SENSORIMOTOR INTERFACES: TOWARDS ENACTIVITY IN HCI

Iain Carson

A Thesis Submitted for the Degree of PhD
at the
University of St Andrews

2022

Identifiers to use to cite or link to this thesis:

DOI: https://doi.org/10.17630/sta/225
http://hdl.handle.net/10023/26445

Sensorimotor Interfaces: Towards Enactivity in HCI

Iain Carson

University of
St Andrews

This thesis is submitted in partial fulfilment for the degree of
Doctor of Philosophy (PhD)
at the University of St Andrews

September 2021

Abstract

This thesis explores the application of enactive techniques to human computer interaction, focusing on how devices following ‘sensorimotor’ principles can be blended with interface goals to lead to new perceptual experiences. Building sensorimotor interfaces is an exciting, emerging field of research facing challenges surrounding application, design, training and uptake.

To tackle these challenges, this thesis cuts a line of investigation from a review of enactivity in the related field of sensory substitution and augmentation devices, to a schematic taxonomy, model and design guide of ‘the sensorimotor interface’; developed from a theoretically-grounded, enactive approach to cognition. Device, interaction and training guidelines are drawn from this model, formalising the application of the enactive approach to HCI.

A readily-available consumer device is then characterised and calibrated in preparation for testing the model validity and associated insights. The process highlights the effects of accessible, easily-implemented calibrations, and the importance of mixed-method approaches in assessing sensorimotor interface potential.

The calibrated device is utilised to conduct a detailed, methodological investigation into how concurrently available sensory information affects and contributes to uptake of novel sensorimotor skills. Robust statistical modelling concludes that sensory concurrency has a profound effect on the comprehension and integration of enactive haptic signals, and that efforts to carefully control the nature and degree of sensory concurrency improve user comprehension and enjoyability when engaging with novel sensorimotor tasks, while reducing confusion and stress.

The work is concluded by speculation on how the presented derivations, methods and observations can be used to directly influence future sensorimotor interface design in HCI.

This thesis therefore constitutes a primer to the principles and history of sensory substitution and augmentation, details the requirements and limitations of the enactive approach in academia and industry, and brings enactivity forward as an accessible, viable and exciting methodology in interaction design.

Acknowledgements

This work was supported by the University of St Andrews (School of Computer Science).

Academic

I would like to thank my supervisors, Aaron Quigley, Uta Hinrichs and Loraine Clarke, for their enthusiasm and support at every stage of this project. Whether face to face or spanning an 11 hour time difference, each of you have dedicated time and attention to helping me realise valuable contributions in this exciting yet very niche, unfamiliar field.

Without Aaron's experience and regular ideation and inspiration, I would not have known where this journey would begin, or end up. Collaborating with Uta was critical in the design and completion of the most challenging user study I have conducted to-date. Her requests for more clarity will forever guide my scientific writing. And finally, Loraine must be commended for her unquestioning engagement with the project in its final stages, offering invaluable support in the completion of this thesis.

I also extend sincere thanks to Miguel Nacenta and Julie Harris, for specific technical advice and assistance towards the practical experimentation, and their patient dedication of time beyond responsibility.

Personal

Thank you, Reda and Ken Carson, my loving parents, for believing in me throughout. In the moments where I most doubted myself, you never did.

Thank you to everyone at St Andrews dept. of Computer Science for the space you created for me. To Alice, Xu, Ryo, Julian, Nnamdi, David and uncountable other peers who inspired, and shared cakes, lunches, tech support, and long afternoons in the lab fixing my terrible code.

To the friends and housemates who joined me on this highly unpredictable ride. Fionnghuala, Montse, Zac, Bjarne, Marta and Haley - together we built a home of shared food, walks and laughter strong enough to withstand a global pandemic.

Finally, the dedication to my future self? This thesis is a simple reminder of what I am capable of when I put my mind to it.

Candidate's declaration

I, Iain Carson, do hereby certify that this thesis, submitted for the degree of PhD, which is approximately 59,000 words in length, has been written by me, and that it is the record of work carried out by me, or principally by myself in collaboration with others as acknowledged, and that it has not been submitted in any previous application for any degree. I confirm that any appendices included in my thesis contain only material permitted by the 'Assessment of Postgraduate Research Students' policy.

I was admitted as a research student at the University of St Andrews in September 2017.

I received funding from an organisation or institution and have acknowledged the funder(s) in the full text of my thesis.

Date 8/03/2022 Signature of candidate

Supervisor's declaration

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of PhD in the University of St Andrews and that the candidate is qualified to submit this thesis in application for that degree. I confirm that any appendices included in the thesis contain only material permitted by the 'Assessment of Postgraduate Research Students' policy.

Date 8/03/2022 Signature of supervisor

Permission for publication

In submitting this thesis to the University of St Andrews we understand that we are giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. We also understand, unless exempt by an award of an embargo as requested below, that the title and the abstract will be published, and that a copy of the work may be made and supplied to any bona fide library or research worker, that this thesis will be electronically accessible for personal or research use and that the library has the right to migrate this thesis into new electronic forms as required to ensure continued access to the thesis.

I, Iain Carson, have obtained, or am in the process of obtaining, third-party copyright permissions that are required or have requested the appropriate embargo below.

The following is an agreed request by candidate and supervisor regarding the publication of this thesis:

Printed copy

No embargo on print copy.

Electronic copy

No embargo on electronic copy.

Date	8/03/2022	Signature of candidate
Date	8/03/2022	Signature of supervisor

Underpinning Research Data or Digital Outputs

Candidate's declaration

I, Iain Carson, understand that by declaring that I have original research data or digital outputs, I should make every effort in meeting the University's and research funders' requirements on the deposit and sharing of research data or research digital outputs.

Date 8/03/2022 Signature of candidate

Permission for publication of underpinning research data or digital outputs

We understand that for any original research data or digital outputs which are deposited, we are giving permission for them to be made available for use in accordance with the requirements of the University and research funders, for the time being in force.

We also understand that the title and the description will be published, and that the underpinning research data or digital outputs will be electronically accessible for use in accordance with the license specified at the point of deposit, unless exempt by award of an embargo as requested below.

The following is an agreed request by candidate and supervisor regarding the publication of underpinning research data or digital outputs:

No embargo on underpinning research data or digital outputs.

Date 8/03/2022 Signature of candidate

Date 8/03/2022 Signature of supervisor

To my parents, and my future self.

Publications

Some of the research in this thesis has previously been published in peer-reviewed journal proceedings:

Carson, I., Quigley, A., Clarke, L., & Hinrichs, U. 2021. *Investigating the Effect of Sensory Concurrency on Learning Haptic Spatiotemporal Signals*. In Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 5(1), 1–30.
DOI=`https://doi.org/10.1145/3448102` **(Chapter 6)**

CONTENTS

LIST OF FIGURES

LIST OF TABLES

DEFINITIONS

Cognitivism Cognitivism refers to the collection of theories of mind which model cognition as the computational manipulation of symbolic representations. In this approach, perception and action are somewhat separated, with environmental stimuli being perceived passively and consciously interpreted before action is taken.

Enactive Interface EIs follow enactivist, engagement-driven (rather than cognitivist or representation-driven) approaches to interface design. Froese defines an EI as 'a technological interface that is designed for the purpose of augmented sense-making' [62, p368].

Enactivism Enactivism (enactivity) refers to the collective approaches and paradigms of cognitive science which argue that cognition is a dynamic process that does not need to rely on representations. In Enactivism, sense-making and the extraction of environmental information by an agent or organism derives from the process of goal-driven interaction and engagement, rather than comprehension of explicit, representative information.

Sensorimotor Contingency (SMC) Under Sensorimotor Theory, SMCs represent the lawful relations between one's own actions, and the resulting changes in sensory signals.

Sensorimotor Interface (Term introduced in this thesis) A sensorimotor interface is a technological interface (device or system) built in concordance with the principles of Sensorimotor Theory and associated, compatible theories of mind, to alter and augment user perceptual experience.

Sensorimotor Theory (SMT) O'Regan and Noë's Sensorimotor Theory (of Vision) [133] is a specific enactivist model for cognition which models perception and cognition as arising from dynamic agent-environment interactions. Sensorimotor theory specifically frames perceptual experience as the goal-directed exercise of mastery over lawful action-sensing relations termed 'sensorimotor contingencies'.

Sensory Augmentation Sensory Augmentation refers to the artificial extension or augmentation of human perceptual or sensory abilities beyond those of natural, unimpaired human capabilities, through extension of the principles pertaining to Sensory Substitution.

Sensory Substitution Sensory Substitution is a technique and mechanism by which environmentally-sourced information usually associated with one sense or modality is received through another.

CHAPTER ONE

INTRODUCTION

> We see with the brain, not the eyes
>
> *Paul Bach-y-Rita*

In the search for ever richer connections between humans and machines, a new generation of interactive principles and associated devices is emerging as a potential solution. These systems derive from principles of enactivity: a philosophical school of thought holding activity central to the emergence of meaningful perceptual experience [48]. Such systems encourage engaged, active interactions for task assistance through augmentation of perceptual experience, rather than defaulting to the presentation of more information, opening up new pathways to interaction design in HCI.

Perhaps the most well-recognised enactive system under continuing utilisation for research and application in HCI is the ‘feelSpace’ belt, a commercialised variant of an experimental device first studied in 2005 [124]. Such a device, in its current, publicly available form [1], provides navigation and orientation information through vibrotactile motors located in a belt worn around the torso. Cyclists wearing the device perceive the silent, abstract vibrations which communicate information relevant to their journey, such as directions. These signals are simple and limited in their capacity to communicate information, and offer at best a minor improvement in activity experience and safety. Investigations into how the belt may be re-purposed and adapted to various other applications, with potentially richer information communication (and associated increases in general utility) are still in progress [93].

However, more generally, systems such as feelSpace demonstrate an approach to interactivity which sits in contrast to the majority of interfaces we encounter. Their interactive principles derive from a different branch of cognitive science, generally termed an ‘enactive’ approach [105], which seeks fundamentally to alter perception via active engagement, rather than adopting the more ‘cognitive’ approach of presenting additional information (usually visually) for interpretation.

Taking an enactive approach to interaction design presents a key advantage. In a world where screens are becoming larger and more detailed, the visual field is increasingly over-utilised, with additional text and graphics being the easy, default option for the presentation of information pertinent to cognitive tasks. Enactive interfaces promise alternative methods for the presentation of salient information, and in HCI, is therefore slowly emerging as a potential technique applicable to scenarios where users are already suffering from information overload, recognised for its suitability across sense-making and multimodal interaction [62].

Enactivity is not new to cognitive science, nor interface design. A collection of closely related devices—Sensory Substitution and Augmentation Devices (SSADs)—has been researched and developed since the 1970s, often cited as beginning with the invention of the ‘Tactile Visual Sensory Substitution device’ (TVSS) [9]. The TVSS, a device invented by Paul Bach-y-Rita (the ‘grandfather of sensory substitution’) simply encoded a live image from a hand-held camera into tactile ‘pixels’ felt on the user’s back. It was devised to substitute vision as an assistive tool for the visually impaired, and Bach-y-Rita’s original study showed that with practice, users would develop a vision-like experience when using the device, reporting the sensation of ‘seeing with the skin’ [169, p23]. However, Bach-y-Rita foresaw applications far more diverse than simple sensory substitution. He was interested in the idea of not just substituting, but *creating* sensory experiences using similar techniques, and he envisioned such interactions as ubiquitous tools for future human-machine interfaces [12]. Given the trend of increasingly visual, complex information representations, it is difficult to imagine what it would be like to experience a world in which the majority of technologically-derived salient information is communicated in a natural, sensory fashion, through enhanced perception. With the current maturity of the technology, such visions are perhaps still limited to the realm of science fiction. In Appendix A, I provide a fictional, creative exploration of what such a world may look like, describing the unique interactions that may be possible through a vibrotactile wearable capable of providing contextually-aware, enactive, spatiotemporal feedback. But, even if we are to imagine such a world, and the application of Sensory Substitution and Augmentation techniques are a pathway to achieving such dreams, what more work needs to be done to realise this future?

For now, enactive systems are limited in application diversity and suitability, and in the depth and utility of information that can be communicated. Reviewing the literature reveals that development and adoption are challenged by a range of issues, from design, through application, to testing and development. In this thesis, I explore the various types of devices that exist in academia and commercially, and confirm that despite their advantages and promises, there is more to be done to develop the technology to a point where daily use is feasible and attractive, and to make progress towards ubiquitous presence.

Of critical relevance to HCI, I show that the underlying interactive principles of enactivity are fundamentally different to most other commonly exercised mechanisms and paradigms of information presentation, and that we do not yet have the tools necessary to design, evaluate and compare these alternative systems.

There is a spectrum between the models and approaches of enactivity being proposed in this thesis, and those encountered more widely in HCI [62][1]. As the main goal of this thesis is to increase feasibility and accessibility of enactive principles in HCI, I will try to bridge this gap through presentation of information across this spectrum.

Ground work must be done to establish how enactivity may be understood and applied from a human-machine interaction perspective. As such, this thesis aims to contribute by using explicit description of enactive models of perception to develop frameworks and guidelines for enactive interaction design. Following their derivation and presentation, the validity of these models and guidelines is reinforced through evaluation both experimentally, and in context of the wider literature. These efforts constitute a robust and consistent approach to formalising the application of enactivity to interaction design.

1.1 Thesis Scope

This thesis seeks primarily to bring the models and approaches of enactive perceptual psychology to the applications and ambitions of human computer interaction, clarifying the fundamental principles of enactivity and how it may be applied to the design of new interfaces. Through this work, I also hope to improve comparability of enactive devices in academia, and enlighten the development of enactive devices for practical applications.

Specifically, this work approaches the creation of new interfaces in HCI through the application of the aforementioned ‘perception as action’ sensorimotor theories found in enactivist (and other post-cognitivist or non-representationalist [105]) psychology. In combination with principles from this approach, my research addresses how consideration of perception as an embodied activity, rather than a passive experience, may offer new ways to interact with computers. This consideration is built upon to create a model which may inform device design, develop comparative metrics and identify suitable applications for enactive principles. The chosen model

---

[1] An example of a device which sits between the two would be the computer mouse. Buttons represent explicit action in a fashion largely cognitivist in nature, while the fluid, responsive movement of the cursor in response to real-world action represents a sensorimotor experience more compatible with enactivity.

is then both utilised and tested in user-experiments focusing on learning and training mechanisms, and the results offer confirmation of the model's effectiveness, and insights into how learning may be optimised. Altogether, this work intends to contribute to the application of enactivity as a technique for use in HCI, suggesting in an introductory (but well-grounded) format how enactive interfaces may be designed, evaluated, compared and presented so as to maximise the chance of their success in experimentation and commercialisation.

1.2 Research Motivation

Broadly, enactivity, as an approach to philosophy of mind, frames and describes mechanisms for how experience and perception arise. Performing research which utilises, expands on or critiques enactivity contributes to its accuracy and effectiveness in such descriptions, improving our fundamental understanding of ourselves and our experiential relationship with reality. This knowledge ultimately contributes to a description of how the human brain functions, and to the nature of experiential existence. While this general concept sounds (and can be) rather abstract and philosophical, specific knowledge surrounding perceptual experience is unarguably important to psychology and computer science, as accurate models of perception and the associated experiential qualities describe user abilities and experiences; key informers of interaction design and developmental and learning practices.

Though the insights describing perception that we may gain in doing so are fundamentally interesting in and of themselves, tangibly, enactive interfaces may offer a pathway to realising new experiential interactions within an increasingly digital world, though some work is required to translate the philosophical perspectives of enactivity into guidelines which may inform application. Specifically and directly, two key challenges motivate the research presented in this thesis: firstly, determining the potential of enactive interfaces (knowing *what* enactive interfaces can be used for), and secondly, designing enactive interfaces (knowing *how* to make them effective in those use cases).

It remains unclear how enactivity may be applied to interaction design to create meaningful experiences. As noted above, in lieu of moving towards ever increasingly dense information in the visual field, enactive engagements offer the potential for communication of salient information in an uninvasive, on-demand fashion.

Enactive (and other post-cognitivist) theories of mind and models of perception that may guide us towards tackling these challenges are still actively under development [105]. Much practical investigation into enactivity through devices in these fields is dedicated to the testing and maturation of these models, mainly through the exploration and development of Sensory

Substitution and Augmentation Devices (SSADs). However, there currently exists no good definition, taxonomy or schematic for the modern SSAD, nor a clear summary of the principles of design and interaction required to allow an HCI researcher to start applying enactive techniques to interface research. Indeed, not all SSADs strictly follow principles of enactivity, leading to a degree of confusion which requires clarification.

Moreover, outside of Sensory Substitution and Augmentation, enactive interfaces are still in the early stages of development, and are far from maturity and ubiquitous use. Understanding which lines of enquiry are most worth pursuing, and the true potential of enactivity in new applications, requires identification and formalisation of their successful application to-date.

Following this, designing, building and introducing enactive interfaces requires addressing several further challenges. At the most basic level, research is required specifically on what types of source data and control or input mechanisms enable an enactive engagement, as well as practical testing of enactive interfaces against models of interaction. Equally pertinently, work is required to understand how such interfaces may be introduced to users, to minimise training time, maximise ease of use, and ultimately increase utility and acceptance. This requires further theoretical and experimental investigations into learning and training mechanisms. Given enactivity's origins in describing perceptual experience, independent of technology, progress made in describing learning mechanisms of enactive interfaces may also contribute to our understanding of how our natural senses form, develop, and result in our unique perceptual experiences. As such, discoveries may be valuable in understanding how wider sensorimotor skills may be learned, independent of the use of technology, and could inform the learning and training of wider sensorimotor skills; an interesting topic in a broad range of fields from sports and dance to music and architecture.

To summarise, improving the knowledge behind how such systems communicate information, and testing and refining them in context for maximum effectiveness, intuitiveness and a quick path to mastery, may open paths to commercialisation and further drive interest in the topic, as well as informing more fundamental practises in cognitive science. Such requirement for clarification on what enactive interfaces are and how they may be built and introduced effectively, forms the basis for the research questions in this thesis.

1.3 Research Questions

Here I will lay out the research questions, which have guided the theoretical and practical research undertaken herein, and may define the scope and contributions of the thesis:

1. How can we define and characterise enactive interfaces?
   Little literature specifies the construction and composition, interactive principles, limiting factors and suited applications of enactive interfaces. There is opportunity for clarification in the theoretical foundation of how enactivity may be applied to interaction design.

2. How can we construct and compare enactive interfaces, and apply enactivity in interface design?
   This question relates to the design and application of enactive interfaces, and the research methods most suited to comparing and improving them.

3. How can we improve the effectiveness of learning to use enactive interfaces?
   This question seeks to address how enactive interactions may be mastered through explanation, application and testing of learning mechanisms consistent with enactive theories of perception and cognition.

1.4 Research Approach

In light of these research questions, the first stage of engaging with enactivity in HCI involved a cross-disciplinary literature review (psychology, neuroscience and computer science) of the systems that most commonly and effectively follow principles of enactivity for information transfer—Sensory Substitution and Augmentation Devices. As part of this literature review I have investigated the history of SSADs and their success as rehabilitative and augmentative interfaces. Through this, I identified the need for greater specificity of models in applying enactivity to interaction design.

A potential key contribution was identified in the unification of terminology and standardisation of research approach in HCI, and the decomposition of the components and interaction principles of an effective, enactive interface. Sensorimotor Theory of Vision [133] (a relatively accessible position within enactivism with well-defined scope and proven applicability to SSADs) is therefore adopted as a suitable cognitive model for theoretical grounding of enactivity as an interaction mechanism. Development of the model in Chapter 4 results in the introduction of the ‘Sensorimotor Interface’ - devices that enable interaction in accordance with sensorimotor theory.

Steps were then taken to apply the principles of the sensorimotor interface to the characterisation and calibration of a commercial games controller, in preparation for its use as an enactive interface. Through a quantitative, controlled user experiment, I also explore perceptual biases and the effectiveness of communicating explicit information using this controller. It was concluded that calibration steps are indeed useful and effective in reducing perceptual biases, and that many metrics concerning the precision of signal interpretation do not provide good insights into the system's potential as a sensorimotor interface.

A longer-term mixed-method experiment was then conducted exploring the effect of visual information on learning of new haptic sensory signals through sensory concurrency. The calibrated games controller was used in a 4-day experiment, to explore whether a new sensorimotor contingency between haptics and character control could be learned faster using carefully controlled visual information in a virtual environment. The results show that, in corroboration with the cognitive models described earlier in the thesis, the effects are significant. This led to the conclusion that an understanding of existing perceptual biases and the integration of new signals into is an important consideration in the development of new sensorimotor interactions.

Finally, through discussion and presentation of hypothetical future experimentation, the impact and conclusions of the research are explored.

1.5 Thesis Contributions

This thesis aims to draw together the fields of HCI (computer science) and enactivity (psychology/philosophy), through application of enactive theories of mind to the design and development of systems for human-computer interaction. To summarise, this thesis contributes the following:

- *The Sensorimotor Interface - Theory contribution*
  The first major contribution of this thesis is a model-derived definition of the 'sensorimotor interface', synthesised from psychology and HCI, with a table of design dimensions to aid the application of enactive principles to interface design (Chapter 4). Deriving insight from enactivist models of perception (specifically, Sensorimotor Theory [133]), I present a clear and consistent human-computer interface model, detailing its components and its interaction principles, and defining the 'sensorimotor interface'. This model allows us to describe many enactive interactions and devices in a manner consistent with literature observations. This additionally clarifies how sensorimotor interfaces may be distinguished from other multimodal interfaces (including SSADs) and specifically addresses issues associated with testing and comparison which stem from inaccurate expectations of the device class. In its presented form, Chapter 4 also acts as a primer to the principles and design of sensorimotor interfaces, highlighting previous and potential issues and ideas in design, comparison, and application of a wide range of enactive principles to HCI. Through identification of the main hurdles and opportunities in the field, this work presents a pathway towards advancing research and wider deployment in both sensorimotor interfaces, and SSADs.

- *Calibrating and Measuring Vibrotactile Sensorimotor Interfaces - Experimental contribution* The second contribution of this thesis is a collection of methods for the calibration and optimisation of signals in enactive interfaces featuring vibrotactile outputs. A process for characterisation and calibration of (primarily ERM) vibrotactile motors is drawn from the literature, and directly applied to an off-the-shelf device in a demonstrated case-study. It is shown that these calibrations and preparations are historically either ignored, unreported, or applied sporadically in existing literature, despite being well-established techniques. This is remedied through a (hitherto unattempted) theoretical explanation and experimental confirmation of the importance of this calibration in the context of Sensorimotor Theory, and a summary of the steps taken for future reference (Chapter 5). This work also represents an experimental demonstration of the limitations of human performance in explicit information extraction from enactively-encoded drive signals, resulting in a criticism of some of the existing methodology used to quantitatively investigate enactive interfaces, and reinforcing the importance of the diverse methodology involved in the final contribution.

- *Investigating Mechanisms for Learning of New Sensorimotor Skills - Experimental contribution* Learnability has been identified as a major barrier to the wider uptake of enactive interfaces. Approaches to improving the learning experience and reducing training time have been addressed in a general sense through literature reviews [19, 114], but these reviews suggest the field requires experimental contributions providing solid evidence for effective learning mechanisms. As such, I draw from principles in ‘Direct Learning Theory’ [75] (a theory compatible with enactivism which aims to explain how new sensorimotor skills are learned) to hypothesise an effective learning condition which leverages vision as a feedback source (utilising sensory concurrency). This condition is then evaluated experimentally against other conditions historically utilised in enactive SSAD training environments. My findings show, in accordance with direct learning theory, that the visual information available during task performance has a profound effect on how well new sensory signals can be interpreted when visual information is removed. This contribution adds experimental weight to the validity and utility of sensorimotor theory, direct learning theory, and my model of the sensorimotor interface, and also highlights the promise of controlled sensory concurrency in improving sensorimotor interface and SSAD training.

1.6 Thesis Structure

Figure 1.1 shows how the various chapters span the fields of cognitive psychology and HCI, colour coded by contribution type. The first two chapters detail the literature and theory from

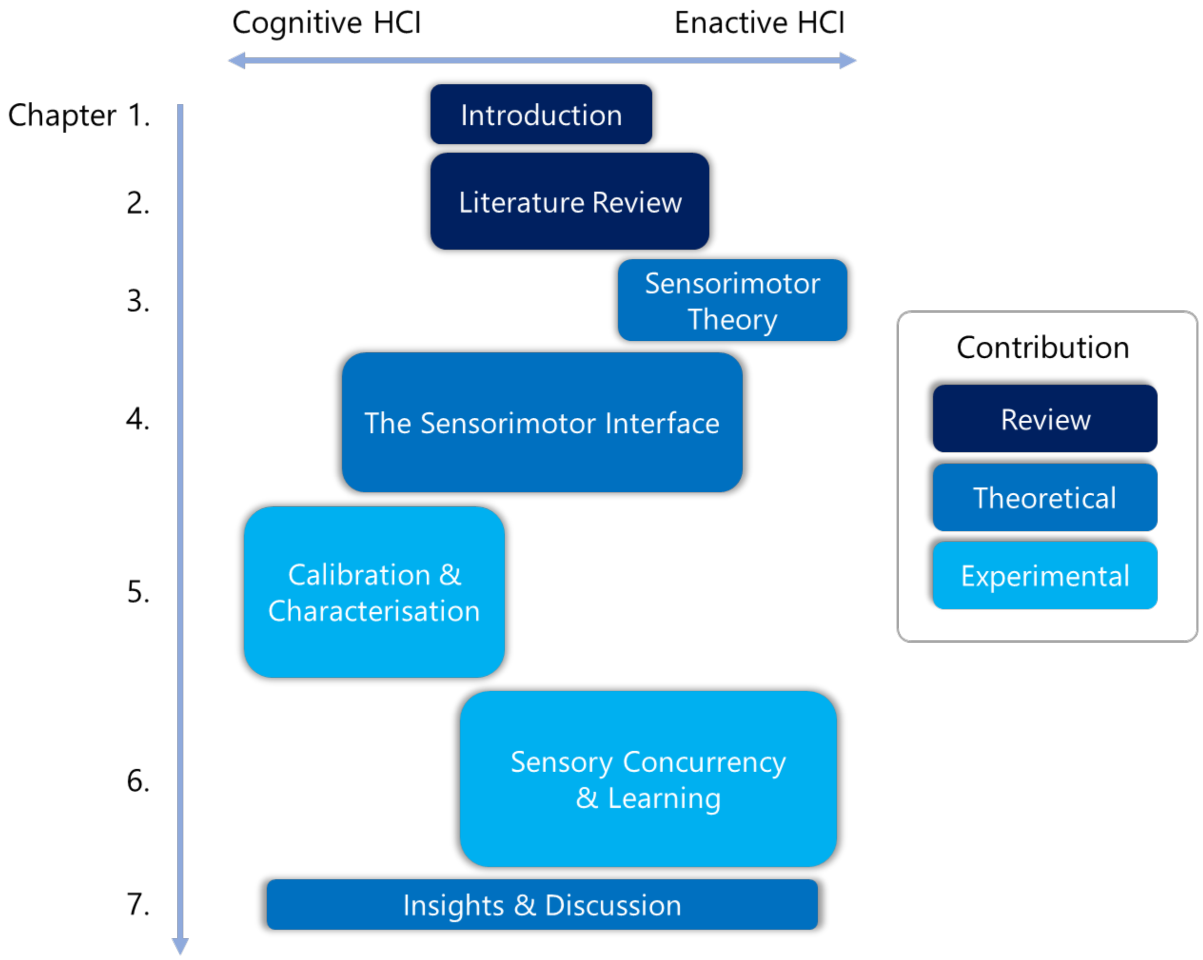


Figure 1.1: Structure of the thesis showing chapter contribution type, in context of their place in the spectrum between cognitivist and enactivist research in HCI.

which these interfaces are derived, forming the groundwork for a taxonomy which spans the field. Experimental work then moves from demonstrating methods more consistent with evaluations familiar to HCI, before more suitable methods for evaluation of enactive interfaces are adapted and implemented in Chapter 6. Contribution develops from early literature explorations, through theoretical contributions and ends with involved theoretical and experimental explorations.

In the final chapter, I further discuss and evaluate how enactive principles may be applied given the insights of the preceding chapters.

The dissertation is split into two parts, theoretical and practical: Following this introduction, the first part (Chapters 2-4) begins an exploration of enactivity in HCI through a review of the history and philosophy of Sensory Substitution and Augmentation devices; highly relevant and influential systems that may be drawn on to inform design of enactive interfaces. The literature review and associated theory of cognition are then synthesised into a model for applying enactive principles in HCI, leading to a definition and detailed taxonomy of the ‘sensorimotor interface’ - a term introduced in this thesis. In the second part (Chapters 5-6) I outline experimental work which has been undertaken to explore emerging aspects and unknowns of sensorimotor interfaces in HCI, contributing to growing knowledge in the field. Chapter 7 concludes the thesis with a discussion of the findings and how they affect applications and scenarios of how this technology may manifest in the future.

Chapter 2 extracts the sentiments and goals of the ‘sensorimotor interface’ through a literature review. A history of SSADs, early principles, and their emerging application in HCI is presented, highlighting their relevance to the research on enactivity with reference to existing devices. At this stage, I highlight how and why enactive interfaces require models of cognition that differ from those for explicit, cognitive interfaces we more often study in HCI, and why a different approach to experimentation is necessary, including new metrics, and a paradigm shift in development and evaluation for successful adoption.

Chapter 3 looks specifically at Sensorimotor Theory and its place among similar enactive theories of mind. The choice of theory is justified, its implications considered and its utility in comparison to existing theories of mind is evaluated. Adoption of Sensorimotor Theory as a framework for cognition allows modelling of enactive interactions, leading to direct principles and recommendation that may be applied to experimentation.

Chapter 4 synthesises the fields of HCI and Sensorimotor Theory from Chapters 2 & 3, resulting in a model for the sensorimotor interface and principles for its use. This chapter also summarises the current state of enactivity in HCI in terms of research metrics, comparisons, training, study design, systems/interaction approach and applications, offering a thorough introduction to the field from a multi-disciplinary setting. Finally, a critical consideration of the evaluation and measurement of sensorimotor interfaces is undertaken before a design document outlining the factors important to application of enactivity in interface design is presented.

Chapter 5 concerns the calibration of vibrotactile systems for use in enactive applications; utilising a combination of established techniques, and the principles discussed in previous chapters. I describe the preparation of an Xbox One controller for future experimentation, and participant trials are conducted to determine the effectiveness of these preparations (and the consequences of failing to) in a comparison of pre-post calibration signal perception. A discussion as to why this is important for enactive interaction design is presented.

Chapter 6 describes a study investigating the effects of sensory concurrency on the learning of a haptic spatiotemporal signal. The study directly addresses and confirms learning mechanisms and predictions drawn from the models built in previous chapters. This chapter contributes new paradigms for training and evaluation of enactive signals while showing the effectiveness of Sensorimotor Theory as a foundation for enactive interface design.

Chapter 7 revisits many of the insights presented in the thesis, exploring and discussing the further potential and limitations of enactivity in HCI. Finally, insights from the design principles and experimentation are outlined and utilised to guide towards future work and thesis conclusions.

CHAPTER TWO

SENSORIMOTOR INTERFACES—LITERATURE REVIEW

To address the first two research questions presented in Chapter 1, the first half of this thesis is concerned with the exploration of literature surrounding enactivity and sensorimotor interactions, and the history of related devices which have led to enactive interfaces in HCI. From the theoretical grounding developed in the following two chapters, I work towards defining in detail the Sensorimotor Interface in Chapter 4, presenting a model and series of guidelines upon which new interfaces utilising the principles of enactivity may be built for human-computer interaction.

With this in mind, the present chapter explores the wide literature surrounding enactive interfaces, starting with their origins in assistive technologies as Sensory Substitution and Augmentation Devices (SSADs). I will show how SSADs are often linked with enactive theories of cognition, and differ fundamentally in their approach to interaction and information transfer to the cognitive, symbolic systems and user interfaces more commonly encountered in HCI. Analysing the history and inspiration, design and operating principles of SSADs is therefore extremely useful in illustrating the consequential differences between enactive and cognitive approaches to interaction design, and establishes firm foundations for later work.

2.1 Introduction

Figure 2.1 illustrates how this literature review contributes to the wider goal of this thesis: to strengthen the link between the principles of enactivity and the practical realms of HCI. This goal is approached through the description of ‘sensorimotor interfaces’, a term introduced in this PhD research to refer specifically to interfaces that communicate information through interaction in a manner consistent with O’Regan and Noë’s ‘Sensorimotor Theory of Vision’ (SMT) [133]. SMT is an enactive model of perception and cognition which will be covered in detail in Chapter 3 before being applied rigorously to the development of new interaction techniques in Chapter 4.

Sensorimotor interfaces derive from and are inspired by devices built on the principles of enactivity. Enactive devices generally feature continuous, sensory-like outputs with implicit meaning, which react dynamically to user action and the environment. An excellent example of this is The Enactive Torch [62], which provides distal touch information through vibrotactile feedback, and will be described in detail in the next section. These can be contrasted with

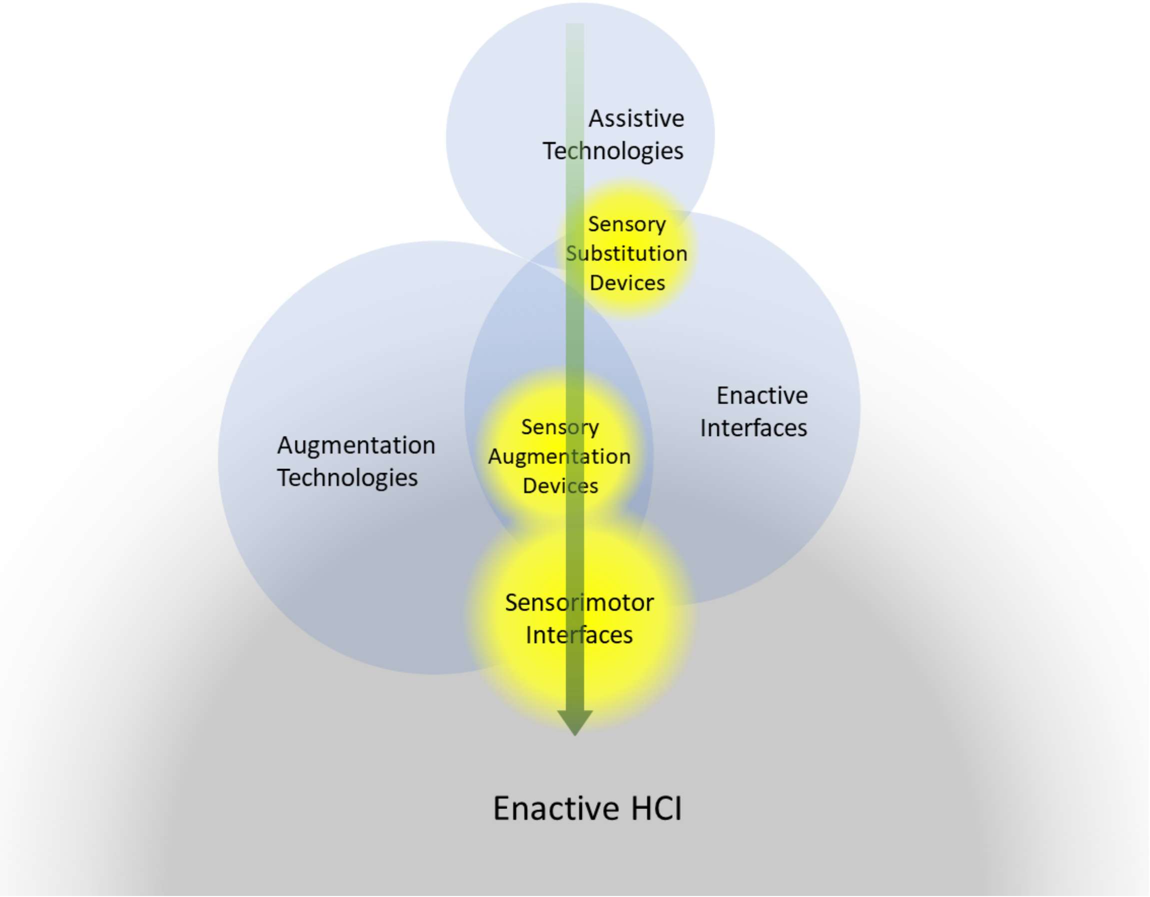


Figure 2.1: Trajectory of the background work of this thesis, through topics of application and research.

the ‘cognitive’ interfaces we are perhaps more familiar with in HCI, which feature symbolic representations of information for cognitive interpretation, explicit comprehension and utilisation. An example of this would be the icons on a PC desktop, which explicitly symbolise and communicate their access to the applications or files ‘within’.

Despite being relatively niche and under-explored, enactive devices manifest as tools in cognitive and perceptual psychology [137], while also finding documented applications in applied medicine (such as Schoonmaker and Cao’s vibrotactile feedback device for communicating surgical tool pressure [148]) and physiotherapy (for example Pan *et al.*’s skin-stretching balance rehabilitation device [134]). Alongside being used as tools for human studies, they are built, reviewed and documented in HCI across fields such as immersive analytics (where use of VR is especially suitable [119]) and industrial or commercial applications for emerging interaction techniques (notably driving [93] and piloting [123, 162]). They are, thus, represented by a wide field of potentially relevant literature.

A large concentration of research related to enactive principles in device and interface design is found in the study of Sensory Substitution and Augmentation Devices (SSADs). Sensory substitution systems are often loosely defined as devices which translate information usually associated with one modality to signals detectable by another [100, 163], with sensory *augmentation* being the term applied when the source of these signals reflects environmental information usually inaccessible through the natural senses (such as distal touch, magnetic fields, or other artificial data). However, the exact definition of ‘information’ and ‘modality’ in this context, and what it means for a signal to be ‘sensory’ are ambiguous and not entirely compatible with widely-held cognitive models of interaction in HCI.

The idea that information which is somehow ‘sensory’ in nature may be communicated through multimodal interfaces is an exciting prospect for the future of HCI, and represents much of the source of inspiration for existing enactive devices in the field. Defining exactly what a modality is, in the context of Sensory Substitution and Augmentation, and thus deriving and modelling sensory interactions more broadly in preparation for their utilisation in interface design, is grounded in the theoretical reviews of Chapter 3, and constitutes the basis of the modelling work undertaken in Chapter 4 to build back up to a full picture of sensorimotor interfaces.

But before I approach the details of *how* to build sensorimotor interfaces and bring enactive approaches to HCI, it is helpful to start with an exploration of the devices that over the past few decades have inspired and influenced their creation, starting with SSADs. Such an exploration will detail the motivation and inspiration for the appropriation of enactive principles into Human Computer Interaction, while revealing the open questions, difficulties and limitations in the field.

2.2 Sensory Substitution and Augmentation Devices

The related, but distinct, techniques of Sensory Substitution and Augmentation arguably originated with the efforts of Paul Bach-y-Rita [11]. Bach-y-Rita noted the ability of blind users to proficiently 'perceive' the world by distal touch, through use of the white walking cane, and, similarly, how Braille enables reading in place of inked letters. Both technologies, though simple, represent a substitution of a vision-like experience through touch. From this, Bach-y-Rita suggested that 'we see with the brain, not the eyes' [13, p285], and questioned whether technology and electronics may be used to enable similar perception-altering experience, with the goal of creating new assistive technologies for the sensorily impaired.

In 1969, Bach-y-Rita demonstrated 'visual substitution' using the Tactile Vision Substitution System (TVSS) (illustrated in Fig. 2.2), a vibrotactile display unit designed to allow visually impaired people vision-like experiences. The TVSS presented camera-captured pixel data directly onto a tactile display unit mounted on the user's back as a spatio-temporal vibrotactile signal. With time and practice, this signal was interpreted by users, and attributed reliably to external stimuli, allowing object recognition and basic appreciation of environmental topography in a fashion similar to vision.

This was surprising to the neuroscience community which, until that point, had largely considered the structure and function of the brain (and, thus, sensory experience) to be fixed and immutable [15]. It was observed that events such as trauma and stroke can cause irreversible damage to the brain, and if such events affected sensory experience, it was widely accepted that restoration or rehabilitation of the experiential component of sensing would require biological rejuvenation of the damaged tissue itself.

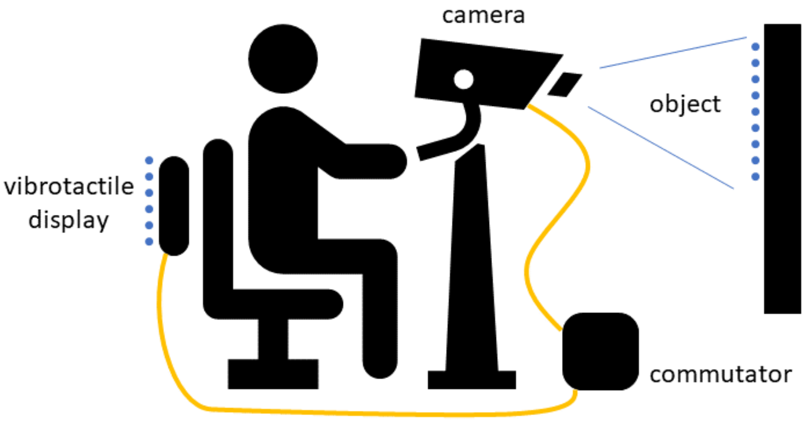


Figure 2.2: Schematic of the original TVSS.

However, the success of the TVSS in allowing visual information to be interpreted through the skin highlighted the ‘plasticity’ of the brain in working with sensory information. Developing this idea through his later work [14], Bach-y-Rita explored how the brain may be able to work with various types of sensory information through such sensory substitutions, and highlighted the technique’s potential for allowing non-invasive sensory compensation, rehabilitation, or even augmentation beyond the natural senses.

This early work in Sensory Substitution and later, Sensory Augmentation, paved the way for the creation of new interaction techniques leveraging neuroplasticity to allow interpretation of new sensory information. Interaction with natural or computerised environments through artificial apparatus, in a fashion resembling or approaching the experience and fluency of natural interactions, is an attractive prospect. However, the details and true potential of such interactive techniques remain a topic of ongoing research, motivating the work presented in this thesis.

Throughout this chapter, I will take a closer look at several prominent SSADs, starting with a select few exemplary systems. Through unpacking existing work on SSADs I explore common definitions and technical approaches, highlighting the potential impact outlined through empirical studies and revealing the gaps and limitations of previous work.

2.2.1 Sensory Substitution

The TVSS established the general technical framework upon which many subsequent sensory substitution devices would be built: a sensor such as a camera captures environmental information, which is fed for processing by an electronic coupler, which coordinates the activation of a stimulator. The stimulator typically targets a biological receptor (such as the retina, ears or a tactile mechano-receptor) to allow the devices’s output signal to be sensed, thus ‘substituting’ the modality usually associated with the input signal (in the case of the TVSS, vision) with the targeted modality (in the case of the TVSS, touch).

However, it was noted early on in users trials that simply presenting the visual image through the tactile array was not enough for perceptual experience to emerge. Bach-y-Rita describes how, in his original study [11], the camera of the TVSS was at first fixed. Users presented with the tactile stimulation from this device displayed very limited ability to discriminate the stimuli. It was not until users were handed control of the camera, performing explorative movements under observation of the changing stimulus, that perceptual experiences started to emerge. This discovery revealed the now widely-accepted belief that ‘action’ is required for users to learn the meaning of substituted sensory signals [6, 96]. I interpret from this that action-centred, ‘enactive’ models of perception may be required to fully understand how sensory substitution techniques

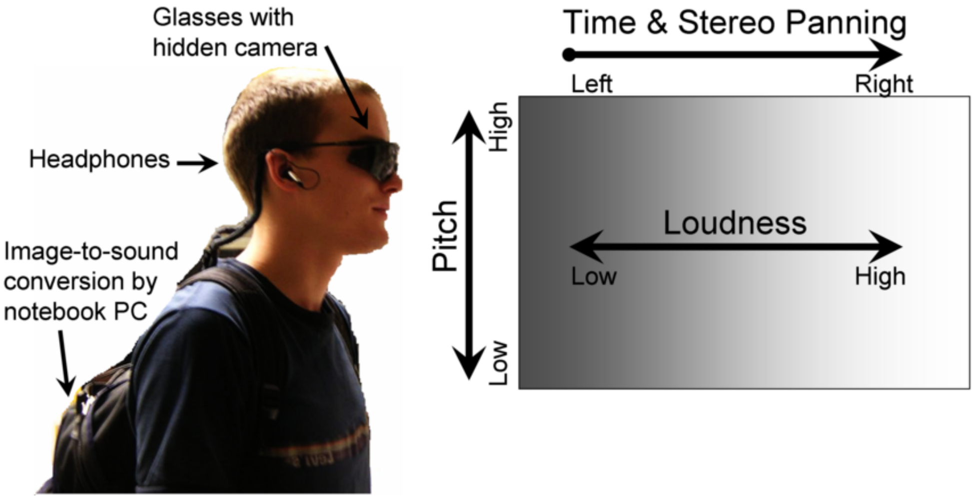


Figure 2.3: From Proulx *et al.* [138]. The vOICe SSAD [117]. (a) The device camera is mounted in the glasses, headphones play the result of the image-to-sound conversion. (b) Encoding principles.

enable communication of environmental information.

