodal Interfaces, Trento, Italy.

Baldwin, C. L., Coyne, J.T. (2003). Mental workload as a function of traffic density: comparison of physiological, behavioural, and subjective indices. Second International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design, Iowa City, University of Iowa.

Bergmann Tiest, W. M., Kappers, A. M. L. (2006). "Analysis of haptic perception of materials by multidimensional scaling and physical measurements of roughness and compressibility." Acta Psychologica **121**(1): 1-20.

Bertelson, P. (1967). "The time course of preparation." Quarterly Journal of Experimental Psychology **19**: 272-279.

Bohn, J., Coroam, V.,Langheinrich, M.,Mattern, F.,Rohs, M. (2004). Disappearing computers everywhere - Living in a world of smart everyday objects. New Media, Technology and Everyday Life in Europe Conference, London.

Bouvrie, J., Ezzat,T., Poggio, T (2006). Inter-voice audio morphing, MIT.

Breitmeyer, B. G. (1984). Visual masking: An integrative approach. New York, Oxford University Press.

Brewster, S., Wright, P., Edwards, A. (1992). A detailed investigation into the effectiveness of earcons. Auditory Display: Sonification, Audification and Auditory Interfaces. G. Kramer, Addison-Wesley**:** 471-498.

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

Brown, L. M., Brewster, S.A., Purchase, H.C. (2006). Multidimensional Tactons for Non-Visual Information Presentation in Mobile Devices. 8th Conference on Human-Computer Interaction with Mobile Devices and Services, Mobile HCI 2006, Helsinki, Finland.

Brown, L. M., Kaaresoja, T. (2006). Feel who's talking: using tactons for mobile phone alerts. CHI '06 Extended Abstracts on Human Factors in Computing Systems, Montréal, Québec, Canada.

Buxton, B., Gaver, B., Bly, S. (1990). The Use of Non-Speech Audio at the Interface. Tutorial Notes: ACM Conference on Human Factors in Computing Systems.

Casto, R., Nguyen, T., Printz, M. P. (1989). "Characterization of cardiovascular and behavioral responses to alerting stimuli in rats." American Journal of Physiology - Regulatory,Integrative and Comparative Physiology **256**: 1121-1126.

Chakravarti, L. (1967). Handbook of Methods of Applied Statistics, vol. 1.

Chan, A., MacLean, K., McGrenere, J (2005). Learning and Identifying Haptic Icons Under Workload. World Haptics, Pisa, Italy.

Chang, A., O'Sullivan, C. (2005). Audio-Haptic Feedback in Mobile Phones. Extended Abstracts of ACM Conference on Human Factors in Computing Systems, CHI 2005, Portland, Oregon.

Cooke, D. F., Graziano, M. S. A. (2003). "Defensive Movements Evoked by Air Puff in Monkeys." Journal of Neurophysiology **90**: 3317-3329.

Cox, T. F. (1988). Multidimensional Scaling. New York.

Craig, J. C. (1982). "Temporal integration of vibrotactile patterns." Perception & Psychophysics **32**: 219 - 229.

Craig, J. C. (1985). "Tactile pattern perception and its perturbations." Journal of the Acoustical Society of America **77**(1): 238-246.

Craig, J. C. (1989). "Interference in localizing tactile stimuli." Perception & Psychophysics **45**(4): 343-355.

Craig, J. C., Evans, P.M. (1987). "Vibrotactile masking and the persistence of tactual features." Perception & Psychophysics **42**(4): 309-317.

Craig, J. C., Quian, X. (1997). "Tactile pattern perception by two fingers: temporal interference and response competition." Perception & Psychophysics **59**(2): 252-265.

Day, J. (2004). Can BMW's iDrive Pass Its Road Test Now? Electronic Design.

Di Lollo, V., Enns, J. T., Rensink, R. A. (2000). "Competition for consciousness among visual events: the psychophysics of reentrant visual pathways." Journal of Experimental Psychology: General **129**: 481-507.

Dixon, S. R., Wickens, C.D. and McCarley, J.S. (2007). "On the Independence of Compliance and Reliance: Are Automation False Alarms Worse Than Misses?" Human Factors **49**(4): 564-572.