An example of a more modern device that shares the principles of the TVSS is the vOICe [117]—a device worn as glasses which encodes a forward-facing monochrome camera's image into an audible *soundscape* for interpretation. The camera's image is encoded and heard column-by-column, panning from left to right over 1s, with bright pixels sounding loud and pixels high in the view being higher in pitch (see Fig. 2.3). This coupling process (i.e., converting the 2D camera image into this audible soundscape) involves a degree of algorithmic dimensional reduction such that, unlike the TVSS, the final output signal is of quite different quality to the original image.

Short-term use of The vOICe (˜15 hours) has allowed simple object identification [5] in the fashion and timescale comparable to early studies with the TVSS [11]. Long-term users of The vOICe have reported phenomenological changes in perceptual experience [167], with self-reported experiences when using the device resembling those of the substituted modality (vision) rather than the substituting modality (hearing).

Importantly, it must be noted that action has again been found to be critical in the learning of the vOICe's presented signal [5]. Separately, Reynolds and Glenney confirm that users trained through active, dynamic control of The vOICe develop significantly greater object identification and orientation skills than users trained through presentation of exemplar images for recall [142].

Despite the differences in form and its more involved encoding processes in comparison to the TVSS, the success of the vOICe renders it a notable vision-substitution device, offering effective independence and vision-like sensibilities to both congenitally and late-blind users.

As with the TVSS and the vOICe, most sensory substitution devices share the goal of replacing or compensating for impaired senses, promising users a degree of functional autonomy and increased independence without significant compromise or invasive surgery. These devices often use a camera as their detector, due to their origins in assisting visually-impaired users. However, as observed by Bach-y-Rita, this is not a strict requirement. For example, tactile-to-tactile sensory substitution has also been demonstrated, aiming to assist sufferers of somatosensory impairments [10].

Indeed, Bach-y-Rita and several other authors [46, 167] have suggested that older, non-technological techniques may also be considered sensory substitutions, such as Braille (a substitution of reading from vision to touch), or even reading itself (as a substitution of the audible word to vision). The clear definition of what a sensory substitution is, and a definitive description of the devices that enable it, is a continuing topic of debate in the field, and one that I address in more detail in Chapter 4.

The perceptual changes experienced by users of the TVSS and The vOICe suggest the brain has astonishing flexibility in adapting to integrate new sensory information sources. Environmental information of high quality and utility is being made available to users of these systems through an implicit, sensory fashion, inspiring wider use of sensory substitution for generalised interface design. However, if we are to generalise such devices to give access to information beyond the natural senses, we must identify what types of signal and information are suitable for perceptual integration, and explore and understand the biological and cognitive rules that govern this process of learning and integration.

This exploration has, to some degree, been a focus in the related field of Sensory Augmentation (see Fig. 2.1), concerning the devices and use cases which feature a source signal not usually accessible to (unimpaired) human perception.

2.2.2 Sensory Augmentation

The past decade has seen the growth in the topic of Sensory Augmentation, a field closely-related to Sensory Substitution which aims to use similar technologies and approaches to expand human perception beyond what is naturally possible.

In 2005, the Feelspace project [124] utilised a belt of vibrotactile actuators to communicate global magnetic field information; a sort of sensory compass. The belt features 13 vibrating elements and delivers continuous global orientation information as vibrotactile stimulation to its wearer by activating the element pointing north.

While such a magnetic sensibility has been noted in the animal kingdom, it has not been reliably observed in humans [96]. The successful integration of such a signal into general perception would therefore represent a ‘sensory augmentation’.

As such, long-term experimentation was conducted to evaluate the utility of such a signal in navigation tasks, while also furthering our understanding of how such novel sensory information integrates with the natural senses to contribute to human perception and consciousness [124]. Users consistently wore the belt for six weeks, conducting simple pointing tasks and outdoor exercise sessions daily. Similarly to long-term users of the TVSS and The vOICe, users of the feelSpace belt who engaged in these tasks reported a change in perceptual experience [124], and later studies also observed corresponding changes in brain activity through MRI scanning [96]. However, unlike the vision-like experience associated with the sensory substitution systems, users of the feelSpace reported the development of a *new* sense of spatial perception. They were further able to utilise the signal in a meaningful way during performance of basic navigational tasks (such as completion of basic paths, and pointing to their origin following random displacement) [124]. The authors claim that these perceptual and behavioural changes are consistent with ‘perceptual integration’ of the compass signal. Such insights further reinforce both the potential effectiveness of such an approach to multimodal interfaces, and the motivation for exploration and adaptation of SSAD-derived technologies for task enhancement in wider applications.

The feelSpace belt has since been re-used in several others studies [17, 92, 93, 164], and indeed commercialised as a general navigation aid [1]. Some studies use the belt to investigate the practicality and effectiveness of sensory augmentation for new applications, such as driving [93]. In other studies, the feelSpace belt has been used to test and develop theories of cognition seeking to explain the natural senses and origins of perception [96]. One such theory is O’Regan and Noë’s Sensorimotor Theory of Vision [133], which, though referring directly to vision, may be described as an enactivist model of general cognition. I will cover sensorimotor theory in detail in Chapter 3, before applying its principles and related literature to create a model for enactive interactions in Chapter 4. For now, it is helpful to note the existing link between technological sensory augmentation and sensorimotor theory.

2.3 Defining the SSAD

Although the TVSS, The vOICe and the feelSpace belt have been described as SSADs, there is no clear, agreed-upon definition of what constitutes an SSAD to be found in the literature.

A search through SSAD-tagged publications returns a broad suite of devices, ranging from digital assistive technologies similar in goal and approach to the TVSS, to tools designed to test and develop theories of enactivity and cognition more generally [62, 124]. Many of the industrial and commercial applications feature devices that have been adapted from SSADs [1, 93], although it is rare for them to be referred to directly as such. The omission of the SSAD label in these cases may suggest that there is a recognition of some difference between an SSAD and a generally SSAD-like human-machine interface, reflecting how different developmental work is required for the successful application of SSAD principles beyond assistive or academic use cases.

Commercial interfaces derived from SSADs typically use continuous, dynamic signals to attempt to augment perception directly, tackling information bottlenecks in high bandwidth applications [93] or striving for increased immersion in augmented and virtual reality [135, 142].

Visell provides a clear definition for sensory substitution typical of the field: *"Sensory substitution refers to the translation of sensory information that is normally available via one sense to another [163, p38]"*. Most researchers share this sentiment; that *Sensory Substitution* requires the presentation of information usually associated with one *sense* or *modality* to another [6, 111, 163]. While *Sensory Augmentation* is far less frequently summarised so concisely, the term usually indicates that the origins of the source information are not naturally sensible by humans [6, 62, 96, 145].

Visell [163] notes that these *operational* definitions do not offer any pretence or promise of the effectiveness of devices, nor do they provide clear insights into *how* or *why* the techniques may be useful or effective in interaction design. More immediately, such definitions are broad, offering much room for interpretation regarding what constitutes a 'sense' or 'modality', and what types of information may be suitable for substitution or augmentation.

Alternatively, Visell comments that SSADs can be defined by their *functional* properties, framing: *"sensory substitution as the provision of information to assist people with sensory impairments [163, p39]"*. Again, the nature of 'sensory' is unclear in this definition, and the inclusion of a purpose introduces ambiguity: Due to their strong overlap with (and origins in) assistive technologies, 'functional' definitions can also include devices which seek to replace sensory information, but are not enactive (such as descriptive audio navigation aids [22]). It can further include almost all devices through which information is communicated in multimodal fashion. Previous work in this area has introduced such interfaces to facilitate demanding tasks, such as sensory impairment aids like the Neosensory Buzz [127] which allows *feeling sounds* on the arm, or related efforts by Zhao [171] which targets language directly, some of which are enactive (e.g. the former) and some of which are not (e.g. the latter).

A more elegant and directly applicable approach to definition is provided by Froese in the introduction of The Enactive Torch, a 'tool for the science of perception' [62, p365]. Froese remarks that early SSADs such as the TVSS, vOICe and feelSpace belt were specifically designed to enhance or modify user's perceptual experience, and, in line with the action-centred observations detailed above, argues that it is the *enactive nature of their interaction* that enables this. The work leads to a definition of the 'enactive interface' as 'a technological interface that is designed for the purpose of augmented sense-making', with augmented sense-making being 'giving a person opportunities to create novel modes or modalities of perceptual interactions' [62, p368]. The enactive interface therefore defines a set of tools and devices featuring the enactive, perception-altering properties shared by many SSADs, but without the burdens of the field's history, and association with assistive technologies.

In similar fashion, in this thesis, rather than endeavouring to find or create a definition for the SSAD, I work towards introducing the term 'sensorimotor interface'. This term is used to refer specifically to the set of interfaces and devices which adhere to 'sensorimotor' principles of interaction, outlined in Chapter 3 and applied through Chapter 4.

The 'sensorimotor interface' I define bears close resemblance to Froese's 'enactive interface', though with more specific definitions and detailed origins. It further offers several expansions advantageous to generalised applicability in HCI, which are discussed through its description, and in Chapter 7.

To conclude, this thesis focuses on SSADs that apply enactivity with the aim to alter user perceptual experience, rather than those simply matching functional or operational definitions.

The rest of this chapter aims to introduce the relevant literature of SSADs, including reviews and individual examples and experiments, with a focus on the devices that exhibit enactivity or make perceptual claims.

2.4 Reviewing SSADs and Enactive Interfaces for HCI

The use of enactivity in HCI is not new, and nor are attempts to formalise and popularise enactivity alongside more classical interaction techniques (such as visualisations, symbol/icon display, gestures, or realistic simulations) as a technique for digital information communication or interaction.

Visell [163] presents a comprehensive review of tactile sensory substitution techniques and devices, in order to provide design guidance and to encourage the application and use of tactile displays in HCI. SSADs are framed as an organising perspective for understanding how enactivity may apply to interface design, beginning a coherent differentiation from the above classical interfaces commonly found in HCI. Alongside providing the aforementioned functional and

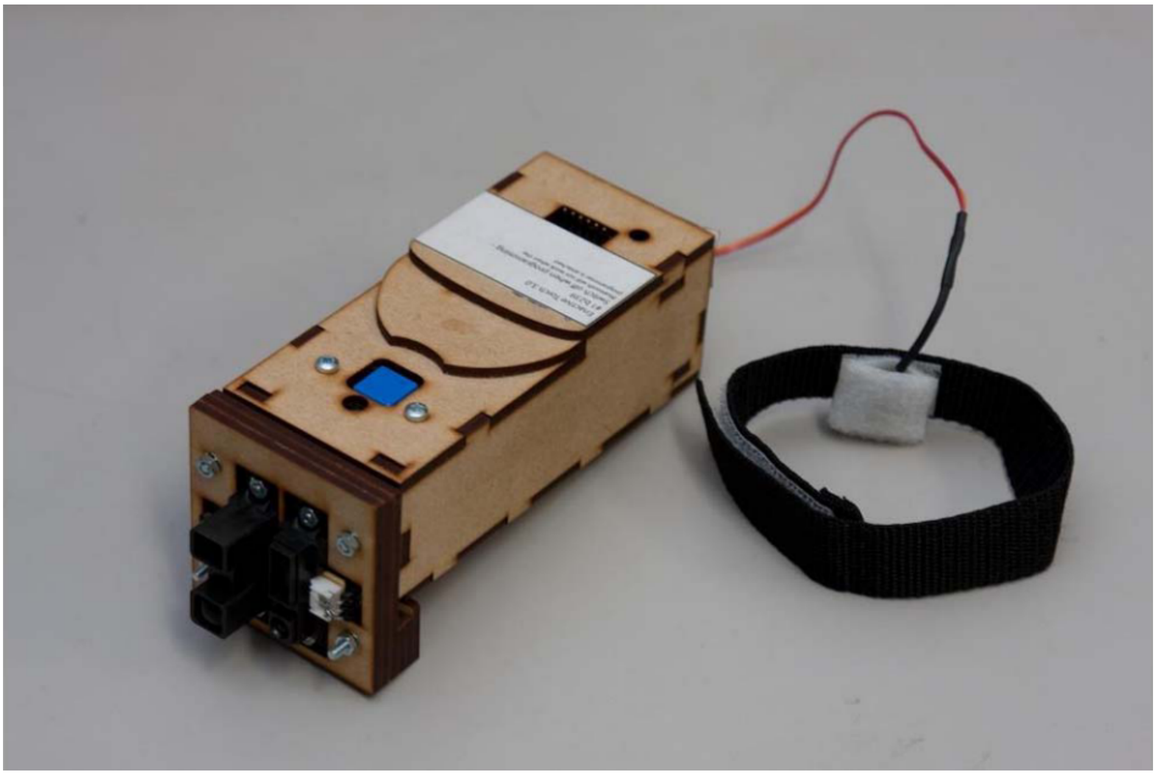


Figure 2.4: The Enactive Torch is a simple, hand-held enactive device, providing distal touch information through vibrotactile feedback. ©2012 IEEE. Reprinted, with permission, from Froese *et. al*, *The Enactive Torch: A New Tool for the Science of Perception*, IEEE Transactions on Haptics, Fourth Quarter 2012 [62]

operational definitions of the SSAD, Visell recognises SSADs as devices suitable for continuous (rather than symbolic) information, and emphasises the importance of action in development of user proficiency. Such requirements immediately highlight that the acquisition of information from enactive interfaces benefits strongly from active engagement, corroborating Bach-y-Rita's observations of TVSS users [11], among others. Yet, Visell highlights that the mechanisms by which these benefits arise are not clear.

Froese's seminal work introducing the Enactive Torch [62] offers both a helpful introduction into the application of enactivity to HCI and an excellent example of a simple enactive interface. The presented version of the Enactive Torch (shown in Fig. 2.4) features a simple forward-facing infrared distance sensor, coupled through a continuous inverse mapping to the intensity output of a vibrotactile motor. In this fashion, when held, the torch simply provides a single measure of distance. It is therefore argued that the determination of an object's location or identity using the torch requires active exploration through user movement, thus rendering the device 'enactive'. Froese describes the use of such interfaces to attain environmental information (such as object distance and identity) through action as 'augmented sense-making'.

In both Froese's [62] and Visell's [163] work, enactive theories of mind are suggested as potential models for how SSADs may be adapted to interaction design in HCI, inspiring much of the work undertaken in this thesis. The work further highlights how SSADs and associated literature may be used to inform such interfaces.

The use of enactive principles in HCI has broad implications, affecting application selection, design and implementation, training, and evaluation. I will now outline these specific areas of impact.

2.4.1 Uptake of Enactive Interfaces

Despite decades of development, the uptake of SSADs and enactive interfaces remains relatively low, both in terms of their academic consideration, and in terms of wider user-acceptance [19, 65, 91, 137]. In reviewing how SSADs may be moved from research laboratories to wider adoption, the two main identified issues are **design** (caused by limitations in design guidelines and models [91]), and user **acceptance** (reflecting difficulty of use, and time taken to achieve proficiency [114]). Here I offer a more detailed look at these challenges and how they have influenced the research undertaken in this thesis.

Design Challenges. Perhaps in reflection of the lack of a good definition, it is hard to find a comprehensive schematic of the SSAD. Visell [163, p39] contributes a commonly-referenced schematic. While several design insights are provided in this illustration, there remains questions on how effective 'functions' of the coupling device may be chosen and implemented, and the true role of action in the learning of these systems. This thesis addresses these challenges by proposing a detailed schematic and interaction principles that are derived from enactive theories of mind.

Practical design considerations for SSADs are carefully detailed by Kristjánsson [91], who outlines the physical requirements of SSADs, with a focus on highlighting usability mistakes that can lead to poor device acceptance. Such insights include consideration of comfort and the avoidance of sensory overload through presentation of unnecessary information. Practical guidelines (such as ensuring comfort and optimising wearability) are useful, however, specifics on interaction principles are limited, leaving a gap between ideation and implementation.

Jones and Sarter [82] summarise the preparation and use of 'tactile displays' (i.e., arrays of multiple vibrotactile motors) for information representation, with specific reference to SSADs. Noting how such displays may be utilised in sensory augmentation, their literature review also contextualises the Tactile Situational Awareness System (TSAS) [123]: a vest-integrated tactile display designed to present information about the intended direction of the pitch and roll of an aircraft. The TSAS presents vibrotactile feedback continuously, and dynamically responds to sensor data from the aircraft, which is under direct control by the user. It therefore exhibits enactive properties, effectively communicating action-driven cues through rich, abstract signals. However, Jones and Sarter's review addresses tactile display communications through reference to tactons (symbolic haptic icons for communication of explicit information [23]), which are largely incompatible with the enactive approach. Such convolution of symbolic and enactive descriptions further highlights that enactive interfaces are often misidentified or misunderstood, and illustrates the need for clearer interaction design models.

Some attempts to tackle design issues in SSADs and to improve uptake by researchers have focused on the design of generalised SSAD toolkits. 'Toolkit for Experimentation with Multimodal Sensory Substitution and Augmentation' (TESSA) [146] and the TongueDuino [50] allow mapping of arbitrary input sources to vibrotactile or electrotactile outputs respectively, enabling prototyping of operational sensory substitution and sensory augmentation systems. However, while these devices tackle hardware issues well, there remains issues in the detailed choices accompanying design of the *interaction* offered by the system, which still inhibit usability and uptake. My work addresses this challenge by deriving interaction principles from enactive theories of cognition as described later in Chapter 4. Such principles could be explored through systems such as TESSA and the TongueDuino.

Learning and Training Challenges. The topic of training, and mechanisms for learning how to use SSADs and similar enactive interfaces, are topics of interest in the literature, as the long training time associated with SSAD mastery represents a significant barrier to uptake [19, 114]. Such exploration also plays a key role in this PhD research.

Several studies present SSADs, and then explore and compare the effects training has on user proficiency [6, 142, 143, 149]. Previous work has also presented findings on the comparison of different learning mechanisms in the context of SSADs, e.g. [62, 171].

However, the aforementioned surveys by Kristjànsson *et al.* [91], Auvray & Myin [6] and Visell [163] all highlight that a systematic approach to training and learning is still missing. Despite their age, these points are still relevant. Bertram [19] points out that there are currently no standards for training nor testing of SSADs, and that it is important to consider the goal of the training before deciding on its nature and duration. Bertram identifies feedback as being critical to the successful apprehension of SSAD signals, and lists a wide variety of explicit feedback types relevant to task execution or outcome. Given that the devices discussed in the review are generally enactive, featuring continuous, implicit outputs, it is surprising that the effect of more continuous, implicit feedback mechanisms (such as would be offered through active exploration) is not considered. As part of my PhD research, I explore the effect of implicit feedback in learning the interpretation of enactive interfaces as outlined in Chapter 6.

Reynolds & Glenney demonstrated that active training in a fast-paced game-like environment was effective even when of limited relevance to the evaluative tasks [142]. In their presented study, users of the vOICe demonstrated the best object-finding capabilities following a training session resembling interactive play. The mechanism by which their action-centred learning environment promotes acquisition of generalised skills warrants further investigation.

Maidenbaum *et al.* [114] investigated how the gap between research and widespread adoption of visual SSADs may be overcome. They suggest that an understanding of the neural mechanisms behind visual restoration in the blind is linked to the development of SSAD proficiency. Focusing on the overlap between neurological rehabilitation and the design of training protocols and models for (self-)learning of SSADs is an exciting and under-explored area of research, and reflects how models of perception, perceptual development and rehabilitation can inform the refinement of models for SSADs.

To this effect, Lobo [125] has also investigated how concurrent senses may affect SSAD learning, experimentally exploring the effectiveness of blindfolding when learning new visual-substitution devices. Blindfolding (visual sensory deprivation) is relatively effective in promoting the learning of new vision-like sensorimotor skills, resulting in greater proficiency than training without blindfolding. Such an approach is also justified through 'Direct Learning Theory' [75], a theory of learning compatible with enactivity which addresses how feedback may be implicitly gained during enactive task performance.

My work builds on these insights through theoretical and practical exploration of how learning theories compatible with enactive principles may inform the mastery of enactive interfaces. I will further detail and discuss direct learning theory in Chapter 3, and start to formalise and evaluate its implications for SSAD and enactive interface training through experimentation in Chapter 6.

2.4.2 Evaluating Enactive Interfaces

There is wide variation in experimental standards, methodologies and device capabilities across SSADs and other enactive devices, which has limited somewhat the coherence and trajectory of the field to-date. This might be partially responsible for the slow adaptation of enactive principles in interface design [19, 65, 91, 137]. An overview of the benefits and individual contributions of some of these studies will help in clarifying the path to establishing enactive approaches in HCI. In the following section, I group and outline studies by their approach to evaluating and proving the *feasibility, effectiveness, or usability* of enactive interfaces, with further notes on some more involved or philosophical approaches investigating how *perception* and *perceptual experience* are affected through device use.

Perceptibility Can I perceive and differentiate distinct signals?

A number of studies on SSADs have been conducted to assess the feasibility of a designed interaction from the perspective of whether its output signals are perceivable and differentiable. As part of this, it is assessed, for example if an SSAD's output signals can be differentiated, or if the designed interaction is otherwise capable of communicating the information as intended. For example, evaluations of the ProximityHat, a head-mounted sensory augmentation system which communicates peripheral distance information [18] have demonstrated that users, with practice, have some ability to use the hat for basic navigation. However, such experiments were only conducted after a detailed series of iterations and refinements to the arrangement of the sensors, to ensure that the navigation tasks were at least feasibly possible.

Similarly, Sohl-Dickstein *et al.* evaluated the feasibility of their echolocation device [154] through a series of signal identification and differentiation experiments. While the ability of participants to identify features from the echolocation signal show that there is the capacity for information transfer in the device, the effectiveness of the device in practical settings has not been explored.

Similarly, the evaluation of Carton & Dunne's Firefighting Glove [32] has shown that participants can make distance judgements or gap identification using the vibrotactile glove's presented distal touch signal, but offer little further insights into application suitability.

Such experiments are valuable as pilot tests for later device designs or iterations, and offer fundamental results related to the field of psychophysics. For example, Lynette Jones summarises much of the common ground between psychophysics and device design in a survey of tactile communication systems [80]. However, while participants may be able to differentiate signals, these results do not reflect the utility of the signal in a practical setting, nor the potential emergence of perceptual experiential changes.

Effectiveness & Efficiency: Task Performance Many SSAD evaluations involve experiments that compare users' task performance with and without the SSAD in question.

For example, Nagel *et al.* evaluated the feelSpace belt [124] using a between-subjects comparison of aided and un-aided users completing locomotive tasks. Maze navigation, path completion and obstacle avoidance were assessed as measurements of proficiency in the use of the device's spatiotemporal navigation signals. The experiment was designed such that the tasks require active attention and comprehension of the augmentation signal for success, and this was demonstrated through superior performance of belt-wearers in comparison to a control group. Disabling the belt also resulted in participant performance dropping to that of the control group, further reinforcing the use of the belt signal in task completion.

The feelSpace was additionally modified and used to augment perception in driving scenarios [92, 93]. Participants using these sensory augmentation systems demonstrated judgement improvements that were consistent with effective use of the information the belt provided. However, as noted in these publications, it is uncertain whether these improvements can or should be attributed to a perceptual augmentation (as was the intention of these specific interfaces), or whether participants are utilising the device signal in a more conscious, cognitive fashion.

User Experience In-the-wild, long-term evaluations of SSADs coupled with assessments of usability are rare, as commercial sensory substitution devices have an extremely small user base, even among the blind community they were originally designed for [100]. This makes it difficult to model and understand how well new SSADs will be received, and how ubiquitous more generalised enactive systems may become. Ward's interviews [166] with long-term users of the vOICe [117] provide useful insight into what living with an assistive sensory substitution device may be like, with participants reporting positive improvements to quality of life through altered perceptual experiences, but also noting the appearance of strange, synaesthetic-like experiences.

Van Erp's torso-mounted vibrotactile vest "TOAST" [162] (not unlike the TSAS [123]) equipped an astronaut on the ISS with an augmented signal reflecting orientation. TOAST was designed to assist with tasks in micro-gravity, and to reduce space-sickness. In a post-flight interview, the astronaut described the device as being useful during complex task execution, but deemed it too uncomfortable and cumbersome for daily wearing.

Long-term usability therefore represents an important facet of research in the development of ubiquitous perception-altering devices.

Altered Perception As noted above with reference to studies of the feelSpace belt [124], it is generally argued in SSAD literature that observed behavioural changes and reported experiential shifts may indicate 'perceptual integration' of the device [126, 145, 149, 164]. However, at the same time it is also often highlighted (usually by those that oppose sensorimotor positions on cognitive psychology) that it is difficult to distinguish such integration from simple cognitive interpretation of the signal [46, 100].

There have been attempts to define and measure perceptual integration, which are worth noting.

Sainz-Martinez' [145] work has started to shed light into how SSAD usage is linked to perceptual integration, through extension of definitions proposed by Nagel [124], and Stafford *et al.* [156]. As well as offering a wide introduction to the field and history of SSADs, the work offers details on how methodology (mixed-method analysis, including performance metrics and qualitative work) can be used to evaluate experiential qualities of SSADs and their effect on perception: Under the proposed definitions, 'cognitive integration' simply reflects an ability to use the SSAD signal. 'Perceptual integration' requires observation of e.g. cross-modal communication or judgement, or the appearance of illusions. A third, 'experiential integration' is proposed as full integration of the SSAD signal into conscious experience, and it is tentatively proposed that 'microphenomenological interview' (an involved method of enquiry designed to estimate phenomenal qualities of experience) be used to determine its success.

Kałwak *et al.* [84] present several measurement approaches as a general set of guidelines for quantitative and qualitative studies of sensory substitution experience. The work concludes that behavioural and neuronal methods generate excellent data on SSAD usage and utility, but that first-person perspective methods provide greater insight on the experience itself.

In contrast to these efforts to measure perceptual integration, Deroy [46] offers through surveying the limitations of SSADs that sensory substitution may be more comparable to reading than it may be to vision or similarly intuitive, perceptual interactions, arguing that expectations of true perceptual change are unrealistic. Ward suggests a description of sensory substitution as a synaesthesia-like phenomenon [167].

The diversity of these attitudes highlights that in the current literature, describing what perception and experience *are* in scientific terms, is as much of a challenge as measuring them. I will now outline how such descriptions of perception are approached and applied.

2.4.3 Models of Cognition for Sensory Substitution and Augmentation Devices

Previous work has focused on understanding SSA and its limitations through principles of psychology [7, 100] or biology/neuroscience [137, 167]. Many of these surveys draw from experimental results from outside of HCI and often do not contain new experimental data or devices, and are therefore easily overlooked by the HCI community [19, 153, 167]. While work such as this thesis seeks to contribute to interaction design, there remains a wealth of experimental evidence to draw from in wider research.

SSADs have been studied from perspectives which both embrace and reject enactivity. Froese's Enactive Torch [62] clearly accepts enactive approaches to cognition, holding action and movement as critical mechanisms for sense-making. However, many other devices have been explored from viewpoints which reject enactivity, and rely on the internalisation of representations of the outside world for sense-making, exploring associated phenomena such as multisensory cue combination [54], and crossmodal correspondences [155]. While these phenomena are important, rejecting enactive models of cognition provides somewhat disappointing answers to several philosophical questions, for example: What modality does perception using an SSAD belong to? [5, 111].

As will be shown in the following chapter, approaching SSADs from an enactive perspective offers a clearer path to answering such questions. My PhD research focuses on solidifying the use of SSADs and derived systems in interface design by tackling cognition and interaction from an enactive perspective, specifically ‘Sensorimotor Theory’ [133]. I further highlight the effectiveness of such an approach, and address potentially incompatible experimental observations, in Chapter 7.

2.5 The Principles of Enactivity in HCI

Through their natural, multimodal nature of interaction, interfaces utilising enactive principles have great scope for application in ubiquitous technologies, from addressing information overload and reducing cognitive strain, to greater immersion and intuitiveness in information-rich environments. In this section, I briefly outline how information, cognition, immersion and intuitiveness are connected to SSADs and enactive interfaces, and what embracing enactivity means from an interface design perspective.

2.5.1 The Relevance of SSADs to HCI

At this stage, it may help to reinforce the relevance of SSADs to interface design in HCI through three points:

Informing Interaction Design. Firstly, the study of SSADs can inform enactive interface design. As indicated by the utility of the vOICe to long-term blind users, sensory substitution and augmentation signals have the potential for communication of ‘abstract and rich’ [167] environmental information. Communication of such information is applicable in demanding applications, and motivates further exploration of SSADs for application in these areas.

SSADs Exemplify Enactivity in HCI. Secondly, user control, or dynamic action, has been identified as critical for the learning of SSAD signals. This suggests enactive models of cognition may be best suited to understanding how using these devices results in changes to perception, linking the fields and opening up the rich literature of SSADs for reference in enactive device design.

Potential Beyond Assistive Technologies. Thirdly, as shown by efforts in sensory augmentation, source signals do not necessarily need to be naturally human-sensible, indicating the possibility of adapting SSADs to meet diverse applications with both real and artificial information sources.

The final paragraphs of this Chapter explores the direct impact and implications of bringing enactive principles (and associated SSAD literature) to the field of HCI.

2.5.2 Externalisation

Early work by Lenay *et al.* [99, 100] looks at how technology, and tool use, affects perceptual experience, and notes how mastery of a tool can result in *device transparency*, that is; the device disappears from conscious focus when in use. Indeed, it has been argued that achieving a level of mastery in the use of enactive interfaces can cause the interface itself to become 'invisible' [62]. In such observations, device signals become associated with stable objects 'out there', i.e. features of the environment are perceived directly, rather than being associated with the device. Lenay *et al.* note for example how a tool such as a pair of glasses becomes invisible the moment they are used [100]. In the context of SSADs, this phenomenon is sometimes described as 'externalisation' [9, 100].

Where most commonly encountered digital interfaces may present symbolic, textual or other explicit task information for consumption and comprehension, a user engaged enactively with a sensorimotor interface may directly perceive task information which may otherwise be inaccessible, or slow to derive or comprehend. To illustrate, a navigational aid system which verbally describes surroundings (such as "What's Around Me?" [22]) may provide explicit location information to equip blind users with the knowledge for independent navigation of well-documented high streets. By contrast, blind users who master the vOICe visual-to-audio sensory substitution device [117] see through the tool with vision-like qualities which may be flexibly applied in both familiar and novel situations. Despite both being labelled as sensory substitution devices, their principles of interactivity and the effect of mastery are vastly different: The first demands attention to provide descriptive, representational information about the environment, whereas the second gives access to environmental information while itself becoming transparent.

It is this 'invisibility' or 'externalisation' that presents opportunities in HCI. In aiming for externalisation in HCI, one can model the interface as a tool capable of extending perception into information previously only accessible to digital sensors [163]. The phenomenon of tool invisibility also highlights the important difference between SSADs/enactive interfaces and other multimodal devices, illustrated through their effect on 'affordances' as further described below.

2.5.3 Affordances, Action and Perception

The notion of 'affordances' is central to enactivity, central to the concept that we perceive the environment in terms of our intentions and abilities to act within it [45]. In post-cognitivist psychology, affordances are framed as *"meaningful opportunities for engagements [48, Chapter 1, p9]"*, and closely resemble the affordances described by Don Norman [129], which may be more widely known across HCI.

de Paz *et al.* have evaluated and demonstrated the modification of affordances through sensory augmentation with the use of body-scale apertures [45]. Users were equipped with a horizontal array of motors (attached to the abdomen) that each vibrated as a function of the distance to the nearest object (similar to the ProximityHat [18] or firefighting glove [32]) and asked to explore a virtual aperture and judge its 'passability'. Modification of the aperture's width resulted in reported changes in passability following user exploration. Thus, the aperture's 'affordance' of 'passability' by the user was perceived and modified through active use of the SSAD. In this experiment, for the user to experience perceptual changes, both the new augmented sensory information and the embodied action were critical.

Enactive principles govern the design of tools and interfaces that are built such that perceived opportunities to act (i.e., affordances) during use are in some way different to those of the unaided user.

2.6 The Sensorimotor Interface in HCI

In comparison to the multitude of SSAD studies built for specific applications, there are very few examples of generally applicable enactive interfaces in HCI, and almost no examples that push the concept of sensorimotor interaction and presentation of virtual, abstract affordances.

Searching for publications which portray devices specifically using enactivity in computer science journals, contributions may be found in the context of driving [92, 93], emerging sports [66] or game-like activities [40].

While it is clear that such interfaces do appear in HCI, the sporadic appearance in the literature and occasionally uncertain methods of application, design and evaluation suggest that it is still a field very much in need of fundamental development.

Lobo [125] evaluates and confirms the robustness of ‘post-cognitivist’ models (specifically ‘ecological psychology’[1]) for description and evaluation of such interactions through several user studies. Her work shows that affordances are fluid, and highly dependent on both presented sensory information and the user’s ability to act within an environment. This process presents perspectives on how SSADs and other perception-altering devices may be well-modelled by approaches such as enactivity (and other, related post-cognitivist frameworks) [105], though little work is undertaken to extend such frameworks into generalised guidelines for application in interface design.

Attempting such an extension forms the basis of my PhD research. In the next chapter, I describe O’Regan and Noë’s ‘Sensorimotor Theory of Vision’ (SMT), a framework for cognition which aims to explain how perception emerges from enactive engagements. SMT has been directly applied to SSAD research previously [96], but it has not been formalised as a framework for interaction design in general. As such, in Chapter 4 I present a model for the sensorimotor interface as the formal application of sensorimotor principles in interface design.

2.7 Summary

In this chapter, I have introduced enactive interfaces through exploration of the history and development of SSADs and assistive technologies.

While SSADs are a good starting point for the implementation of enactive principles in interface design, the term “SSAD” is somewhat broad (as it covers assistive technologies) and there is no clear model for their construction. However, many SSADs offer an exciting insight into what may be possible through enactivity: interactions that enhance or modify perception with the promise of externalisation through mastery.

We now have a clear differentiation between the cognitive interfaces found in HCI and the enactive interactions targeted by this thesis. Where many interfaces in HCI communicate affordances through symbolic representation, an enactive interface changes perception to reveal environmental properties without presentation of symbolic or representative information.

---

[1] While ecological psychology and the enactive approach are technically separate philosophies, they share many principles regarding the dynamic linking of action and perception through sensorimotor loops. However, while ecological psychology models affordances as information acquisition, enactivity poses affordances as a sense-making phenomenon arising from goal-driven application of sensorimotor interactions [48]. Work on reconciling these frameworks is beyond the scope of this thesis, though conveniently, the philosophical difference of the models at this depth has little effect on their use in development of novel interfaces.

In order to progress further, we need an enactive theory of cognition which attempts to explain how externalisation and general cognition and sense-making arises from natural, un-augmented interactions. As such, in the next chapter, I introduce Sensorimotor Theory as an enactive cognitive model that claims to explain the emergence of phenomenal experience through action. Extension and application of sensorimotor theory to interaction design leads to a model and explanation of the thesis' titular 'Sensorimotor Interface', and may be used to inform both design of new interfaces, and the choice of new applications in HCI.

CHAPTER THREE

SENSORIMOTOR THEORY

In the previous chapter, I outlined the roots of enactivity in HCI, and described its relation to interface design through an exploration of SSADs; devices which (typically) attempt to assist or aid users through augmentation or alteration of perceptual experience. However, some SSADs are more successful than others in achieving such perceptual augmentations. If we are to create interfaces which target the modification or creation of perceptual experience, then further work is required to understand the principles that explain *how* and *why* perceptual experience originates, before extending and applying these principles to interface design.

In this chapter, I outline the issues that manifest when trying to apply cognitivist models to the task of designing new perceptual experiences. I then outline O'Regan and Noë's *Sensorimotor Theory of Vision (SMT)* as a suitable model for the task. SMT formalises the principles of how and why natural perception (initially, vision, though later effectively generalised) arises from enactive engagements, offering explanations for many observed properties of SSADs. It is specific enough to act as a framework for furthering the deliberate design of devices supporting new perceptual experiences, while being flexible enough to reveal a range of exciting prospects in this process.

Following the explanation and introduction of SMT in this chapter, I will pursue the extension and application of SMT to interface design through the derivation of the 'Sensorimotor Interface' in Chapter 4, constituting the first contribution of this thesis.

3.1 Introduction

In Chapter 2 I detailed how enactive interfaces may augment perception, changing environmental affordances and revealing previously inaccessible information without the presentation of symbols or explicit properties. While the aim of such devices should now be clear, we still need to discuss the method by which enactive interfaces can achieve this change in perception.

In order to reliably create interfaces which allow new perceptual experiences, I propose starting from an *action-centred* model of perception and cognition capable of explaining how natural perceptual experience arises, accounting for the observation that both natural and artificial perceptual experiences require active engagement (rather than passive observation) to develop. The chosen model should be able to define and differentiate sensory modalities, and to account for the different qualities of perceptual experience associated with them. Finally, as learning and training have been identified as barriers to uptake of SSADs [114] and require further research [19] in enactive interface design, our model should also accommodate methods by which the use of enactive interfaces can be learnt.

To this end, this chapter introduces and explains O'Regan and Noë's Sensorimotor Theory (of Vision) (SMT) [133] and its specific approach to enactive cognition. SMT is an enactive model of cognition and perception which has been developed and tested in parallel with the wider field of Sensory Substitution and Augmentation, and has since been widely studied. Previous work suggests that this model is helpful in the understanding of both natural and artificial perceptual experiences.

This chapter therefore starts with an outline of the problems that SMT was proposed to tackle, and a description of its place in wider cognitive psychology. I will then deconstruct SMT in the context of natural sensory perception through a description of its components, before then reframing these components in the context of SSADs and the artificial creation of new perceptual experiences.

3.2 Models of Cognition

The nature of cognition, and the search for a theory of mind capable of accurately modelling perception and sensory experience has been the subject of philosophical debate for millennia.

The issue of reconciling how perception and consciousness arise from the physical matter that makes up our bodies, minds, and the world we observe, is made apparent by the thought experiments of Chalmers [35], which can be traced back to Levine [101]. Levine noted the 'explanatory gap' which lies between descriptions of the biological or physical processes involved in experience and descriptions of the phenomenal quality of experience. Essentially, while it may be possible to identify the brain region, neural circuitry, or underlying cellular structures which correlate with behavioural accounts of vision, the mechanism by which visual perceptual

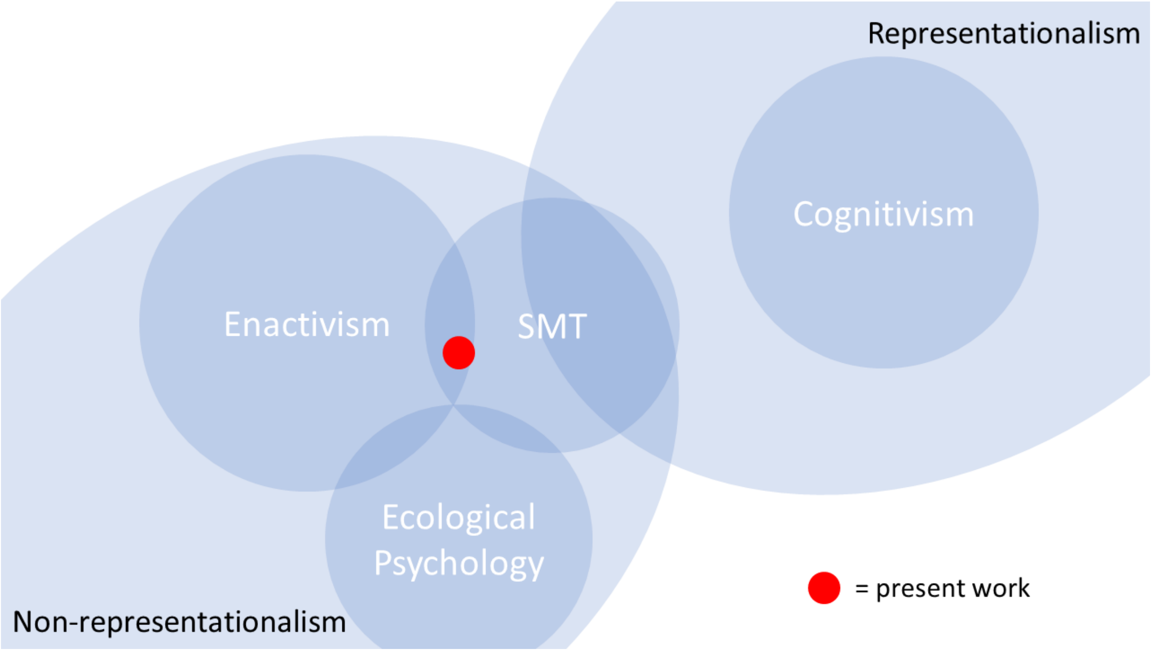


Figure 3.1: Schematic for Cognitive Science, based on Di Paolo [48]. Most interfaces in HCI build on ‘Cognitivism’-derived principles. Current work (red dot) approaches enactivity in HCI through application of sensorimotor theory, here labelled ‘SMT’.

experience rises from these structures remains mysterious.

Such observations are expanded upon by Chalmers, noting that standard scientific enquiry will in time provide adequate explanations for observable properties of the mind, such as integration of information, focus of attention or deliberate behaviours, through a reductionist approach of ascribing these properties to measurable, replicable neurological patterns [35]. However, as above, scientific enquiry through such reductionism falls short of an explanation as to *why* conscious experience should accompany such observable phenomena, thus leaving an ‘explanatory gap’ between physical observation and experiential consciousness which evades explanation. This is often described as ‘the hard problem’ (of consciousness) [84].

The explanatory gap is rarely a problem in HCI, as most interfaces may be designed according to cognitive paradigms corroborating reductionist and ‘representationalist’ approaches. Such paradigms follow branches of cognitive psychology which treat perception and the senses as largely passive ‘information in’ processes. For example, if a user interface is designed to deliver information pertaining to a computerised system (e.g., a car fuel gauge), a solution would be to visually represent that system using symbolic representation (such as a needle on a dial) which must be labelled. The user does not need to *act* upon the system to acquire this information: its meaning can be effectively understood passively, through observation of the symbol (the needle’s position) of the real-world property it reflects (the actual amount of fuel in the tank).

This philosophy works well with a wide range of interaction techniques. On-screen icons 'represent' information, explicit gestures 'symbolise' the spatial dimensions of system control and navigation they enable, and labelled buttons demark points of action which the user may take to issue commands. Outputs are presented explicitly, usually textually, numerically or graphically, and affordances are communicated through symbolic representation.

In all of these cases, the user is able to comprehend the signal presented by the system through conscious consideration of what the output represents. Explicit presentation of information in this fashion is therefore consistent with representationalist philosophy (see Fig. 3.1). Presentation of information for easy, intuitive comprehension is a key component of interface and visual design, and thus occupies a large space in device and interaction design principles.

However, as noted above, such approaches are not suitable for explaining the origins of natural perceptual experience, nor the observations of users fluently using SSADs or tool use. If these models cannot account for the origin of perceptual experience, how can they be used to identify the properties of the interfaces which allow new perceptual experiences to develop?

In other words, representationalist paradigms cannot account for why action would be required for perception to occur. Under strictly representationalist models, simply presenting sensory information to an attentive user would result in a representation in the brain which may be perceived, considered, and acted upon accordingly. The observation that interaction is required to develop proficiency with such interfaces [11, 163] contradicts the notion that perception is a passive process.

There have been attempts to include 'action' as an important component in the acquisition of SSAD proficiency by framing it as a sort of feedback mechanism which allows a representationalist understanding to form. Several authors therefore tentatively append the requirement of action onto classic models of perception [6, 19, 114]. However, close inspection has revealed that these explanations, while superficially helpful, are unable to offer satisfying answers to key practical (and philosophical) questions of enactive device use: Why are action-based SSAD training paradigms superior to passive ones [106, 143]? Furthermore, why and how does learning-through-action result in unique sensory qualities, such as the 'vision-like' experience reported by users of the TVSS? For my PhD research I have adopted O'Regan and Noë's Sensorimotor Theory of Vision (SMT) [133]; an enactive account of perception and consciousness introduced specifically to address these questions.

3.3 Introduction to Sensorimotor Theory (SMT)

SMT is part of the wider 4E, ‘post-cognitivist’ approach to cognition [75, 105], so called as they reject some of the posits of the widely popular models of ‘cognitivism’, the core ideology shared by the representationalist frameworks described above. The ‘E’s stand for *embodied, embedded, enacted, and extended.* The post-cognitivist approaches are not a unified or single theory, but all hold action as central to the emergence of perceptual experience and wider cognition. Under 4E approaches, perceptual experience and wider cognition are not considered passive, information-handling processes, but adaptive and behavioural activities. That is, sensory experience is not based on environmental representations, but emerges from the exercise of sensory-motor ‘skills’ by a goal-driven organism (or ‘agent’) in its environment.

The 4E approaches differ in their approach to defining and explaining what these skills are, and how they are acquired and exercised. SMT falls under the wider umbrella of the ‘enactive approach’, a perspective in cognitive psychology which holds the organism–environment relation as the basis of cognition, modelling cognition as a dynamic process that does not rely on environmental representations. Thus, enactivity and associated devices and interfaces pursue the goal of communication through dynamic, tool-like engagement, rather than the use of informational representation.

Under SMT specifically, these ‘skills’ are structured rules governing the relation between sensory stimulus and agent-driven action, called ‘sensorimotor contingencies’ (SMCs). Each SMC therefore reflects the set of rules by which specific sensory stimuli change in response to specific organism actions. Figure 3.2 schematically describes such a relationship between action and sensing. In the following sections, I explain this action-centred approach to cognition in more detail, focusing on how its components and structures lead to a practical definition of natural perception.

3.3.1 Components of Sensorimotor Theory

In its originally presented form, SMT describes three components to the emergence of perceptual experience—two types of SMC and the idea of ‘awareness’. To illustrate these, let us consider O’Regan and Noë’s original example of visual experience through SMT.

Apparatus-SMCs. These laws of interaction are properties of the sensory apparatus through which the agent perceives the world. For example, in the context of vision, there are laws which cause distortion as the eye moves: the stimulation on the retina shifts and distorts in a very particular, but consistent way, determined by the eye’s motion, and ocular optics (see [133, p941,

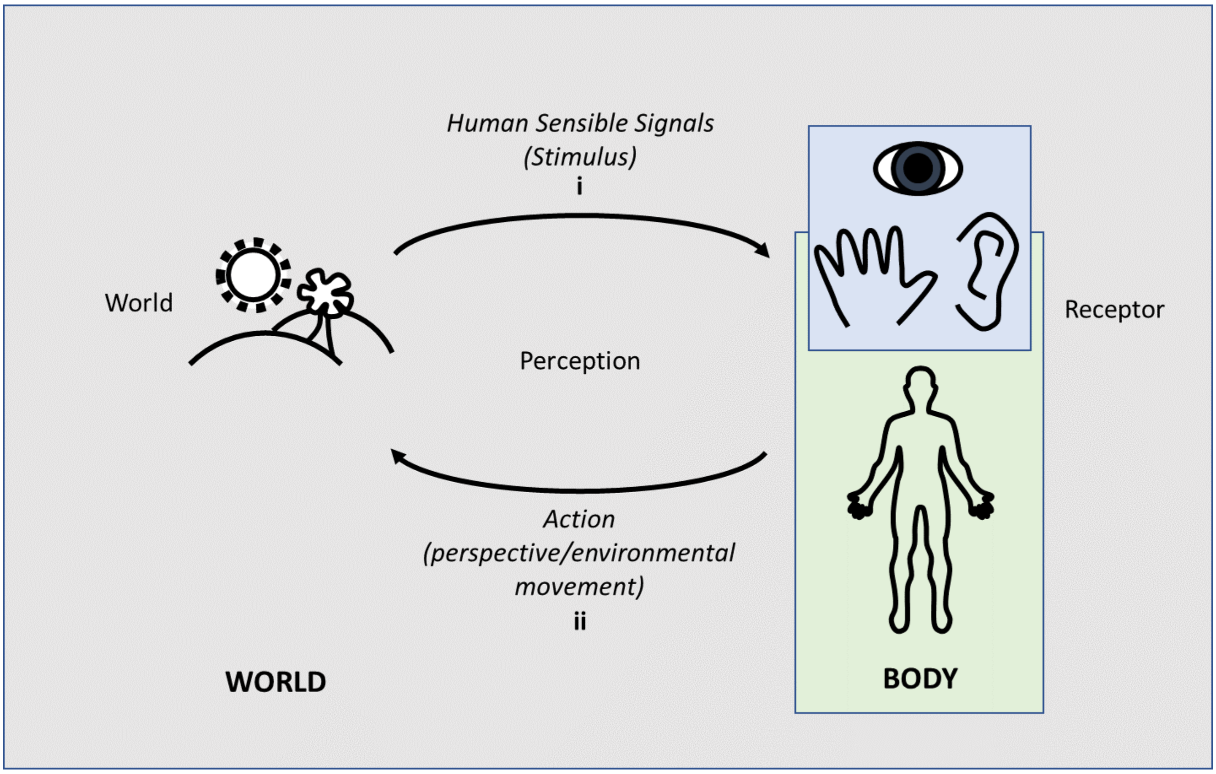


Figure 3.2: Sensorimotor Theory (SMT) models perceptual experience as arising from the dynamic relationship between sensory stimulation (i) and agent-driven action (ii). As an *enactive* model of cognition, perception is not a passive 'information-in' process, but instead requires active engagement; an observation consistent with the emergence of new perceptual experiences in SSAD use, and central to enactive interface design.

Fig.1] for a full explanation). These consistent relationships between the changing retinal signals and associated movements constitute the apparatus-SMCs of vision.

Object-SMCs. These are laws that govern how object attributes and structure affect the physical properties we are able to perceive. Shape, for example, is a visual property belonging to visible objects, but it is also a tactile property. However, while rotating a coin creates an elliptical projection on the eye, this physical rotation does not result in elliptical projections on the skin. Thus, while shape is a property belonging to both modalities, the laws describing how shape manifests through interaction with visual or tactile apparatus are different. Knowledge of how an object's attributes affect visible light under manipulation is reflected through the object-SMC of vision.

Awareness. Finally, SMT notes that even if an agent is attuned to both the apparatus- and object-SMCs of vision, it is possible for objects or events in their visual field to go *unperceived*, if their attention is not focused on that particular object or event. Perception therefore requires at least a degree of awareness of that sensory stimulus.

3.3.2 Perception from a Sensorimotor Theory Perspective

To summarise, SMT models perception and perceptual experience as arising from the combination of three components, exemplified by the visual sense:

1. Apparatus-SMCs must be mastered. User's movement is critical for learning and exercise of apparatus-SMCs. *Vision depends on mastery of the eyes.*
2. Object-SMCs must be mastered. The presence of environmental features with object-SMC attributes is required for exercise of the object-SMC. I.e. The SMC will only result in perception if there is something to be perceived. *Vision depends on mastery of light's interaction with objects in the environment.*
3. Attention and goal-directed exercising of the SMCs is required. *Vision requires attention to be paid to visual stimuli.*

With these components in mind, in the following subsections, I explain several (distinct) further ideas, observations and consequences of SMT, highlighting the utility it provides in defining and understanding perceptual experience.

3.3.3 Defining Sensation, Perception and Modalities

SMT allows us to recognise the fundamental difference between sensation and perception. To summarise and paraphrase O'Regan and Noë's account of these terms [133]:

'Senses' are determined by the biological "hardware" the sensing agent possesses. The senses are affected by external, real-world stimuli. Stimulation of the sensory apparatus causes a *sensation*. A signal or stimulus may be said to be sensed once an electrical signal has reached the brain. This definition includes no argument of sense-making, or external attribution.

By contrast, 'perception' happens when the results of stimuli correspond to the categorisation of objects and events in the environment. As described above, perception requires both knowledge of the appropriate object- and apparatus-SMCs, and awareness of the stimulus.

With this, perception can now be formally understood as the activity of aware exploration of the environment, in ways mediated by knowledge of the relevant sensorimotor contingencies. In other words, perceiving is the active exercising of mastered knowledge of sensorimotor contingencies. Each sensorimotor contingency, that is, the relationship between action and changes in stimulus, has a specific structure, and thus supports an associated unique perceptual experience. This experience is specific and unique to each different sensory modality, such as vision, hearing, and touch.

Thus, specific ‘modalities’ are therefore defined not by the biology that accommodates them, but instead by the lawful relations between action and sensing which allow the modality to be wholly and uniquely described. A modality is defined as the perceptual experience associated with an individual or specific set of sensorimotor contingencies [133].

3.3.4 Practically-Acquired Knowledge

Knowledge of sensorimotor contingencies is a practical, not a propositional form of knowledge [133], i.e., sensorimotor skills are represented by knowledge of *how* to do things, rather than descriptions of states and objects. For example, we are not able to accurately describe all of the distortions that an object would go through as upon moving our eyes around. However, the presentation of a distorting lens immediately in front of the eyes (such as a new pair of reading glasses) would be immediately noticed, and the experience of vision through this lens would noticeably change.

That perception depends on the extraction and integration of apparatus- and object-SMCs as knowledge is critical to the application of SMT to interface design. As SMCs cannot be accurately, explicitly explained or described, the knowledge must therefore be exercised actively, and acquired practically.

3.3.5 Learning New Sensorimotor Skills

Modelling knowledge of SMCs as practical, rather than propositional, may also help to explain why users of enactive systems presented with explicit training information (i.e. specific, descriptive feedback on performance or method) rarely perform better than those with no ‘training’ at all [62, 143].

Progress may be made knowing that the mastery of SMCs would benefit from practical, rather than propositional feedback [19], under specifically designed, action-centred training circumstances. As part of my research, I focus specifically on the use of concurrent sensory information as a practical feedback mechanism during sensorimotor task learning. This focus is illustrated through the experimentation described in Chapter 6, where I evaluate the role of contextual sensory information as a practical feedback mechanism.

To complete the requirements of a ‘suitable model’ laid out in Section 3.1, an explanation for how new perceptual experiences may be actively learned and developed must be proposed. Higueras-Herbada *et al.* [75] provide such an explanation in their introduction of ‘Direct Learning Theory’ (DLT) for learning of sensorimotor skills. They describe learning processes in enactive interactions as mechanisms which *change* perception while sensorimotor contingencies are being exercised; i.e. learning-by-doing. DLT formalises learning-by-doing by modelling and structuring the learning process around three elements; *intention, attention, calibration*, which will be introduced in detail in the next chapter. Such elements describe the practice of a conscious, goal-directed effort in an environment where existing perceptual experience allows some degree of sense-making and task completion, but where development of new sensorimotor skills is useful.

3.3.6 Mastery of SMCs

One final definition worth exploring is that of ‘mastery’, as defined by O’Regan and Noë in the context of mastering the knowledge of SMCs: if perceptual experience is changed, rather than created, ‘mastery’ represents the increasing degree of knowledge that the agent possesses regarding apparatus- and object-SMCs [133].

O’Regan and Noë define ‘mastery’ as an attunement or knowledge about the sensorimotor contingencies that govern an interaction, the knowledge being a modelled internalisation of contingencies, and the degree of mastery reflecting the accuracy of this model. Alternatively, Di Paolo suggests that ‘mastery’ refers collectively to sensorimotor skills, referring to it as ‘progressive growth and refinement of an agent’s sensorimotor repertoire’ [48, Chapter 4].

In any case, ‘mastery’ is not an absolute measurement or discrete requirement, but a progressive, never-ending dynamical process in the organism-environment relationship. Such a concept is important in the design and measurement of perceptual experience, as it allows perception to be considered and measured as a spectrum of competencies and associated experiences, rather than a binary phenomenon which either arises, or does not.

Now that I have introduced the requirements of perception according to SMT, I will outline how this theory has been, and may be, applied to the design of new perceptual experiences through interaction design.

3.4 Sensorimotor Theory in HCI

The application of SMT to interface design is primarily exhibited, as described in the previous chapter, through development and testing of SSADs [17, 85, 96, 102]. The following sections will outline how experimental evidence found in such studies have allowed SMT to be refined and developed to account for new observations in phenomenal consciousness.

3.4.1 SMT in SSAD literature

SSADs have been utilised to test the viability of SMT as a theory of mind since its conception [17, 48, 62, 85, 96, 124]. The original feelSpace studies stated from the outset that their goal was to create new sensorimotor contingencies [124], and that their creation would result in a new quality of perception. The premise of the experimentation and observations was to offer a new modality through basic mapping of navigational information to a tactile belt, and subsequently to evaluate whether perception itself changed as this new modality was mastered. However, it is not entirely clear whether this was achieved (due, perhaps, to methodology, metrics, or participants simply not having enough time to master the presented contingencies). While some observations (e.g. anecdotal recounts of changes in perceptual experience [124, 166]) fall in support of SMT, others (such as the limited degree of improvement in task performance [96]) are less supportive.

Some work has claimed to have directly observed the ‘emergence of apparatus-based SMCs’ [17] through comparison of the exploration techniques employed by users of a three different SSADs: a minimalist head-mounted SSD, a hand-held echolocation device, and the vOICe [117]. Bermejo argues that the measurable differences in exploratory sweeping movements employed while completing a shape recognition task reflect the different apparatus-SMCs of the different devices. Similarly, object-SMCs are observed emerging from differences in *how* individual shapes are identified and differentiated through exploration, such as sensing corners vs rounded edges.

To summarise, the emergence of perceptual experience using SSADs can be modelled under SMT through description of the device as an extension of bodily apparatus-SMCs, with the environmental parameters they encode being described by novel object-SMCs. Mastery and exercise of these device-mediated SMCs gives rise to new, SSAD-specific perceptual experiences [96].

Numerous other examples of the application of SMT to the context of SSAD research exist, serving to support the utility and viability of SMT for modelling enactive interactions. However, there currently does not exist a structured synthesis of SMT into design parameters and experimental guidance for the creation of new SSADs and interfaces.

3.4.2 Applying SMT to Advance Enactivity in HCI

The evidence and utility of SMT and enactive models of cognition in HCI is growing [49, 62]. With this, so does the requirement for more structured guidance on the design and development of these devices from a theoretical foundation.

In my research—as described in this and the previous chapters—I have presented enactivity as an approach to interaction design which aims to augment perceptual experience, rather than presenting information to be cognitively interpreted. In this chapter, I have presented an argument as to why SMT is a suitable framework for achieving this, and outlined the principles upon which it models perception. To conclude the chapter, I will re-summarise the implications of SMT in the context of HCI.

As outlined above, SMT explains *how* and *why* action and perception are linked, supporting the early observations of Bach-y-Rita with users of the TVSS [11] and providing answers to the questions posed in later reviews of the field of Sensory Substitution and Augmentation [6, 19].

SMT's particular utility in the design of new interfaces derives from the idea that perception is dependent on lawful action-sensing relations (sensorimotor contingencies) rather than a *biologically* defined physiological or neurological structure (as would be expected in cognitivism or another reductionist viewpoint).

Where representationalist models of cognition offer weak explanations for how novel sensory experience may emerge through presentation of signals to an existing modality, SMT was specifically proposed to fill this gap [133]. By considering SMT, we may build devices which link previously un-sensible stimuli with human-sensible signals. Provided these links are lawful and consistent, SMT allows for the extraction and mastery of these laws through active exploration. Exercising mastery of these laws to complete goal-driven tasks gives rise to new perceptual experiences, subject to and supported by existing perceptual experience.

In the next chapter, I present the the details of devices, training programs and applications specifically derived from the practical application of SMT, and present a thorough schematic and derivation of the sensorimotor interface.

3.5 Summary

In the quest to design new perceptual experiences through technology, SMT offers a good starting point as a model of cognition, providing a foundational explanation of perception largely compatible with observations of SSADs and other enactive interfaces.

In describing perception as a dynamic, pliable process of discovery, SMT specifically offers promising perspectives on how the modification of perceptual experience may be levered in the design of new interfaces. Furthermore, through its compatibility with wider principles in enactivity and other 4E approaches to cognition, SMT can be supported by well-established associated concepts and theories, such as ‘affordances’ and ‘direct learning theory’.

However, SMT must be further developed into a more detailed interaction design model before new modalities may be deliberately created or designed. In the next chapter I therefore attempt to specify in concrete terms the nature of devices designed to support new SMCs, and by doing so, detail the requirements for building and using a sensorimotor interface.

CHAPTER FOUR

THE SENSORIMOTOR INTERFACE

This chapter synthesises the findings from the previous two chapters to create a new framework for sensorimotor interface identification, design and application.

As shown in Chapter 2, application of enactive principles in HCI remains sporadic and continues to emerge and develop. There are several key issues left to resolve in the scope, design and application of enactive interactions outside of assistive applications, and in this chapter I seek to synthesise the work of the previous two chapters to build a framework towards this.

Most interfaces in HCI are built upon a symbolic collection of icons and abstractions designed to provide access to information through representation. This type of information requires cognition and comprehension before explicit decisions may be made. It is understandable, yet unfortunate, that the design, evaluation and choice of application for all interfaces has evolved from such a paradigm. While most interfaces built are examinable in a fashion consistent with cognitivist psychology, and prove highly effective and intuitive for the majority of applications, such approaches are demonstrably limited in the domain of perception-altering, sensorimotor interfaces.

In Chapter 3, I suggested that an ‘enactivist’ approach to perception through SMT is more suitable than ‘representationalist’ models as a theoretical foundation for building devices which augment perception. This chapter uses a schematic taxonomy to illustrate how cross-disciplinary literature reviewed in previous chapters may enable formal definition of the sensorimotor interface, and illustrate direct application of these principles. Here, I arrive at describing the design of devices that specifically strive to enable new perceptual experiences.

The review summary reveals three evolving practical device paradigms, constituting effective technical guidance while explaining measurable leaps in success, interest and viability of SSADs. From these paradigms, and the principles of SMT, I establish concrete device identification and design guidelines, leading to clear definitions of what enables and constitutes a 'Sensorimotor Interface', allowing differentiation between them and other devices. Further experimentation and training guidelines are presented throughout, with the goal to enable greater usability and working to close the gap between research and real-world application of sensorimotor interaction principles.

Finally, a series of tables detailing various devices found in the literature is presented, overall constituting a primer to existing SSADs and sensorimotor interfaces, uniting disciplines and providing an access point to the field for researchers from HCI and beyond.

4.1 Introduction

As demonstrated in Chapter 2, the use of devices and interfaces utilising enactive principles to augment perception and overcome information and attention limits are growing in prominence in HCI, yet the theoretical foundations of design and application of enactive techniques remain unclear.

To remedy this, in this Chapter I collate a literature summary and present a well-developed taxonomy based on the discoveries, techniques and insights from Chapters 2 and 3, linking SSAD and enactive device history with post-cognitivist models of enactivism.

Starting from the principles of O'Regan and Noë's Sensorimotor Theory [133] as described in Chapter 3, and with reference to the schematic presented in Figure 4.1, the following taxonomy breaks down the sensorimotor interface into constituent components, and details the principles of its function and interaction. The following subsections discuss each of the associated concepts in detail, highlighting the contribution to the nature and integration of the sensory information.

This chapter thus draws together the history and study of SSADs and other enactive devices, with use of theoretical principles from SMT to generate a detailed, labelled taxonomy for the sensorimotor interface. The organised, concise presentation of this information bridges the gap between cognitive psychology and interaction design through practical application of SMT.

This derivation is further explored to draw clear distinctions between devices conforming to the enactive approach, and other multimodal interfaces; an understanding which may be extended to inform comparisons between studies, and improve replicability.

With the integrated literature base presented here, the field of engineering and HCI may find such a resource facilitates prototyping of sensorimotor interfaces to be more accessible and lead to the design of devices that are easier for end users to adopt. I anticipate that wider adoption of sensorimotor interfaces as tools to modify or enhance perception in more general interface design may be applied successfully to data-rich and smart environments. Moreover, such systems show promise in virtual and augmented reality environments, where the distinction between what is ‘real’ and what is not can be further enhanced by such interfaces [115].

This chapter therefore makes the following contributions: 1) A concise taxonomy of device construction from a systematic review of research literature and commercial applications, 2) Novel insights and approaches drawn from this taxonomy and associated literature, 3) Device and experiment design principles and concepts, in design-document form, for use in the design of future sensorimotor interfaces and associated studies.

4.1.1 Approaching the Literature

I have approached the literature by exploring Sensory Substitution and Sensory Augmentation Devices together (as SSADs), with the ultimate goal being to identify devices, and underlying principles, capable of altering or enhancing perception through interaction.

However, perceptual claims (or goals) are not ubiquitous among SSADs, with Visell [163] noting that some authors choose to define devices by ‘operational’ or ‘functional’ characterisations.

SSADs could be defined or grouped by other shared attributes, such as the effort ‘to deliver information to a subject via a sense that does not usually deliver that information’ [112], in the vein of an ‘operational’ definition. However, this is not sufficient for the development of interaction principles, as it does not give insights into the nature of the ‘information’ or the result of its delivery. In addition, a definition of ‘sense’ beyond reference to the five Aristotlian senses [111] remains evasive. Broadly, it has also been noted that individual *modalities* within such Aristotlian senses may be capable of supporting different phenomenal and perceptual experiences [52]. This is especially true in the case of Sensory Augmentation, where the source data may have artificial, naturally imperceivable origins.

Similarly, ‘functional’ SSADs such as Facebook’s TLC [39], designed specifically for compensation of sensory impairment, prove exceptionally useful in the context of communication. Though, as this device employs a degree of symbolic representation, it is unclear how such a system (in its present form) may give rise to new perceptual experiences if such experiences emerge from enactive interactions.

Bertram notes how the applications between Sensory Substitution and Sensory Augmentation converge at *"providing perception beyond a user's normal sensory capabilities [19, p234]"*, though also remarks that the true extent of their similarities and the degree of this convergence has not been expressed formally.

While this set of requirements was sufficient for identifying relevant systems through literature review in Chapter 2, the need for clear objectives, interaction principles and design guidelines remains. I can now, with reference to SMT as introduced in Chapter 3, provide greater clarity in the application of enactivity to interaction design from a theoretical foundation.

To describe and design devices which target changes in perception, I here present a rigorous and coherent taxonomical schematic and detailed definition of such devices, applying the label 'Sensorimotor Interface' due to their theoretical derivation from SMT. Then, I introduce design and implementation guidelines and form the foundations for guidance on the metrics and mechanisms suitable for the characterisation and comparison of perception-augmenting interfaces in HCI. Together, this approach aims to take enactive interaction beyond SSADs and to formalise enactive interaction principles for wider application.

4.2 Formalisation of the Sensorimotor Interface - A Taxonomy

While this section constitutes a taxonomy in the classification and overview of sensorimotor interface components, the detail and literature referenced in each description contributes to the deeper goal of the introduction and familiarisation of the *enactive* paradigm of interactivity. As discussed in Chapter 3, interactivity is central to the emergence of perceptual experience with the SSADs designed under enactive paradigms. As such, the schematic (Fig. 4.1) therefore both accurately refers to the construction of such SSADs, while simultaneously introducing the more generalised and explicitly-defined 'Sensorimotor Interface'. A sensorimotor interface is defined not by its components and aims, but by the enactive fashion in which it is used: as part of a dynamic, goal-driven process consistent with SMT [133], in which the nature and philosophy of action and sensing (which together describe the quality of perception) play as significant a role as the individual components of the device.

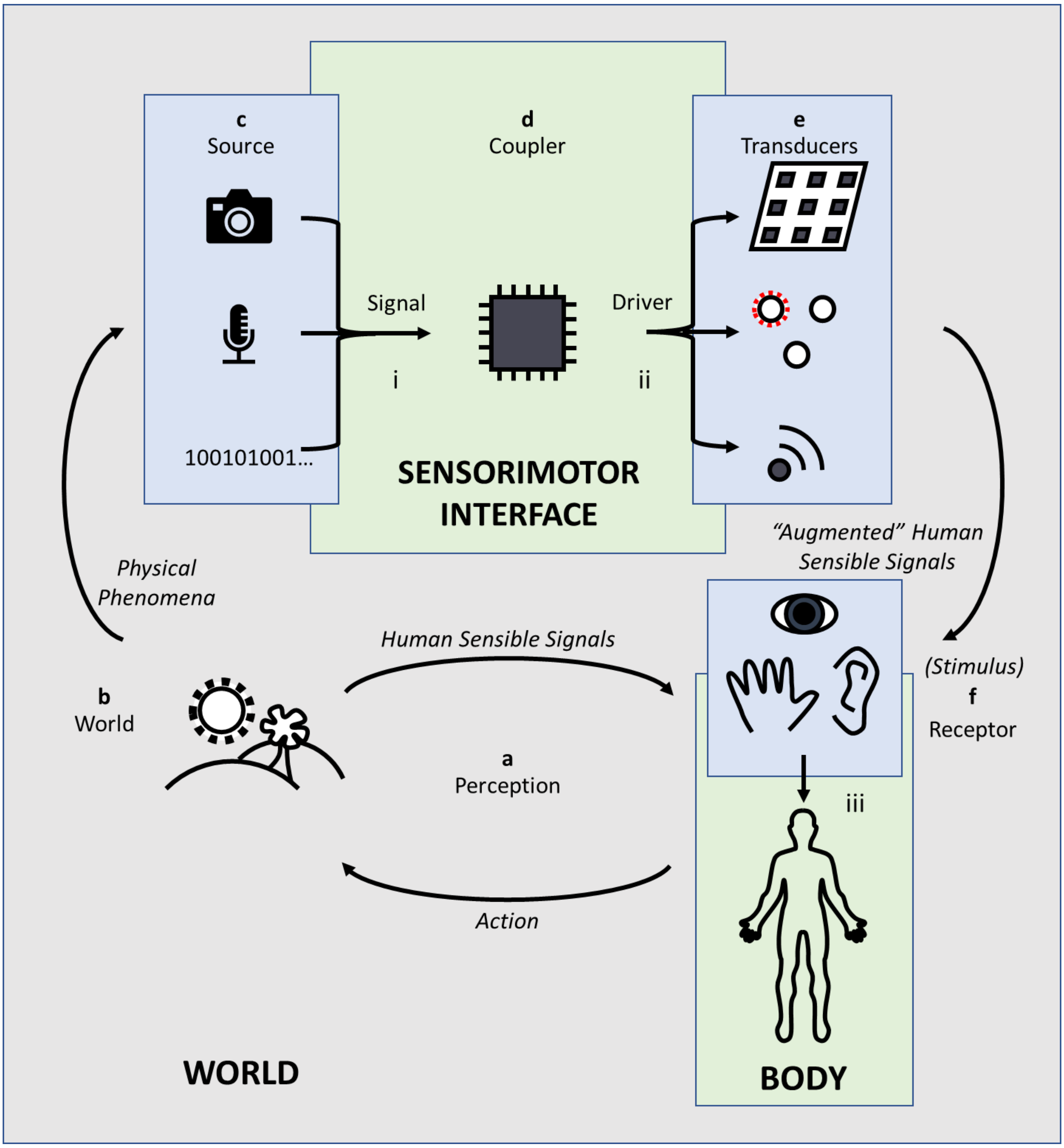


Figure 4.1: Detailed schematic of a 'Sensorimotor Interface' and associated world/human interaction. The interface itself is illustrated as three components (*c, d, e*) allowing an interaction supporting an artificial action/sensing loop between the human and world, resulting in new perceptual experiences *a*.

4.2.1 Taxonomy

4.2.1.1 Figure 4.1 a, Perception

As discussed in detail in Chapter 3, cognitivist psychology (representing the bulk of models assumed for interface design), models cognition as a one-way information-processing experience, where information is presented to the user for digestion and consideration before deliberate action is undertaken.

However, Auvray & Myin note that: *"users' movements seem crucial for learning to properly perceive with a Sensory Substitution device [6, p1047]"*, a phenomenon previously noted by Paul Bach-Y-Rita in his original studies [9, 14] and confirmed in most other SSAD studies since [124]. This is a parallel drawn with natural vision as early as 1963 [72] - by many accounts, action and user control is a necessity for perception, and vice-versa.

In contrast to conventional cognitivist psychology, '4E' approaches to cognition (the "E"s stand for embodied, embedded, enacted, and extended) provide an ideal starting point for description of devices where action plays a central role in perception. In 4E models, cognition is not passive information-processing, but adaptive behaviour: it is action and active embodiment within the environment that leads to perception.

SSADs have proven themselves as useful tools for further developing and refining 4E theories of perception, with many designed specifically for this purpose [62, 70, 124] subsequently being re-utilised for further perceptual studies [25, 108]. Improving skilful task execution through altering, widening or enhancing the user's perception (rather than augmenting or presenting explicit knowledge) represents a pillar principle in the differentiation between sensorimotor interfaces and other multimodal interfaces. The requirements and implications of such an approach, introduced in Chapter 3, are further discussed below.

Many users that achieve device mastery through enactive interaction report the experience of "externalisation", where the device becomes 'transparent' and the user is left with a novel perceptual experience [6]. This potential further highlights a conceptual difference between the role of sensorimotor interfaces and most multimodal interfaces. In most multimodal interfaces encountered in HCI, a key question in design is how to present system 'affordances' [48, 129] such that system capabilities are obvious and intuitive. However, when building a sensorimotor interface, the goal instead is that through development of proficiency with the device, perception changes to reveal new environmental or task affordances while the tool or interface itself becomes transparent.

To clarify, 'action' refers to both change in perspective (where sensory Stimulus changes due to Receptor moving relative to the environment, as demonstrated by the original movement of the head-mounted camera in the TVSS [9]), and change in environment (where objects or the Source are actively changing state due to extended action by the agent [75]).

Only once the lawful relation between Action and Stimulus is mastered can changes in Stimulus sensed in the absence of deliberate movement (i.e. world/environmental events un-initiated by the agent) be Perceived as changing properties of the World. To summarise: relevant task information, affordances, and overall advantage of a sensorimotor interface is accessed through action and engagement, rather than passive presentation.

4.2.1.2 Figure 4.1 b, World

The World is not directly experienced. Per SMT; perception, and the experience of embodiment in the world, arises as a by-product of goal-driven action by independent, wilful agents. Wilful intent drives experimental action, and perception arises as the lawful relations between these actions and changes in sensory stimuli are *mastered* and *exercised*. Together, mastery of the multitude of lawful relations, dubbed 'Sensorimotor Contingencies', that our bodies support, leads to cognition and experience.

We are therefore limited not only to the properties of the World which may be detected by our natural sensory apparatus (often philosophically referred to as our "Umwelt" [95]), but also by the facets of the world upon which we may take meaningful action.

Some Sensory Substitution devices have been proposed which may allow 'sensing' of static or non-actionable data (such as stock market data [51]). While the meaning of patterns or information in such interfaces may potentially be learned through provision of dedicated explicit feedback [19], a sensorimotor contingency may not be formed if the agent is unable to act on the data, and such systems would therefore not qualify as sensorimotor interfaces.

However, if motor actions are also *lawfully* extended to act upon this data (such as through the use of digital controllers), sensorimotor contingencies may be mastered and the possibility of perception remains. One could conceive of synthesising abstract, artificial 'motor' actions allowing influence over stock market data, allowing the extraction of lawful relations, and the subsequent comprehension and perception of the data's meaning.

Through data gathered by computerised systems, and extension of motor actions through digital controls, sensorimotor interfaces offer the promise of sensorimotor connection to properties of the World, real and virtual, which are currently outside our perception, but which may be helpful, or enjoyable to perceive.

4.2.1.3 Figure 4.1 c, Source

A sensorimotor interface is a tool allowing interaction with a data Source. In the case of Sensory Substitution, the data source is representative of real-world sensory information that is normally available via a natural sense, for example, vision or audio (Table 4.3). Sensory Augmentation can enhance or expand human perception by using extra-perceptual information as the data source. This data can be representative of real physical objects or information (such as a compass [124], chemical signatures [3] or a distance measurement [33, 62]) or specific features or properties of an actively utilised tool [148] or machine [92, 93, 123]. The Source could also represent naturally sensed information, though in an enhanced or modified fashion which somehow improves perception accuracy or scope [86] or enables direct interaction with information that would otherwise be derived from more deliberated interpretation [92].

Although this measured signal can be processed and encoded by the Coupler, the resolution, complexity and bandwidth of the final output may be limited by the Source data. As such, the Source should be chosen with the final requirements of the interaction in mind.

4.2.1.4 Figure 4.1 d, Coupler

The Sensorimotor Coupler is not necessarily a single physical processor, but rather represents the system which maps the Signal data to a Driver, preparing it for presentation to the user. The mapping or encoding which takes place in the Coupler is dependent on the Source and Transducer, but is also subject to wide design decisions, some of which will be addressed in Chapter 5.

Given the continuous (as opposed to discrete) nature of human sensory input, and SMT's modelling of perception requiring 'lawful' action-sensing relations, sensorimotor interfaces are *"most relevant to continuous (as opposed to symbolic) methods of interaction [163, p39]"*. With continuous input data, this may be a simple process of linear mapping [62] or dimensional reduction [80]. The Coupler plays a critical role in defining the SMC that the device enables. SMC's are concisely described as *"lawful relations between motor actions and associated sensory stimulations [85, p47]"*, so to support the formation of a new SMC, the Coupler must lawfully map World/Source and Transducer/Stimulus.

However, there is an interesting and critical debate as to what constitutes a 'continuous method of interaction', and whether it is truly required, both spatially and temporally. For example, users report that the vOICe may enable formation of a smooth vision-like sensorimotor experience, and detection of hyper-resolved motion [166], despite the presented soundscape more resembling a scrolling sequences of images at $\tilde{1}$Hz, rather than a continuous optical flow as in natural vision. Equally, while it has been shown that humans are capable of interpreting symbolic data through tactile [23, 81, 90] and audio [24] interfaces, the discrete nature of symbolic outputs in response to action may prevent formation of an SMC. On this principle, Deroy rejects the argument that mastery of an SSAD results in a truly perceptual experience [46], instead suggesting that the experience is more akin to reading. This is argued by framing speech perception through the visual interpretation of written material as ‘piggyback’ onto an existing, mastered, speech perception. The piggybacking essentially offers an alternative, visual pathway to information naturally accessed through audition. However, note that while it is of course possible to develop high proficiency in reading, words are themselves a symbolic representation of world properties and thus the activity is perhaps best described through cognitive, representationalist models. By rejecting symbolic and other non-sensorimotor SSADs, I define a set of devices for which there is a strong argument, through SMT, for the creation of new perceptual experiences.

In this model, the Coupler also enables steps to be taken to compensate and account for Transducer and Stimulus limitations, improving intuitiveness of device signals when presented to new users. Essentially, adjusting Coupling to take into account the limitations of fully characterised Transducers ensures that the output Stimulus is representative of the Source data. Similarly, some sensory experiences (such as vibrotactile touch) are perceived non-linearly, potentially skewing interpretation of signals in early device us and thus slowing learning. The necessity and details of this calibration process are further discussed, and tested, in Chapter 5.

4.2.1.5 Figure 4.1 e, Transducer

A Transducer converts the electronic signal output by the Coupler into a physical signal.

Given the limitations of human bio-receptors, the most common Transducers in SSADs and other enactive interfaces (see Tables 4.1 and 4.2) are audio speakers and visual or haptic displays, for Stimulation of auditory, retinal or tactile receptors respectively. Haptic displays can be versatile, allowing interface with signals conveying temporal and spatial information [80, 81, 83]), but they do have limitations in terms of bandwidth [130] and perception [83]. Visual displays are rarely used for SSADs, as the user’s centred or focused vision is typically either the modality requiring substitution [117], or otherwise occupied by the task. However, they are perhaps underutilised, with stimulation of peripheral vision [52, 132] offering the possibility of augmentation without general impedance of the visual modality.

Finally, electrotactile displays offer both good resolution and bandwidth, with Danilov’s Brainport [43] and Dublon’s Tongueduino [50] proving highly effective SSADs. Their greatest disadvantage is the occupation of the tongue, which changes facial aesthetics (perhaps undesirably) while also preventing speech.

4.2.1.6 Figure 4.1 f, Receptor (Stimulus)

I define the Receptor as the human sensory receptor type which the interface’s Transducer is designed to Stimulate. Successful stimulation of the sensory Receptor by the Stimulus results in the transduction of a physical (World) signal to an action potential (electrical signal) within the nervous system, for subsequent transmission to the brain for sensory processing.

SMT suggests that lawful stimulation of the chosen Receptors in response to action is sufficient for perceptual experience. However, for adults with existing mastered modalities and Sensorimotor Contingencies, there are both biological and perceptual considerations associated with doing so:

Biologically, the action potentials created in response to mechanical stimulae are noisy, and of a given 'resolution'. The Stimulus emitted by the Transducer should be designed with these limits in consideration. Further, the physical connection between the Receptor and brain may be associated with neurological structures with less plasticity than may be required for development of new sensorimotor interactions. An example of this would be optical grid cells [121], which seem to be structurally reflective of environmental stimulae, pointing towards a representational comprehension of space. Attempting to build new contingencies around visual phenomena that stimulate optical grid cells may prove difficult, given their structural properties and specialisations.

Perceptually, knowledge of related or prior modalities associated with the chosen Receptor may aid or hinder formation of a new Sensorimotor Contingency, for example the non-linear perception of vibrotactile intensity [148], or existing perceptual associations via cross-modal correspondences [69, 155] or bio-mimetics [149].

It is hypothesised that the signal could be scrambled [133] and still allow formation of a Sensorimotor Contingency, provided the interactions are consistent and lawful. However, this seems intuitively at odds with regards to how easy it would be to learn such scrambled signals, and the potential structures or analogue natures of neurological connections (such as the aforementioned grid cells). Such implications are explored in experimental context in Chapter 5.

4.2.2 Discussion of Taxonomy

Until now, the most commonly referenced publication including a schematic of the SSAD is by Visell [163]. While action does play a part in this model, it must be noted that the overall structure and importance/labelling of action and the SSAD as a whole resembles that featured in Figure 1. of Froese [62, p366]; a largely 'cognitivist' perspective. By contrast, the above taxonomical schematic extends 'enactivist' principles outlined by SMT, resulting in a detailed schematic consistent with the outline Froese presents in Fig. 2 [62, p336].

As noted shortly before the taxonomy, the labels 'sense' and 'modality' are absent from the schematic descriptions. In many studies, rather than specifying the Receptor (and associated Stimulus) specifically, a 'target modality' is named [163, 166]. However, doing so creates a nuanced but important problem, which I discuss here.

Receptors Piggybacking onto a target modality means accepting and accounting for the limitations and learned perceptions associated with that modality, and may account for phenomena which are more similar to reading than natural perception [46]. However, while there are certainly interactions between existing modalities and new SSAD signals, under SMT there is nothing prohibiting the formation of a new modality on top of or alongside existing modalities. For example, a child with intact auditory organs, but who has not yet 'learned to hear', may through consistent exposure to audible Stimulus from a sensorimotor interface learn to perceive the world through this sensorimotor pathway which would not necessarily be described as 'hearing' if the contingencies describing the modality were not equivalent to hearing. i.e if the Source was not itself sound.

As such, I use the word Receptor to specify the biological receptors that are receiving the new signal, without making claims as to the utility of the signal or the effectiveness of the device in allowing formation of a new modality. Specifying only that the interface signal will be Transduced to a Receptor not usually associated with that signal is consistent with Visell's 'operational' SSAD definition. However, knowledge that most (if not all) biological receptors will, for sensorily unimpaired adults, have one or more associated modalities, is important for considering how the signal may be initially interpreted and subsequently learned. This will form a major point of discussion in Chapter 5.

This therefore also means that further specification of device and signal is required to meet the requirements of a sensorimotor interface - in aiming to allow and enable perceptual changes, existing biological and perceptual limitations must be taken into account.

Senses and Modalities A Modality is not the same as a sense. Under SMT, a Modality corresponds to a defined set of one or more Sensorimotor Contingencies, reflecting Action, Stimulus, and the associated lawful relations between them. 'Sense' is usually interpreted from the Aristotlian perspective, as a less scientific collection of perceptually similar experiences attributable to specific biological hardware. For example, foveal and peripheral vision are demonstrably different modalities, though they belong to the visual sense [52]. Similarly, musical frequency and volume are Transduced by the same auditory hardware, yet their perception and differentiation classify them as different modalities in accordance with their separate SMCs. Such separation allows for engagement of multiple modalities through the common sense of hearing, with each additional modality representing a new channel upon which signals can be piggybacked. For example, piggybacking onto frequency differentiation as well as volume and spatial separation in Visual-Audio systems [6] allows communication of colour [2].

Further Implications At this point, a summary example may clarify the terminology, specifically why the label Receptor (in the context of a perceptual augmentation) is a more specific and suitable word than Sense, Modality, or Perception for the interaction with the human in the diagram. Let us consider the creation of the new modality of distal touch, as in the Enactive Torch [62]. Distal touch is the ability to feel an object's distance (or perhaps even its texture or shape) without physical contact with the body. This may be achieved through creation of a simple device which, when pointed directly at a surface, vibrates proportionally to the distance to that surface. Closer surfaces produce strong vibrations, while farther objects, weaker vibrations.

Upon first engagement, a user sweeping the device before them would experience a change in vibrotactile intensity as the distance to objects changes throughout the motion. They can be said to perceive changes in the strength of the vibration, however, they would not know what the signal means *in the context of distal touch*, as they have not yet mastered this modality. Therefore, they do not perceive the change in vibration as a change in the distance it reflects, and cannot attribute the signal to objects 'out there'. To summarise, the user is not perceiving through distal touch, though Receptors in the Sense of touch are being stimulated, and touch sensation is being perceived.

In order to claim mastery of the new modality, active movement of the device in the environment is required. With time, the user procedurally learns to extract the laws that govern how movements correspond to changes in vibration caused by environmental topography. Only after mastering this process may it may be argued that the action, the world property, and the corresponding sensation in my hand are linked through the modality of distal touch.

This simple process highlights the importance of the distinction between Receptor, modality and sense, and how modalities are not fixed, but may be learned through acquisition of procedural knowledge through sensorimotor interfaces.

4.2.3 The Evolution of SSADs - Paradigms of Design

While the above taxonomy provides some insights into the unique construction of a sensorimotor interface, in this section I seek to enable further understanding of their origins and interaction principles, and to generate insight into how such a device may enable perceptual changes. Further, by using the literature to highlight the properties that differentiate sensorimotor interfaces from other multimodal approaches, I narrow down to a clear set of guidelines which may be used to inform the design, training, testing and evaluation approaches of sensorimotor interfaces in emerging interactions.

Here, I detail through the presentation of three paradigms how discoveries and developments of SSADs have helped progress understanding and development of sensorimotor interactions, and summarise the origins of the field.

4.2.3.1 Paradigm 1: Make it Valuable (1972-)

The TVSS [9] was a revolutionary SSAD designed for blind users, and was built on Bach-y-Rita's idea that the 2D surface of the skin could behave like the 2D surface of the retina. Having retinal data presented through the skin allowed users to perform tasks which usually rely almost entirely on sight, revealing a wide range of useful affordances in the natural world. Thus, a device may be considered 'valuable' if it provides access to environmental or task information which is not easily directly accessible using an existing sensory modality, or if such information would require conscious or cognitive processing (such as logical derivation or calculation e.g. cardinal directions) to acquire. The success of the TVSS demonstrates the power of brain plasticity, and that similar approaches can yield practically positive results in a range of tasks generalised as widely as the natural senses. Even when such SSADs are slow to learn (as is often the case), the independence gained through proficiency with such devices creates incentive for their uptake and encourages persistence in the learning process.

SMT models cognition as arising from deliberate, skilful action and perception by independent, goal-driven agents. In the possession of relevant Sensorimotor Contingencies, agents may take action (and understand consequential environmental changes) in order to achieve these goals. In other words, SMC's enable the skilful and deliberate achievement of agent-driven goals. As such, I reason that the SMC, or action-sensing loop, that an SSAD supports must be related, relevant and *valuable* to agent-driven goals.

In the field of HCI, additional challenges arise. SSADs offer a novel way to connect with data and machines, however, unlike Sensory Substitution Devices for rehabilitation, such as a visual-to-tactile device as above, they are rarely a mechanism enabling independent daily practicality. For this reason, incentives and advantages of using a sensorimotor interface must be carefully considered, and the affordances which are revealed through device mastery must be clearly and carefully linked to the task at hand. In specialist fields where dedicated training is required anyway (such as piloting aircraft, or sports), the advantages of including such interfaces amongst other facets of training may be sufficient. However, in consumer applications, even if incentives and advantages are apparent, it is unlikely that mainstream success may be achieved without further consideration for how mastery of the device may form without significant investment.

4.2.3.2 Paradigm 2: Make it Practical (2000-)

Since the millennium, the minification and cheapening of batteries, transducers and processing power has improved the widespread adoption of SSADs through two avenues. Firstly it has reduced the costs of providing them as practical daily aids: compare Meijer's original work [117] to the current implementation of The vOICe [116]. This resulted in publicity and awareness that such systems exist, essentially exposing the technique to a wider audience of researchers. Secondly, the lower cost has increased accessibility to a wider community of tinkerers and engineers, explaining the explosion in devices built for research in the past decade (see Table 4.2). Most post-2000 devices follow engineering principles established in published guides [14, 100] and have some appreciation for the need for practicality, allowing conceptual consideration of their use during in-the-wild testing and beyond. Maidenbaum [115] suggested that visual sensory substitution device uptake has historically been limited by expense and unwieldiness, and the currently limited daily practical advantages of equipping a sensory augmentation system further exacerbates this disincentive. Hindrance may be caused through limitation of movement (such as a physical impedance in performing the task), or limitation of attention (through physical discomfort, irritation, or continuous excessive stimulation). By adopting a form such that the interface does not hinder, even when being learnt, such barriers may be minimised.

4.2.3.3 Paradigm 3: Make it Sensorimotor (2010-)

Bach-y-Rita, the inventor of the TVSS [9], considered Sensory Augmentation long before the first real explorations were executed [124]. Evidence concluded that the unnatural sensations are learnable, however, it was also confirmed that the sensation must be learned 'actively'. This has been previously observed in SSDs [9] and even earlier, in the natural senses [72]. This evidence provides strong support for SMT [133], which, as detailed above, ties motor actions and associated sensory stimulae together through lawful relations known as Sensorimotor Contingencies (SMCs) [85]. Modern SSADs lean heavily on SMT (and related 'post-cognitivist' theories of cognition [62, 107]), and it is now widely accepted [6] that modelling around such action-perception relations is integral to the successful design and use of SSADs. A device or system may be considered 'sensorimotor', if it allows formation of a sensorimotor contingency, as guided by SMT (Chapter 3).

A system may therefore be considered sensorimotor not by its form or application, but by the nature of its interaction. Users must first and foremost be able to experience the signal (i.e. it must in some fashion produce a Stimulus detectable by the target Receptors). Further, the user must be able to take actions which cause that signal to change in lawful response. This lawful response allows the formation of a new sensorimotor contingency between the source (even if virtual) and target modality, with the interface bridging between (see Figure 4.2). If the user is unable to take action that causes the received Stimulus to change (either through passive change in perspective, or through environmental effect), then a Sensorimotor Contingency cannot form. Conforming to the requirements of a sensorimotor interface therefore requires supporting a lawful, enactive, dynamic connection between an actionable Source and a sensible Stimulus.