Driver, J., Spence, C. (2004). Crossmodal Spatial Attention: Evidence from Human Performance. Crossmodal Space and Crossmodal Attention. Crossmodal Space and Crossmodal Attention, Oxford University Press**:** 141-177.

Duncan, J., Martens, S., Ward, R. (1997). "Restricted attentional capacity within but not between sensory modalities." Nature **387**: 808-810.

Engstrom, J., Aberg, N., Johansson, E. (2005). Comparison between visual and tactile signal detection tasks applied to the safety assessment of in-vehicle safety information systems. Third International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design, Rockport, Maine.

Enns, J. T., Di Lollo, V. (1997). "Object substitution: A new form of masking in unattended visual locations." Psychological Science **8**: 135 – 139.

Enns, J. T., Di Lollo, V. (2000). "What's new in visual masking?" Trends in Cognitive Sciences **4**: 345-352.

Enriquez, M. (2002). A Study of Haptic Icons. Computer Science. Vancouver, University of British Columbia**:** 119.

Enriquez, M., MacLean, K. (2003). The hapticon editor: A tool in support of haptic communication research. 11th Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems (HAPTICS'03), Los Angeles.

Enriquez, M. J., MacLean, K. E. (2004). Impact of haptic warning signal reliability in a time-and-safety-critical task. 12th Annual Symposium on Haptic Interfaces for Virtual Environments and Teleoperator Systems, IEEE-VR2004, Chicago, USA.

Enriquez, M. J., MacLean, K. E. (2007). "Backward and Common-onset Masking of Vibrotactile Stimuli." Brain Research Bullettin(Special Issue on Robotics and Neuroscience.).

Enriquez, M. J., MacLean, K. E. (2008). "Backward and common-onset masking of vibrotactile stimuli." Brain Research Bullettin, Special Issue: Robotics and Neuroscience **75**(6): 761-769.

Enriquez, M. J., MacLean, K. E. (2008). The role of choice in longitudinal recall of meaningful tactile signals. Symposium on Haptic Interfaces for Virtual Environments and Teleoperator Systems, Reno, Nevada.

Enriquez, M. J., MacLean, K. E. (2008). Visualization Tools for Analysis of Multidimensional Scaling Dissimilarity Data TR2008-12. Vancouver, BC, University of British Columbia.

Enriquez, M. J., MacLean, K. E., Neilson, H. (2007). Interplay of tactile and visual guidance cues under multimodal workload. Vancouver, Department of Computer Science, University of British Columbia.

Enriquez, M. J., MacLean, K. E., Chita, C. (2006). Haptic phonemes: Basic building blocks of haptic communication. Eighth International Conference on Multimodal Interfaces (ICMI'06), Banff, Alberta, Canada.

Evans, P. M. (1987). "Vibrotactile masking: Temporal integration, persistence, and strengths of representations." Perception & Psychophysics **42**(6): 515–525.

Evans, P. M., Craig, J. C. (1986). "Temporal integration and vibrotactile backward masking." Journal of Experimental Psychology: Human Perception and Performance **12**: 160 - 168.

Exner, S. (1888). "Über optische Bewegungsempfindungen [On optical motion sensation]." Biologisches Zentralblatt **8**: 437-448.

Forsyth, B., MacLean, K. E. (2006). "Predictive haptic guidance: Intelligent user assistance for the control of dynamic tasks." IEEE Transactions on Visualization and Computer Graphics **12**(1): 103-113.

Frei, P., V. Su, et al. (2000). Curlybot: Designing a new class of computational toys. Conference on Human Factors in Computing Systems (CHI 2000), The Hague, Netherlands.

Ganz, L. (1975). Temporal factors in visual perception. Handbook of perception. C. M. P. Friedman. New York, Academic Press. **5:** 170-231.

Gaver, W. W. (1986). "Auditory icons: Using sound in computer interfaces." Human-Computer Interaction **19**(1): 74.

Gaver, W. W. (1988). Everyday listening and auditory icons. Computer Science. San Diego, University of California.

Gaver, W. W. (1989). "The SonicFinder: An interface that uses auditory icons." Human Computer Interaction **4**(1): 67-94.

Gerofsky, S. (2008). Haptic interfaces and design-based software for teaching functions. http://cust.educ.ubc.ca/faculty/gerofsky.html.

Gescheider, G. A., Bolanowski, S.J., Verrillo R.T. (1989). "Vibrotactile masking: Effects of stimulus onset asynchrony and stimulus frequency." Journal of the Acoustical Society of America **85**(5): 2059-2064.

Gescheider, G. A., Migel, N. (1995). "Some temporal parameters in vibrotactile forward masking." Journal of the Acoustical Society of America **98**(6):

3195-3199.

Gray, R., Tan, H.Z. (2002). "Dynamic and predictive links between touch and vision." Experimental Brain Research **145**: 50-55.

Green, D. M., Swets, J. A. (1966). Signal Detection Theory and Psychophysics. New York.

Greenwald, A. G., Draine, S. C., Abrams, R. L. (1996). "Three cognitive markers of unconscious semantic activation." Science **273**: 1699 - 1702.

Griffiths, P., Gillespie, R. B. (2004). Shared control between human and machine: Haptic display of automation during manual control of vehicle heading. 12th International Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems (HAPTICS'04).

Hall, G. (1991). Perceptual and Associative Learning. Oxford, Oxford University Press.

Haller, R. (2003). "The display and control concept iDrive - Quick access to all driving and comfort functions." ATZ/MTZ Extra (The New BMW 5-Series)(August): 51–53.

Heller, M. A., Schiff, W. (1991). The Psychology of Touch. Hillsdale, New Jersey, Lawrence Erlbaum Associates.

Ho, C., Tan, H.Z., Spence, C. (2005). "Using spatial vibrotactile cues to direct visual attention in driving scenes." Transportation Research Part F: Traffic Psychology and Behavior **8**: 397-412.

Hollins, M., Bensmaia, S., et al. (2000). "Individual differences in perceptual space for tactile textures: Evidence from multidimensional scaling." Perception & Psychophysics **62**(8): 1534-1544.

Hollins, M., Faldowski, R., Rao, S., Young, F. (1993). "Perceptual dimensions of tactile surface texture: A multidimensional scaling analysis." Perception and Psychophysics **54**: 697-705.

Howard, J. P., Rogers, B.J. (1995). Binocular Vision and Stereopsis. Oxford, Oxford University Press.

Huggins, W. H., Entwisle, D. R. (1974). Iconic communication: An annotated biography. Baltimore.

Hupe, J. M., James, A. C., Payne, B. R., Lomber, S. G., Girard, P. & Bullier, J.. (1998). "Cortical feedback improves discrimination between figure and ground by V1, V2, and V3 neurons." Nature(394): 784-787.

Immersion (2005). Immersion VibeTonz System. http://www.immersion.com/mobility/.

Iqbal, S. T., Horvitz, E. (2007). Disruption and recovery of computing tasks: Field study, analysis, and directions. ACM Conference on Human Factors in Computing Systems (CHI '07), San Jose, CA.

Jones, L., Sarter, N. (2008). "Tactile displays: Guidance for their design and application." The Journal of Human Factors and Ergonomic Society **50**(1): 90-111.

Just, M. A., Carpenter, P.A., Keller, T.A., Emery, L., Zajac, H. and Thulborn, K.R. (2001). "Interdependence of nonoverlapping cortical systems in dual cognitive tasks." NeuroImage **14**(417-426).

Klatzky, R. L., Lederman, S., Hamilton, C., Grindley, M. (2003). "Feeling textures through a probe: Effects of probe and surface geometry and exploratory factors." Perception & Psychophysics **65**(4): 613-631.

Kruskal, J. B. (1964). "Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis." Psychometrika **29**(1): 1-27.

Leung, R., MacLean, K. (2007). Evaluation of haptically augmented touchscreen GUI elements under cognitive load. Proceedings of 9th International Conference on Multimodal Interfaces (ICMI '07).

Macaluso, E., Driver, J. (2001). "Spatial attention and crossmodal interactions between vision and touch." Neuropsychologia **39**: 1304-1316.