4.2.4 Application of the Paradigms

To demonstrate the effect of these paradigms on the approach to building an SSAD, consider for example the evolution of the communication of language through vibrotactile interfaces. Initial endeavours by Geldard [63] used encoded symbols, essentially teaching word by word.

More recently, the heavily encoded intuitive syllable-based approach of Zhao [171] led to further increased generalisability and usability. However, the limited symbol set and static, non-reactive learning mechanism directly impacts learnability. Many symbols may be learned, but without scope for dynamic user-guided action it is unclear how new sensorimotor contingencies may develop.

Instead, Eagleman and Novich [51, 130] took a more direct, acoustic approach to encoding. The dynamic, sensory nature of this early research allowed deaf users to develop an appreciation for audio through a directly mapped vibrotactile sensation, and eventually led to the commercially available Neosensory Buzz [127]. Unlike the more symbolic solutions, users proficient with the Buzz demonstrate unguided (self-taught) improvement in competence, and the device allows a degree of generalisability to non-verbal sounds. The Source and Stimulus/Receptor of each case is comparable, with the difference in practical limits attributed to the overall approach to valuable signal, practical design and sensorimotor interaction.

Creating effective new SSADs therefore depends on the design of a practical device, the creation of a strong coding, and the enabling a system of sufficient complexity and practicality for valuable substitution or augmentation.

In embracing the paradigms of interaction, the Buzz therefore demonstrates how the paradigms may be applied to transform a device from a niche SSAD (under functional or operational definitions) into a device that allows genuine changes in perceptual experience: a sensorimotor interface.

Here, I will expand on the significance and implications of the need for robust action-perception relations, and derive principles for the application and implementation of sensorimotor interfaces.

4.2.5 Sensorimotor Interfaces in HCI

Despite the clear and ambitious dreams and expectations of Paul Bach-y-Rita, SSADs have not seen significant expansion and implementation. Studies are often sporadic, and comparability between SSADs in general has been widely recognised as an issue in the field [70, 107].

Following its initial creation as a tool for exploration of Sensory Substitution and expansion of perception (as discussed in detail in Chapter 3), the feelSpace belt [124] was commercialised as a navigation device to allow hands-free provision of directions for cyclists [1]. In its capacity as such a device, user action does not cause changes in signal, so it is difficult to see how this could affect perception. By contrast, the feelSpace belt has also been used in detailed investigations into how more sensory signals may affect perception [85, 96].

More recently, in HCI, the feelSpace belt was used in vehicle driving applications which aim to expand perception through provision of dynamic, sensory signals, but without specific reference to the field of Sensory Augmentation. The two studies (both by Krüger et. al [92, 93]) are further noteworthy for their demonstration of the potential for sensorimotor interfaces in realistic HCI applications, and for highlighting how solid implementation and realistic testing may reveal device and system potential without the need for significant long-term studies. While ‘Sensory Augmentation’ is provided as an author keyword, the reports avoid this terminology and the justification for presentation of such a dynamic signal, and note that similar statistically significant experimental results may arise from more simple ‘notification-like’ techniques. Both training time and the testing methodology are highlighted as issues which may cause this.

There are several devices which are labelled as SSADs, but which do not allow formation of an SMC, either as they present symbolic data (and thus cannot be mapped to continuous World properties) or static information (unaffected by user action). As discussed in the above taxonomy and paradigms, while potentially effective assistive devices, these systems do not constitute sensorimotor interfaces. Prominent examples of such ‘functional’ SSADs would include the Optacon [104] or Tactaid [88].

It would seem there are two major pitfalls to sensorimotor interface exploration in research: design (including application choice) and testing (including training/introduction).

Design. As many approaches to interface design, testing and training are optimised to evaluate whether affordances have been well-presented, they may not always be successfully applied to sensorimotor interfaces, where the purpose instead is to highlight *new* affordances through altering perception. Instead, I here utilise insights from SMT to highlight a pathway to improved learnability and usability and achieve such interactions. In combination with improved recognition of suitable applications and more informed methods of device evaluation, sensorimotor interfaces have the potential for impact through comparable laboratory studies, and further, wider usage and uptake in real world applications.

Testing. A reason for limited comparability is perhaps that work is performed independently (rather than through collaboration or derivation) using inconsistent metrics against varying standards. Devices are built, evaluated and reported on by a single group, with a very small minority of devices being reused despite overlaps in capabilities (see Table 4.3). Each group builds its SSADs for a specific purpose or application before evaluation through unique metrics and requirements. The resulting incomparability between reports hinders both device generalisability and study findings, and limits impact on the development of new devices and theory around the subject. A methodological approach to evaluation (addressing the unique qualities of sensorimotor interfaces as tools for *perceptual alteration*) is considered and adopted in this work.

4.2.6 Cognitivism and Enactivism

As per Chapter 3, the strong link between SSADs and action-perception models of cognition (such as the previously discussed SMT [133]) is arguably the key to differentiating such systems from the majority of multimodal interfaces found in HCI. Froese [62] poses that SSADs and similar sensorimotor interfaces may be differentiated from more traditional multimodal devices through the consideration of their approach to interactivity, which is broadly categorised as being either ‘cognitivist’ or ‘enactivist’.

The ‘cognitivist’ approach (as summarised by Lobo [105]) models cognition as a process separate from perception and action, arguing that cognition consists of information-processing so as to form, store and act upon symbolic representations of the outer world. Such a model works well for both single- and multi-modal interfaces which present explicit information for user digestion and careful consideration. The traditional calculator is a simple example of such an interface for which this model works; real-world quantities and mathematical operations are represented on both input and output through symbols which must be understood before meaningful work can be performed. In their design, such interfaces require consideration of usability factors such as those described by Nielsen’s [128] usability characteristics. Such factors are designed to maximise intuitiveness through behaviour consistent with user expectations and prior symbolic associations, advertising system ‘affordances’ to invite explicit actions.

By contrast, ‘enactivist’ approaches commit to the idea that cognition is a dynamic process that does not need to rely on representations. The design and use of SSADs leans heavily on enactivist principles, where such dynamic processes are modelled by theories such as SMT [133], or related post-cognitivist fields of environmental psychology. The consequence of this approach is that the extraction of information from such interfaces (and perception more generally) derives from the process of goal-driven interaction and engagement, rather than the comprehension of explicit, representative information [163]. The information and interaction associated with such interfaces may more closely resemble riding a bicycle than it may engagement with a traditional computational interface. Explicit, symbolic information such as speed, elevation or altitude may not be directly accessible, however qualities such as balance, terrain texture, and the structure of the bicycle itself must be understood and utilised while engaged in the activity, allowing meaningful work (forward propulsion) to be performed. The effect of engaging skilfully with such a system is that the system itself becomes invisible while environmental affordances change - a hill, corner or distant feature will look and feel different when riding a bicycle than when on foot.

Expanding and applying the perceptual models of enactivism (and more specifically, SMT) to the domain of HCI and SSADs requires translation of the model's principles into engineering and interaction guidelines. To complete our set of requirements for the modern sensorimotor interface, let us clarify the details of such an interaction.

4.2.7 Lawful Interactions

Ward asks, in context of vision-like experience arising from use of The vOICe, whether *"any consistent mapping between vision and audition may lead to these kind of experiences? The latter is predicted by sensorimotor accounts provided the user has control of the camera [166, p499]"*. In coordination with the suggestions of Visell [163], in its simplest form [163], the Coupler constitutes a homeomorphic mapping of input to output signals, with interpolation and dimensional reduction accounting for discrepencies in the number of sensory and actuation dimensions.

I have detailed that an SMC is the lawful mapping of Action to change in Stimulus. By the model and insights derived above, the argument summarised is that if the sensorimotor interface is to support new perceptual experiences, then the interaction with the environment *through* the sensorimotor interface must itself be similar in nature to natural apparatus-SMCs, lawfully and dynamically coupling Stimuli reflecting World properties to user Action. This concept is illustrated in Figure 4.2, which shows the sensorimotor interface as part of the natural action-sensing perceptual loop previously described in Figure 3.2. The resultant coupling between Action and Sensing *through* the sensorimotor interface may be thought of as a composite apparatus-SMC, where the natural sensory laws and the sensorimotor interface laws are combined into a single lawful relation. It can then be argued that the perceptual experience which emerges from exercised mastery of this composite apparatus-SMC is the sensory augmentation experience that we are seeking to design.

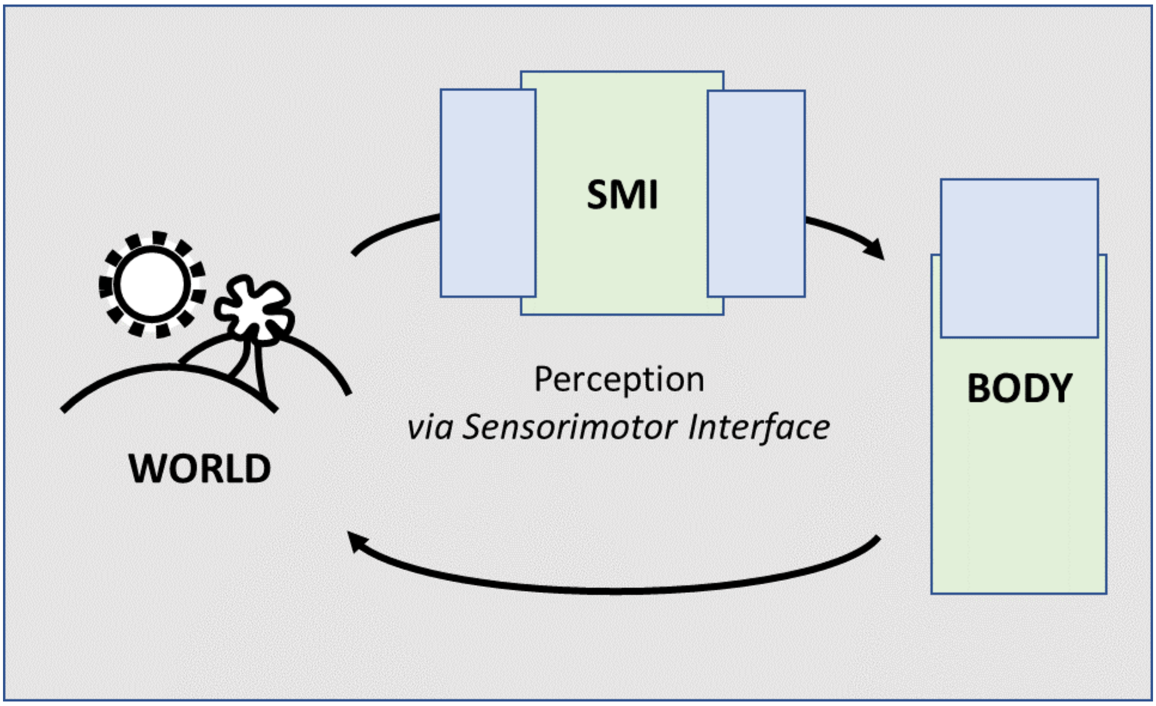


Figure 4.2: Augmentation of perception using a sensorimotor interface. The lawful mapping between Action, World and Stimulus, consistent with sensorimotor theory, is a critical requirement for a sensorimotor interface.

All components of the interface are critical in this process. The Source must be capable of sampling such that it continuously and accurately measures its chosen World property. Optimal Transducer performance is achieved when the output Stimulus is of a resolution and dynamic range which matches the biological capabilities of the associated Receptor; a process which may be optimised as detailed in Chapter 5. Finally, the Coupler must enable a fast and consistent mapping between the Source and Transducers.

As I have defined the composition, and the nature of interaction required for building a sensorimotor interface, it is now important to address how the introductory or training approaches that may be suitable for cognitive interfaces (such as the presentation of affordances) may not be suitable for sensorimotor interfaces. The final part of this chapter constitutes an exploration of the factors that influence learning and development of proficiency in the use of SSADs and sensorimotor interfaces.

4.2.8 Feedback and Learning of Sensorimotor Interfaces

Although there is strong evidence that the nature, time and purpose of training play important roles in the outcome of training [19], to-date there is a deficit in quantitative research clarifying the effects of these attributes. Furthermore, there is disagreement as to both the potential for and measurement of perceptual integration achievable by SSADs by any mechanism [46, 167]. Some efforts have been made to develop regimes allowing categorisation and measurement of perceptual integration [124, 145, 156] though experimentation enabling classification of devices and training protocols into recognised regimes is not yet established.

The effects of learning through deliberated use and experience is to-date an under-explored mechanism, given the lack of long-term SSAD studies outside those of the Feelspace [124] and vOICe [117] projects. However, given the resource- and time-intensive requirements of supervised training, such sustainable, user-driven learning mechanisms may eventually prove key to the widespread adoption of sensorimotor principles in industrial and commercial environments [12].

How sensorimotor interfaces are learned, mastered and utilised is therefore a topic of continuing research, driven both by the insights they may reveal regarding formation and mastery of natural senses and perceptive skills, and through a desire to increase penetration and uptake in assistive technologies markets [19, 114]. The difficulties and dedication associated with long-term exposure for mastery (in some cases continuing to evolve over many years [166]) remains a barrier to uptake for commercial SSADs, even in high incentive applications such as audio or visual sensory aids.

Active training has been identified as playing a critical role in the development of SSAD mastery [91], though research investigating how the nature and content of training may be optimised is ongoing, and models for mastery which can be adapted to reveal learning mechanisms are limited.

From one perspective, guided learning and explicit feedback are suggested as effective mechanisms for training [5, 19], through either simple confirmative or more complex fine-grained information presented during or after task performance. However, this typically requires a prolonged period of expert input during periods of dedicated training, representing a major barrier to uptake [114]. Additionally, given the herein presented non-representationalist approach to modelling sensorimotor systems, the utilisation of explicit (declarative) feedback to assist in the learning of new continuous, perceptual (procedural) skills seems somewhat counter-intuitive, if not directly incompatible, with sensorimotor principles [75].

Instead, Higueras-Herbada offers a dynamical systems-derived approach to learning of sensorimotor skills labelled 'Direct Learning Theory' [75], expanding on models of learning new sensorimotor tasks in an ecological psychology fashion largely compatible with SMT. Essentially, if we are to model sensorimotor interactions using post-cognitivist methods, then the path to mastery of such interactions also requires an approach consistent with post-cognitivist psychology: *"a theory of learning from the approach should be consistent with the commitments to describe cognition [75, p393]"*. In other words, if cognition is modelled as an enactive process, then the methods of learning pursued should also be consistent with enactive principles.

Direct learning theory also provides a satisfying mechanism for the acquisition of perceptual skills and experiences through 'learning by doing', rather than necessitating the existence of external feedback, or any form of representationalist or explicit understanding of the task at hand. By direct learning theory, an agent learning new sensorimotor skills needs to be engaged in a goal-driven task, and exposed to changes in relevant sensory stimuli in the context of their actions in order to refine their sensorimotor engagement.

Such theories, when formalised and summarised, may prove useful for frameworks of learning in enactive interface design. While the direct learning theory detailed in Higueras-Herbada's publication pursues a rigorous mathematical approach to modelling the learning of a simple sensorimotor task, it also allows wider, qualitative consideration of enactive learning as being comprised of three key components, which I summarise here:

Intention The user is a goal-driven agent attempting to execute sensorimotor skills to achieve their desired goals, in terms of the 'action they aim to perform, or what property they aim to perceive' [75]. In the development of a new sensorimotor skill, the user is learning to improve correlation between their *intended* action and perceived environmental properties. For effective learning to occur, the user must therefore have clear goals which can be expressed in terms of sensorimotor actions and perceptions, and the success of these goals will contribute to the education of the user in future choices and sensorimotor mastery. Truly effective tasks will have clear goals and provide feedback guiding the user to greater proficiency through subsequent execution, as well as providing direct sensorimotor feedback during task performance.

Attention The user is learning to focus their *attention "toward the environmental variables most relevant to task completion [75,]"*. With practice, users fine-tune their attention to sensory activity and action which contribute to task progress. Sensorily, this includes both filtering out environmental noise and increasing sensitivity and comprehension of the signals that carry relevant task-related information, while actions become more specific, accurate and less wasteful (visualised in various figures by Higueras-Herbada [75]). In the context of application-specific sensorimotor interfaces, careful provision of signal reflecting variables rich in task-relevant information offers efficient pathways to learning.

Calibration The user utilises feedback (in the form of action/sensing relations and intent/outcome differences) to *calibrate* sensorimotor activity. Feedback may take many forms, though without it, the user is unable to correct for judgemental errors and cannot gain sensorimotor proficiency. While feedback in its simplest form (commonly applied during training) may be explicit and quantum as mentioned above [19], such approaches may suggest a requirement for representational models of perception which this taxonomy seeks to leave behind. Instead, a truly enactivist approach to learning requires adaptive mechanisms that derive directly from the experience of task performance [48], suggesting a normative evaluation or continuous process of behavioural adaptation is ongoing during the sensorimotor learning process.

Following direct learning theory, I argue that adaptation of the training environment to optimise these three mechanisms should prove an effective pathway to improving (or at least enabling) development and learning of new sensorimotor skills through sensorimotor interfaces. As such, when creating new sensorimotor experiences, pursuit of direct learning theory requires a system which enables progress towards a clearly-defined goal (intention), using a low-noise, highly-relevant signal (attention), with readily-available feedback on action and sensing through multiple channels (calibration).

4.3 Summary: Design Dimensions of Sensorimotor Interfaces

The above sections provide a detailed yet verbose account of the application of the enactive SMT to the creation of novel sensorimotor interfaces in HCI, from the origins of their interactive principles in sensory substitution and augmentation. This approach lays the theoretical groundwork required to define and build sensorimotor interfaces consistent with enactive models of perception, constituting a hereto unique design primer and field introduction.

At this point, in summary and for future reference, I condense the above information into a table split by five identified design dimensions, representing entry points for the design and implementation process.

4.3.1 Design dimensions

The choosing and description of the following design dimensions began with a careful exploration and consideration of numerous possible dimensions which may describe the interface; concerning the user, technological limitations and decisions, interaction features, applications, learning and teaching mechanisms and the resulting perceptual experience and sensorimotor skill sets. These starting dimensions were largely sourced or inspired by the key considerations for design presented in SSAD reviews by Kristjànsson et. al [91], Visell [163] and Maidenbaum et. al [114].

The implications of choosing these dimensions were subsequently detailed and compared in context of the above paradigms and learning principles. They were then iterated through combination and refinement to minimise redundancy and overlap, and to maximise utility and insight from a design perspective.

One dimension commonly considered in such design frameworks, but omitted from this implementation, is a ‘user’ dimension, which would generally detail the role or profile of the user independently from the technology. However, as the function and effectiveness of enactive interfaces are dependent on their utilisation and mastery (rather than simple design intuitiveness), attempting to de-convolve the user from the device and interactions presented a large degree of redundancy in the table. Instead, the ‘user’ manifests as an integral component of the Interaction, Learning and Perception dimensions. Presenting the dimensions from a holistic user-device perspective streamlines communication of the most important concepts.

Further, as ethical considerations of altering or enabling perceptual experience (while acknowledged as important) are largely out of scope of this thesis, such dimensions have been omitted in this summary, resulting in presentation of design principles from a largely practical, technological standpoint.

In summary, the following five design dimensions were chosen in consideration of the critical points addressed above, to provide a clear and complete high-level overview of the factors affecting sensorimotor interface design and implementation.

From a practical perspective, the **Application** (Design Dimension A) is a key starting influence in the design process. Any **Device** (Design Dimension D) limitations will also drive the system design process. As this and the previous chapter show, the method of **Interaction** (Design Dimension I) and the unique enactive nature of sensorimotor interfaces in contrast to traditional devices in HCI imposes constraints and considerations on design. Alongside this, sensorimotor interfaces are designed to reveal new environmental affordances through augmentation of **Perception** (Design Dimension P), therefore existing perceptual factors must be considered. Finally, when moving away from representationalist cognitive models, training and passive **Learning** (Design Dimension L) of sensorimotor signals is a non-trivial matter which demands integration with other design components.

These dimensions are further defined as follows:

Device	The physical device, including source/transducer apparatus and coupler algorithms. Device, in this sense, is a critical dimension referring to all physical, technological components of the interface.
Interaction	The mechanism or paradigm of engagement. Under SMT, the dynamics of interaction define to a large degree the Sensorimotor Contingencies which give rise to perceptual experience.
Application	The task and environment in which the system is applied. Sensorimotor Interfaces enhance performance by augmenting perception to reveal new environmental affordances.
Learning	The conditions under which users may gain familiarity and competence in device use. Consideration for the conditions and mechanisms allowing development of device proficiency must be taken at the design stage.
Perception	Mechanisms and factors describing perception, considering both existing and emerging perceptual experience.

4.3.2 Design summary

In presentation of Table 4.3, dimensional influences are intended to be read across-then-up. The row header represents a fixed, constrained or hypothesised parameter (driven for example by hardware availability, desired/required application, research interests or even design decisions arising through the process). Reading across, the impact of the constraint on other dimensions is summarised at the intersecting cell, where further detail may also be considered through reference to the relevant (linked) paragraph:
p1 = Paradigm 1: Make it Valuable
p2 = Paradigm 2: Make it Practical
p3 = Paradigm 3: Make it Sensory
l1 = Learning: Intention
l2 = Learning: Attention
l3 = Learning: Calibration

* = Further information illustrated through experimentation in Chapter 5.
** = Further information illustrated through experimentation in Chapter 6.

Pursuing such a sequence of constraints and decisions until the designed system reflects all summarised dimensions and their intersections is intended to enable guidance towards a reliable, grounded sensorimotor interface through our structured, flexible framework.

4.4 Discussion - Designing Sensorimotor Interfaces

Sensorimotor interfaces are niche, and not without limitations. However, as understanding and use of the herein detailed techniques grows, their principles will be applied to an increasingly wide range of applications. In this section, I outline how Table 4.3 may be used to guide the building of new sensorimotor interfaces in HCI for maximum effectiveness, using the design and implementation of a modern sensorimotor interface, 'The Lateral Line' [93], as an example. The Lateral Line is a system which dynamically encodes directions and temporal proximities towards approaching objects into a vibrotactile belt, implemented to augment perception in a driving environment with the goal to demonstrate performance benefits in terms of enhanced driving safety.

Most situations where the building of an SI will be considered begin with such an application or task requirement. However, before applying enactive principles to the creation of a new interface, it is wise to ascertain whether this approach is truly suitable for the application. Simple symbolic output from interfaces which respond to specific user commands (such as button clicks or text entry) are often the optimum choice in applications where explicit, precise information must be carefully considered and interpreted by the user, for example the settings on precision machinery, or the output from a calculator. Interfaces which demand semantic, explicit or cognitive understanding are best offered symbolically, while SIs may be described as best suited to procedural, implicit or sensory interactions. Ascertaining whether the source of information in the application is best reflected by representationalist or enactivist psychology [48], through cognitivism or enactivism [62], is the first stage in system design, and may be approached through understanding of the application and the nuances of the cognitive models described above. However, there are some rules of thumb and further observations which may prove beneficial. For example, some indication of the application choice comes from a need to plan and anticipate. If one understands the consequences of ones actions in a dynamically shifting application, and senses those changes, one can act more dynamically. The Lateral Line is a system designed to provide information about traffic approaching on out-of-vision trajectories, and therefore begins with an application setting and goal. A rudimentary symbolic notification system could be used effectively to suitably warn drivers of threats, however, with access to additional salient information a competent user may be able to plan head or enact nuanced judgement, benefiting from look-ahead in addition to the safety improvements of a potentially simpler system. The application is therefore highly suited to the implementation of a sensorimotor interface.

Given SMT's emphasis on the link between enactivity and perceptual experience, considering supported device interaction (and whether it is even compatible with enactive principles) is critical to determining application suitability. SIs are unsuitable for communication of passive information which does not change with perspective or Action (such as weather forecasts, or stocks and shares), unless the interface also extends motor skills into the domain, allowing active exploration.

By contrast, SIs are suitable where the continuous, dynamic nature of the signal enables nuanced decision making in demanding applications requiring high bandwidth interactivity in a range of potentially novel circumstances, but where specific understanding of the exact figures is not critical. Applications for such interfaces are still emerging, though games [142, 143], sports [40], navigation [96] and driving or piloting vehicles [93, 123] have shown promising suitability. Virtual and augmented reality provide especially well-suited opportunities for experimentation and prototyping [37]. In our exemplary application, users of the Lateral Line have direct control of the vehicle from which proximity data is continuously calculated, rendering the signal dynamic and continuous, and offering a high-bandwidth perspective on nearby hazards to assist in decision making.

Once it has been determined that a sensorimotor interface is suitable for the application, the device itself must be designed: the Source, Coupler and Transducer must be specified. The Source for a Sensory Augmentation Device may be chosen from an almost unlimited domain - any World property which may be measured may be utilised as a Source. Many devices Source from single, simple World properties such as magnetic fields [124] or distances [62]. Novel Sources may also be formed from artificial values, application-specific metrics [46], or user parameters [66]. For consistency with the other design dimensions, it is essential that the Source follows Paradigm 1 (reflected by the learning principle of Intention) and offers valuable new task-relevant information the user, as superfluous task information may be ignored. Further of the impact and nuances of Paradigm 1 will be discussed in Chapter 6, where the value of a novel augmented signal relative to existing stimulae is compared and explored experimentally. The Lateral Line combines several measured World properties [92, 93] to calculate information pertaining to impending collisions, in a fashion similar to the pressure waves detected by creatures employing its namesake biological sensory apparatus. Under this design, fast-approaching traffic that may result in a collision is communicated through intense vibrations, while drivers travelling at the same relative speed as other traffic do not receive any vibrations. Such a source was biologically inspired, though note that it also carefully follows Paradigm 1 and optimises user

Attention towards the most important information, providing signals pertinent to decision making while minimising the communication of static information.

The natural senses suitable for engagement with sensorimotor interfaces are essentially limited to touch, sight, hearing and electrostimulation. However, within these senses there are multiple Receptors which may be engaged, and within Receptors there are multiple (natural) modalities which contribute to existing perceptual experience. Similarly, as modalities reflect collections of SMCs, they may be supported by more than one Receptor (such as balance, which depends on sensory signals from vestibular and visual systems). In choosing a suitable Stimulus, the information density of the proposed modality, its match with the chosen Receptor, and its impact or interference on other information related to the application must be considered. Consider Visell's notes [163] discussed in Chapter 3 regarding the choice of output signal. In order to ensure linear values were perceived linearly, Krüeger et. al exponentially scaled their device's vibrotactile signal to compensate for the body's logarithmic perception of vibrotactile intensity, testing several mapped exponents to minimise irregularity and resulting in a linear perception. The importance of such a strategy has been remarked and implemented in detail in previous studies [148], though the exact details of the procedure are not always well-documented or well-implemented. Further information on how, and why, the calibration and scaling of devices for linear perception is critical, is detailed experimentally in Chapter 5.

Finally, to maintain comfort and minimise distractions, the device form and signal must honour Paradigm 2: the user must be able to utilise the device with minimal physical or distraction impairment to the task. This is again well-considered in the Lateral Line, owing to its communication of 'temporally impending collisions', rather than simple proximal information. By only communicating dynamic information, the device minimises presentation of vibrotactile signal in circumstances where the user does not need to consider taking action.

There are several further points of interest represented by intersections in Table 4.3. For example, cell LP suggests that to improve uptake, the system should provide benefits from the get-go. Results from The Lateral Line specifically note that the nature of the encoded signal provides benefits that may be represented or replicated by symbolic, notification-like interactions which simply draw user attention towards incoming vehicles. However, within the signal there is more information which over time may be interpreted and utilised as the user gains experience through regular use. In such cases, while the notification-like qualities provide immediate advantages, improvements in signal comprehension will lead to more nuanced decision making, look-ahead and ability to plan: a true sensory augmentation. Similar principles have been noted in other applications, such as the vOICe vision to audio sensory substitution device allowing users to understand if there is something nearby or far, above or below, within minutes. With time, more complex motion etc. can be perceived [166], yet here, I argue that the advantages provided from the get-go represent an important pathway to mastery through presentation of incentive-linked action.

Table 4.3 cells AD and ID warn that, through careful implementation of form and signal, the use of a sensorimotor interface should not impair the user in comparison to performing the task unaided. Such consideration should apply to both the form/location of the device, and the nature of the signal. In the Lateral Line, the belt-worn device does not impair user senses or actions critical to driving (such as the hands or head). Furthermore, as noted above, communication of impending collision, rather than passive proximal information, reduces the vibrations users experience in circumstances where action may not need to be taken, helping to avoid desensitisation or frustration associated with constant vibrotactile stimulation.

Finally, as the torso is also unlikely to rotate in comparison to the vehicle, note that the consistent orientation is likely also beneficial to allowing the user to attribute signal changes to external environmental events.

4.4.1 Building Sensorimotor Interfaces

Prior to building new devices that explore whether novel augmentations are feasible, one should first consider whether reuse of an existing device or framework may be of greater benefit to the community, where the contribution lies in developing a more generalisable or otherwise superior encoding (Coupler) algorithm, or connection of an existing device to a previously unutilised Source. This, in conjunction with solid experimental design (see next section) would also serve to homogenise the field, and in time establish well-founded standards.

Several SSADs and toolkits already exist which have the potential to, or already been demonstrated to, implement sensorimotor principles in diverse applications. Most of these feature APIs or open-source access to the Coupler, reducing effort required in low level implementation and providing fast access to hardware proven to support sensorimotor interactions. Notable examples include the feelSpace Belt [1], TongueDuino [50], and vOICe [116], with more recent devices of high potential and flexibility including the NeoSensory Buzz [127].

4.4.2 Experimentation

Table 4.3 may also help identify literature gaps and opportunities for impactful research. In circumstances where device constraints impose limitations on other dimensions, the literature may not offer satisfactory solutions, thus highlighting groundwork to be performed before concrete design decisions can be made. For example, with fixed Device requirements, performing investigative studies into Application suitability calls for optimal Perceptual integration, and an efficient training (Learning) environment to allow rapid evaluation of device suitability. Such constraints led to the development of studies focusing on the effect of existing perceptual biases on new signal interpretation (as in Chapter 5), and the role of concurrent sensory information in feedback (Chapter 6).

The type of experiment performed with an SSAD for greatest contribution can be deduced from a combination of the above principles and the *intent* of the work. Given the infancy of the field, a focus on re-utilisation of devices and techniques across publications will help consolidate standards, while further work exploring the fundamentals of perception and models will establish which of these standards have been derived from robust experimentation, and which are simply self-reinforcing [19]. The following two chapters will explore in greater detail the current standards and limitations of evaluation methodologies in the field.

4.5 The State of the Art

With clarity on enactivity and sensorimotor device design now established, I may now outline the state of the art through a brief recap of the latest efforts in the field.

The current state of the art in enactivity and sensorimotor interface design shows progress in two major directions: further understanding of human perception and enactivity as a viable theory of mind, and innovation in device design and application.

4.5.1 Human Perception

Alongside ongoing philosophical debate, 4E approaches to cognition continue to evolve and develop through empirical contributions. In similar nature to the early work by Froese et. al [62], sensorimotor interfaces continue to prove critical tools in such practical empirical work. Lobo directly explores 4E approaches through utilisation of customised sensorimotor devices for full-body movement [107, 108], gaining insights into both the effectiveness of 4E theory, and yielding practical insights that may directly inform human physical education [8].

4.5.2 Devices and Applications

Regarding applications and innovation in enactive interfaces, the majority of developments embracing and utilising sensorimotor principles fall into three broad categories, each with highly correlated design and use:

In the first instance, they manifest as navigation or orientation aids, featuring a compass or gravity-oriented accelerometer, presented through some audio or haptic signal. Such devices are designed to augment or assist human perception of space, through provision of extra-sensory orientation information. The feelSpace belt [96, 124] is an early examples of this, directly leading to more modern developments such as the Auditory Compass [149]. This approach has also inspired devices featuring wide-area or whole-body vibrotactile stimulation, with wider acceptance of the breadth of potential sources for the data. However, despite their now well-established design and history, key open questions remain surrounding the richness of the perceptual experience these devices may offer [149], and drive innovation on how best to assist users in achieving device proficiency [37].

In the second category, devices aim to offer vision-like experiences through presentation of encoded audio or haptic signals. Focus remains centred on development of new algorithms for efficient information encoding [2], and the investigation of effective training regimes for improving user skills [27]. Given sensorimotor theory's origins in explaining vision-like experiences, many contributions developing from sensorimotor principles and the application of associated theories stand to directly inform the design of such sensory aids and their training programmes. Broader efforts to augment sensory experience holistically are demonstrated through devices such as the Neosensory VEST and Buzz [127].

The third category of devices offers an augmentative distal touch (or similar external object/spatial-derived) experience through a haptic signal. Almost all early efforts in some capacity resemble the Enactive Torch [28,32,62]. These devices present distal touch information directly, though rely on high levels of task concentration and user focus for signal utilisation. By contrast, more modern approaches have begun to consider carefully how distal touch or related proximity information may be carefully encoded for direct task relevance and lower attention requirements [87]. Krueger et.al. [92,93] demonstrate such encodings and their effectiveness in their augmentative driving experiments.

It is likely that there is merit and reason behind taking either approach to encoding spatiotemporal proximity information, and that both will benefit from greater understanding of learning and perception in lock-step with the evolution of 4E theory.

Despite representing the majority of sensorimotor interfaces demonstrated to-date, these categories are not exhaustive of the potential applications for enactivity in HCI.

It is possible that further innovation is at least partially limited by the accessibility of knowledge in the field, requiring the organisational groundwork that this thesis has attempted to rectify to this point. Further education and understanding of enactive principles and human perception

Table 4.1: List of commercial sensorimotor interfaces

Device Name	Year	Source	Stimulus	Purpose
vOICe	1991	Visual	Audio	SS Blindness
TVSS	1969	Visual	Haptic	SS Blindness
BrainPort Vision Pro	2005	Visual	Electrostimulation	SS Blindness
FeelSpace Belt	2005	Compass	Haptic	SA Navigation
BrainPort Balance Plus	2017	Orientation	Electrostimulation	Vestibular Correction
Neosensory Buzz	2018	Audio	Haptic	SS Deafness

may therefore be enough to encourage innovative thinking on this front [91, 114].

With each innovation, sensorimotor principles are evolving to meet the modern desire for greater bandwidth in interactivity, greater intuitiveness, reduction of cognitive overload, and spanning the attention gap.

4.6 List of Sensorimotor Devices

To conclude this chapter, I present the following two tables in summary of the devices most relevant to the field, serving as a brief referential overview, while also highlighting that there is yet much work and innovation to be pursued in the field, and setting the scene for the contributions presented in following chapters.

Table. 4.1 provides an overview of existing devices commercially available for use as SSADs, all of which have some degree of potential for wider sensorimotor application.

Literature demonstrates that many devices are developed for lab experimentation which are not available for commercial use. As such, Table. 4.2 details these systems with a primary focus on devices which aim at real world application or purpose, grouped into the broad categories described in this section.

These lists have been compiled to highlight influential and successful systems, unique implementations and highly-cited or useful exemplars in the field. They have been populated through a combination of distillation from other reviews (e.g. [6, 80, 163]), and methodological consideration of publications referencing SSADs and/or experimental investigation of enactivity. Critically, all devices have in some fashion through their associated publications (referenced in the table) demonstrated design, evaluation or interaction principles consistent with *enactive* models of cognition.

Table 4.2: List of scientific sensorimotor interfaces - grouped by general category of application, ordered by date of publication

Study	Year	Source	Stimulus	Purpose
TSAS [123]	2004	Orientation	Haptic	SA
FeelSpace Belt [124]	2005	Compass	Haptic	SA
EyeMusic [2]	2014	Visual	Audio	SS Blindness
Neosensory VEST [130]	2015	Multi	Haptic	SA
TESSA [146]	2015	Multi	Multi	Toolkit
Sonic Eye [154]	2015	Distance	Audio	SS Blindness
Auditory Compass [149]	2017	Compass	Audio	SA
Skin Stretch [134]	2017	Orientation	Haptic	Vestibular Correction
TLC [171]	2018	Audio	Haptic	SS Deafness
Sonic Pathfinder [74]	1984	Distance	Audio	SS Blindness
Upton Eyeglasses [52]	1991	Audio	Visual	SS Deafness
PSVA [29]	1998	Visual	Audio	SS Blindness
The Vibe [70]	2010	Visual	Audio	SS Blindness
Tongueduino [50]	2012	Multi	Electro	Toolkit
Echo Distance [126]	2017	Distance	Audio	SA
Enactive Torch [62]	2012	Distance	Haptic	SA
Firefighting Glove [32]	2013	Distance	Haptic	SA
Tactile-Sight [28]	2013	Distance	Haptic	SA
Superhearo [86]	2016	Audio	Haptic	SA
The Lateral Line [93]	2020	Temporal Proximity	Haptic	SA

4.7 Summary

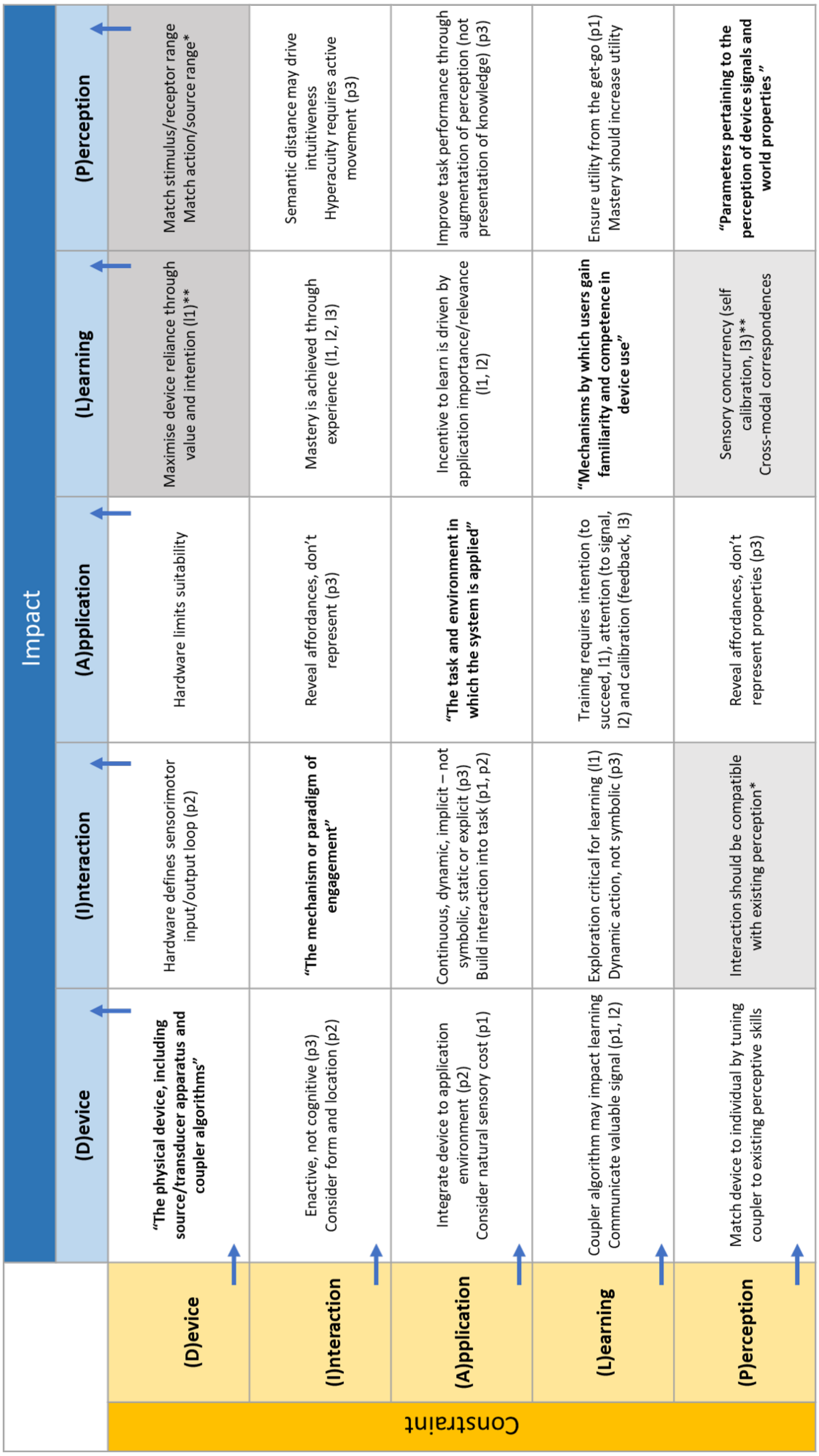
Through this chapter I have outlined the importance, applications and potential of sensorimotor interfaces; devices derived from the enactive approach which represent and apply the principles of SMT in interaction design. Although the field is niche, it is growing from a wide and multidisciplinary literature base, where many insights from Sensory Substitution and Augmentation, perception science and psychophysics are relevant. The tables, paradigms and guidelines herein bring together this literature, highlighting potential points of contribution, outlining pitfalls and facilitating collaboration and action in HCI. With this, I hope to see an increase in the quality and quantity of sensorimotor interfaces and associated experimentation and theory, strengthening the path to widespread adoption.

The following two chapters now represent the efforts I have taken to explore and test the creation and application of vibrotactile sensorimotor interfaces through practical experimentation, drawing on and reinforcing the insights highlighted in this chapter. At this stage, I focus specifically on

Table 4.3: SSADs organised by Input/Output

	Source					
Stimulus	Visual	Audio	Distance	Orientation	Compass	Multi
Visual		[52]				
Audio	[117] [2] [29] [70]		[126] [66] [74]		[149]	
Haptic	[9]	[127] [86] [171]	[62] [32]	[134] [123]	[124]	[130]
Electro	[159]			[158]		[50]
Multi						[146]

how existing perceptual experience in adults limits and guides experience using new vibrotactile systems, through an investigation of sensorimotor interfaces through the metrics of traditional psychophysics in Chapter 5. I then utilise these findings to proceed with an exploration of learning paradigms through a structured, multi-day experiment outlined in Chapter 6.



Constraint	Impact: (D)evice	Impact: (I)nteraction	Impact: (A)pplication	Impact: (L)earning	Impact: (P)erception
(D)evice	**"The physical device, including source/transducer apparatus and coupler algorithms"**	Hardware defines sensorimotor input/output loop (p2)	Hardware limits suitability	Maximise device reliance through value and intention (l1)**	Match stimulus/receptor range Match action/source range*
(I)nteraction	Enactive, not cognitive (p3) Consider form and location (p2)	**"The mechanism or paradigm of engagement"**	Reveal affordances, don't represent (p3)	Mastery is achieved through experience (l1, l2, l3)	Semantic distance may drive intuitiveness Hyperacuity requires active movement (p3)
(A)pplication	Integrate device to application environment (p2) Consider natural sensory cost (p1)	Continuous, dynamic, implicit – not symbolic, static or explicit (p3) Build interaction into task (p1, p2)	**"The task and environment in which the system is applied"**	Incentive to learn is driven by application importance/relevance (l1, l2)	Improve task performance through augmentation of perception (not presentation of knowledge) (p3)
(L)earning	Coupler algorithm may impact learning Communicate valuable signal (p1, l2)	Exploration critical for learning (l1) Dynamic action, not symbolic (p3)	Training requires intention (to succeed, l1), attention (to signal, l2) and calibration (feedback, l3)	**"Mechanisms by which users gain familiarity and competence in device use"**	Ensure utility from the get-go (p1) Mastery should increase utility
(P)erception	Match device to individual by tuning coupler to existing perceptive skills	Interaction should be compatible with existing perception*	Reveal affordances, don't represent properties (p3)	Sensory concurrency (self calibration, l3)** Cross-modal correspondences	**"Parameters pertaining to the perception of device signals and world properties"**

Figure 4.3: Table of Design Dimensions. Reading across then up allows identification of design influences and constraints for the effective evaluation or creation of a sensorimotor interface. References p1-p3 reflect paradigms of sensorimotor interface design. References l1-l3 reflect learning principles from DLT. Grey cells indicate insights reflecting experimental contributions described in *Chapter 5 and **Chapter 6

CHAPTER FIVE

CALIBRATING AND MEASURING VIBROTACTILE SENSORIMOTOR INTERFACES

Specific stimulation of the skin's mechanoreceptors (touch receptors) through vibration, known as the vibrotactile modality, is an interesting target modality for sensorimotor interfaces for several reasons. Tactile displays lend themselves particularly well to novel interaction interfaces, allowing a large area for interaction without necessitating impairment of vision and audio, while simultaneously optionally leaving hands free to perform tasks as usual [82].

The challenges and opportunities presented with vibrotactile systems have been extensively explored in prior work. Vibrotactile feedback is already experienced by people on a daily basis through controllers [67, 139], off-the-shelf mobile devices [59], and smartwatches [34, 73]. Utilisation of vibrotactile stimulation has been explored for several interaction tasks attempting perceptual augmentation, such as navigation [39], explicit information transfer [26,130], language (e.g. letters [109] or phonemes [171], and perhaps most relevantly, for research into models of enactivity itself [62, 163].

In this chapter I focus on devices using continuous, dynamic, vibrotactile signals for *enactive* interactions consistent with sensorimotor theory. Such interactions are far less-explored in the literature than symbolic encoding principles, and there exists outstanding questions with regards to how these continuous signals should be designed and optimised for enactive applications. If we are to communicate information consistently and accurately to users through dynamic vibrotactile signals, we must understand how the signals are perceived and interpreted. With this understanding, we can consider how to account for these perceptions, and explore the implications of failing to do so.