MacLean, K., Enriquez, M. (2003). Perceptual design of haptic icons. Eurohaptics, Dublin, Ireland.

MacLean, K. E. (1996). The haptic camera: A technique for characterizing and playing back haptic environments. The 5th Ann. Symp. on Haptic Interfaces for Virtual Environments and Teleoperator Systems, ASME/IMECE, Atlanta, GA.

Maltz, M., Shinar, D. (2003). "New alternative methods of analyzing human behavior in cued target acquisition." Human Factors **45**(2): 281-295.

Manzey, D. (1998). "Mental Performance in Extreme Environments: Results from a 438 day Spaceflight." Ergonomics **41**(4): 537-559.

Marcel, A. J. (1983). "Conscious and unconscious perception: experiments on visual masking and word recognition." Cognitive Psychology **15**: 197 - 237.

Merikle, P. M., & Daneman, M. (2000). Conscious vs. unconscious perception.

The new cognitive neurosciences, Cambridge, MA: MIT Press**:** 1295 - 1303.

Michelitsch, G., M. Osen, et al. (2004). Haptic Chameleon. EuroHaptics 2004, Munich, Germany.

Miller, G. (1956). "The magical number seven, plus or minus two: Some limits on our capability for processing information." Psychological Review **63**.

Miura, T. (1990). Active function of eye movement and useful field of view in a realistic setting. From Eye to Mind: Information Acquisition in Perception, Search, and Reading. d. Y. a. R. P. R. Groner. Amsterdam: North-Holland**:** 119 -127.

Mowrer, O. H. (1940). "Preparatory set (Expectancy) - Some methods of measurements." Psychological Review Monograph **52**(233).

Nolin, J., Stemniski, P., Okamura, A. M. (2003). Activation cues and force scaling methods for virtual fixtures. Proc. of 11th Int'l Symp on Haptic Interfaces for Virtual Environment and Teleoperator Systems, Los Angeles.

Oulasvirta, A., Tamminen, S., et al. (2005). "Interaction in 4-second bursts: The fragmented nature of attentional resources in mobile HCI." Proceedings of ACM Conference on Human Factors in Computing Systems (CHI '05) **CHI Letters, vol. 7**(1): 919-928.

Oxenham, A. (2001). "Forward maskIng: adaptation or integration?" Journal of the Acoustical Society of America **109**(2): 732-741.

Parker, D. M., Lishman, J.R., Hughes, J. (1992). "Temporal integration of spatially filtered visual images." Perception **21**(2): 147-160.

Pasquero, J., Luk, J., Little, S., MacLean, K. (2006). Perceptual analysis of haptic icons: An investigation into the validity of cluster sorted MDS. Haptic Symposium 2006, Washington, DC.

Patten, C., Kircher, A., Ostlund, J., Nilsson, L. (2004). "Using mobile telephones: Cognitive workload and attention resource allocation." Accident Analysis & Prevention **36**(3): 341-350.

Paulhus, D. L. (1991). Measurement and control of response bias. Measures of personality and social psychological attitudes. P. R. S. J.P. Robinson, & L.S. Wrightsman. New York Academic Press., Academic Press.

Payandeh, S., Stanisic, Z. (2002). On application of virtual fixtures as an aid for telemanipulation and training. Tenth Symposium on Haptic Interfaces for Virtual Environments and Teleoperator Systems, Orlando, Florida.

Potter (1976). " Short-term conceptual memory for pictures." Journal of experimental psychology: Human learning and memory **2**(5): 509 -522.

Poupyrev, I., Okabe, M., Maruyama, S. (2004). Haptic feedback for pen computing: Directions and strategies. ACM Extended Abstracts of the ACM Conference on Human Factors in Computing Systems - CHI 2004.

Rantanen, E., Goldberg, J. (1999). "The effect of mental workload on the visual field size and shape." Ergonomics **42**(6): 816-834.

Rinker, M. A., Craig, J. C., and Bernstein, L. E. (1998). "Amplitude and period discrimination of haptic stimuli." J. Acoust. Soc. Am. **104**(1): 453–463.

Rinker, M. A., Craig, J. C., Bernstein, L. E. (1998). "Frequency and amplitude discrimination of haptic stimuli." Journal of the Acoustical Society of America **104**: 453-463.

Rosenberg, L. (1994). Virtual fixtures: Perceptual overlays enhance operator performance in telepresence tasks. Mechanical Engineering, Stanford University.

Rupert, A. H. (2000). "An instrumentation solution for reducing spatial disorientation mishaps: a more 'natural' approach to maintaining spatial orientation." IEEE Engineering in Medicine and Biology: 71-80.

Sherrick, C. A., Cholewiak, R. (1986). Cutaneous Sensitivity. Handbook of Perception and Human Performance. e. a. K. R. Boff. New York, John Wiley and Sons. **12:** 12-57.

Sherrick, C. E. (1953). "Variables affecting the sensitivity of the human skin to mechanical vibration." Journal of Experimental Psychology **45**(5): 273-282.

Sillito, A. M., Jones, H.E., Gerstein, G.L., West, D.C. (1994). "Feature-linked synchronization of thalamic relay cell firing induced by feedback from the visual cortex." Nature(369): 479-482.

Slaney, M., Covell, M., Lassiter, B (1996). Automatic audio morphing. Procceedings of the 1996 International Conference on Acoustics, Speech, and Signal Processing, Atlanta, GA.

Snibbe, S. S., MacLean, K. E. (2001). Haptic metaphors for digital media. ACM Symp. on User Interface Software & Technology (UIST 2001), Orlando, FL.

Spence, I., Domoney, D. W. (1974). "Single subject incomplete designs for nonmetric multidimensional scaling." Psychometrika **39**(4): 469-490.

Srinivasan, M. A. (1989). "Surface deflection of primate fingertip under line load." Journal of Biomechanics **22**(4): 343-349.

Srinivasan, M. A., Whitehouse, J.M., Lamotte, R.H. (1990). "Tactile detection of slip: Surface microgeometry and peripheral neural codes." Journal of Neurophysiology **63**: 1323–1332.

Srinivasan, S., Wang, D. (2005). Modeling the perception of multitalker speech. Interspeech, Lisbon, Portugal.

Steele, M., Gillespie, B. (2001). Shared control between human and machine. Human Factors and Ergonomics Society, Minneapolis, MN.

Stein, B. E., M.A. Meredith (1993). The Merging Of The Senses. Cambridge, MA, The MIT Press.

Swindells, C., MacLean, K.E. (2007). Capturing the dynamics of mechanical knobs. World Haptics Conference, IEEE Press.

Swindells, C., Maksakov, E., MacLean, K.E., and Chung, V. (2006). The role of prototyping tools for haptic behavior design. Symposium on Haptic Interfaces for Virtual Environments and Teleoperator Systems, IEEE Press.

Tan, H., Lim, A. and Traylor, R. (2000). A psychophysical study of sensory saltation with an open response paradigm. In the 9th Ann. Symp. on Haptic Interfaces for Virtual Environment and Teleoperator Systems, ASME/IMECE,, S. S. Nair.

Tan, H. Z., Charlotte M. Reed, Lorraine A. Delhorne, Nathaniel I. Durlach, and Natasha Wan (2003). "Temporal masking of multidimensional tactual stimuli." Acoustic Society of America **114**(6): 3295-3308.

Tan, H. Z., Durlach, N.I., Reed, C. M., Rabinowitz, W. M. (1999). "Information transmission with a multifinger tactual display." Percept Psychophys **61**(6): 993-1008.

Tan, H. Z., Gray, R., Young, J. J.,Traylor, R. (2003). "A haptic back display for attentional and directional cueing." Haptics-e **3**(1).

Tan, Y. (2003). Validity of haptic cues and its effect on priming visual spatial attention. 11th Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems.

Tang, A., McLachlan, P., Lowe, K., Saka, C. R., MacLean, K. (2005). Perceiving ordinal data haptically under workload. Seventh International Conference on Multimodal Interfaces (ICMI 2005), Trento, Italy.