Performing classification and perception experiments as established in the literature may lead us to methods that ensure signals are accurately interpreted *cognitively*, allowing extraction of explicit information. However, it is not clear what these results may tell us in the context of *enactive* interactions. We therefore must evaluate how useful existing metrics and methods are for gaining insight into the effectiveness of an enactive signal, and the information which may be extracted from such interactions.

In this chapter, I perform an experimental exploration of how the vibrotactile signal from an off-the-shelf device, the Xbox One controller, is perceived and interpreted. This reveals several limitations of the device, which feeds directly into an exploration of technical measurements and calibrations that allow many of these limitations to be overcome. The experimentation also reveals that methodology commonly employed for evaluating *cognitive* interpretation may not be the most useful or relevant in understanding *sensorimotor* interactions.

5.1 Introduction

Encoding information into effective and expressive vibrotactile signals to enable adequate perception and interpretation is non-trivial [163]. The design space is large and depends on factors such as the device's capabilities (e.g., the output range and nature of vibration) and the intended usage scenarios (e.g., value judgements vs dynamic interaction). While many vibrotactile interfaces have been proposed and tested [82], device capabilities and characteristics are not always reported, and the encoding choices and resultant physical signal presented to participants is not always clear. This makes it difficult to replicate studies or to generalise findings beyond the test scenario.

The presence and importance of device characteristics and choices of encoding in the design, implementation, and adoption of new sensorimotor interfaces is often overlooked in the literature [163]. To illustrate directly, of the 17 publications thus far cited in this thesis which detail the use of haptic motors for the communication of signals in sensorimotor interfaces, only three detail the characterisation and calibration of their motors to take into account technical and human factors [30, 80, 148]. A further 6 describe characterisation and their encoding regime to a degree sufficient for replication, though fall short of justifying their encoding, or in describing the final sensible qualities of the system output [28, 32, 45, 49, 66, 93]. The remaining 7, which unarguably contribute useful insights to development of both SSADs and enactive interfaces, do not touch on characterisation or calibration at all [11, 17, 20, 33, 36, 86, 162].

As discussed in detail in Chapter 4, there are two loose encoding paradigms for information communication in vibrotactile interfaces—symbolic ('tactons': spatio-temporal patterns of vibration with explicit meaning [23]) and continuous (fluid, reactive sensations dynamically linked to user action or input [62]).

Symbolic encodings are most often applied in applications that focus on the communication of language [38, 171] and explicit information [49, 81], and must be learned, memorised and interpreted, much like the characters in a written language. Vibrotactile devices utilising such tactons for communication of task information may classify as SSADs under Visell's 'functional' definition—if they perform the function of compensating for a sensory deficit. However, as tactons are discrete and symbolic, they cannot respond fluidly and dynamically to user action, making it difficult, if not impossible, to imagine how they may allow the formation of new sensorimotor contingencies. Such symbolic tactile displays therefore do not meet the sensorimotor interface requirements laid out in Chapter 4, and, thus, their comparison and metrics do not fall within the scope of this thesis. With this said, it is useful to note at this stage that an evaluation of the speed and accuracy by which users may classify or recognise tactons is widely considered a good estimate of their effectiveness [23, 130].

With regards to continuous encodings, even when considering the simplest, homeomorphic one-to-one mapping of input domain to target stimulus, the science is not well-established: there are significant questions left to answer both generally in the field, and for researchers designing for specific applications.

- Does a mapping have to be linear to be effective? If so, should the mapping be linear in objective quantity (e.g. an output frequency), or in the values perceived by existing modalities?
- Are such perceptions important for signal learnability?
- Should the signal mapping be static or adapt to the task at hand?
- Does training remedy or exacerbate problems with nonlinear or inconsistent signal interpretation, and if so, under which conditions?

While these are interesting questions purely from a perceptual research perspective, they also highlight gaps in existing literature which must be addressed to facilitate practical progress in the application of sensorimotor interfaces. Naturally, creating publications with generalisable findings and the potential for replication is also desirable, in addition to demonstrating successful device use under test conditions. Equally, in cases where devices have failed to show significant improvements in performance, without insights into the signal presented to the user and its encoding, it is difficult to determine the source of the issues that led to these negative results; be it the interaction principle, the chosen encoding, or the metrics used to evaluate performance.

As discussed in Chapters 2 and 4, the vast majority of sensorimotor interfaces evaluated in academic literature are SSADs. Success metrics described in SSAD studies vary widely, due to many factors stemming from the lack of a single definition for what an SSAD is (as functional and operational SSADs are often conflated with those targeting perceptual augmentation), and the unique application and approach required for each device [163]. Such a history has led to a confusion of mixed evaluation standards.

Studies which evaluate SSADs in real-world contexts, both in the light of behavioural and experiential qualities are in many ways good for estimating a device's potential and the suitability of an SSAD application. However it can be difficult to set up and execute such studies, as they are complex and hard to generalise, and often present funding or other resource limitations. Furthermore, due to their specificity, the extraction of generalisable results from device-specific, tailored experimentation can be difficult.

As in most HCI device or system evaluations, SSADs therefore benefit from lab-based methodologies, which yield comparable estimations of device capabilities under controlled conditions. In SSAD literature, it is common to measure participants' performance in judgement tasks, testing their ability to extract the source signal (e.g. distance or direction) from the output signal (e.g. vibrotactile intensity). The accuracy of these judgements as a comparable metric is deemed a suitable reflection of the device's potential for real-world applications [18, 19, 32, 65, 146]. These judgement tasks can resemble the absolute magnitude perception tasks typically utilised in traditional psychology, but feature some differences. In order to minimise unintentional biases, traditional psychology dictates that users be allowed to determine their own scale upon which to gauge response signal [64, 79]. However, in SSAD studies—because the point of interest is (in these metrics) the extraction of the original data signal from the output sensation—participants are, instead, often required to judge their perception against the original scale following pre-training [32, 130].

Such judgement tasks give a picture of the entire system that sits between the source data and the decision made (see Fig. 5.5). These tasks do not necessarily provide insight into the degree of mastery that the participant has acquired to that point with the SSAD, instead only reflecting the interpretability of the presented signal encoding under specific training and testing conditions, rather than overall device potential. The overall effect of these individual components to the final measurement is unclear, and the need for further insight forms the basis for the investigations presented in this chapter.

In the project presented in this chapter, I start by measuring and visualising the perceived vibrotactile output of an off-the-shelf games controller (the Xbox One controller, designed for general vibrotactile feedback in video games) in line with established judgement task methodology. The limited technical capabilities of such off-the-shelf devices can impact their performance and usability as SSADs, and in taking perceptual measurements, I note that these limitations may be manifesting in the results of the experiment. I then explore, suggest and demonstrate how technical measurements and calibrations may help to consider both device limitations and human perception when encoding data as vibrotactile feedback. Following a re-calibration of the device based on these measurements, I confirm its effectiveness through further perceptual experimentation and discuss differences in the outcome. This pre-post comparison of human performance in the interpretation of vibrotactile feedback highlights the importance of the technical measurements and interventions, while also raising questions regarding the metrics themselves, and their place in SSAD literature.

The work presented in this chapter therefore serves three purposes:
Firstly, I determine the general accuracy with which participants can make judgements on communicated SSAD data values following brief familiarisation periods, as is common in the literature. I note that such measurements and the expectation of extracting explicit information in static context is neither realistic nor relevant to the understanding of a vibrotactile sensorimotor interface's potential.

Secondly, I perform a series of simple characterisations and calibrations, demonstrating steps that can (and should) be taken to improve a vibrotactile sensorimotor interface's reliability in communication of source data. I then demonstrate the effect this has on the interpretation of the source signal. This reinforces the importance of the characterisation and calibration steps in the development of reliable devices. The steps are summarised and presented for future reference and digestion.

Thirdly, the work presented here sets the groundwork for future experimentation, as this process results in a fully characterised and reliable vibrotactile device suitable for the sensorimotor interface experimentation conducted in Chapter 6.

5.2 Measuring Perceived Vibrotactile Feedback from the Xbox One Controller

Research on the creation of perceptible vibrotactile feedback is well-established. Techniques and methods that rely on haptic actuation most commonly employ vibration motors with a moving mass (e.g. eccentric rotating mass (ERM) [39, 109, 110] or voice-coil actuators (VCA) [98, 150, 152]). While other actuators exist (e.g., material deformation through piezoelectric [61] or electrovibration [16]), these are far less common in off-the-shelf consumer devices, and are usually integrated to meet requirements which simpler vibrotactile motors, such as the ERM, cannot fulfil.

This section details the experimental efforts taken to understand the technical limitations of two ERM motors in the Xbox One controller, and how the vibrotactile signals they produce are perceived using off-the-shelf settings. Starting with the aforementioned magnitude estimation tasks, an exploration of the results highlights that the classification-style judgements that participants are asked to perform do not necessarily reveal the quality of information users may use in wider task context. Noting that dynamic, enactive interactions also allow comparative judgements, I question whether an adaptation of the magnitude estimation task (in the form of a *difference estimation task*) may provide further insights. Such tasks are undertaken in an effort to understand device and signal suitability for sensorimotor applications.

Both experiments reveal technical device limitations which could conceivably lead to perceptual problems. I therefore undertake a literature investigation into previous efforts to alleviate some of these technical limitations, and highlight the key measurement and calibration steps taken to overcome many of these limitations. Finally, through repeating the magnitude estimation experimentation with a proof-of-concept group, I briefly demonstrate that these calibrations affect perception, and are useful in eliminating some concerning observations, but re-iterate that this methodology cannot reliably yield a full picture of device potential in sensorimotor applications. I finish this chapter with suggestions on how device and signal evaluation may be optimised through different metrics and study designs, and prepare to do so experimentally using the calibrated controller.

5.3 The Xbox One Controller as a Vibrotactile Interface

I chose the Xbox One controller[1] (see Fig. 5.1) for the studies described in this chapter and also in Chapter 6. The Xbox controller is a versatile vibrotactile feedback device that is relatively

---

[1] https://support.xbox.com/en-US/xbox-one/accessories/xbox-one-standard-controller-info

low-cost and readily available off-the-shelf. While popular as a gaming input device capable of providing vibrotactile "rumble" feedback to facilitate more immersive gaming experiences, the Xbox controller has not been explored in scenarios where vibrotactile feedback is used to encode quantitative information (e.g., in immersive analytics scenarios [68] or navigation tasks [85]). I therefore decided to use the Xbox controller as a vibrotactile device both to explore its potential for use in such scenarios and to illustrate the measurements I have considered as part of our device profiling.

Figure 5.1: Microsoft Xbox One Controller showing the location of the two small (1, 2) and two large (3, 4) ERM motors for haptic feedback.

The Xbox One controller features four ERM motors (described as generic "rumble motors") that provide haptic feedback. It connects to a PC via Bluetooth, and each vibration motor can be independently controlled using the Windows.Gaming.Input [2] namespace in the Universal Windows Platform API. There are two large motors (one in each palm grip), and two small motors (one in each trigger button). The two large motors (Fig. 5.1.3 & 4) are fixed to the main controller body, so while they operate over different frequency ranges, the resulting vibration when activated is experienced nearly equally in both hands and thus differentiation of signals is very difficult. However, the two trigger motors (see Fig. 5.1.1 & 2) are vibrationally well-isolated from each other, and can therefore be easily distinguished and intuitively associated with the left and right hands. These features suggest the trigger motors of the Xbox One controller may be suitable for communication of the continuous information spaces relevant to enactive, sensorimotor interactions.

---

[2] `https://docs.microsoft.com/en-us/uwp/api/windows.gaming.input`

In the following experimentation, I explore how well these continuous information spaces may be communicated and perceived, starting with a perceived absolute magnitude estimation task.

5.4 Magnitude Estimation Tasks

As described earlier, magnitude estimation tasks are a form of controlled study task utilised in traditional psychology [64]. They are relatively simple to set up experimentally, and can generate metrics of the accuracy and consistency of users' interpretations of sensory stimuli. Here, I explore how variations on the magnitude estimation task may be used to evaluate the potential of SSADs.

5.4.1 Perceived Absolute Magnitude Estimation

In traditional cognitive science, the perceived absolute magnitude (PAM) estimation task is used to create a profile of perceived signal magnitude by asking users to assign numerical values to presented signals on an arbitrary, per-experiment, user-defined scale [41]. This task often reveals the non-linearity of human perceptual response to otherwise objectively linear scales, allowing the modelling of human perception in response to a given physical stimulation.

In the pursuit of reliable metrics for measuring and comparing the effectiveness of SSADs, I ask how encoded sensory signals are cognitively interpreted (i.e., how accurately and reliably the original numerical values can be extracted from the signal, as demonstrated in SSAD studies featuring single, discrete vibrotactors [32, 130]). As such, I explore PAM accuracy as a measurement that indicates the accuracy with which people perceive and are able to interpret the magnitude of a vibrotactile signal, *given a particular quantitative scale*. The ability to accurately interpret a vibrotactile signal quantitatively can be important. For example, in immersive analytics scenarios, where abstract data represented by virtual artefacts distributed in a 3D space a user may want to explore the artifact in search of trends and patterns using a VR headset [89]. In such an application, while enactive engagement through a sensorimotor interface is an attractive alternative to sifting through tabulated or even visualised data, an appreciation for the accuracy of non-visual, augmented perception of the underlying data would remain important. Previous studies have evaluated the accuracy of magnitude estimations of vibrotactile intensities [165]. However, it is much more common for device effectiveness to be evaluated qualitatively and in-context [56].

5.4.1.1 PAM Estimation on the Xbox One Controller

In methodology resembling the familiarisation or training phases of several SSAD studies (such as Carton and Dunne's distance perception task [32], or Novich's single-motor interpretation task [130]), the PAM accuracy of a vibrotactile device requires that participants be introduced to the encoding of quantitative values to vibrotactile signals first, in a learning phase where they can actively explore and familiarise themselves with the scale of values. In a series of trials, participants are then exposed to vibrotactile signals to which they will assign a quantitative value as per the original scale. The accuracy and repeatability of these judgements over numerous trials is then used to gauge the reliability of information communication through this channel.

During pilot testing, it was noted that there were potentially some differences between the left and right trigger motors in terms of vibrotactile output in response to specific API commands. I was therefore also interested to know how accurately participants would then compare magnitudes between the left and right motors, and if asked to match values, whether their matching strategy would reflect the values of the single motor scales they were initially presented with (the original values), or the perceived vibrotactile intensity. This measurement was conducted in the form of a simple matching task, as described in more detail below.

I therefore measured the accuracy of the PAM estimations and matching tasks for the left and right front trigger motors of our Xbox One controller (see Fig. 5.1), across 12 participants (5 female, all right handed, aged between 20-40), as per the following study procedure.

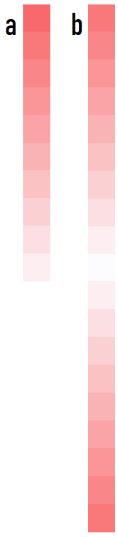


Figure 5.2: The scale presented to the participants during a. the perceived absolute magnitude estimation and b. the perceived difference estimation task

Participants were handed the Xbox One controller (described above) and instructed to hold it comfortably, with their index fingers resting on the trigger buttons so as to feel the vibration on the button directly. Participants were then presented a visual scale on a computer screen that represented an encoding of numbers from 1 to 10 vibration intensities (see Fig 5.2.a). Dragging a pointer on the scale using the control stick on the Xbox One controller allowed participants to explore and familiarise themselves with vibration intensities and the mappings to the quantitative scale. This training session was followed by a number of trials where participants were presented with a vibration intensity on the controller, and were then asked to provide a quantitative estimation of this intensity by dragging the slider on the visual scale to the desired position.

The trials were presented in 5 blocks. Each block consisted of 40 trials (10 intensity levels x 4 repetitions) and focused on a particular task condition as specified below. This results in a total of 200 trials. The order of trials was randomised within each block. I describe the conditions represented by each block of trials below.

Vibration Left Only, Right Only, and on Both Motors. In Blocks 1-3 participants were asked to estimate the absolute magnitude of vibration intensity on either the left, the right, or on both motors together and to specify their estimates using a slider on the quantitative visual scale shown in Figure 5.1, Right.a.

Match Right to Left, Match Left to Right. Blocks 4 & 5 consisted of the matching tasks, where participants were presented with a vibration intensity on one motor (e.g., in their left hand) and asked to use the control stick on the Xbox to adjust the vibration intensity of the other motor (e.g., in their right hand) until it matched the first. For these two blocks no visual scale was displayed. These blocks were introduced to explore how participants would compare the vibrations of the two motors, and to reveal differences in perception or interpretation between them.

5.4.1.2 Data Collection and Analysis

For each trial, the presented and reported intensity values were recorded alongside presentation and decision timestamps.

As the experiment was performed as an explorative study, there were no statistical analyses planned for the results. All insights were drawn from visualisations, which could be used to suggest further lines of enquiry and inform controller and participant characteristics which may be helpful in determining the limitations of the Xbox One controller as a sensorimotor interface in subsequent studies.

The data was visualised using confusion matrices and line charts, see Fig. 5.3. These visualisations are well-suited to showing discriminatory accuracy for both intensity and spatial

Exposure	% Correct
both	68
left	68
right	70
match to left	42
match to right	56

Table 5.1: Percentage correct judgements made allowing off-by-one (+/-1) in the Absolute Magnitude Estimation task using the off-the-shelf controller.

vibrotactile encodings [76, 130], but may also be generalised to other waveform parameter encodings. Here, I investigate what insights may be drawn from the study.

5.4.1.3 Findings

The results of the experiment can be visualised for each motor in the form of a confusion matrix for each task condition (see Fig. 5.3.top).

Confusion matrices have been used to analyse this type of experiment in previous work [130,145]. Each confusion matrix in Figure 5.3 shows the presented intensity levels as rows and the recorded intensity levels as columns across each of the different task conditions. The matrix cells are colour-coded according to the frequency of presented-estimated intensity pairs—the lighter the cell value, the higher the frequency. Frequencies are presented as a fraction of all responses across all participants. A perfect absolute magnitude estimation across all participants would result in a yellow diagonal from the top left to the bottom right in each matrix.

The results can also be plotted on a line chart showing % of correct judgements per presented vibrational intensity (see Fig 5.3.bottom). This chart assists in identifying overall accuracy and consistency of estimates, as discussed below.

While the charts are helpful in discerning the exact accuracy of user judgements for each intensity value, calculating the percentage of all judgements made which differ from the presented value by up to +/-1 may be compared between motors [32]. Such groupings can also give insights into how many truly differentiable classification categories may be presented at a given accuracy. In a set of 10 intensities, allowing +/-1 judgements is equivalent to creating 4 non-overlapping categorical intensity ranges (at values 1, 4, 7 and 10). The percentage of +/-1 correct judgements for each condition in this task were as described in Table 5.1.

Analysis of Magnitude Estimations on Individual Motors. The confusion matrices can provide useful insights into participants' experience of the encoding of quantitative values into

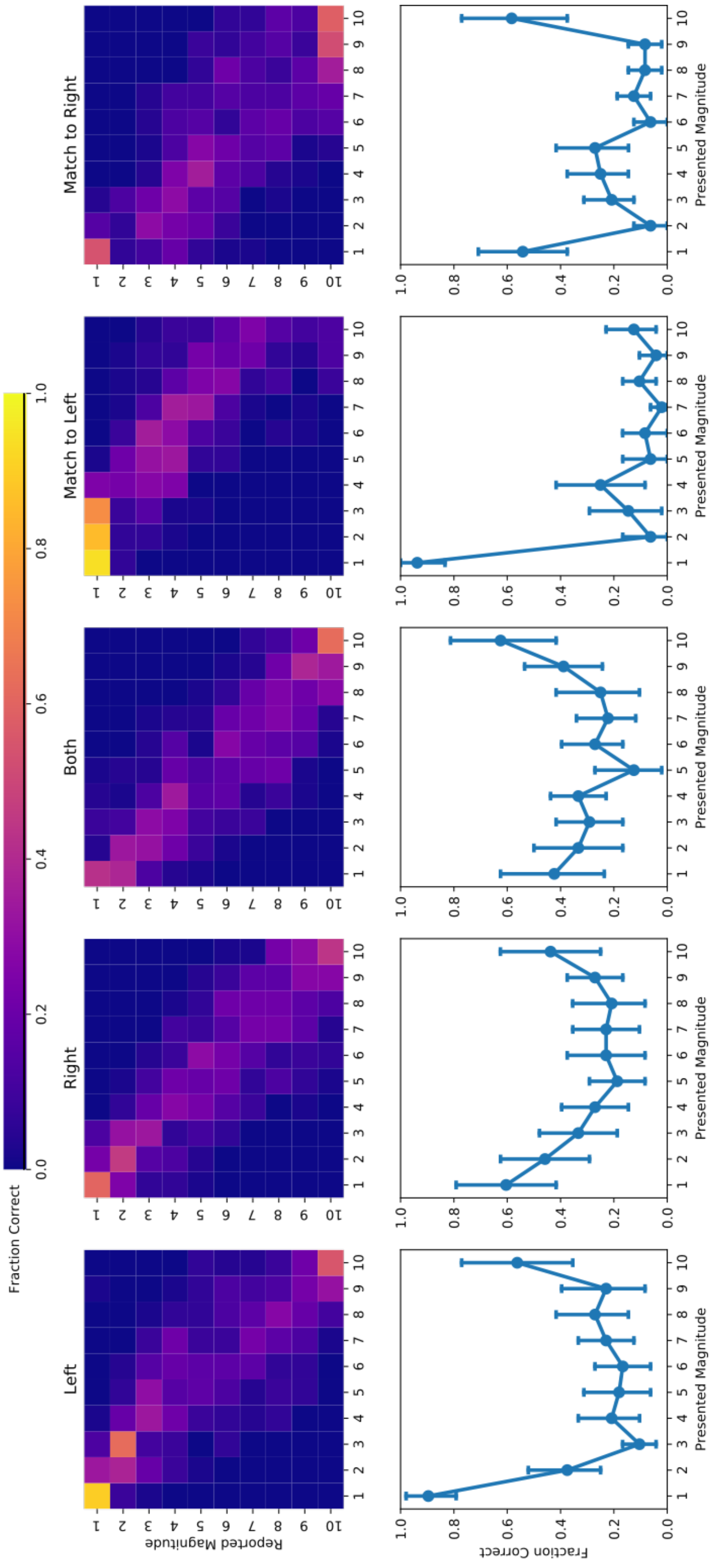


Figure 5.3: Confusion matrices showing presented vs recorded vibrotactile magnitude values, above intensity classification performance, for each task condition, averaged across all participants.

vibrotactile intensity, more specifically biases or trends, and ability to discriminate between encoding levels. The absolute magnitude estimation charts for the ‘Left’- and ‘Right’-only conditions suggest that participants were able to estimate vibrotactile intensities of the two motors with comparable accuracy, and, while there is some deviation from the ideal perfect diagonal, responses are, generally, linear. The lighter cells in the corners of these two matrices suggests participants had greater certainty and accuracy in matching intensity vibrations at the upper and lower ends of the intensity range, though with deviation towards intensity underestimation in the left motor shown by the brighter square at 2,3.

The line charts with error bars shown in Figure 5.3.bottom, provide a complimentary perspective to the corresponding confusion matrices. In these, the y-axis shows the fraction of correct estimations for all responses across all participants; the x-axis shows the presented intensity levels. As there are 10 discrete levels in the scale, performance at chance would correspond to an accuracy of 0.1.

The line charts for the ‘Left’- and ‘Right’-only conditions confirm that intensities at the edges of the scales were more accurately judged than those at the centre.

The figures in Table 5.1 shows % accuracy allowing +/-1 judgement errors is comparable between individual motors, and for both motors together.

Analysis of Matching Task Results. Looking at the rightmost matching task summaries in Fig. 5.3.top, the asymmetric, curved deviation from the ideal diagonal in the matching tasks suggest differences in the perceived intensities of the motors. For the Match to Left task, participants were more likely to set the intensity of the right motor lower, while in the Match to Right, a similar skew in the other direction was observed. The intense squares in the top left of the Match-to-Left and bottom right of the Match-to-Right charts suggests that participants were unable to select vibrational intensities perceived to match at the edges, and thus selected the minimum or maximum value setting where relevant.

The lower % of correct +/-1 judgements shown in Table 5.1 for the matching experiments suggests that there is a discrepency in judgement between the motors of more than +/-1 step, around half the time.

5.4.1.4 Discussion of Absolute Magnitude Estimation Results

It is unclear whether the biases in selection towards the edges of the intensity scales (shown in Fig. 5.3) are caused by cognitive decision making, or whether these may be compensated for through adjustment of the controller output. Ideally, for reliable information communication, all intensity values would be judged with similar accuracy and consistency.

Similarly, with regards to the matching experiment, in a scenario where individual motors represent comparable quantities yet feature asymmetrical vibration profiles, there may be biases caused by differences in perception in contradiction of the learned scale. It is unclear whether participants are attempting to match each motor to their absolute perceived intensity, or the learned value, and this contradiction likely caused confusion. It is impossible to know whether these results arise from controller asymmetries, or by perceptual differences or biases attributable to human factors, such as a difference in sensitivity, familiarity or comfort between hands.

Table 5.1 shows that participants can accurately report judgements of intensities within +/- 1 68-70% of the time. As described above, in a usage scenario where explicit magnitude estimation is important, the Xbox One controller therefore may provide four differentiable intensity levels at an accuracy of up to 70%. This confirms findings from previous studies assessing the explicit classification of vibrotactile intensity levels [148]. Greater accuracy could conceivably be achieved through encoding via more sophisticated spatio-temporal patterns [130]. Our observations therefore raise an important question: If the continuous encoding results in so few differentiable levels, does this mean that attempting to communicate rich, environmental information using our controller's vibrotactile feedback is futile?

Absolute value estimation is not the only mechanism by which users may extract data from numerical interfaces. For example, while users may not be able to accurately estimate the *absolute strength* of a vibrotactile signal on a quantitative scale, they may be able to interpret the *relative strength* of one vibrotactile signal compared to another one.

Novich [130] uses an absolute magnitude estimation task to calculate the potential of certain quantitative encodings in information communication, and concludes that continuous intensity encodings are inferior to categorical spatiotemporal patterns in communication applications, owing to weaker performance in the extraction of explicit values.

However, such estimate tasks do not consider the ordinal, comparable nature of numeric encodings, and may not offer a reliable metric in this context. For example, Carton and Dunne's study [32] notes how in several cases, participants were able to correct their distance estimates using information extracted from the relative difference of their previous to current guess. Lobo [107] also notes that hyperacuity (the increase of sensory precision beyond native sensory capabilities through movement) may only arise through iterative, enactive movement, which is not possible through classification, and not measured through passive, static magnitude judgements. In this sense, it would seem that an exploration of participants' ability to discriminate and compare between intensities may also yield insights into how information may be communicated in continuous vibrotactile encodings.

In the next section, I describe how we may attempt to characterise the user's ability to perform these comparisons and explain how users may extract information from changes in the vibrotactile signal during task performance.

5.4.2 Perceived Difference Estimation

Some usage scenarios such as navigation tasks do not require the accurate estimation of absolute vibrotactile intensities, but, rather, a relative estimation of the magnitude of one vibration level in comparison to another (e.g., has the vibration signal intensity increased and, if so, by how much?). The more accurately a user can estimate differences in magnitude, the more nuanced the reactions to these changes can be. Accurate comparisons of changing quantitative data represented by vibrotactile signals may be critical.

More importantly and pertinently to this thesis, from an enactive perspective, in applications where vibrotactile signals reflect environmental parameters, user actions causing environmental changes will cause respective shifts in vibrotactile intensity. This changing signal represents a critical point of feedback in the formation of new sensorimotor contingencies. Accurate and repeatable estimation of the magnitude of this shift therefore reflects the quality of feedback available, and drives the user's ability to respond proportionately to changes in the environment. In well-designed sensorimotor interfaces, observations of dynamic changes of the device signal, caused by, e.g., changes of the user's position enable observation of relations between action and perception. It is widely recognised that such control and action is key to the learning of encoding of the signal [6, 9], and is a firm requirement for the emergence of perceptual experience in sensorimotor theory [133]. Thus, signal strength comparisons and observation of dynamic changes in a signal in tandem with user action is a critical component in the learning of sensorimotor interactions.

In order to characterise the usefulness of this signal, the next section explores the accuracy with which users can judge the difference between vibrotactile signals. This is measured through experimentation designed to discern whether a vibrotactile signal is perceived to have changed, and if so, whether the difference represents an increase or decrease, and by how much. This exploratory investigation was designed to provide insights into how accurately participants may extract and utilise the information encoded in sequentially different output signals.

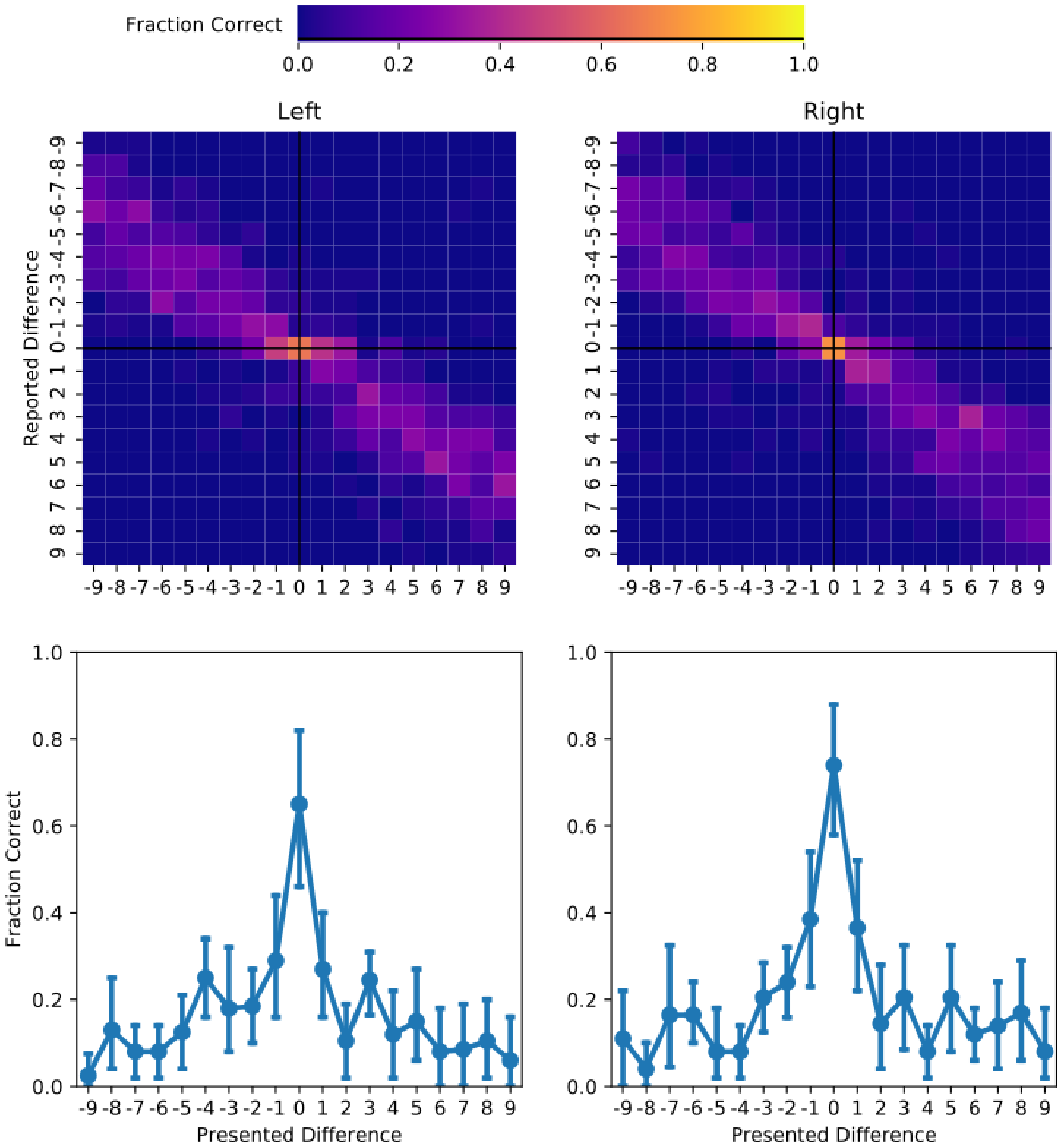


Figure 5.4: Confusion matrices showing presented vs recorded differences in vibrotactile magnitude values, above corresponding classification performance, for each task condition, averaged across all participants.

5.4.2.1 Perceived Difference Estimation on the Xbox One Controller

Measuring the perceived difference between two signals requires presentation of two sequential intensity values. Such difference measurements are commonly employed in the literature when measuring ‘Just Noticeable Difference’ values, designed to gauge user ability to differentiate between small changes in stimulae, such as intensity [41]. As we are interested in how well users may extract useful information from the difference in sequential signals (i.e., whether they may be taking advantage of how the signal changes when it is changing dynamically), we require participants to specify whether the signal has changed, and to attempt to quantify *by how much*. Evaluating how accurately participants can do this may provide insights into how participants use the information of a dynamically changing signal.

I measured the perceived difference between the given vibrotactile signals for the left and right trigger motors of our Xbox One controller (see Fig. 5.1), across 10 participants (3 female, one left-handed, aged between 22-31).

Participants were handed the controller and instructed to hold it comfortably, with their index fingers on the trigger buttons so as to feel the trigger button vibration directly. Participants were then presented the same visual scale (Fig. 5.2.a) as the above absolute magnitude estimation experiment for exploration and familiarisation with the vibration intensities and their mappings to the quantitative scale.

This experiment was presented in two blocks. Each block consisted of 95 trials (difference levels from -9 to +9 = 19 levels, x 5 repetitions) with a specific task condition (total 190 trials). The order of the trials was randomised within each block. The two task conditions were *Left only* and *Right only*: participants would experience all vibrations in a block on the same trigger motor, either left or right. In each trial, participants were presented two consecutive vibrotactile signals (each presented for one second), separated by a gap of half a second. After the second signal was presented, participants were asked to estimate the perceived difference in magnitude between these intensities through indication on a bi-directional scale (see Fig. 5.2.b).

5.4.2.2 Data Collection and Analysis

For each trial, the presented starting intensity, the actual difference and the reported difference were recorded, alongside presentation and decision timestamps.

5.4.2.3 Findings

Similarly to the measurement of perceived absolute magnitude, the results of this experiment can be visualised for each motor in the form of a confusion matrix (see Fig. 5.4.top).

The point charts shown in Figure 5.4.bottom correspond to the above confusion matrices. The y-axis shows the fraction of correct estimations for all responses across all participants; the x-axis shows the presented difference levels (with a difference of 0 in the centre). As there are 19 discrete levels in this system, performance at chance would correspond to an accuracy of

˜0.05. This shows that participants are extracting information from the difference channel for most values, with higher accuracy at up to +/-3 steps.

Analysis of Difference Magnitude Estimation Task Results. Both matrices indicate roughly linear diagonal trends in responses to the presented vibration, though the intercepts appear at +/-6, rather than the expected +/-9. Estimated judgements are within 1-2 squares at the centre, though spread horizontally up to 6-7 squares nearer edges.

The line charts show that participants are able to recognise that the signal has not changed with around 70-80% accuracy on both motors, though a change of +/-1 is correctly identified just 30-40% of the time.

5.4.2.4 Discussion of Difference Magnitude Estimation Results

As chance estimations for these difference judgements would represent just 5% accuracy, the ability of participants to correctly recognise one-step changes in intensity at 30-40% accuracy indicates that comparative judgements even between small changes can communicate some useful information. This insight is especially interesting when considering that small changes (and comparisons between them) would not be possible if the system were restricted to 4 discrete intensities, as suggested by the results of the absolute magnitude estimation task above.

However, considering how the accuracy drops off for large magnitudes, saturating at around +/-6 steps for both motors, users should not be expected to use such encodings to accurately estimate large differences of vibrotactile signals.

As briefly mentioned in Section 5.4.2.1, assessing the discriminability of vibrotactile intensity [60, 157] and location [44] is an established methodology performed as a psychophysical test to assess humans' sensitivity to vibrotactile signals. However, response to these discrimination tasks is limited to selection between two choices, typically "greater/lesser" or "left/right". Such measurements are featured in, for example, Carton and Dunne's study with the distal firefighting glove [32]. Here, we assess user ability to judge the magnitude of the difference between values, as it reflects the importance of this channel for interaction with continuous data, while also highlighting how user-controlled movement which causes signal changes can contribute to the information users are able to access through interaction.

We note though that aside from observing the potential extraction of information from relative differences, these metrics provide little further insight on the effectiveness of continuous encodings when applied to sensorimotor interfaces. The task reflects a cognitive decision making or judgement process in an abstract environment, following a brief exposure to a predefined signal with no contextual or task-related meaning. As well as highlighting that there are uncertainties in the control, encoding and design process, this brings into question the applicability of the absolute and relative magnitude estimation tasks to the comparison of SSADs or sensorimotor interactions in general. Such issues are clearly important in the comparison and design of sensorimotor interfaces, and will be addressed in detail later in this chapter, and in Chapter 6.

5.5 Accounting for Devices and Humans in Vibrotactile Communication

The absolute and relative magnitude estimation tasks detailed above have highlighted that several inaccuracies arise when extracting information from the vibrotactile signals encoded into the Xbox controller, including bias and inconsistency (indicated by skew, and horizontal spread in the confusion matrices respectively). However, it is not clear to what degree these are factors reflecting human abilities, or artefacts of the device and encoding. The underlying causes of these inaccuracies must be understood and addressed before meaningful vibrotactile interactions can be designed.

Several studies have undertaken processes to minimise issues in vibrotactile communication by considering the factors between the source data and cognitive interpretation [30, 148]. These factors are illustrated in Fig. 5.5. Here, I outline how each factor contributes to interpretation and detail its importance. Following this, I explore principles for how they may be accounted for through calibration, to achieve improvements in perception and interpretation with relatively low effort. Applying these calibrations to the Xbox One controller, I repeat the human experimentation detailed in section 5.2 and evaluate the effectiveness of the method taken.

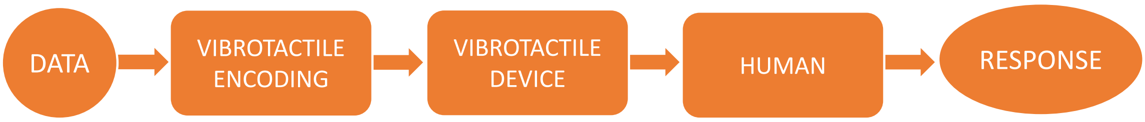


Figure 5.5: Schematic representing the factors which lie between source data and cognitive interpretation of a vibrotactile signal by the user

The schematic in Fig. 5.5 starts with the raw data. This is encoded into a signal which is presented to the vibrotactile device. The next stage considers the technical characteristics of the vibrotactile device itself, such as the intensities (frequencies and amplitudes) that its hardware (typically motors) can support and manifest. Next, we acknowledge that human factors also play an important role; people have to be able to first perceive, and then to interpret the data values represented as vibrotactile signals, and to learn the significance of the encodings, so the human perceptual system must be considered. The first stage (the data encoding) is typically the only stage in this process that may be controlled completely (as hardware response is not always adjustable). Thus, the more accurate the model of the other two stages (device and human), the more effectively the encoding may be optimised such that the source data is perceived accurately.

In this section I outline approaches for profiling the vibrotactile device technically, and accounting for the human perceptual perspective. I then indicate how characterising such interfaces can inform design decisions for accurately encoding quantitative information into vibrotactile feedback, and point towards why this is also important for sensorimotor interactions.

While the technical measurements and calibration processes I outline here are well-known, I found that they are not consistently applied in the areas of multimodal interaction [36] or sensory substitution and sensory augmentation [62, 108], sometimes even when the necessary parameters are measured [76]. By contrast, detailed characterisations [49] and efforts to account for objective measurements in-situ [103] represent thorough understanding of the benefits of preparing vibrotactile hardware for information communication. At a minimum, utilising manufacturer-provided measurements or technical specifications [165] is superior to ignoring the need for calibration entirely.

5.5.1 Technical Measurements - Profiling the Vibrotactile Device

When encoding quantitative data into vibrotactile feedback, it is important to know how the output vibrotactile signal changes depending on the values of the drive signal (the input of the vibrotactile device). This is essential for making an informed mapping of quantitative data to a vibrotactile signal range. The process requires the technical measurement of (1) the relationship between the drive signal (e.g., input value to the device, or the voltage that drives a motor or other tactor) and the output signal (e.g., the intensity of the vibrotactile signal) and (2) the response speed of the output signal to changes in drive signal (bandwidth).

Knowing or estimating these relationships allows us to make an informed encoding between the data signal and the output signal.

In the absence of a reliable product datasheet (as is typically the case with off-the-shelf or consumer devices), establishing a quantitative relationship between drive signal and output signal requires the measurement of the output signal over the range of the drive signal. Usually, the parameters of interest of the output signal correspond to what can be perceived by human skin, that is, the dominant frequency of the output signal and its amplitude or displacement.

The dominant frequency can be measured by attaching an accelerometer or microphone to the point of vibration, and identifying the peak frequency in the Fourier transform of the measured signal. Attaining an accurate measurement for amplitude requires the careful calibration of the measurement equipment, and is therefore more difficult. If a motor's technical datasheet describes this relationship already, this can be considered and referenced in lieu of testing. However, the datasheet is of limited use, since the response will depend on how the tactor is integrated within the device and even how the device makes contact with the human body.

In principle, it would be possible to independently control frequency and amplitude separately. However, in the most commonly found type of tactor (ERM motors) dominant frequency and amplitude are coupled into a single output parameter. This parameter, usually referred to as the "intensity" [23, 130], means that as voltage (the input) increases, both frequency and amplitude increase in a monotonic relationship. Therefore, a measurement of frequency alone is a proportional representation of intensity, so optimisations pertaining to the frequency output of the device will result in an optimisation of intensity as well [130, 136]. In the empirical parts of this chapter I assume a single output parameter (intensity) controlled by a single input signal, however, generalisation of our approach to multiple outputs should be possible.

When the input signal changes level, the frequency and amplitude (intensity) of the output signal do not change immediately. This is because transducers (and ERM Motors specifically), are physical objects that have inertia, which results in a progressive adaptation of the output to its stationary state (e.g., the frequency of the output measured for a continuous signal at that level). If this adaptation is too slow it means that changing the input quickly will not result in full changes of the output, reducing the ability of the overall system to transmit information. This kind of slow adaptation may be referred to as a low *bandwidth*[3].

The bandwidth can be approximated as the inverse of the rise and fall time. The output's rise time is the time measured for the output to change from 10% to 90% in response to a step change from 0 to 100% in the input. The output's fall time is the time that the output takes to change from 90% to 10% in response to a step change from 100% down to 0%. As a commonly applied rule of thumb in signal characterisation, the bandwidth (the highest significant information frequency the signal may carry) is 0.35 divided by the greater of the rise or fall time [118]. Data signals that change faster than the bandwidth will not result in changes in the output.

---

[3] Note that this bandwidth is not directly the bandwidth of the output signal as measured by the characterisation sensor, but instead the bandwidth of the system when considering the level of *intensity* (as defined above) as the signal

5.5.1.1 Measurements of Frequency

For our Xbox Controller I profiled the two small ERM motors embedded in the left and right trigger buttons as described above (see Fig. 5.1.1 & 2). In order to measure the frequency of the ERM motors, a microphone was connected to each trigger, and each motor's vibration was increased from the minimum value that produced a vibration, to the maximum value that the API allowed, in steps of 10%. The microphone recording was taken at each step and the peak frequency of each step was identified and recorded from the microphone signal's Fourier transform.

The results of the measurements are shown in Figure 5.6, revealing clear differences between the left (blue) and right motor (orange) in terms of both range of expressible frequencies (the left motor outputs at an overall lower frequency band) and the specific output frequency in response to a given drive signal (the left motor output frequency aligns with the right motor when the command value is increased by 20%).

Scaling the Output of the Xbox Controller for Intensity Linearity. Plotting the intensity (or frequency, as an alias for intensity in the case of ERMs) against the drive signal and finding the best-fit equation (in our case, a 2^{nd} order polynomial as shown in Fig. 5.6) produces a model to allow control of specific output intensities. As a first stage, the output of the device may be scaled to be objectively linear. Secondly, differences in the motors may be considered, and accounted for if desired. In applications where symmetry is desirable, calibration of the motors such that their output signal is identical is necessary, and is only possible following this characterisation. The individual technical characterisation of vibrotactile devices can be particularly important as even devices of the same brand and technical specification may still vary. For example, I found that measures of frequency of a second Xbox One controller revealed even a greater discrepancy between the vibration intensity of the two front ERM motors.

The method of measuring the intensity of a vibrotactile device as described above is specific to devices using ERM motors which is, overwhelmingly, the output encoding of choice in vibrotactile signals (owing to the common use of ERMs and issues in differentiating between amplitude and frequency [23]).

I performed this scaling on the Xbox controller, and re-measured the vibrational output frequency in response to the linearly scaled API commands. Fig. 5.7 shows the Xbox controller's ERM response curves following a scaling to linearity, and matching of motor frequency output ranges. I note residual inaccuracies and deviations from the commanded linear response. Given that off-the-shelf devices are not designed for such precise control, these deviations may be attributed to technical limitations such as drive circuitry and manufacturing precision.