Tanner, W. P., Jr. and Swets, J. A. (1954). "A decision-making theory of visual detection." Psychological Review Monograph **61**: 401-409.

Taylor, M. M., Creelman, C. D. (1967). "PEST: Efficient Estimates on Probability Functions." Journal of the Acoustical Society of America **41**(4A): 782-787.

Ternes, D. (2007). Building large sets of haptic icons: rhythm as a design parameter, and between-subjects MDS for evaluation. Computer Science. Vancouver, University of British Columbia.

Ternes, D., & MacLean, K. E. (in review) (2008). Building large sets of haptic icons: rhythm as a design parameter. Submitted to ACM CHI '08, Florence, Italy.

Tipper, C. M. (2003). Is inhibition of return truly a reflexive phenomenon? An investigation of the role of temporal predictability in a classic inhibition of return paradigm. Department of Psychology. Vancouver, BC, University of British Columbia.

Turvey, M. T. (1973). "On peripheral and central processes in vision: Interferences from an information-processing analysis of masking with patterned stimuli." Psychological Review **80**: 1-52.

van Erp, J. B. F., Spapé, M. M. A. (2003). Distilling the underlying dimensions of tactile melodies. Proceedings of Eurohaptics, Dublin, Ireland.

Van Veen, H. A., van Erp, J.B.F. (2003). Providing directional information with tactile torso displays. Eurohaptics 2003.

Verillo, R. T. (1992). "Vibration sensation in humans." Music Perception(3): 281-302.

Ward, L. (1977). "Multidimensional scaling of the molar physical environment." Multivariate Behavioral Research **12**: 23-42.

Ward L., R. J. (1981). "Cognitive Set and the Perception of Place." Environment and Behavior **13**: 610-632.

Weber, D. L., Green, D.M. (1978). "Temporal factors and suppression effects in backward and forward masking." Journal of the Acoustical Society of America **64**(5): 1392-1399.

Weisenberger, J. M. (1994). "Tactile temporal masking: Effect of multiple maskers." Journal of the Acoustical Society of America **95**: 2213-2220.

Wickens, C. D. (1984). Engineering Psychology and Human Performance. Columbus, Ohio, Charles E. Merrill.

Wickens, C. D. (1992). Attention, time-sharing, and workload. Engineering Psychology and Human Performance. New York, Harper Collins**:** 364-411.

Wickens, C. D. (2002). "Multiple resources and performance prediction." Theoretical Issues in Ergonomics Science **3**(2): 159-177.

Yazdani, M. G., D. (1990). Iconic Communication. Department of Computer Science, Exeter University.

Yeomans, J. S., Li, L., Scott, B.W., Frankland, P. W. (2002). "Tactile, acoustic and vestibular systems sum to elicit the startle reflex." Neuroscience and Biobehavioral Reviews **26**: 1-11.

Appendix A

Ethics Forms

This research was approved by the Behavioral Research Ethics Board at The University of British Columbia under certificate number B01-0470. Dr. Karon MacLean acted as the faculty advisor for the group of user studies involving human participants titled:

Physical and multimodal user interfaces - usability & psychophysics.

Copies of the Certificate of Approval Form and Participant Consent Form are included in this Appendix.

The following page contains a copy of the Certificate of Approval dated 22 August 2005 from the Behavioral Research Ethics Board.

The University of British Columbia
Office of Research Services and Administration
Behavioural Research Ethics Board

Certificate of Approval

PRINCIPAL INVESTIGATOR MacLean, K.E.	DEPARTMENT Computer Science	NUMBER **B01-0470**

INSTITUTION(S) WHERE RESEARCH WILL BE CARRIED OUT
UBC Campus ,

CO-INVESTIGATORS:
Enriquez, Mario, Computer Science; Luk, Joseph, Computer Science; Pedrosa, Ricardo, Computer Science; Smith, Jocelyn, Computer Science; Swindells, Colin, Computer Science; Yohanan, Steve, Computer Science

SPONSORING AGENCIES
Natural Science Engineering Research Council

TITLE :
Physical and Multimodal User Interfaces - Usability & Psychophysics

APPROVAL RENEWED DATE	TERM (YEARS)	AMENDMENT:	AMENDMENT APPROVED:
AUG 22 2005	1	August 19, 2005, Co-PIs / Research method	AUG 22 2005

CERTIFICATION:

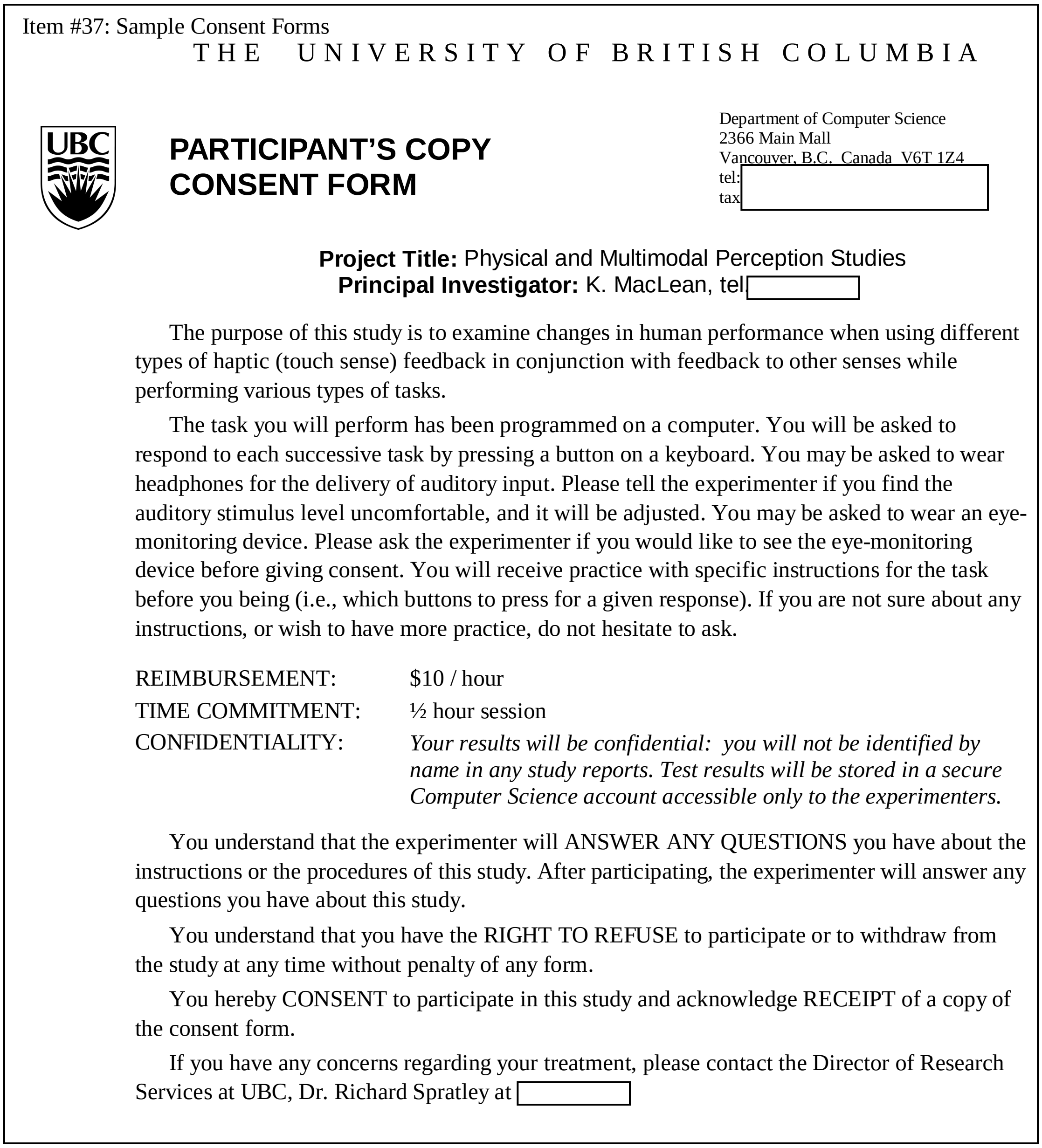
The protocol describing the above-named project has been reviewed by the Committee and the experimental procedures were found to be acceptable on ethical grounds for research involving human subjects.