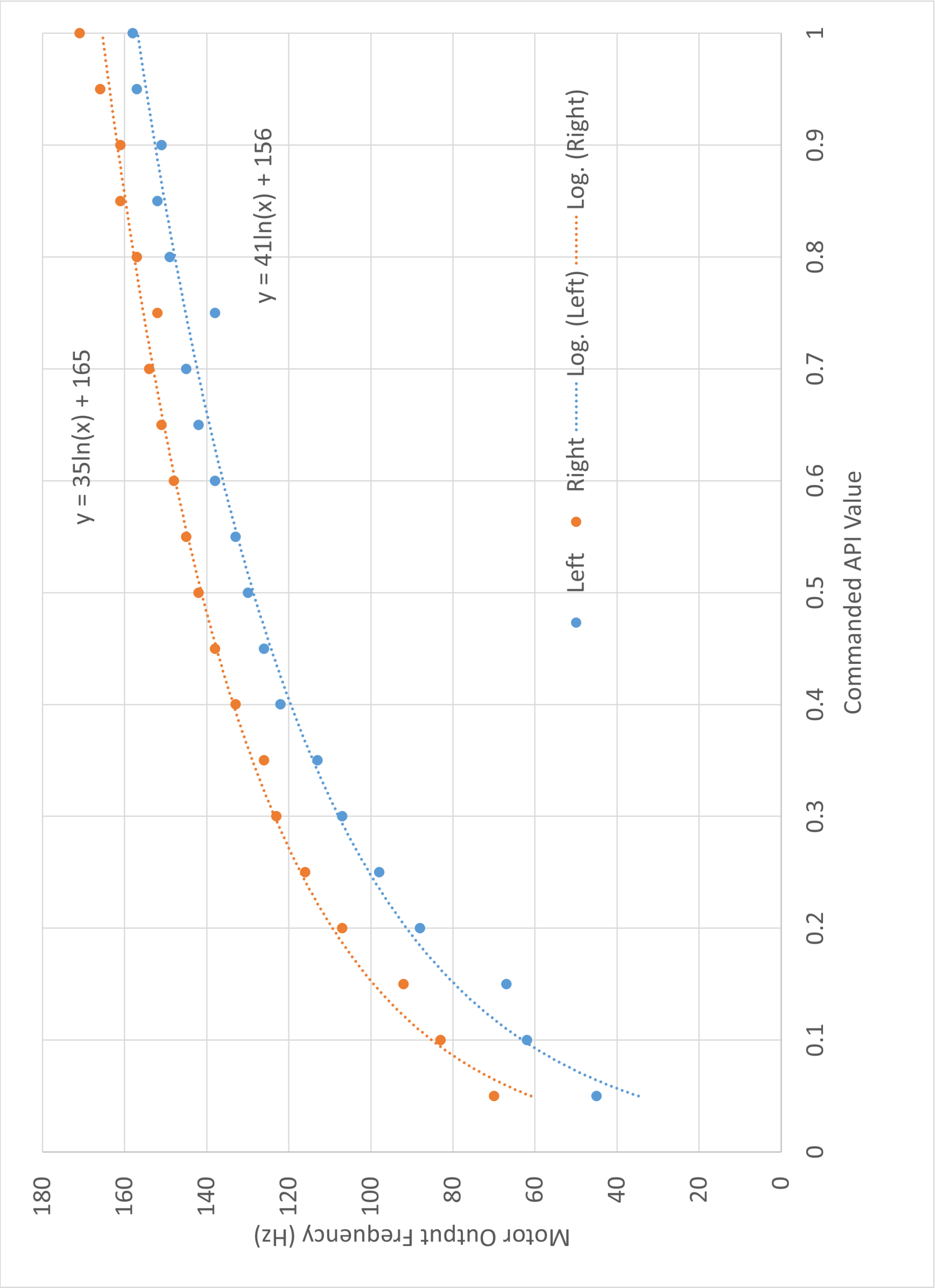


Figure 5.6: Measured frequency of vibration output by the front left (blue curve) and front right (orange curve) ERM motors of the Xbox Controllers in response to the API signal without motor characterisation. Such naive use of provided API controls may result in asymmetric, nonlinear physical output, leading to unpredictable perception.

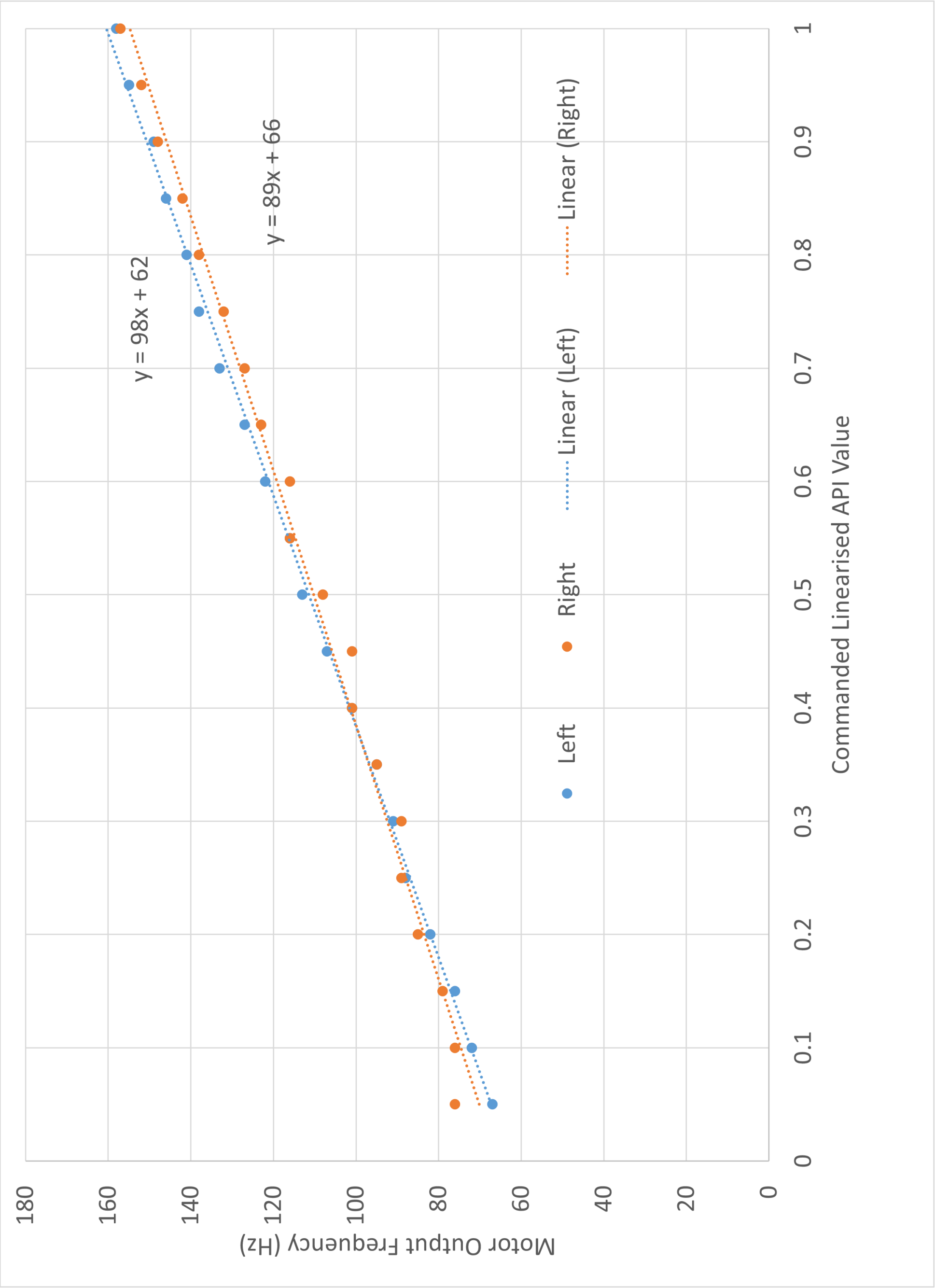


Figure 5.7: Measured frequency of vibration output by the the front left (blue curve) and front right (orange curve) ERM motors of the Xbox Controllers, following motor frequency range matching and scaling for linear mapping between input signal and output frequency.

The measurement of technical device characteristics as described above should be considered the minimum standard for driving the encoding of quantitative information and their vibrotactile representation. The reporting of these models and measurements enables generalisabilty and replication of novel vibrotactile devices and related studies.

The device characteristics and limitations gleaned through technical measurements, in combination with the results obtained from perceptual experiments as described above, can be used to inform effective and expressive vibrotactile representations of quantitative data. Furthermore, even without perceptual experiments, the output signals of multiple motors can be calibrated: Differences between the range and response time of individual motors can be corrected by restricting all motors to the range of frequencies they are all able to output. Individual outputs may be objectively linearised as above, by applying the model's equation to the drive signal. The reduction in the overall range of frequencies available will result in a drop in absolute magnitude estimation accuracy (as the individual levels will be closer in frequency, and thus less differentiable), but it will enable more accurate perception of differences and symmetry between motors. At this stage, it is important to consider which calibration steps are most important, also considering the intended usage scenario. In vibrotactile displays where motors act as identical "pixels", this type of calibration is critical. However if signals are intended to represent unrelated variables which will never be compared, then such homogeneity may not necessarily be advantageous.

5.5.2 Just Noticeable Difference & the Weber-Fechner Law - Profiling Human Perception

As per the second stage in the presented schematic (Fig. 5.5), several human factors affect how vibrational intensity is perceived. Cognitive interpretation by the human involves an interaction between the incoming physical perturbation and an already-mastered sensorimotor contingency, sometimes described as "piggybacking" onto a target modality [149, 163]. The chosen modality will have physical and neurological associations and non-linearities which must be taken into account when encoding signals for information transfer. As such, here I present a simple model describing vibrotactile perception, for the informed presentation of haptic signals.

The Just Noticeable Difference (JND, or ΔT) of a stimulus is a psychophysical metric defined as *"the smallest detectable change in a stimulus or difference between two stimuli that can be reliably detected, often defined as the difference for which the percentage of correct discriminations is 75 per cent [41, p206]"*. This value provides an indication of the "sensitivity" of a sensory channel [83] and allows us to encode the output signal of a vibrotactile device for optimal perception.

The JND is typically measured by presenting pairs of stimuli and recording the user's indication as to whether the second stimulus was higher or lower than the first. The magnitude of the difference between stimuli within a pair is reduced until the rate of incorrect responses drops below the desired accuracy level (typically 75%), yielding a figure for the user's threshold of differentiability. Adaptive algorithms (such as the staircase method or Quest [168]) which dynamically adjust the presented stimulus pairs based on previous responses are now considered standard [160] and are relatively straightforward to implement [168].

The JND of a stimulus (including vibrotactile intensity [60]) is usually found to be roughly proportional to the magnitude of the stimulus; a phenomenon described by Weber's Law [41] (see Jones [82] for an opposing view). At sensory extremes, this fractional value is no longer accurate, and Weber's law no longer applies. Within the range that Weber's Law does apply, perceived intensity will be proportional to the logarithm of the actual measured intensity, according to the Weber-Fechner law: $p = k \ln \frac{S}{S_0}$

By this law, to generate increasing levels of stimulus which will be perceived as linear steps in intensity, the intensity of a given step S_n may be determined by the following formula:

$$S_n = S_0 (1 + \frac{JND}{S_0})^n$$

This scaling, which has been implemented by the most robust and thorough studies noted above [130, 148, 165], seeks to maximise the participants' ability to differentiate and classify signal magnitudes, and to perform accurate difference estimations.

Measurement of the JND at several intensities allows confirmation of the range within which Weber's law applies. The limits of the minimum/maximum output encoding value should be limited to this range, ensuring that the above points hold and reducing the risk of unpredictable interpretation.

In some recent publications, standard values for the JND have been used in lieu of measurements to optimise for signal perception as described here [165]. However, the JND for vibrotactile stimulus has been shown to vary depending on the frequency and type of vibration [79, 136], and the size and location of the site of stimulation [76]. Where feasible, it should therefore be measured and reported, especially for new devices or stimulus locations for which the literature cannot provide reliable JND values.

Evidence of measurement of the JND followed by optimisation for linear perception can be found in perception-based studies [113], with exceptional works by Cardin et. al [30] and Schoonmaker and Cao [148] in the context of deliberate customisation at a per-user level.

Naturally, the success of the device in further comparative measurements and the associated statistical tests in these studies.

As a first step, the JND is a useful indicator of the required resolution of the output signal. If the impression of a continuous output signal is desired, the device output signal must have an expressive step size of the JND measured at the least intense value. This ensures that incremental changes in output signal will typically be unnoticeable across the usable range. Being unable to tell the difference between adjacent output values means the signal will always be perceived as continuous.

The JND can be cumbersome to measure, and measurement at several intensities across multiple participants is a time-consuming endeavour. As the JND deviates towards sensory extremes, it is most representative at the midpoint of the sensory range. As such, it is reasonable to measure a single JND at the midpoint between the lowest intensity which is barely detectable, and the highest intensity which is tolerable without discomfort. Literature may also be used as a guide for the range within which Weber's law applies.

Using a comparable literature value [82, 113, 136] to guide scaling and resolution of output signal is superior to applying the drive signal naively.

Scaling the Output of the Xbox Controller for *Perceptual* Linearity Literature values for the JND of vibrotactile feedback on the fingertip vary [136] widely, though are typically reported between 10-20%. I chose to adopt a JND estimate of 12% in a compromise of literature estimates. The left and right trigger motors of the Xbox One controller were parameterised from the technical measurements discussed in Section 5.2, and the outputs were scaled by equation 5.5.2 (defined in Section 5.5.2) to achieve an exponentially increasing intensity output that would be *perceived* linearly.

Using the same methodology as in section 5.2, I re-measured the frequency output of both controller trigger motors in response to a linearly increasing input signal. The results of this scaling are seen in Fig. 5.8, which shows how the output frequency (Y-axis) now rises exponentially (exponential trendlines Expon.(L) and (R)) in response to commanded API signal (X-axis). The figure further highlights the limitations of the controller, as the effects of reducing the effective range of the motors are visible as discrepancies between the two motors at the same API value, and deviations from the overlaid trendlines.

Following this characterisation and calibration process, I repeated the user-studies detailed in section 5.4. In the next section, I describe and discuss the effects of the re-calibration process on the aforementioned absolute and relative magnitude estimation tasks.

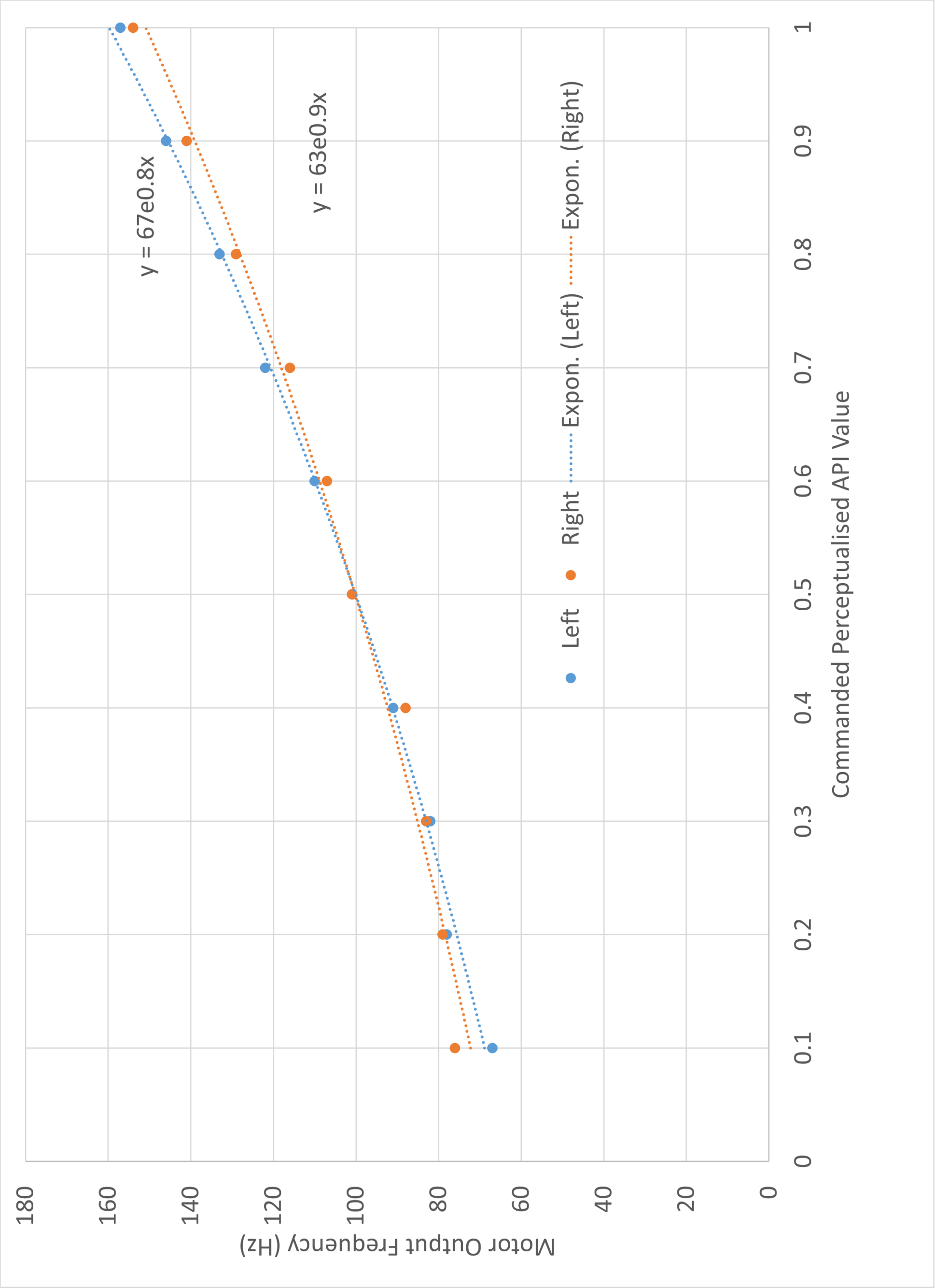


Figure 5.8: Measured frequency of vibration output by the front left (blue dots) and front right (orange dots) ERM motors of the Xbox Controller, following scaling for linear perception as per the motor frequency response parameterisation and the Weber-Fechner model for vibrotactile perception. Exponential best-fit curves (Expon.) are shown by respectively coloured dotted lines.

5.6 Post-calibration Measurements

Here, I outline the repeated studies following the calibration of the Xbox controller: the results of taking the steps detailed in the previous section to achieve perceptually linear feedback.

5.6.1 Post-calibration Magnitude Estimation Experiment

The Xbox One controller used in the initial study was characterised and calibrated as per the steps described in the previous section. Participants in the following studies were therefore exposed to vibrational frequencies from 70-160Hz, in 10 discrete steps of magnitude each separated by slightly less than one JND.

5.6.1.1 PAM Estimations on the Calibrated Xbox One Controller

The conditions of the study were identical to the original study (detailed in section 5.4), except for the scaling of the output vibrations to correspond to a perceptually linearly increasing scale as detailed in the previous section.

I measured the perceived absolute magnitute (PAM) estimations for the left and right trigger motors of our Xbox One controller across 5 participants, two of which had partaken in the pre-calibration experiments. The same block structure and methodology as the pre-calibration study was followed, exposing participants to 5 blocks of trials. Each block consisted of 40 trials (10 intensity levels x 4 repetitions) and focused on a particular task condition as before; Vibration Left Only, Right Only, Both Motors, Match Right to Left, and Match Left to Right. The order of trials was randomised within each block.

5.6.1.2 Data Collection and Analysis

For each trial, the presented and reported intensity intensity values were recorded alongside presentation and decision timestamps.

The analysis of this data is conducted through exploration of comparable visualisations to the first iteration of the study (confusion matrices and line charts), with supplementary observations on the off-by-one error summaries.

5.6.1.3 Findings

The results of the post-calibration modified PAM experiment are visualised in the confusion matrices and line charts shown in Figure 5.9.

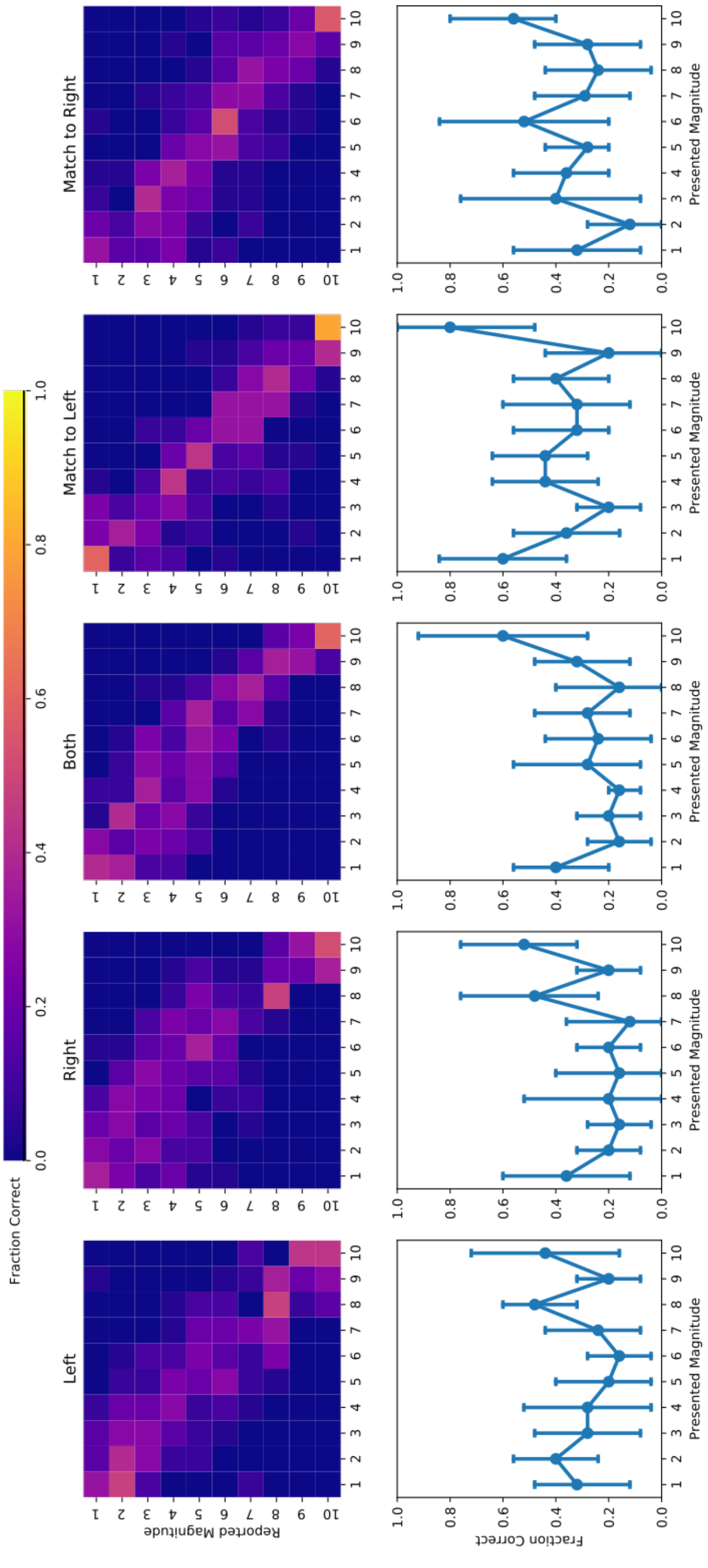


Figure 5.9: Confusion matrices showing presented vs recorded vibrotactile magnitude values, above intensity classification performance, for each task condition, averaged across all participants, following calibration.

Exposure	% Correct
both	70
left	72
right	61
match to left	79
match to right	70

Table 5.2: Percentage correct judgements made allowing off-by-one (+/-1) in the Absolute Magnitude Estimation task using the Xbox One controller *after calibration*.

The confusion matrices for the Left, Right and Both motor magnitude estimation tasks lie relatively close to the diagonal, suggesting a slight improvement in consistency. The line charts show less deviation (between 20-60% accurate) than in pre-calibration tests (10-90% accurate), supporting this.

When performing this tasks with the uncalibrated controller (see Fig. 5.3) biases were apparent towards the ends of the intensity ranges as bright yellow squares offset from the diagonal. In contrast, the calibrated device results in responses more evenly spread across the available range, with fewer bright pixels, especially away from the diagonals. This is especially well-pronounced in the ‘Both motors’ condition. This evenness suggests that scaling for perceptual linearity has reduced the influence of innate perceived magnitude on participants’ matching of vibrational intensity to the trained scale.

This suggestion is strengthened by the line charts, which show more even structure across the range of perceived intensities. The previously high spikes at the edge values for individual and both motors have been reduced to a level comparable with other intensities. While a reduction in accuracy may seem superficially negative, in this case, it indicates that participants are no longer biased towards choosing extreme values.

The charts reflecting the matching tasks indicate that calibration has had a strong positive effect on the symmetry of the device. This is again confirmed by the line charts, which show a more even and overall higher accuracy in judgement.

Calculating the off-by-one errors as shown in Table 5.2 indicates that overall accuracy of participant estimates for the magnitude estimation tasks was similar to the pre-calibration. The improvements in symmetry noted from the confusion matrices and line charts are confirmed by the increase in off-by-one judgement percentages for the Match to Left and Match to Right tasks, from 42% and 56% to 79% and 70% respectively.

Analysis of Post-Calibration Estimation Tasks. The effect of calibration on participant perception of the output signal, in comparison to pre-calibration results, is most noticeable in two areas. *Firstly*, biases towards choosing extreme values have been reduced, while the corresponding off-by-one accuracy has been largely unaffected. This indicates that calibration allows participants to extract a similar amount of information overall, but are experiencing each intensity with more similar and even accuracy. *Secondly*, the accuracy of the matching tasks has been improved, and a larger degree of symmetric perception between the motors is achieved. The importance of this symmetrical perception will be discussed in further detail below.

5.6.2 Post-calibration Perceived Difference Experiment

For completeness of analysis, I repeated the perceived difference estimation task as detailed in section 5.4.

5.6.2.1 Perceived Difference Estimations on the Xbox One Controller

This post-calibration experiment involved the same 5 participants (all male, age between 22-32) as the above absolute magnitude estimation task, 2 of which had previously undertaken the pre-calibration difference magnitude estimation task. I measured the perceived difference between the given vibrotactile signals for the left and right trigger motors of the calibrated Xbox One controller under the same methodology as outlined above. Trials were presented in two blocks each of 95 trials (difference levels from -9 to +9 = 19 levels, x 5 repetitions) with a specific task condition (total 190 trials). The order of the trials was randomised within each block. The two task conditions were again Left only and Right only.

5.6.2.2 Data Collection and Analysis

As before, for each trial, the presented starting intensity, the actual difference and the reported difference were recorded, alongside presentation and decision timestamps.

The data were again analysed through visualisation in confusion matrices and line charts.

5.6.2.3 Findings

The results from the post-calibration difference estimation task are shown in Fig. 5.10. The charts show participants are typically able to identify that the signal has not changed intensity with an accuracy of around 70%. However brighter squares running vertically from the centre point on both matrices indicates that participants often perceived no change in stimulus even when presented with a change of several steps.

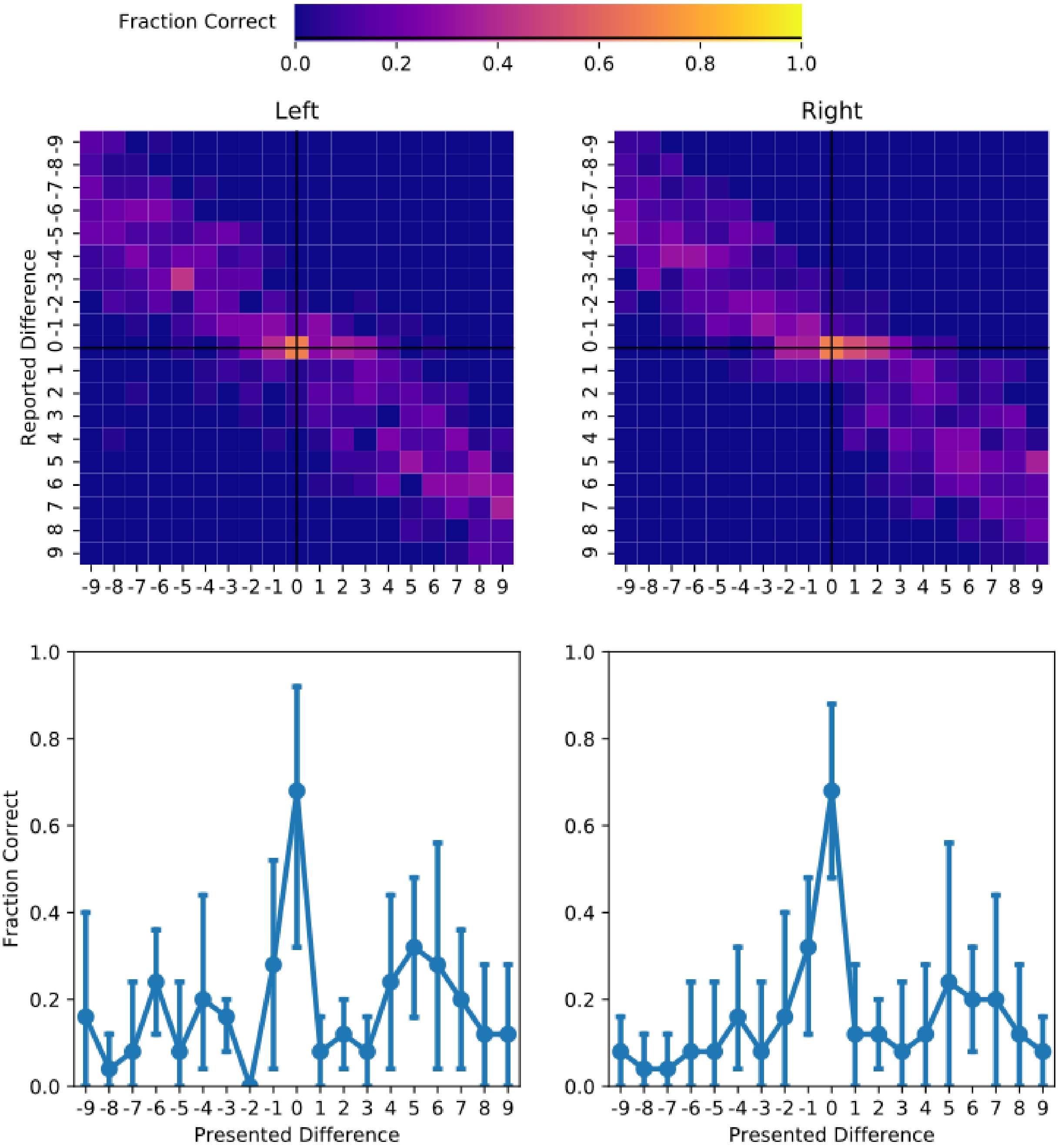


Figure 5.10: Confusion matrices showing presented vs recorded differences in vibrotactile magnitude values, above corresponding classification performance, for each task condition, averaged across all participants, following calibration as per above.

Analysis of Post-Calibration Difference Magnitude Estimation Task Results. There are 8 JND steps of 12% between the motor frequency lower limit of 70Hz and the upper limit of 156Hz, so following calibration, the resultant 10 steps in magnitude are each separated by just under 1 JND. This is reflected by the difficulty participants experience in perceiving smaller presented differences.

When compared to the pre-calibration perceived difference measurements, the results are interesting from two key perspectives: *Firstly*, as seen in the confusion matrices, participants are still biased towards reporting smaller difference magnitudes than presented, and would rarely report differences of more than +/-6 intensity steps in a 'saturation' of difference reporting. This may be down to several reasons, which will be suggested in the next section.

Secondly, participants still appear to have difficulties interpreting small changes in the signal, often reporting no change in response to presented differences of several intensity steps. This may seem counter-intuitive, given that around 70% of absolute magnitude estimates were accurate to within +/-1 (as per Table 5.2, suggesting that participants should be able to judge differences of +/-2 or greater very accurately.

5.6.2.4 Discussion of Post-Calibration Difference Magnitude Estimation Results

Overall, it seems the calibration of the absolute magnitude for linear perception (which was, per technical measurements, largely successful) did not have a significant impact on the perception of difference magnitude.

This may be an artefact of the task itself: as participants are asked to choose from a scale which requires moving the slider from the centre (zero) region towards the outermost values, there may be a psychological factor causing resistance against choosing the highest values in change.

Alternatively, a potentially interesting interpretation of this result is that the judgement of differences (the magnitude of signal change in some context) is itself being perceived *directly*, and is therefore subject to Weber's law. This would suggest that *differences* between stimuli are also perceived logarithmically. If this is the case, accounting for these biases and underestimations through further calibration would affect the perception of the absolute magnitude estimation task.

To illustrate with an example, an experiment could be performed in which participants are asked to compare the acceleration they experience while sat in a vehicle. When asked to judge the perceived acceleration, participants may respond non-linearly to the sensation of increasingly aggressive acceleration, even if they are able to respond accurately, and linearly, with estimates of the starting and ending velocities involved.

To summarise, the difference estimation task proved useful in guiding *expectations* of the vibrotactile modality in the presentation of continuous, comparable numeric source information. I conclude that difference judgements may be perceived non-linearly, with comparisons of small values (under static presentation) being hard to perform accurately. However, it does not immediately provide answers as to the viability of the encoding or device presented for SSAD or sensorimotor utilisation, which I will discuss in detail in the next section.

5.7 Characterisation and Calibration Summary

Figure 5.11 summarises the steps detailed above to characterise and calibrate the Xbox One controller. Below, I generalise these steps for application to a wider range of vibrotactile systems.

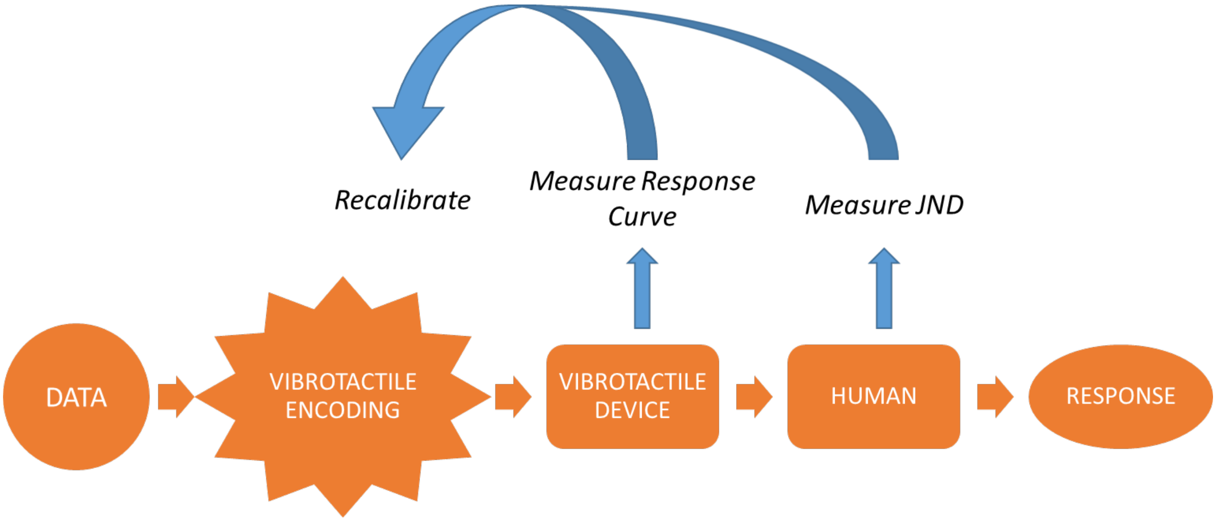


Figure 5.11: Schematic representing the factors which lie between source data and cognitive interpretation of a vibrotactile signal by the user, with further reference to the system characteristics which may be used to calibrate the device for reliable perception.

1. Firstly, in order to understand the range and nature of vibration intensities available, the output signal response curve of each vibration motor of interest must be measured in response to the presented command signal. This may be for example a voltage, PWM, or an API value afforded for control through programmatic interface (as was the case for the Xbox controller). The measured output must be a directly perceivable value such as vibrotactile energy or 'intensity', or a technology-dependent proportional proxy for such, for example in the case of ERMs, frequency (see Section 5.5.1.1). The control encoding may then be scaled such that the output value is linearised with respect to the drive signal, allowing predictable, accurate representation of the raw data in the output signal.

2. Next, the Just Noticeable Difference (JND, or ΔT) of the target modality must be measured, or appropriated from a suitable literature source. This value should be parameterised into an exponential function according to the Weber-Fechner law (see Section 5.5.2). The control encoding may now be scaled by this function, such that the (now known) output is perceived to be directly, linearly proportional to source data.

5.8 Discussion

To conclude this chapter, I will look deeper into the implications of the methodology and results described above.

Firstly, I will address the schematic presented in Figure 5.5 through a re-vision of the steps involved in device preparation.

To begin, technical measurements allow modelling of the output signal in relation to the drive signal. Motors can then be characterised and calibrated with consideration for symmetry, homogeneity, and limitations of the rate of change of the data signal. Following this, scaling the signal using a relevant value for the JND (measured or from literature) enables the output signal to be optimised for greatest perceptual differentiation. After these steps have been taken, experimentation relevant to the final application may provide genuine insight into the success of the tested interaction, with confidence that the effects of device limitations and human perceptual factors have been accounted for.

While individual components of the described schematic may be addressed and measured in isolation, optimal results are achieved when the motivation behind all steps in the framework is understood, and appropriate action is taken to remedy potential sources of device limits and nonlinear perception.

5.8.1 Calibrating Devices for Linear Perception

The methodology applied above for the calibration for linear perception was derived through reference to several thorough studies in the literature [30, 148]. In Schoonmaker and Cao's work [148], nearly three quarters of the study is dedicated to describing and detailing the per-user characterisation undertaken. It was effected such that a linear increase in their source data (in this case, force applied through a surgical tool) was experienced as a linear increase in vibrational intensity. The argument behind undertaking such a characterisation and calibration process is not heavily detailed, however, Schoonmaker and Cao infer (reasonably) that following such steps will enable the creation of more intuitive information encodings in vibrotactile feedback, potentially leading to greater usability, and more reliable and comparable results in academia.

The degree and detail of individual calibration and customisation described in Schoonmaker and Cao's study exceeds our own, perhaps to the point of inaccessibility. Schoonmaker and Cao present specialist measurement of several advanced vibrational features (such as spring and electro-mechanical constants), and then performs customisation of the output response curve to cater for a per-user measured JND at the site of contact. By contrast, I present simple acoustic measurements, and utilise a literature-sourced value for the JND to perform perceptual scaling. While our methodology may be less robust compared to Shoonmaker's, given the limited accuracy of the device response (as exhibited by the deviation from commanded values shown in Fig. 5.8), it is highly unlikely that more accurate measurements or calibration would have a significant effect on the results of the magnitude estimation tasks.

The pre-calibration estimation studies highlighted biases and asymmetries that I was able to address. The presented step-by-step process mimics to some degree the efforts undertaken by Schoonmaker and Cao, albeit with a much lower fidelity device and simplified measurement and calibration process. Our findings therefore suggest that taking the time to measure and calibrate vibrotactile devices for sensorimotor interface studies even at a basic level is valuable, and should be considered a standard for future experimentation.

A further limitation (manifested as issues in discrimination between values in the post-calibration tasks) may have arisen from the reduction in output frequency range of each motor. As the range of each motor was reduced to accommodate for symmetry, there remained just 8 JND gaps from 70Hz to 160Hz. However, in order to keep results comparable between the pre- and post-calibration tasks, participants were again exposed to 10 different magnitudes. Each adjacent step is therefore separated by less than 1 JND, which understandably leads to inconsistent perceptual estimations.

As such, the success of the herein presented method for measurement and calibration offers several options on how accessible compromise may still achieve notable improvements in perception and performance.

In addition to highlighting the utility and importance of this methodology, the steps taken through this study served the purpose of preparing the Xbox One controller for the presentation of continuous information, in which symmetry and a linear perception of the scale is critical (see Chapter 6). The experimentation has confirmed that the assumptions of perceptual linearity following calibration are valid.

However, Sensorimotor Theory (summarised in Chapter 3) suggests that we can form a new modality with any abstract set of rules, provided they are robust and consistent. If sensory modalities can be learned regardless of the linearity or consistency of the rules, is this calibration useful?

Learning, and academic study, of new sensorimotor interactions would be simple if novel sensorimotor interfaces were presented and available as children, before modalities are robustly developed, and before a complete understanding of our own perceptual skills and limitations evolves. Observing how the new modality integrates into consciousness, and how the world may be perceived with additional affordances, would play a significant part in testing and informing sensorimotor theory and the potential for associated interactions. However, this is not the case. In most applications, new interfaces are offered to and utilised by adults, able to interpret and estimate sensations, and whose sensory toolkit is essentially fully-formed. Adults have an understanding of space, having occupied and existed in space our entire lives. We understand movement and perspective, and can differentiate between static objects, dynamic objects and even objects with agency [7]. Contradicting or ignoring the existing perceptive skills with which we experience the world would inevitably hamper the learning of new vibrotactile encodings. We need to ensure that the mappings we present to adults are complementary to their experience and perception of the world [137].

5.8.2 Quantitative Metrics and 'Interpretation' Comparisons

Visualising the perceived vs. actual absolute vibrotactile intensities in the diffusion matrices as shown in Figure 5.3 indicates the reasonable expectations regarding the accuracy of users' estimation of vibrotactile intensities, which, in turn, can inform the design of vibrotactile encodings of SSADs. For example, these findings can guide the choice of intensity ranges if accuracy is an important requirement.

Importantly, in agreement with Novich [130], our results suggest that continuous encodings are not well-suited for the communication of explicit data. Even after calibration, for 10 levels of vibration, presented independently with unlimited time for interpretation, when allowing +/-1 of the presented value the typical accuracy of estimates in the absolute magnitude estimation tasks was around 70%. This is significantly weaker than the values reported by Novich for 10 discrete spatiotemporal tactons, which approached 70% before allowing for off-by-one errors. That said, owing to their intuitive nature and comparability, homeomorphic mappings are still well-suited to sensorimotor interfaces.

My experiments also allow the modelling of the relationship between the input data and the perceived signal across participants. Measuring the perceived vibrotactile intensities after the technical device characterisation and calibration as described in Section 5.5.1.1 establishes the relationship between the input data and the perceived signal, providing confirmation of correct calibration, interpretation of device specifications or technical assumptions. In the absence of technical data, the results of this study highlight potential perceptual limitations of the device if utilised as-is.

The results of this test therefore reflect both device and human factors. A consideration of the (relatively poor) accuracy of participants in this task should guide the task encoding design: biases in cognitive processing and judgement manifest even when primary issues such as the linear perception of intensity have been addressed. This is not unique to the vibrotactile modality. Consider the mastered visual modality. Cognitive judgements on high-order visual properties such as area or volume often result in non-linear perception [122] despite mastery of the modality and an ability to compare lower order elements, in this case length. It is therefore reasonable that proficiency with a sensorimotor interface may develop, and enable meaningful interactions, despite poor performance in cognitive interpretation of higher order or comparative elements of the modality.

5.8.3 Re-thinking Metrics for Sensorimotor Interfaces

While asking participants to estimate or judge source data or environmental variables using SSADs is a relatively common task in the literature, many of the above insights indicate the incompatibility of the 'enactive' or 'sensorimotor' approach with that of the 'cognitive' paradigm derived from wider device design and comparison.

Building on Kałwak's insights [84], this leads us to ask how we may truly tease apart the device from the human, and the cognitive estimation from the utility or existence of a new sensorimotor contingency. When evaluating sensorimotor interfaces, the field needs metrics which allow meaningful comparisons of the devices and interactions we have built, not just proxies for psychometrics or a re-evaluation of the limits of human perception in a new format.

Unmentioned thus far, and common to all of these metrics, is the important and fundamental difference between the presentation of static data for estimation (as above), and the availability of such data as an on-demand, dynamic signal during task performance. Given that dynamic control is a key requirement of Sensorimotor Theory, it seems reasonable to accept that utilisation of sensorimotor signals is dependent on context and control. Participants may be able to make absolute judgements of static signals in the vibrotactile modality (such as how vision allows fine discrimination of spatial objects or patterns) but being able to perform such judgements likely depends on their already-mastered vibrotactile modality, and some cognitive interpretation, rather than on a new, emerging sensorimotor skill.

Further work is therefore required in order to address reliably whether the success or failure of a given experimental sensorimotor interface can be attributed to the application design, the human body, or the device itself.

While continued development and evaluation of the techniques, methodology and principles described above may allow incremental improvement of signal reliability, I propose that pursuit of such concrete comparison measurements may be a dead-end route in the evaluation of sensorimotor interfaces for HCI. Simply put, the traditional approach of HCI studies in the form of a quantitative comparison of performance in simple perception or task-performance studies, especially with minimal effective training and device familiarity, is not suitable for meaningful comparisons of sensorimotor potential.

That said, creating new metrics which *do* reflect device potential, is of critical interest and relevance to the literature. Results which allow frequentist statistical evaluation to significance are technically favourable in the literature, especially in HCI, as they are directly comparable. Adaptations of the experimentation presented above may yield such results, or alternatively, may provide complementary or confirmatory data following more qualitative testing of new sensorimotor interfaces, as will be discussed in the next chapter.