[redacted]

Approval of the Behavioural Research Ethics Board by one of the following:
Dr. James Frankish, Chair,
Dr. Cay Holbrook, Associate Chair,
Dr. Susan Rowley, Associate Chair

This Certificate of Approval is valid for the above term provided there is no change in the experimental procedures

The following pages contain a copy of the consent form signed by user study participants.



Item #37: Sample Consent Forms

THE UNIVERSITY OF BRITISH COLUMBIA

Department of Computer Science
2366 Main Mall
Vancouver, B.C. Canada V6T 1Z4
tel: [redacted]
fax

PARTICIPANT'S COPY
CONSENT FORM

Project Title: Physical and Multimodal Perception Studies
Principal Investigator: K. MacLean, tel [redacted]

The purpose of this study is to examine changes in human performance when using different types of haptic (touch sense) feedback in conjunction with feedback to other senses while performing various types of tasks.

The task you will perform has been programmed on a computer. You will be asked to respond to each successive task by pressing a button on a keyboard. You may be asked to wear headphones for the delivery of auditory input. Please tell the experimenter if you find the auditory stimulus level uncomfortable, and it will be adjusted. You may be asked to wear an eye-monitoring device. Please ask the experimenter if you would like to see the eye-monitoring device before giving consent. You will receive practice with specific instructions for the task before you being (i.e., which buttons to press for a given response). If you are not sure about any instructions, or wish to have more practice, do not hesitate to ask.

REIMBURSEMENT:	$10 / hour
TIME COMMITMENT:	½ hour session
CONFIDENTIALITY:	*Your results will be confidential: you will not be identified by name in any study reports. Test results will be stored in a secure Computer Science account accessible only to the experimenters.*

You understand that the experimenter will ANSWER ANY QUESTIONS you have about the instructions or the procedures of this study. After participating, the experimenter will answer any questions you have about this study.

You understand that you have the RIGHT TO REFUSE to participate or to withdraw from the study at any time without penalty of any form.

You hereby CONSENT to participate in this study and acknowledge RECEIPT of a copy of the consent form.

If you have any concerns regarding your treatment, please contact the Director of Research Services at UBC, Dr. Richard Spratley at [redacted]

Department of Computer Science
2366 Main Mall
Vancouver, B.C. Canada V6T 1Z4
tel: ([redacted]
tax: (

RESEARCHER'S COPY
CONSENT FORM

Project Title: Physical and Multimodal Perception Studies
Principal Investigator: K. MacLean, tel. [redacted]

The purpose of this study is to examine changes in human performance when using different types of haptic (touch sense) feedback in conjunction with feedback to other senses while performing various types of tasks.

The task you will perform has been programmed on a computer. You will be asked to respond to each successive task by pressing a button on a keyboard. You may be asked to wear headphones for the delivery of auditory input. Please tell the experimenter if you find the auditory stimulus level uncomfortable, and it will be adjusted. You may be asked to wear an eye-monitoring device. Please ask the experimenter if you would like to see the eye-monitoring device before giving consent. You will receive practice with specific instructions for the task before you being (i.e., which buttons to press for a given response). If you are not sure about any instructions, or wish to have more practice, do not hesitate to ask.

REIMBURSEMENT:	$10 / hour
TIME COMMITMENT:	½ hour session
CONFIDENTIALITY:	*Your results will be confidential: you will not be identified by name in any study reports. Test results will be stored in a secure Computer Science account accessible only to the experimenters.*

You understand that the experimenter will ANSWER ANY QUESTIONS you have about the instructions or the procedures of this study. After participating, the experimenter will answer any questions you have about this study.

You understand that you have the RIGHT TO REFUSE to participate or to withdraw from the study at any time without penalty of any form.

You hereby CONSENT to participate in this study and acknowledge RECEIPT of a copy of the consent form:

NAME ____ DATE ____________________

(please print)

SIGNATURE __________________________________

If you have any concerns regarding your treatment, please contact the Director of Research Services at UBC, Dr. Richard Spratley at [redacted]