5.9 Limitations

There are three major limitations to the work presented in this chapter which warrant attention and discussion here. First and foremost, the methodology by which the Xbox controller was measured is not as robust and detailed as that presented in similar literature studies [148]. Additional work could have been conducted to measure the JND for our controller specifically, and to further customise the output to cater for individual requirements. As noted above however, given limited device fidelity, further customisation and fine tuning of the controller output would result in effects smaller than the errors in the consistency of device output, and much smaller than the reliability of human interpretation in the performed experimentation. From this perspective, the simplicity of the methodology, especially given its effectiveness in reducing biases and improving symmetry of the off-the-shelf device, should instead encourage understanding and attempting to perform such calibrations in future studies, and highlight the consequences of failing to do so. In any case, as manifested by Schoonmaker and Cao [148] and Cardin et. al [30], the measurement and calibration steps outlined here have proven benefits in SSAD and sensorimotor studies, yet are still not well-represented in the literature. This can be addressed through simple consideration of the methodology presented above.

As a second point, the methodology undertaken for the Absolute Magnitude Estimation task may have introduced biases through presentation of the initial fixed scale. The setup was designed to explore how participants can extract raw data, rather than as a simple perceptive experiment, but this setup can cause biases, affecting the value participants report through a convolution of cognitive decision making (learned against the presented scale) and the natural perceived intensity. I argue that this more realistically reflects the type of decision making participants would pursue when learning to engage with new sensorimotor experiences. However, it is also safe to conclude that these particular investigations are simply reflective of controller capabilities, and do not give broad or deep insight into sensorimotor experiences, not least because of the lack of user control (action) in learning the signal. At this point, we must seek other methodology for determining encoding and interaction effectiveness in the context of actual sensorimotor engagement, as will be discussed in the next chapter.

Finally, the studies performed suffer from low participant numbers and asymmetry between study tasks. While this creates limitations in the concrete conclusions that can be drawn from the experimentation, the process undertaken constituted an informative exercise in perceptual experimentation and presented opportunities to learn first-hand where improvements to the controller and signal could be made. Completing this calibration process and repeating experimentation, even with low numbers, confirmed that the device was suitable for communication of symmetric signals that would be perceived linearly. This represents an effective proof-of-concept for the simple calibration methodology outlined. Similar experimentation could certainly be undertaken which is statistically robust, symmetrical and more suitable for pre-/post-comparison, and the pursuit of such metrics (as motivated this work) is helpful for the developing field. However, we must be careful to form hypotheses that lead to information about device potential, rather than striving to take utility from statistical significance in narrow lab tests. It is difficult to imagine what such a hypothesis would look like, for such experiments, in the context of sensorimotor interfaces.

5.10 Conclusions

In this chapter, I have drawn three conclusions:

Firstly, a key limitation of sensorimotor interfaces has been confirmed and demonstrated: in scenarios where participants must make quantitative absolute evaluations or comparisons between dynamic signals, explicit estimations will be inaccurate. In concordance with Visell [163] and Novich [130], I confirm that such estimations and the extraction of specific or numerical information, are therefore far better suited to cognitive interfaces featuring symbolic outputs (such as tactons) and explicit commands (such as button clicks).

Secondly, I have shown that measurement, calibration and scaling of ERMs is a critical first step when building a sensorimotor interface using these technologies. Perceptual interpretation will occur when presenting vibrotactile signals, and these interpretations likely play a part in the intuitiveness of new interfaces, thus affecting the formation of effective sensorimotor contingencies. Without such efforts, the system's response to the drive signal is unknown, the user may not interpret the output signal in a fashion consistent with the source data, and there is no way of separating the effectiveness of the device from the output signal, or from the human.

Thirdly, performing measurements and calibrations to account for existing perception is realistically accessible and achievable with limited resources. There is little reason why an awareness of and mitigation for such calibrations should not be displayed in future work involving vibrotactile devices for sensorimotor interactions.

Despite the focus on metrics in this chapter, the limited conclusions I may draw highlight that a more well-rounded approach to evaluating devices, across multiple scientific and experiential metrics, would be more informative and better contribute to long-term understanding than a simple execution of the metrics presented above. Indeed, in the most robust of SSAD studies such as the long-term evaluations of feelspace [124] or the vOICe [138], this sentiment and the benefits of diverse metrics are reflected.

In the next chapter, I proceed with the evaluation of our characterised vibrotactile device in a multimodal application, and perform the application comparison using a wide range of metrics, inspired by several previous studies. In addition to this, I focus on how feedback and other existing sensory modalities play a role in the intuitiveness and learning of more complex sensorimotor decision making. It is intended that performing such a wide evaluation over a prolonged period of time, in task context, will give novel insights into the application of sensorimotor techniques going forward.

CHAPTER SIX

SENSORY CONCURRENCY AND LEARNING OF A HAPTIC SPATIOTEMPORAL SIGNAL

The research presented in this chapter was published in 2021: "Investigating the Effect of Sensory Concurrency on Learning Haptic Spatiotemporal Signals" [31].

Learning to use sensorimotor interfaces presents unique challenges, and the difficulty in overcoming these challenges has long been considered a barrier to uptake [19].

As described in Chapters 2 to 4, and in line with my derived model of sensorimotor interaction, perception through a sensorimotor interface requires *extraction* of the lawful interactions between action and stimulus, termed sensorimotor contingencies. Exercising knowledge of such contingencies forms the basis of perceptual experience under this model.

As this knowledge is procedural (rather than declarative), it stands to reason that optimising the process by which this skill is acquired (and therefore, developing efficient training programs for such devices) requires insight from theories of mind compatible with the acquisition and exercise of procedural knowledge. These models may then be tested through experimentation, and eventually, training guidelines from these theoretical foundations may be derived.

In this chapter, I derive such guidelines for learning based on insights from direct learning theory, a post-cognitivist theory of learning which encourages formation of procedural knowledge and development of sensorimotor skills in enactive environments.

I outline a study undertaken to test the validity of the model's claims, and to explore how available perceptual information and task-related stimuli contribute to the development of new sensorimotor skills. These experiments are performed in virtual mazes, using the Xbox One controller characterised and calibrated in Chapter 5.

Taking guidance from previous works and the insights of the models derived in this work, I introduce a mechanism for learning sensorimotor skills which leverages the balancing of intention, attention and calibration as components of procedural learning. The mechanism is compared with more commonly-implemented procedural learning conditions (such as sensory deprivation and cross-modal skills-transfer). User learning experience and performance under these conditions is evaluated using a combination of qualitative and quantitative methodologies in a multi-day study.

This chapter therefore demonstrates an experimental evaluation of the models and principles discussed throughout this work, as well as contributing new paradigms for training and evaluation of enactive signals. The results show the effectiveness of SMT as a foundation for enactive interface design, and the scope and implications of its generalisation to new sensorimotor interfaces.

6.1 Introduction

Leveraging the haptic sense to represent continuous and fast-changing information (e.g., the proximity to a target or speed of movement), is a challenge. As human beings we are experts in the visual interpretation of such information, but efforts to accurately interpret haptic feedback have proven a difficult and time-consuming process which requires both intensive training and/or long-term exposure [5, 114]. Due to the abstract nature of the information in the signal, there is often a steep learning curve associated with these devices, long-recognised as a barrier to uptake [19, 65, 91, 137]. Common consumer interfaces (e.g., games controllers, mobile phones, or smart watches) that make use of haptic feedback typically avoid this problem by (1) limiting the complexity of the signal and the information to be represented (e.g., discrete information about system status represented as a small number of distinct haptic symbols or tactons [23]), or by (2) using the haptic modality as a subordinate channel to improve immersion rather than communicating crucial information (such as the Nintendo Switch's HD controller rumble [94]).

In Chapter 5 I discussed the development and calibration of vibrotactile devices to enable "intuitive" mappings between source information and perceived haptic sensations. Such mappings have been shown to facilitate and speed up learning processes of haptic encodings [38, 69]. However, even with optimised mappings, a broad review of existing training strategies for sensory augmentation devices by Bertram and Stafford has shown that we still lack systematic approaches to teaching information encodings in enactive haptic signals. Specifically, we lack understanding of how feedback and learning conditions contribute to the development of competence and people's ability to effectively interpret encodings in-situ [19].

To ground the development of my hypothesis, the area of focus is on the role that concurrent sensory information plays in enabling and encouraging such learning. To illustrate, full "sensory concurrency", where duplicate information is received across more than one sense, provides a feedback channel which may be critical for the learning of dynamic haptic signals commonly used by SSADs [19, 78]. However, "sensory deprivation", where task-critical senses such as vision are inhibited, forces user attention towards the new signal and creates incentives which also promote learning [107]. These opposing paradigms have both been shown in isolation to contribute to successful learning of SSAD signals. Understanding and optimising the role of both sensory concurrency and sensory deprivation in procedural, sensorimotor learning is therefore key to improving the training process. As a result, I hypothesise that both rich feedback and an incentive to learn can be combined through careful consideration of both paradigms, and that sensorimotor signal learning may be accelerated when training under such conditions.

In this chapter I start addressing these questions by actively investigating how the visual sense can be leveraged to facilitate the decoding of vibrotactile information as part of spatio-temporal tasks. In particular, I report the results of an experiment comparing combinations of visual and haptic signals conveying task-relevant information to participants. The findings demonstrate how the various combinations of multimodal concurrency affects the interpretation of a vibrotactile spatio-temporal information in a maze navigation task. This interaction between the visual and haptic sense points towards mechanisms that can be employed to improve decoding and interpretation of vibrotactile information, ultimately reducing the overheads for widespread adoption of sophisticated haptic interfaces.

I detail the systematic comparison of the performance and sentiment of users executing a dynamic, virtual maze navigation task using only a novel haptic feedback signal. The users first execute the task unguided and unsupervised over a medium-term study, under one of several visual and haptic feedback combinations. Each combination reflects a specific hypothesised mechanism for learning through sensory concurrency and deprivation. Users in all conditions were unaware that the final task required navigation using solely haptic feedback, maximising their efforts to complete the task, rather than to learn the haptic signal. Therefore, the results provide an indication of learning effectiveness of each combination of modalities independent of deliberated attention.

The presented experimental design gives insights into how comprehension of new dynamic spatio-temporal vibrotactile signals is acquired. I develop these insights into principles that promote effective procedural learning during normal, productive task execution. The presented principles offer an alternative to dedicated training (such as feedback on training task completion, or dynamic user-controlled feedback systems [78]). By leveraging sensory concurrency to create effective feedback which promotes unsupervised 'learn-as-you-do' approaches, users may develop rich new multi-modal skills without direct supervision, helping to remove barriers to uptake of new expressive interfaces.

Finally, the effects presented accentuate the importance of considering study duration and training mechanisms when evaluating the viability of new, SSAD-like interfaces. I distill such findings into guidelines on the design of devices, signals and conditions to promote the creation of expressive interfaces which may be learned through self-guided, productive task engagement. If applied, such principles promote the creation of comparable user studies in literature, and result in improved reliability of studies evaluating the effectiveness of the abstract, dynamic spatiotemporal mappings unique to these interfaces.

6.2 Studying Vibrotactile SSADs - Related Work

As described earlier, a number of SSADs have been developed in recent years to support a wide range of user interactions and experiences in different contexts, such as independent navigation using the vOICe [117], FeelSpace [96, 124] or EyeCane [115], aiding communication or hearing via the skin [127, 171], and augmenting awareness during driving [92, 93] or other demanding tasks [32].

The literature demonstrates that participants are able to utilise SSAD device signals through exploration [9], that explicit explanation of device signals and introductory training under some conditions does not necessarily improve performance [62], and that a change in perception [167] may develop with sufficient experience and dedication to the use of the device in context.

SSADs are built on the principles that symbolic information presented by traditional interfaces (especially haptics) is often suitable for simple warning or notification data, but that there may be advantages to encoding more information into dynamic spatio-temporal signals: a sufficiently detailed output signal may allow users to plan and prioritise actions.

Although it is generally accepted that exposure time and device use plays a large part in developing proficiency (and therefore reaching device potential), systematic studies rarely have scope or resources for long-term in-the-wild evaluation of device use, leading to a trend in shorter-term studies and measurements designed to give quick insight into device potential.

In developing such insights, studies face two major challenges: 1) Study metrics should accurately estimate long-term device potential, and 2) Participants should be introduced to the device under conditions that enable and promote proficiency during the limited training time. A strong understanding of how device usage and training contributes to proficiency would allow short-term studies more confidence in testing and training, helping to tackle these challenges.

6.2.1 Study Tasks and Metrics

Most SSAD studies performed either in-situ or in lab experiments focus on object recognition and/or locomotive tasks. *Object recognition* tasks often feature in studies involving devices that substitute human vision, producing highly comparable metrics concerning recognition accuracy and speed [5, 9]. In short-term studies of this nature, a generalised sensorimotor skill may be differentiated from a rote-learning of responses to symbolic stimuli by the addition of a generalisation task, requiring that users identify objects in free response or previously untrained conditions [171]. By contrast, *locomotive tasks* (such as maze navigation, path completion or obstacle avoidance) are used to measure proficiency in the use of navigational or spatio-temporal SSADs. Such tasks require active attention and comprehension of the SSAD signal for success. It is reasoned that superior proficiency in the interpretation of the device signal results in superior task performance, allowing an estimate of device potential.

Due perhaps to the experiential component of using an SSAD, metrics for success in SSAD studies vary considerably. While finding statistical significance in performance metrics is an important part of determining device or interaction effectiveness, quantitative results may also be supported (or refuted) by more qualitative assessment of participant experience and device utilisation.

6.2.1.1 Quantitative Measures

Task Performance: Definite metrics collected using methods derived from conventional HCI experimentation, such as task completion times or object identification accuracy. These metrics provide comparable insights into the utility of a device in a specific context.

Error Rate: Errors in complex task performance (such as the number of collisions measured during maze navigation [62] or attempts in a search and locate task [138]) can provide useful insights into participant awareness and competence, as well as device limitations and fidelity of task information in the SSAD signal.

Brain activity scans, sleep patterns or other biological measurements of learning and perceptual integration: While out of scope for most studies, such measurements have been linked to procedural learning during long term SSAD use [96].

6.2.1.2 Qualitative Measures

User Experience: Data collected through questionnaires capturing Likert/NASA TLX scale data, or interviews, providing sentiment and subjective experience when using SSADs. In a unique study of SSAD utilisation, Ward interviews [166] long term users of the vOICe [117]. Weekly supplementary interviews conducted in König's long-term evaluation of the FeelSpace Belt [96] supported factor analysis of quantitative questionnaires, bringing increased validity to experiential conclusions.

Process/approaches to the task: Exploration of participant mental models, sense-making and task completion/device use strategies can provide insights into how specific devices present user-environment affordances. Lobo [106] specifically analyses route selection in a navigational task for insight into how various devices contribute to users' perception of space in a target approach task using the Enactive Torch.Further, understanding participants' task approach also gives insight into the experiential, or "what it is like" component of SSAD utilisation; a concept at the heart of enactive device design [105] and associated theories of cognition [133].

6.2.1.3 Mixed Method Approaches

Mixed-method approaches combining user experience data with performance metrics provide insights between condition groups where performance metrics are unable to differentiate, and may lead to longer-term validity of study results.

Nagel [124] and König's works [96] are critical literature contributions not only for the uniqueness of the long-term investigations, but also for the confirmatory nature of their qualitative results, which reinforce the suitability of some of their chosen training schemes and metrics (the real-world homing task) in the study, while showing the device and training limitation via others (a virtual navigation task speculated to be too dissimilar to the training environment for the device signal to prove useful).

Given the historical context of the measurement and comparison of SSADs detailed in Chapter 2 and the insights gained at the end of Chapter 5, the present study uses a detailed mixed-method approach to reach a robust conclusions as to the effectiveness of specific sensory variables in the SSAD learning process. Participant behaviour, process and strategy (through path analysis) all assist in differentiating deliberate user approaches from random, mechanical or uncertain efforts in task completion, where performance metrics may not provide a full picture.

6.2.2 Training and Study Duration

While robust metrics are important, the so-far un-discussed nature of SSAD training, the duration of SSAD utilisation and the feedback provided during use have also been identified as key components in the SSAD learning process [19, 138], and to-date there are still many opportunities for systematic studies breaking down the individual contributions of these factors in enabling device proficiency [19].

Here I evaluate how specific findings from short- and long-term studies contribute differently to the literature, but may both benefit from a systematic approach to feedback and evaluation.

6.2.2.1 Short Term Studies

Contemporary HCI studies are typically performed over the course of minutes or hours, and are designed to evaluate device effectiveness in specific use cases or contexts. The Lateral Line [93], Feeling Uncertain [92] and ProximityHat [170] allowed users a familiarisation time of 5-15 minutes in the context of a simplified or example use-case, before performing a series of device-dependent navigation tasks. As this length of exposure is significantly too short for the development of new sensorimotor skills, these studies may be relying on a degree of "cross-modal skills transfer"; where participants' familiarity with the task enables quick adaptation to receiving comparable sensory information through a new modality [102]. The original study of the Enactive Torch [62] uses several metrics to offer insight into the long-term potential of SSADs despite the short study duration. Though simple in design (a single-handed device housing a motor which vibrates with intensity proportional to the distance to the object it is pointed at), many users developed *"an experience of objects located at a distance in the world [62, p372]"* when using the device to navigate a physical maze, suggesting some degree of integration into the participants' perception of space. Despite offering users 40-45 minutes of "training" in a maze navigation task which closely resembled the main study task, the study's mixed-model metrics revealed no difference between trained and untrained users in performance nor sentiment (captured via Likert scale questionnaire). It is unclear whether this was due to the duration or the nature of training.

In these cases, the study goal and associated metrics were focused on device utility with naive users. The application of such measurements alone, combined with the short study durations, highlights specific limitations.

6.2.2.2 Long-term Studies

Two long-term systematic studies follow the training, use and evaluation of (specifically) vibrotactile SSADs. The seminal, now well-cited study of the Feelspace Belt [124] evaluates how wearing a vibrotactile belt (which "points" towards magnetic north) during waking hours and simple training tasks for 7 weeks affects participants' navigation abilities. Following training, participants engaged in mixed-method tests which revealed both significant effects in performance (under conditions similar to training) and reported experiential differences in comparison to control participants who did not wear the belt. Several years later, König et. al explore the effects of long-term use of the belt using fMRI and EEG [96], demonstrating that long-term use of the belt (under similar nature and duration to the original study) leads to brain activity typically coincident with procedural learning and navigation. König's results show that brain activity indicative of learning effects during SSAD use return to near baseline levels after 4 nights of daily use, suggesting that medium-term exposure to SSADs may prove advantageous over single session experiments.

While both studies reveal interesting details about the effects of long-term SSAD use on perception, it is not clear how the shared duration of 7 weeks was chosen, nor which elements of training and exposure contributed to the development of device proficiency.

6.2.3 Training Strategies for Sensory Augmentation Devices

To summarise, "familiarisation" typically involves using the SSAD in a free exploration environment where natural senses and device signals are available concurrently, while "training" consists of performing a device- and context-relevant task, often in combination with sensory deprivation (such as blindfolds). However, the effects of the contrasting sensory approaches remain entangled. A derived approach to designing SSAD training and familiarisation mechanisms is required. To date, there are very few systematic studies evaluating such mechanisms [19, 91, 100].

6.2.3.1 Implicit vs. Explicit Training

Bertram and Stafford's detailed review on learning of SSAD signals [19] identifies feedback as a minimum requirement for learning of SSADs, and is presented in two distinct forms:

Implicit Feedback - By nature, the signal from an SSAD is itself a form of "implicit" feedback in response to user action. It has been demonstrated since SSAD conception that user control is a necessity for learning; users must be able to observe the effects of their actions on the signal [9]. Changing the signal outwith the context of user action is not sufficient for the development of proficiency.

Explicit Feedback - Additional “explicit” or external feedback may be provided in the context of a goal-oriented task. This may be as simple as the success/failure to complete the task, or more specifically through measures of performance.

“Familiarisation” in shorter studies [18, 62] and longer-term belt wearing in the FeelSpace studies [96, 124] points towards a general awareness of the importance of time spent with implicit feedback: it is expected that users will develop proficiency simply through exposure to the signal, even when not performing task-specific actions. Equally, most studies seem to agree that “dedicated training” must feature an explicit, task-specific feedback component.

6.2.3.2 Intention, Attention and Calibration

As discussed in Chapter 4, Section 4.2.8, if the action and perception of using an SSAD is modelled through enactive approaches such as sensorimotor theory, then the process of learning such enactive experiences also requires an approach consistent with post-cognitivist psychology: a theory of learning from the approach should be consistent with the commitments to describe cognition [75]. In other words, if cognition is modelled as an enactive process, then the methods of learning pursued should also be consistent with enactive principles. While the direct learning theory detailed in Higueras-Herbada’s publication pursues a rigorous mathematical approach to modelling the learning of a simple sensorimotor task, it also allows wider, qualitative consideration of enactive learning as being comprised of three key components, which I summarise for application here:

Intention - The user is learning to connect sensorimotor skill with *intended* environmental outcome. The user must have a goal which requires sensorimotor mastery to accomplish. For comparison in experimentation, the (measured) degree of task success will proportionally reflect the degree of sensorimotor mastery.

Attention - The user is learning to focus *attention* toward the environmental variables most relevant to task completion. This includes both filtering out environmental noise and increasing sensitivity and comprehension of the information-carrying sensorimotor signal.

Calibration - The user is taking feedback to *calibrate* subsequent judgements. Feedback may take many forms, though without it, the user is unable to correct for judgemental errors and cannot gain sensorimotor proficiency.

Under this model, I argue that adaptation of the training environment to optimise the ease of these three mechanisms should prove an effective pathway to improving (or at least enabling) development and learning of new sensorimotor skills with SSADs. As such, when creating new sensorimotor experiences, I aim to provide a signal which enables progress towards a clearly-defined goal (unambiguous intention), is highly valuable, efficient and low-noise (focused attention), with readily-available feedback on judgements (fast, rich calibration).

Here, I explore whether a degree of sensory concurrency may assist in offering such feedback; where sufficient overlap between the information available in the augmented and natural senses allows users to self-calibrate through active exploration [107]. However, providing too much concurrent information during training may be detrimental to signal integration, as users simply retain dependence on the natural senses, ignoring the superfluous signal in favour of more efficient environmental and sensory variables to direct attention toward, thus forming their sensorimotor contingencies around these variables instead. This warning aligns with other observations in the literature, such as the guidance hypothesis [147].

As such, the objective of an optimised self-guided learning condition would be to balance sensory concurrency and sensory deprivation to encourage self-directed mastery of new SMCs during regular task performance.

6.2.3.3 Sensory Modalities as Part of Training

There is little experimentation surrounding how implicit feedback via sensory concurrency (i.e. which senses are available during in-situ or training environment device usage) affects the speed and effectiveness of learning to use new SSADs. Of those found, sensory concurrency has been shown to play an important part in learning effectiveness [78, 107, 138], a sentiment shared in the SSAD community [84].

In both FeelSpace projects' training sessions, participants wearing the belt (and controls not wearing the belt) were asked to make navigation judgements while blindfolded, then permitted to open the blindfold for feedback. Thus, the individual contributions of implicit and explicit feedback during the everyday wearing of the belt are unclear.

Proulx [138] begins an investigation into how the effects of familiarity and training may be detangled, assessing the effects of "learning-by-doing" in a 3-week study, finding that participants using an SSAD during unrelated daily activities outperformed those who used it during training and testing only. However, long-term sensory deprivation (blindfolding) did not notably influence performance. By contrast, in a pretest-practice-posttest study, Lobo [107] showed sensory deprivation (blindfolding) to be more effective than full sensory concurrency (no blindfolding) when learning to use a leg-mounted obstacle detecting SSAD.

Between the measurements of application-specific devices and the results of longer-term studies exploring human perception, it is clear that vibrotactile SSADs are capable of enabling rich interactivity; often offering measurable improvements in decision making and allowing planning at a pre-cognitive level. However, as research interest grows, it is critical that participants are introduced to these devices via grounded methodology which allows comparable evaluation of long-term potential.

There remains questions of how combinations of modalities available during SSAD usage affects learning and comprehension of the introduced signal, and how such combinations may be best utilised for more effective feedback. In the following presented experimentation, I describe a mixed-method approach in a medium-term study designed specifically to address this.

6.3 Studying the Effect of Visual Feedback on the Comprehension of Haptic Signals

In order to investigate how visual feedback affects the comprehension of haptic signals as part of digital navigation tasks, I designed a mixed-method study that combines quantitative measures of performance with qualitative measures of navigation strategy and user experience. Below I provide details of the study design and procedure.

6.3.1 Study Design

Traditional video games often require the navigation of a digital character a form of digital landscape or obstacle path. Frequently, this requires the coordination of hands (via a games controller) and eyes (viewing a screen). In simple games, proficiency is quickly acquired, owing to the versatility of the visual sense and high dexterity of the fingers. The following study is based on a maze navigation task designed to help us explore how various combinations of visual and haptic feedback provided during such a task affect the comprehension and utilisation of the haptic signal when all visual feedback is removed. The digital maze navigation task can be considered a proxy for other real-world navigation task scenarios where haptic feedback is used to augment perception.

The study follows a between-subject design. Each participant performed four sessions of a maze navigation task across four days in a single work-week. In order to compare how different combinations of visual and haptic feedback facilitate the accurate interpretation of haptic feedback over time, participants were divided into four groups - one for each study condition. Participants in each group were introduced to a specific combination of visual and/or spatio-temporal vibrotactile feedback to facilitate their maze navigation. All participants were simply asked to progress through each maze as far as they could. In the first three sessions participants experienced feedback as per the group's condition. Each condition featured a unique combination of visual and/or haptic feedback (detailed below). In the fourth and final session, participants across all groups had to navigate the maze relying on haptic feedback only; direct visual feedback on their position in the maze was removed completely. Accurate and effective interpretation of the haptic signal was essential to successful and effective maze traversal. Participants were not informed that this final session would involve maze navigation using haptic feedback only, to minimise the confounding effects of participants prioritising learning the haptic signal over performing the task. Before I describe the four study conditions, I provide an overview of the apparatus used as part of the study and how haptic and visual feedback was provided during maze navigation.

6.3.2 Study Apparatus

During each study session, participants were presented with a series of simple, procedurally-generated mazes (see Fig. 6.1) to navigate using an off-the-shelf official Xbox One game controller. Mazes were presented on a 24” monitor. Below I describe the maze design, how maze navigation was supported through the game controller and how the maze topology was represented in the form of visual and/or haptic feedback.

Maze Design & Character Navigation Each maze consisted of a single vertical path of infinite length. Maze paths were procedurally generated by carving a series of square, passable path blocks through a grid of impassable wall blocks. More specifically, two path blocks are placed on top of each other vertically, followed by arranging one to four blocks horizontally to the right or left of the previous blocks, and repeating (see Fig. 6.1 for two example mazes generated in this way). This enables the automatic generation of different mazes that still bear similarities in terms of topography and difficulty.

In all study conditions and sessions, participants were asked to navigate a space-ship-like digital character forward through these mazes (see Fig. 6.4.b). Rotation of the digital character in 2D space was supported through the game controller’s primary analogue joy stick. The other buttons of the controller supported moving the character forward in the direction it was facing at a rate of 1 unit per second. Each wall block and corridor had side length 1.5 units. The rotation angle of the character was limited to a range between -90 to +90 degrees (9pm to 3pm on a circular clock). This enforced a forward movement within the maze, and prevented participants, especially in the haptic-only condition (see below), from navigating backwards by accident. When attempting to continue rotating while at either limit, visual (the walls turning red) or haptic feedback (low frequency two-hand rumble, distinct from the trigger motor feedback) would be provided at a saturation (visual) or intensity (haptic) proportional to the rate at which they were trying to turn

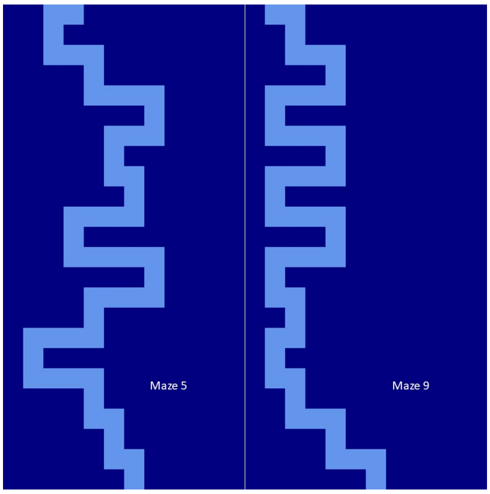


Figure 6.1: Two examples of the first 50 blocks of the infinitely-long procedurally-generated mazes used as part of the present study: Dark blue areas represent impassable walls; light blue areas represent passable corridors.

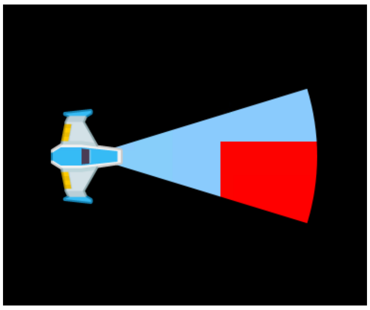


Figure 6.2: Illustration of the graphical feedback presented at the digital character rotation limit of +90 degrees when the control stick remains pushed fully to the right.

(see Fig. 6.2).

Figure 6.3: Microsoft Xbox One wireless controller, showing position of embedded trigger (1,2) and palm (3,4) ERM motors, and button layout.

Game Controller I decided to use an Xbox One controller for this study as it is a low-cost, off-the-shelf device that enables versatile vibrotactile feedback on a continuous scale. The four embedded Eccentric Rotating Mass (ERM) motors can be independently controlled using the Windows.Gaming.Input namespace in the Universal Windows Platform API[1]. Two large ERM motors are fixed to the main controller body and located roughly within each palm grip (see Fig 6.3.3&4). While these motors operate over different frequency ranges, the resulting vibration when activated is experienced similarly in both hands and, thus, differentiation of signals is difficult. However, two smaller ERM motors are located in each trigger button of the controller (see Fig. 6.3.1&2). In terms of the emitted vibrotactile signal, these are well-isolated from each other, and their signal can be easily distinguished and associated with the left and right hands. Their signal also differs from the low-frequency vibration of the largest motor. I therefore decided to use these two trigger motors to produce continuous vibrotactile signals in response to participants' navigation of the digital character in the maze (e.g., direction and proximity to walls, described in detail below), while using the largest palm ERM motor to provide feedback if participants over-stretched the rotation angle of the digital character.

As vibrotactile intensity is typically perceived non-linearly [79, 130], the trigger motors required calibration to support a linear mapping of vibrotactile signal to target information (in this case, the character's distance to a wall). To achieve a linear, symmetrical feedback perceived through left and right controller triggers (see Fig. 6.3.1&2), three calibration steps were taken:

1) Each trigger motor's vibrational intensity profile in response to commanded API values was mapped using a microphone in contact with the trigger. The motors were found to be slightly asymmetrical.

2) The range of output signal intensities was restricted to those that both motors were capable of outputting.

3) The API values were parameterised, and re-mapped to a perceptually linear scale according to the Weber-Fechner law [57], using a literature sourced estimated just-noticeable-distance of 12% [42, 58]).

Following this calibration, pilot testing with 10 participants who performed a magnitude estimation and an intensity matching task confirmed that the adjusted left and right trigger motor outputs were perceived linearly and symmetrically in response to linear input commands.

Design of Visual & Haptic Feedback In the conditions involving *visual feedback*, the digital character was visible to participants (see Fig. 6.6.b–d). In addition to the digital character's position in the maze, participants would get visual feedback on walls in close proximity via a visual cone that indicated the peripheral vision angle and viewing distance of the character (see

---

[1] Windows UWP applications API, Windows.Gaming.Input Namespace

Fig. 6.4.b and 6.6.b–d). This cone's angle was 54 degrees in total (see Fig. 6.4.b). The length of the visual cone (viewing distance) varied across conditions where visual feedback was provided (see description of conditions below).

Figure 6.5 shows two examples of the visual character navigating the maze in the visual conditions: walls are invisible until they come into view through the digital character's visual cone. The figure illustrates the character approaching a corner in the maze, as indicated through the dark blue space intersecting with the character's visual cone.

In the conditions involving the *haptic feedback* mechanism, continuous vibrotactile signals were emitted by the trigger buttons to provide information about distance and direction of walls. This haptic feedback was implemented using two "whiskers" which extend from the character at the same angle as the edges of the visual cone in visual conditions (see Fig. 6.4.a) and a length of 1 unit. The individual intensity of the left and right trigger motors was determined by the distance to the nearest maze wall within range of the whisker. A whisker touching a wall at their outer periphery resulted in a minimum vibration, while walls touching the character itself were considered a collision, and produced the maximum vibration. If there was no wall within reach of the whisker, the trigger motors did not vibrate.

The maze dimensions were designed so that the character's visual cone and/or whiskers would not touch the maze's walls when centred in a corridor (the default, starting state in each task). Participants would therefore only see or feel a wall if moving forward or turning in a way that would bring them off the maze path centre. That is, one navigation strategy could involve simply minimising vibration, though the optimal trajectory (and therefore the greatest progress) is achieved by allowing some proximity to walls and corners.

In all conditions, additional visual feedback showed the number of corridor blocks that had been passed, indicating progress through the maze (white text in each of Fig. 6.6). This visual

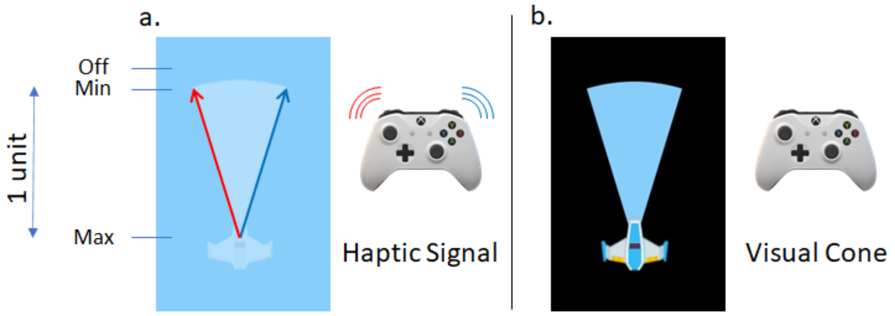


Figure 6.4: Haptic signal implementation. Haptic signal (a.) allows independent detection of left (red) and right (blue) walls at a distance identical to full visual cone (b.)

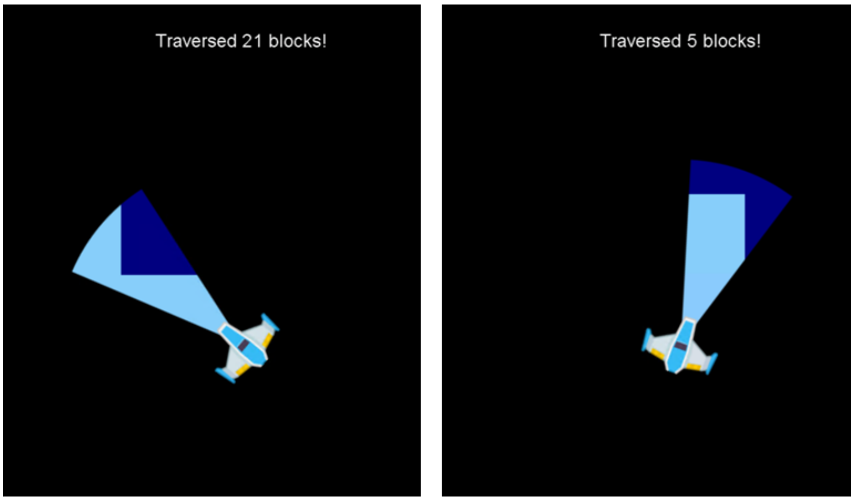


Figure 6.5: Examples of maze corners coming into view in the visual condition.

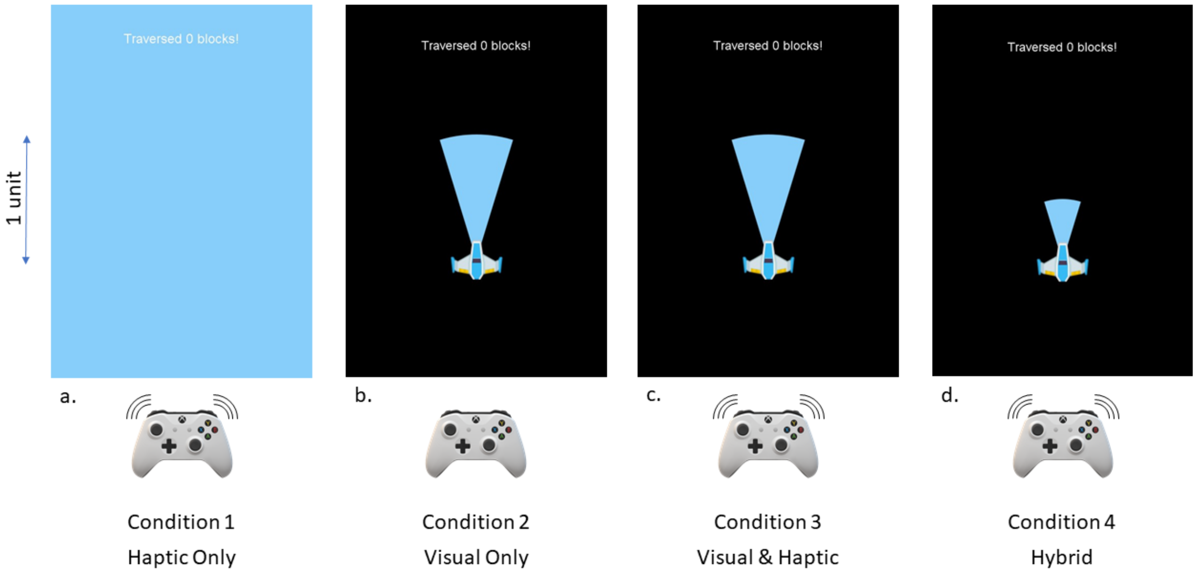


Figure 6.6: Illustration of feedback provided as part of the four conditions.

information constituted the only form of explicit feedback in the study.

6.3.3 Study Conditions & Hypotheses

The four visual/haptic feedback conditions I tested in this study are designed to resemble some of the familiarisation and training mechanisms utilised in the vibrotactile SSAD navigation studies [62, 93, 96, 106, 124] detailed in the related work section. Figure 6.6 shows an illustration of the four designed study conditions. I describe each condition and its role in the study below.

Condition 1: "Haptic Only" Participants in this group received feedback on their position in the maze via haptic signal only. Neither the maze walls nor the digital character was visible in this condition. In order to progress through the maze, participants had to learn how to interpret the haptic feedback by focusing on the changes of vibrotactile signals received through the game controller, in response to their control inputs. The closer participants navigated the digital character toward a wall, the stronger the vibrotactile signal emitted by the games controller. As described above, the signal included directional information: Figure 6.4 shows the haptic whiskers used to map wall proximity to the intensity of vibrotactile feedback experienced on the controller's trigger motors. The visual 'sensory deprivation' in this condition resembles the blindfolded training conditions commonly utilised in several other navigation-focused haptic SSAD studies [32, 33, 62, 106, 107]. Similarly, parallels may be drawn between this condition and the utilisation of a novel SSAD by sensorily impaired individuals, whose sensory apparatus may offer little to no overlap with the information presented by the device. Device signals must be learned through task performance and careful attention to explicit feedback.

Condition 2: "Visual Only" In this condition, participants saw the virtual character's position and orientation in the maze (see Fig. 6.6.b) as well as its viewing cone, in which any maze walls present are rendered. The viewing cone was 1 unit long. Haptic feedback was not provided. This visual feedback conveys more information than the haptic feedback in Condition 1: It shows the rotation of the digital character and allows some appreciation for maze corner topography. After navigating the maze under this condition, participants will begin the final session naïve to the meaning of the haptic signal. It is possible that the visual skills gained in the first three sessions may allow participants to tactilely recognise features in the maze in the final session. Such 'cross-modal skills transfer' in the context of SSADs has been explored across several modalities (see Deroy & Auvray for an overview [47] or Levy-Tzedek for an applied example [102]). Alternatively, a cognitive understanding of the maze structures, controls and shape of the visual cone may allow for fast uptake of the novel haptic signal during navigation.

Condition 3: "Full Visual & Haptic Feedback" Participants in this condition were provided with haptic and visual feedback simultaneously as they navigated through the maze (see Fig. 6.6.c). Visual and haptic feedback in this condition combines the feedback mechanisms provided in Conditions 1 and 2: there is a direct correlation between the visual and haptic feedback. As participants approach a wall in the maze, they see the wall coming closer, while experiencing a stronger haptic signal in the controller. Both feedback mechanisms therefore work independently to facilitate participants' navigation through the maze, but, in contrast to Conditions 1 and 2, this condition allows participants to learn how to interpret the haptic signal, based on the visual feedback they receive simultaneously. The effectiveness of such

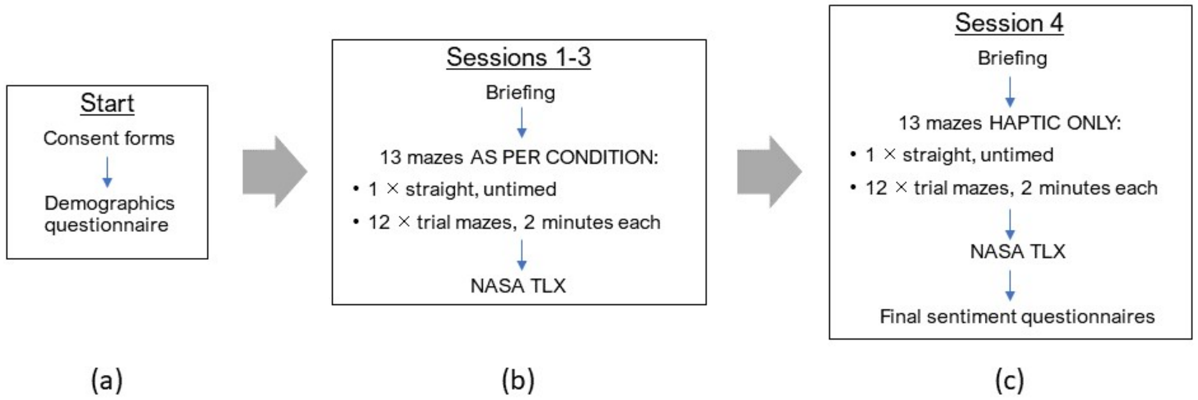


Figure 6.7: Flow diagram showing experimental progression across sessions

'sensory concurrency' in learning new feedback signals has been studied in comparison to sensory deprivation [107] and active or user-controlled feedback [78]. This sensory combination is often described as a 'familiarisation' mechanism in short-term SSAD studies [18, 32].

Condition 4: "Visual & Haptic Hybrid" In this condition, the length of the cone of vision was reduced to 0.5 units, providing 50% of the viewing distance compared to Conditions 2 and 3 (see Fig. 6.6.d). That is, participants could not see as far in the maze as in the two previous conditions. At the same time, participants received the same haptic feedback as in Conditions 1 and 3. Condition 4, thus, provided reduced visual feedback, but information on the maze topology in relation to the digital character was still available through haptic feedback. While participants did not need to pay attention to the haptic feedback to navigate the maze, correct interpretation of the haptic signal while navigating the maze provided an advantage. This Hybrid condition was designed to combine and optimise the beneficial effects of sensory concurrency and sensory deprivation, and was therefore hypothesised to enable most effective learning of the introduced signal (as measured by performance metrics).

6.3.4 Participants & Study Procedure

24 participants (5 males, 19 females) were recruited to take part in the study. Participants' ages ranged between 18 and 31 years ($M_{\text{age}} = 21.4$ years, $SD = 3.5$ years). Participants all reported normal or corrected vision. To minimise potential confounding effects of prior experience with navigation tasks in video games and interaction with the physical game controller itself, we only recruited participants who self-identified as having "little to no video gaming experience". Participants were provided with a £20 book voucher for a local bookstore to compensate them for their time.

Participants were randomly divided into four groups—one for each condition, resulting in a 4x6 between-subjects design (four study conditions as described above; six participants per condition group). A between-subject experimental design was chosen to eliminate learning effects, which were anticipated to be significant. Each participant attended four separate subsequent study sessions that were spread out over a maximum of five days with a maximum of one session per day. This was to ensure that participants were able to sleep after each session—an important component when it comes to the acquisition of procedural task knowledge [96] which the maze navigation task falls under. The gap between each session was never less than 18 hours and no more than 48 hours.

The number of participants and conditions was chosen as a reasonable compromise between allowing the 4-day study to be conducted within a feasible timeframe, and allowing adequate data collection for statistical analysis of repeated measures.

As described above, participants attended four study sessions in total. During the first three sessions they navigated the infinite mazes, making use of the feedback depending on the condition they were assigned to. As part of the fourth and final session, participants across all condition groups were asked to navigate the mazes with only haptic feedback only (identical to Condition 1. See Fig. 6.7 for an illustration of study sessions).

Before beginning the first session (Fig. 6.7.a), participants signed the consent form and completed an initial questionnaire to determine basic demographic information (age, reported gender, handedness and any visual impairments). Following this, the first session started (Fig. 6.7.b).

All sessions began by verbally briefing participants that they would be performing a maze navigation task and that the objective was simply to make as much progress through each maze as possible. Participants were ask to hold the game controller so that their index fingers made contact with the trigger buttons. They were then given time to explore an introductory training maze consisting of a straight corridor of nine blocks to familiarise themselves with the game controller and to ensure that they had understood the task. Participants were asked to navigate to the end of this maze in their own time, to give them the opportunity to explore the functionality of the game controller. The feedback condition during this practice task matched the condition group and session to which participants were assigned.

Participants wore headphones playing brown noise during the experiment to mask the sound of the motors.

They were then given 12 mazes to navigate in a session. 48 mazes were procedurally generated prior to the study as described above; no maze was the same but their level of difficulty was comparable. All participants across conditions and study sessions navigated the same mazes in the same order; 12 mazes per study session. Participants were given two minutes to make as much progress through each maze as they could. Participants were given a 20-second break between each maze. After each study session, participants were asked to complete a NASA TLX survey [71] to coarsely capture their perceived efforts, including mental load, temporal load, effort, frustration and a self-estimate of success. The same procedure was repeated across all four study sessions.

In the final Session 4 (Fig. 6.7.c), participants navigated mazes using haptic feedback only. At the end of the session, participants were asked to complete an additional questionnaire (detailed below) designed to capture their overall experience of the navigation task, considering the two different feedback conditions they had experienced.

6.3.5 Data Collection & Analysis

As part of this mixed-method study, a range of quantitative and qualitative data were collected to assess participants' maze navigation performance. This data focused on their measured performance, their navigation strategies across the four study sessions, and the qualitative experience and perceived difficulty of the tasks.

Quantitative Data. In terms of performance measures, for each maze navigation (trial), the number of blocks participants managed to traverse within two minutes was recorded. We also measured the number of times they collided with a maze wall. This quantitative data was analysed using standard statistical methods (see more details in the results section).

As described above, we also collected data about participants' perceived performance and task effort via a NASA TLX questionnaire and custom questionnaires that focused more on participant sentiment, attention, and intuition during the task through Likert scales (analysed using standard statistical methods, presented with participant responses in Fig. 6.17) and open-ended questions, which contributed to the qualitative data collected.

Qualitative Data. In order to investigate participants' navigation strategies, we computationally logged participants' interactions with the digital character for every trial, so we could accurately replicate and analyse participants' traces through each maze. These navigation traces were analysed visually and qualitatively coded with regard to visual patterns that would provide insights about participants' navigation strategies. After the final session, participants were invited to answer open-ended questions with additional prompt for commentary, to supplement this analysis: *"What did you find most enjoyable about the final session (if you found it enjoyable at all). What did you find the most frustrating about the final session? Any final/further comments,* *thoughts or ideas you would like to communicate?".*

There were no issues, errors or missing values detected during data collection and analysis. Trial 0 (the straight corridor training maze) at the start of each session has been excluded from further analysis. All presented results are based on the analysis of this quantitative and qualitative data. The following section describes our findings.

6.4 Results

The study results are presented in three parts. First I report on the findings regarding participants' performance in the different conditions derived from the quantitative data collected during the study (number of blocks traversed, collisions and a calculated combination of the two). This is followed by an outline of qualitative findings on maze navigation strategies, in relation to these performance measures. Finally, I report on participants' perceived experience of effort, enjoyability and frustration during the maze navigation under the different conditions, based on my analysis of the post-session surveys and post-study questionnaires.

A summary of participant quantitative data across all sessions is presented for completeness in Figure 6.20. Data collected during the first three sessions cannot be directly compared with final session data (outlined green) due to differences in task conditions, and will not be further analyzed.

6.4.1 Effects of Training Modalities on Performance

The following analysis of participants' performance with regard to the maze navigation included three aspects. (1) A quantitative analysis of the number of blocks traversed during the provided two minutes per trial in Session 4, (2) an analysis of the number of wall collisions in the context of the number of blocks traversed, and (3) patterns that emerge from these analyses in terms of potential maze navigation strategies.

Number of Blocks Traversed In order to investigate whether the combination of visual and/or haptic feedback presented during the first three sessions has learning effects when navigating the maze using only haptic feedback, we first analyzed participants' maze navigation performance during Session 4, comparing the number of maze blocks traversed.

Using SPSS, we performed a Generalized Linear Mixed Model (GLMM) [120], in similar fashion to Rossi [144]. This model retains information on trial repetitions and participant-specific variability when estimating marginal means for each condition. We performed this as an

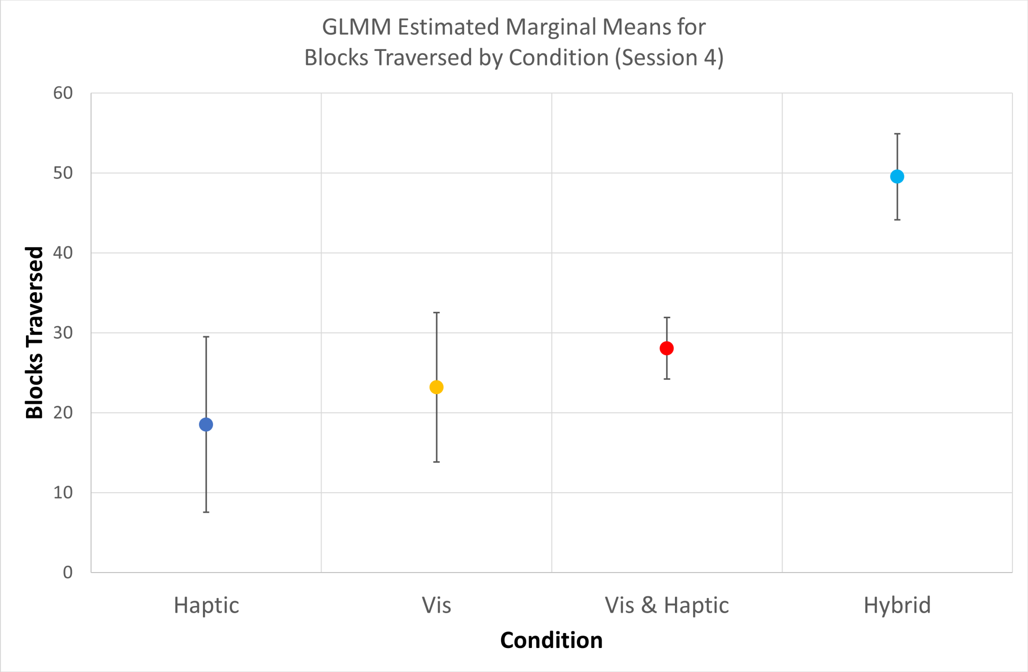


Figure 6.8: Estimated Marginal Means and 95% Confidence Intervals for Number of Blocks Traversed in Session 4. Calculated by GLMM on Blocks with Condition as a fixed factor, and trial nested within participant as a random factor.

analysis of the blocks traversed, with condition as a fixed factor and trial nested as a repeated measurement under participant as a random factor. This test showed a significant effect of condition on blocks traversed (F(3,15) = 19.394, $p < 0.001$).

Estimated marginal means and associated confidence intervals from the model are displayed in Figure 6.8. Estimated population means for the number of blocks traversed were 18.5 (Haptic Only), 23.2 (Visual Only), 28.1 (Visual & Haptic) and 49.0 (Hybrid). Pairwise comparisons (after Bonferroni correction) showed that the Hybrid condition blocks traversed was statistically significantly higher than the other three conditions, but there were no further significant differences.

Overall, the results described therefore suggest that the conditions have a significant effect on the number of blocks traversed in the final session, perhaps surprising given the limited number of participants in the study. The Hybrid condition enabled the most blocks to be traversed. The Visual & Haptic (full sensory concurrency), the Visual (skills transfer), and the Haptic (sensory deprivation) resulted in fewer blocks traversed, though the differences between these conditions were not significant.

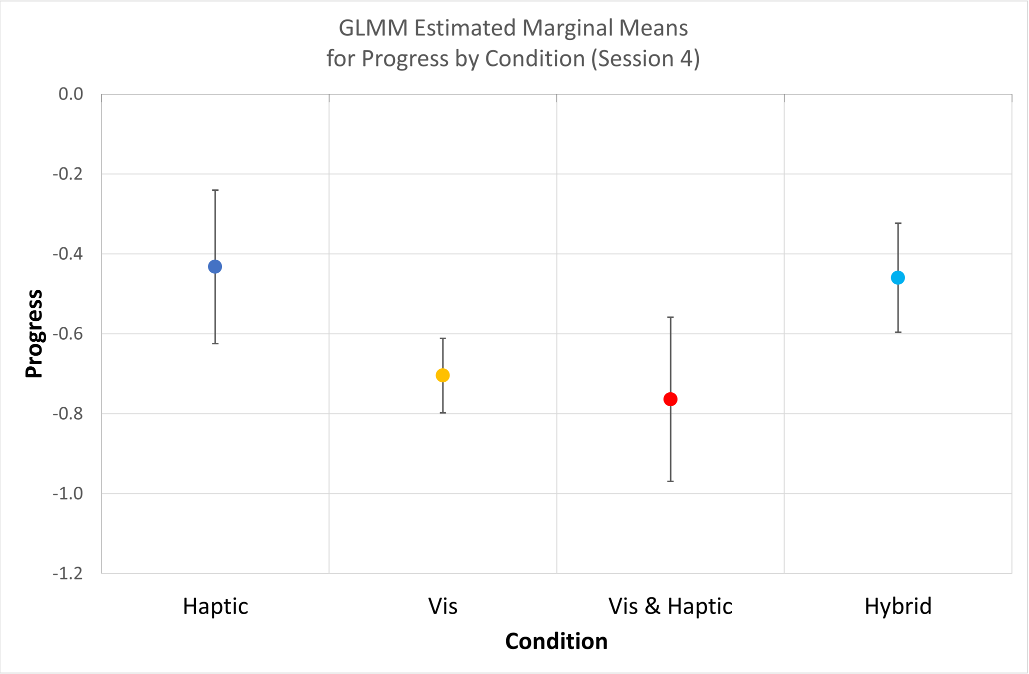


Figure 6.9: Estimated Marginal Means and 95% Confidence Intervals for Progress in Session 4. Calculated by GLMM on Progress with Condition as a fixed factor, and trial nested within participant as a random factor. Progress under the Hybrid condition was found to be statistically significantly higher than the Visual condition.

Number of Wall Collisions The number of traversed blocks alone (as described above) is not sufficient to describe participants' maze navigation. For example, some participants may have traversed only a few maze blocks within the given time because they were careful to avoid wall collisions, while others may have traversed few blocks as they were unsure how to progress. Linear regression analysis between blocks and collisions (grouping by condition) showed positive correlations for the haptic ($R2 = 0.39$, $p < 0.01$), visual ($R2 = 0.60$, $p < 0.01$), and visual & haptic ($R2 = 0.65$, $p < 0.01$) conditions, but a non-significant correlation for the hybrid condition. This confirms that typically, participants who travelled further through the maze on average collided with more walls. As such, no direct quantitative tests were performed on the number of wall collisions alone.

Overall Progress Through the Maze Similar to Froese [62], a "progress score" was calculated for each trial by dividing the number of collisions by the number of blocks traversed. We then inverted this such that a higher score reflects better performance:

$$progress = -collisions/blocks$$

A progress score of 0 indicates perfect performance (no collisions) while progress of less than -1 indicates that participants hit more walls than they traversed blocks.

A GLMM analysis of progress with condition as a fixed factor and trial nested as a repeated measurement under participant as a random factor reached significance (F(3,22) = 5.114, $p < 0.01$). Estimated marginal means from the model are visualised in Figure 6.9. Pairwise comparisons (after Bonferroni correction for multiple comparisons) showed that the Hybrid condition (-0.432) was superior to the Visual condition (-0.704), but that no other pairwise comparisons were significant. Again, the small number of participants in the study must be considered.

Individual Participants' Performances Across Conditions Quantitative analysis of performance in maze navigation can only provide a high-level, aggregated picture of differences in navigation by condition. This limitation encouraged an analysis of individual participants' performances in Session 4. Figure 6.10 shows a scatterplot where the y-axis represents "progress" and the x-axis represents "number of blocks traversed". Circles represent individual participants' trials in Session 4 with colour indicating the conditions to which they were assigned (blue = Haptic, orange = Vis, red = Vis & Haptic, turquoise = Hybrid). The four different quadrants loosely point to different maze navigation strategies in terms of performance:

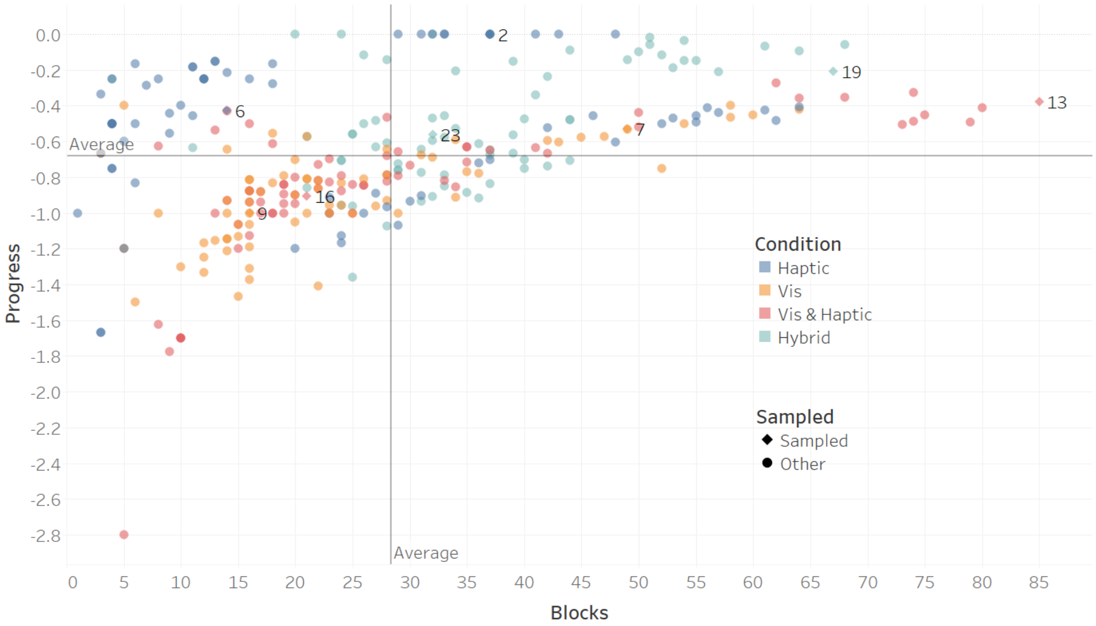


Figure 6.10: Session 4 Progress vs. Blocks. This scatterplot shows the number of blocks traversed (x-axis) vs. the amount of progress made (y-axis) in Session 4 (all trials navigated in Session 4, using haptic feedback only). "Progress" is calculated as $-collisions/blocks$. Each mark represents a single maze navigation (trial), coloured by condition.

- Quadrant 1 (bottom left): Few blocks, low progress scores (i.e., high number of wall collisions).
- Quadrant 2 (bottom right): Traversed many blocks, but low progress as they collided with walls often. Could be characterised as reckless—navigation was fast, rather than careful.
- Quadrant 3 (top left): Traversed few blocks, but received a high progress score—very few wall collisions. Likely cautious when navigating the maze. More focused on staying on the path than quick maze navigation.
- Quadrant 4 (top right): Traversed many blocks and also received a high progress score. These trials demonstrate efficient navigation strategies.

These quadrants loosely reflect the quantitative differences between conditions. For example, participants in the Haptic condition predominantly fall into Quadrant 3 (top-left; blue circles) representing relatively low scores in term of blocks traversed, but with good progress scores. Participants from the Hybrid condition are largely present in Quadrant 4 (top right; turquoise circles) representing good performance scores. Participants in the Visual (orange) and Visual & Haptic (red) conditions largely fall into Quadrant 1 (bottom left).

The scatterplot also indicates a wide spread of performances with several outliers where some participants outperformed the average of their condition group. For example, participant P2 (Haptic condition), compared to other participants in this group, performed extremely well; their trials can be found in Quadrant 4 (top right). Similarly, the trials by P7 (Visual condition) and P13 (Vis & Haptic)—all in Quadrant 4 (top right)—suggest a higher performance compared to the those trials by fellow participants in the respective group conditions. These findings led us to conduct a qualitative analysis of selected participants' navigation strategies as described below.

6.4.2 Maze Navigation Strategies

The following qualitative analysis of maze navigation strategies focuses on individual participants who seemed to have outperformed fellow group members in their respective conditions. These participants were identified through the scatterplot as shown in Figure 6.10. I also identified and selected sample participants from each condition that seemed to represent more typical performance score as reflected by the quadrants of the scatterplot. I then compiled the final maze traces of these identified participants and conducted a qualitative visual analysis on the first few blocks to identify different maze navigation strategies, in terms of utilisation of the haptic feedback. I compare the traces across conditions to see whether the conditions corresponded with particular methods for signal interpretation and maze navigation strategy.

Note that in the presented maze trace visualisations (Figs. 6.11 to 6.16) only the first ~20 blocks of each (infinite) maze are shown for conciseness.

6.4.2.1 Correlation of Traces with Performance Metrics

Here, I compare the traces in Figs. 6.11, 6.12, 6.13 and , 6.14 with their coordinates on the scatter chart 6.10 in order to evaluate the effectiveness of the blocks traversed and progress metrics in reflecting maze traversal competency, thus clarifying the importance of the statistical analysis.

Participants 2, 7, 13, 19 and 23 traversed more than the average number of blocks, suggesting that all conditions enable participants to develop the skills required for fast maze traversal. However, the differences between the estimated marginal means across conditions suggests that the condition had an effect on how quickly participants were able to traverse the maze, with participants in the Hybrid condition traversing fastest. Block count decreased across the Vis&Haptic, the Visual Only, and the Haptic Only conditions respectively.

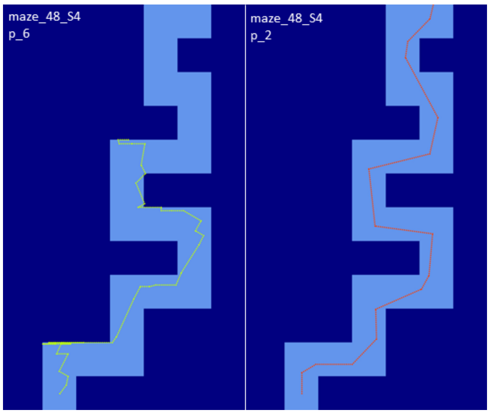
The participants with the lowest progress scores (participants 9 and 16) display a non-dynamic, repeating or mechanical behaviour in Session 4. Conversely, the other traces show signs of deliberate action, indicating that the progress score is a useful metric for deliberate haptic signal utilisation. In combination with the statistically significant differences in progress scores across conditions derived in section 6.4.1, we can conclude that the Hybrid and Haptic Only conditions were more likely to teach participants to proceed cautiously, while the Visual condition encouraged more reckless traversal of the maze in the final session. The Vis&Haptic condition resulted in the lowest estimated progress score, but was not statistically significant to other conditions due to wide participant variance, suggesting that the training condition was highly subjective.

As mentioned above, Quadrant 3 6.10 predominantly featured participants from the Haptic condition, three of whom never traversed more than 20 blocks in any session, despite consistently achieving a progress score >-0.5. This would suggest that the Haptic condition, while effective at teaching participants to avoid walls and navigate the maze deliberately, did not guarantee that participants would become proficient at traversing blocks. This behaviour is reflected in the final trace of P6 (Fig.6.11).

6.4.2.2 Individual Strategy Shifts

Most participants' data were clustered in small areas for each session, reflecting a degree of consistency in strategy. However, some participants display step changes or a wide distribution of performance metrics throughout the final session. In endeavouring to understand how behaviour

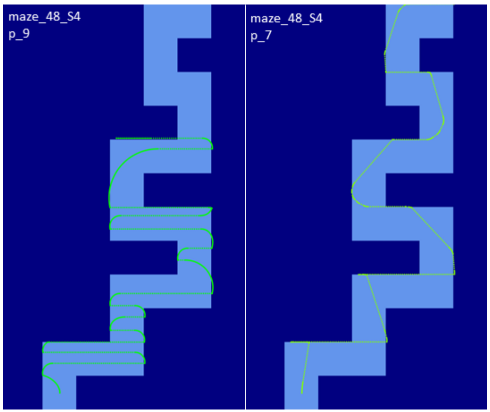
Haptic Only



	P6 (average)	P2 (competent)
Blocks | progress	14 | -0.43	37 | 0.0
Characteristics	Short movements, sharp turns, hitting several walls, retracing steps	Longer movements before deliberate, decisive turns. No collisions
Interpretation	Complexity suggests some uncertainty of the controls or feedback	Central-corridor traversal reflects simple control approach based on minimising vibrations

Figure 6.11: Comparison of P6 and P2 Session 4, trial 12 traces - Haptic Only

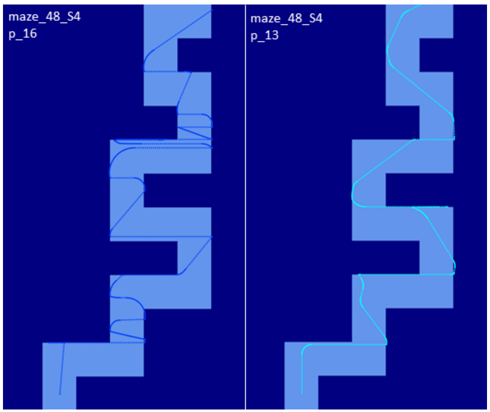
Visual Only



	P9 (average)	P7 (competent)
Blocks | progress	16 | -1.0	49 | -0.53
Characteristics	Repetitive, curved pattern	Some back-tracking, long straight lines
Interpretation	Participant consistently moving forward in the maze, strategically turning the control stick when haptic feedback stops changing	Deliberate and dynamic (though not optimal) response to haptic signal

Figure 6.12: Comparison of P9 and P7 Session 4, trial 12 traces - Visual Only

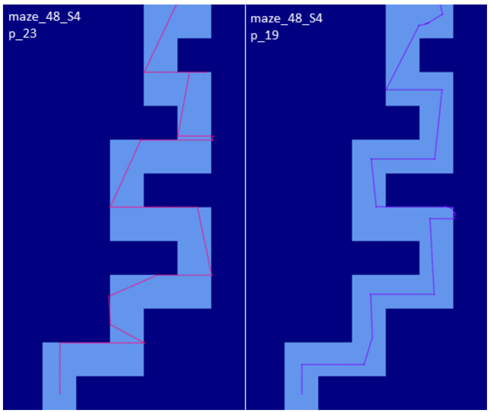
Visual & Haptic



	P16 (average)	P13 (competent)
Blocks | progress	21 | -0.905	85 | -0.38
Characteristics	Inconsistent, dynamic trace with multiple wall collisions	Deliberate turning angles
Interpretation	Some effort to interpret the signal, yet unable to make effective decisions	Strategy similar to P7 (see Fig. 6.12)

Figure 6.13: Comparison of P16 and P13 Session 4, trial 12 traces - Visual & Haptic

Hybrid



	P23 (average)	P19 (competent)
Blocks | progress	32 | -0.56	67 | -0.21
Characteristics	Series of long, straight movements	Efficient, cautious behaviour
Interpretation	Inefficient course. Waited for collision before turning, indicates methodological or mechanical aspect to control	Corridor-centred trace quite similar to P2 (see Fig. 6.11)

Figure 6.14: Comparison of P23 and P19 Session 4, trial 12 traces - Hybrid

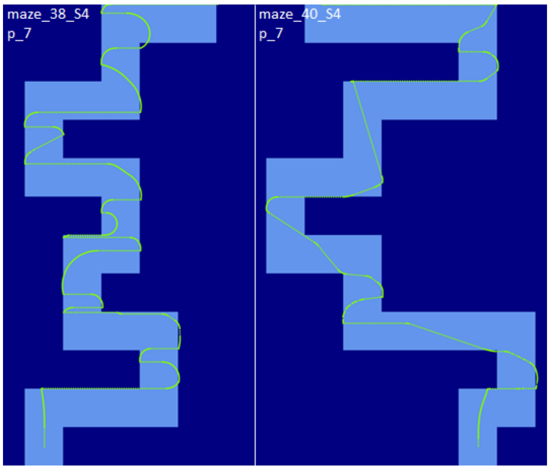


Figure 6.15: Participant 7's traces through mazes 38 and 40 (in session 4).

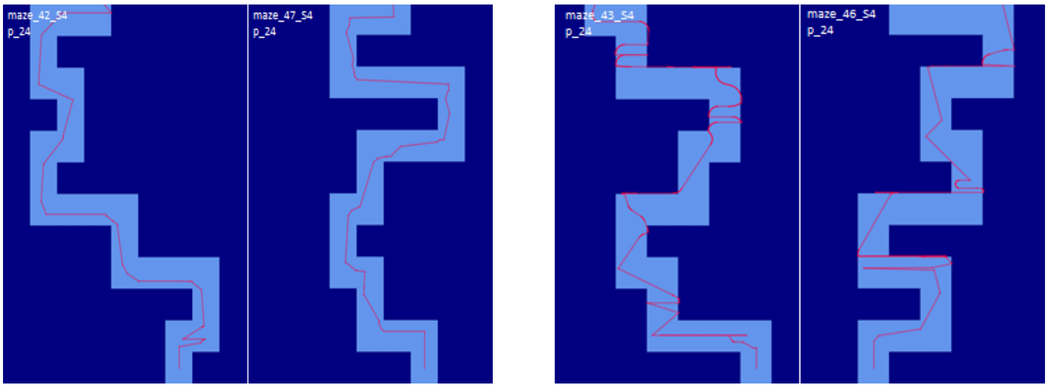
and strategy are linked to the chosen performance metrics, it is important to know whether such performance differences may be explained through changes in strategy, and whether such strategies could be identified and differentiated through analysis of the maze traces. Trial-by-trial examination of the data identified two participants with anomalous distributions of blocks vs progress in session 4:

P7 exhibited significant differences in performance across session 4, achieving 28 blocks at progress -0.93 in trial 2 (maze 38), but 64 blocks at progress -0.43 in Trial 4 (maze 40) as shown in Fig. 6.15. The large differences in block and progress values are confirmed by clear differences in strategy. Maze 38 (Fig. 6.15.left) shows a signature curvy side to side pattern, produced when systematically alternating between turning left and then right while holding the advance button. By contrast, Maze 40 (Fig. 6.15.right) shows more deliberate choice of trajectory and results in a much higher block traversal rate and progress score.

P24's session 4 marks were clustered in two distinct bands, one at Progress -0.6 and one >-0.1, both with traversed block counts of 20-40 blocks. Fig. 6.16.left shows the traces with progress>-0.1, reflecting deliberate and nuanced decision making and good wall avoidance typical of participants in the Hybrid condition. However, Fig. 6.16.right shows more erratic movement and less regard for wall collisions.

6.4.2.3 Implications of Quantitative Metrics and Maze Traces

Addressing the performance metrics and maze traces, we can conclude that the Hybrid condition enabled quick yet deliberate maze navigation, the Visual and Vis&Haptic conditions promoted moderate speed but more reckless traversal, and the Haptic Only condition encouraged deliberate performance of the task but did not guarantee quick maze traversal.



(a) progress > -0.1 **(b)** progress -0.6

Figure 6.16: Shows Participant 24's high progress traces through mazes 42 and 47 (left) and low progress traces through mazes 43 and 46 (right). Two distinct strategies are employed, resulting in very different progress scores despite similar block count.

6.4.3 Likert Responses

At the end of the study, participants were invited to complete a questionnaire designed to capture sentiment and understanding of the game task. The questions and participant responses are visualised in Fig. 6.17.

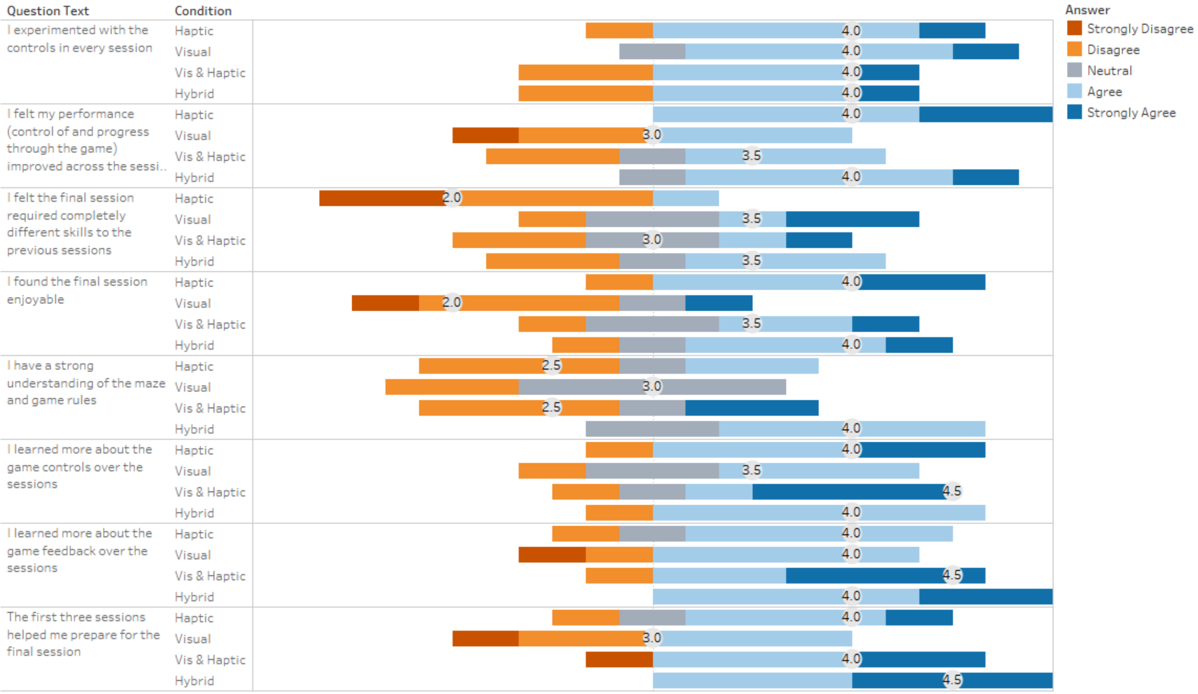


Figure 6.17: Shows distribution of responses on diverging stacked bar chart, and overall median response (grey circle).

A Kruskal-Wallis test comparing the Likert responses across conditions indicated that the only question showing significant difference between conditions was “I felt my performance improved across the sessions”. The sample size is small (6 data points per condition) so achieving statistical significance is generally not viable. However, overall trends in sentiment correlate with the quantitative results, so are included here for completeness.

Figure 6.18: NASA TLX results showing proportion of participants from each condition responding in each binned measurement group. Greater areas of red show increased load in the corresponding index.

The greatest discrepancies grouped by condition were in response to the following statements:

"I felt my performance improved across the sessions": Some participants in the Visual and Visual&Haptic conditions disagreed with this statement. Participants in the Haptic and Hybrid conditions either agreed or were neutral.

"I found the final session enjoyable": Participants in the Visual condition seemed not to enjoy the final session. Participants in the other categories were largely neutral or positive about the final session.

"I have a strong understanding of the maze and game rules": Participants in the Hybrid condition felt on average most confident in their understanding of the game rules and environment, though two participants in the Visual&Haptic condition strongly agreed with the statement.

"The first three sessions helped me prepare for the final session": The Visual condition left participants feeling that the first three sessions were not helpful in preparing for the final session. All participants in the Hybrid condition agreed or strongly agreed that the first three sessions were helpful.

6.4.4 NASA TLX

The results from the NASA TLX scale were binned into seven equally distributed intervals from "Very High" to "Very Low" and the results for each index (split by condition) are presented in Figure 6.18. Similarly, summing scores across participants for each category in each session allowed a comparison of overall scores as seen in Fig. 6.19. As expected, scores for the Haptic

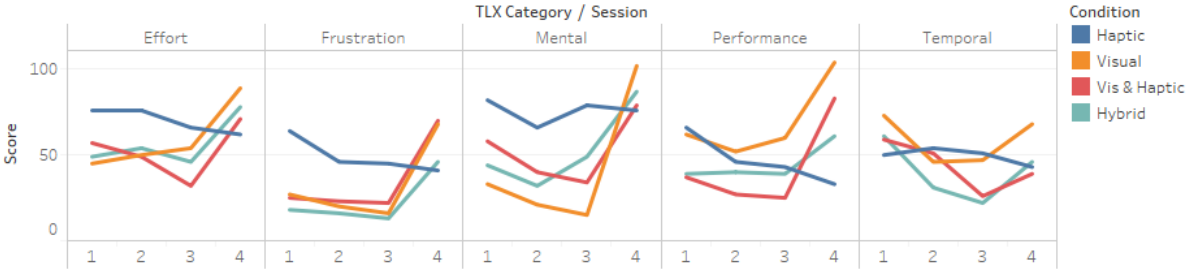


Figure 6.19: Sum of NASA TLX scores across all participants, split by session, condition and category.

condition decreased gradually across the four sessions. Scores in all categories for all other conditions increased from Session 3 to Session 4, though as with the Likert questionnaire results, the small sample size of TLX data rends achieving statistical significance unfeasible, so no quantitative tests were performed.

6.4.5 Open Questionnaire Responses

Through the questionnaire presented at the end of the study, participants were invited to comment on their experience during the final session, some insights of which are summarised here:

“What did you find most enjoyable about the final session (if you found it enjoyable at all)? What did you find the most frustrating about the final session?” From the *Haptic* condition, P2 described the final session as boring and repetitive, while P1, P3 & P5 specifically reported enjoying the exploration and discovery process as enjoyable. P1, P4 & P6 reported frustration from uncertainty of the controls, and P5 & P3 from being rushed. In the *Visual* condition, P10, P11 & P12 described the exploration and novelty of the haptic signal as the most enjoyable aspect, while all participants reported that the lack of progress and not understanding the haptic feedback was frustrating. P13, P14, P17 & P18 in the *Visual & Haptic* condition reported the challenge and attention required in the final session as being the most enjoyable aspect. P13, P14 & P17 said that getting stuck was the most frustrating part of the session, while the other three reported frustration at not being able to see the vision cone. Participants in the *Hybrid* condition reported seeing progress (P20 & P21), engaging in skilful action (P22 & P23) and trying to beat the scores they had achieved in previous sessions (P19) as the most enjoyable parts of the final session. P19, P20, P22 &P24 reported being frustrated by ‘getting stuck’, while P21 suggested that losing the visual cue was most frustrating.

6.5 Discussion

The experimental conditions and study design aimed to address two research questions. (1) Does the nature of sensory concurrency (sensory deprivation, cross-modal skills transfer, or full concurrent feedback) presented to participants in the first three study sessions affect performance, strategy and user experience when the visual feedback is removed (Session 4)? If so, how? (2) In light of our hypothesis, what potential do “hybrid” mechanisms, conveying valuable haptic information while retaining confirmatory visual feedback, have in comparison to single modalities?

I now discuss the findings in the light of these research questions, with reference to the literature.

6.5.1 Maze Navigation Strategies Using Haptic Feedback

As outlined in the previous section, not all participants were able to demonstrate effective maze navigation strategies during the haptic-only Session 4. However, several participants traversed significantly more blocks than others, colliding with few walls (Fig.6.14.Right) or none at all (Fig.6.11.Right), demonstrating that deliberate, effective navigation of the virtual maze using haptic feedback alone is indeed possible.

6.5.1.1 Sensory Concurrency and Maze Navigation

Bertram’s review [19] identified ‘training/session duration’, ‘feedback’ and ‘application similarity’ as key influences in learning to use new SSAD signals. By controlling for training session/duration and application similarity, we isolate feedback as the key variable in the comparison of final session haptic-only performance, and thus through demonstration of significant differences in block traversal and progress, conclude that the information available in the visual channel during the first three sessions has a profound effect on haptic-only maze navigation.

However, contrary to Lobo’s findings [107], where Lobo observed participants performed better following training with a blindfold on than without, here it is observed that full sensory deprivation (condition 1, Haptic Only) may be a less effective learning environment than providing concurrent visual information (condition 3, Visual & Haptic), or even skills transfer (condition 2, Visual Only). The primary differences between the presented experiment and Lobo’s are that in our experiment, participants in the haptic-only condition (visual sensory deprivation) were not at any point provided with an explanation or visualisation of the task, and all other participants were not informed about the haptic-only final session. In Lobo’s pretest-practice-posttest design, participants were aware of both task context, and the upcoming haptic-only testing session. One explanation for the difference in observed effects may be task comprehension— without the additional context of the visual feedback, participants in our haptic-only condition seemed to find learning difficult, consistently reporting high mental load in the Nasa TLX. This difficulty is apparent in the data visualised in Figure 6.20: Participants in the Haptic Only condition continued to improve slowly well into Session 4, while nearly all other conditions' participants had reached a performance plateau in earlier sessions.

Figure 6.20 also reveals that the performance and improvement rate of participants remains remarkably continuous from the end of one session and the start of the next, despite gaps of 18-48 hours between. This observation is consistent with the acquisition of procedural knowledge, as it suggests participants' performance does not degrade rapidly upon resting, unlike declarative knowledge through recall tasks [161]. Such a multi-session training approach seems not to be detrimental to rate of skill acquisition, while avoiding potential issues concerning concentration or fatigue which might arise from attempting the study in a single 4-hour session.

Participants in the Visual Only condition were able to traverse more blocks in Session 4 than those in the Haptic Only condition, suggesting that visualisation of the task contributes significantly to comprehension. Nevertheless, this result is somewhat surprising, as participants in the visual-only condition had no prior exposure to the haptic signal until they encountered it in Session 4. The degree of skill shown during the final session by participants in the Visual Only condition demonstrates the expected competence of untrained participants performing a familiar task using an unfamiliar SSAD. Participants in the Visual condition performed weakly in the haptic-only task (relative to those in the Hybrid and Visual & Haptic conditions); so while it is reasonable that asking users to utilise a novel SSAD to perform a familiar task could give a useful indication of the intuitiveness of the device, we can be confident that this is not an effective mechanism for enabling or demonstrating the long term potential or effects of proficient use of the SSAD.

Participants who performed the first three sessions under the Visual Only condition in general did not enjoy the final session, also reporting consistently high Nasa TLX scores across load index categories (Fig. 6.18). However of the other participants, as two thirds said they enjoyed the final session, this suggests that such a control scheme or mechanism may be of interest in future gaming or entertainment applications. As participants in the Hybrid condition unanimously agreed or strongly agreed that "The first three sessions helped me prepare for the final session", this mechanism could also be a viable method for introducing to novel haptic control systems in such applications.

6.5.1.2 Hybrid Sensory Concurrency

As concluded in Section 6.4, the Hybrid condition was the most successful condition at preparing users for quick maze traversal, as indicated by the overall number of blocks traversed. Furthermore, it was on par with the Haptic Only condition (and perhaps the Visual&Haptic condition) in enabling and encouraging deliberate use of the haptic signal as indicated by the high overall progress scores. This correlates with the response that all six participants in the Hybrid condition agreed or strongly agreed that “The first three sessions helped me prepare for the final session” (Fig. 6.17). These results suggest that the Hybrid case was indeed effective as a learning environment: to answer the original hypothesis (Section 6.1), there may be an optimal level of sensory concurrency which enables sensory integration at a rate and depth superior to both sensory deprivation and full sensory concurrency.

6.5.2 Feedback Mechanisms in SSAD Studies

In parallel with Lobo [107], I remark that the herein findings align with the Guidance Hypothesis [78], which suggests that the more learners rely on some type of feedback during practice, the more they come to depend on that feedback. By reducing the information in the visual channel as applied in the “Hybrid” condition of this study during Sessions 1 to 3, participants may have been encouraged and incentivised to pay attention to and learn the haptic signal, which seems to have helped them in Session 4 in the haptic-only setting.

Similarly, studies by Huet [78] have found user-controlled confirmatory feedback (where users are able to request easily comprehensible task feedback at will) to be more effective than scheduled, incomplete or ambiguous feedback. Based on the study findings, I argue that by retaining an element of the visual, the Hybrid condition provides confirmatory feedback through sensory concurrency which seems to be beneficial for learning how to interpret haptic feedback. These findings may suggest that the Hybrid approach combines aspects of both user-controlled confirmatory feedback and the Guidance Hypothesis; dissuading users from developing a disadvantageous reliance on the visual channel during the task, while providing easily comprehensible feedback when users approach walls.

The simple nature of the novel hybrid approach has some further advantages over user-controlled confirmatory feedback. Firstly, providing a minimalist yet familiar sensory signal alongside a useful new sensory augmentation signal allows the augmentation to be learned without impairment of usual task performance. Further, being a passive condition which does not require users to actively request additional information, it is simpler both in terms of implementation and utilisation. Incidentally, such a hybridised concurrency may be introduced to users in stages:

future work may consider the gradual reduction of visual signal in order to guide users towards greater dependence on the haptic signal.

Given the advantages that task context brings to the integration of the haptic signal, it is conceivable that further explicit feedback and guidance on the nature and design of the haptic signal, even a simple explanation of how the whiskers were implemented, may further enhance the effectiveness of the Hybrid approach to sensory feedback. Equally, it is arguable that explicit feedback such as the number of blocks traversed may not be required at all in real-world goal-driven tasks with user-defined metrics of success (such as safe, efficient driving or remote communication).

6.5.3 Application of the Hybrid Learning Condition in Sensorimotor Studies

Conditions comparable to the Visual & Haptic condition of our study are often utilised in SSAD studies through a 'familiarisation' process as described in section 6.2.2. In short term studies, participants are typically briefly exposed to the novel SSAD signal with full sensory concurrency (uninhibited visual and proprioceptive cues) before performing the task using the novel signal only (e.g. [170]). Similarly, both long-term FeelSpace belt studies [96, 124] have utilised this mechanism: Participants were asked to wear the FeelSpace belt during waking hours for several weeks and received belt information alongside visual, auditory and proprioceptive navigational cues, before the degree of signal understanding was tested. However, as in our Visual & Haptic condition, the haptic cues in the short term and FeelSpace studies are largely superfluous to the spatiotemporal information still available in visual (and proprioceptive) cues, which may limit development of proficiency as it did in the present study. Implementing the principles of the Hybrid condition in such studies may encourage more effective signal utilisation, improving study reliability and revealing longer-term potential of the SSADs.

The Hybrid condition was designed under the principle that the novel modality must form a key part in decision making: providing highly task-relevant information unavailable in other senses, while retaining confirmatory information in existing senses. These principles may be applied to SSAD studies in several ways; for example, by partially depriving some sensory cues such that the novel SSAD signal is more relevant, or by designing training tasks which may be completed without the SSAD, but where an understanding of the haptic signal provides an advantage.

Revisiting Krüger's Lateral Line study [93], note that such Hybrid principles are present by nature of the system design, potentially explaining why successful signal utilisation was demonstrated by participants despite the short study duration and lack of dedicated training: At will, users of

Kruger's system are able to refer to visual feedback present in the rear-view mirrors, receiving information concurrent with the novel haptic feedback. While the haptic feedback is not necessary for task completion, users able to utilise it gain a decisive advantage, thus the user self-guides towards proficiency in comprehending the novel signal.

6.5.4 Limitations and Open Questions

As with many studies in the field of SSADs, the number of participants recruited for investigation limits the conclusions that can be drawn from observed phenomena, particularly limiting the power of statistical tests. Applying a rigorous methodology and gathering multiple data points per participant go some way to combating this, as does conducting the study over a medium term of 4 days. While the consideration of participant as a random variable in the analysis methodology offers robustness against this, expanding on this study with a larger sample size may assist in clarifying the contributions of individual differences to the effectiveness of each condition.

In this study, participants were kept naïve to the haptic-only nature of the final session, in order to reduce the confounding effect of directed focus on the haptic signal. Given the aforementioned contradictory findings of this study against Lobo's [106], addressing the role of deliberated attention on the learning process in future work may assist in explaining why this is the case.

Finally, I argue that the Xbox One controller itself has also proven a suitable tool for the research of vibrotactile signals, and similar game or VR controllers (such as the Oculus Rift Touch [131] or HTC Vive Cosmos [77]) may also be useful off-the-shelf for exploration of vibrotactile navigation systems and similar SSADs.

6.6 Conclusions

In the present study, I have shown that the nature of haptic and visual sensory concurrency presented to users during a virtual navigation task has a strong effect on how users perform when the visual element is removed. Users initially presented with haptic feedback alongside a limited visual signal respond and perform better than users initially presented with haptic feedback only, visual feedback only, or concurrent haptic and visual feedback.

In enabling users from all conditions to navigate the virtual maze using haptic feedback alone, the work has also demonstrated the effective use of the Microsoft Xbox One controller as a simple Sensory Substitution and Augmentation Device, and detailed the process utilised for its preparation as an intuitive and versatile interface for such tasks.

While there is still much work to be done, practical guidance can be taken from this study.

Whether or not the results are definitive and highly generalisable, it is clear that available visual feedback plays a role during device learning, and that learning continues over several days for many of the conditions.

The main contribution here is of practical relevance, and points towards methods and mechanisms for effective introduction of SSADs or other sensorimotor interfaces, longer-term studies, and the effectiveness of hybrid metrics and assessment of acquired procedural knowledge.

More subtle in terms of contribution, but important in context of this thesis, is the demonstrated effectiveness of the model described in Chapter 4. The insights derived from this model have in this chapter been reinforced experimentally, demonstrating its suitability for the design of effective, focused learning conditions for introduction of sensorimotor interfaces to new users. Such insights are difficult to derive from cognitive approaches to cognition, and reinforce the value and importance of applying appropriate cognitive models (and, specifically sensorimotor theory) to enactive interaction design.

Having now demonstrated the calibration and testing of a sensorimotor interface, and experimentally tested some of the hypotheses derived from the model of the sensorimotor interface, I now move on to the final chapter of this thesis, discussing the impact of this work and the future of sensorimotor interactions in HCI.

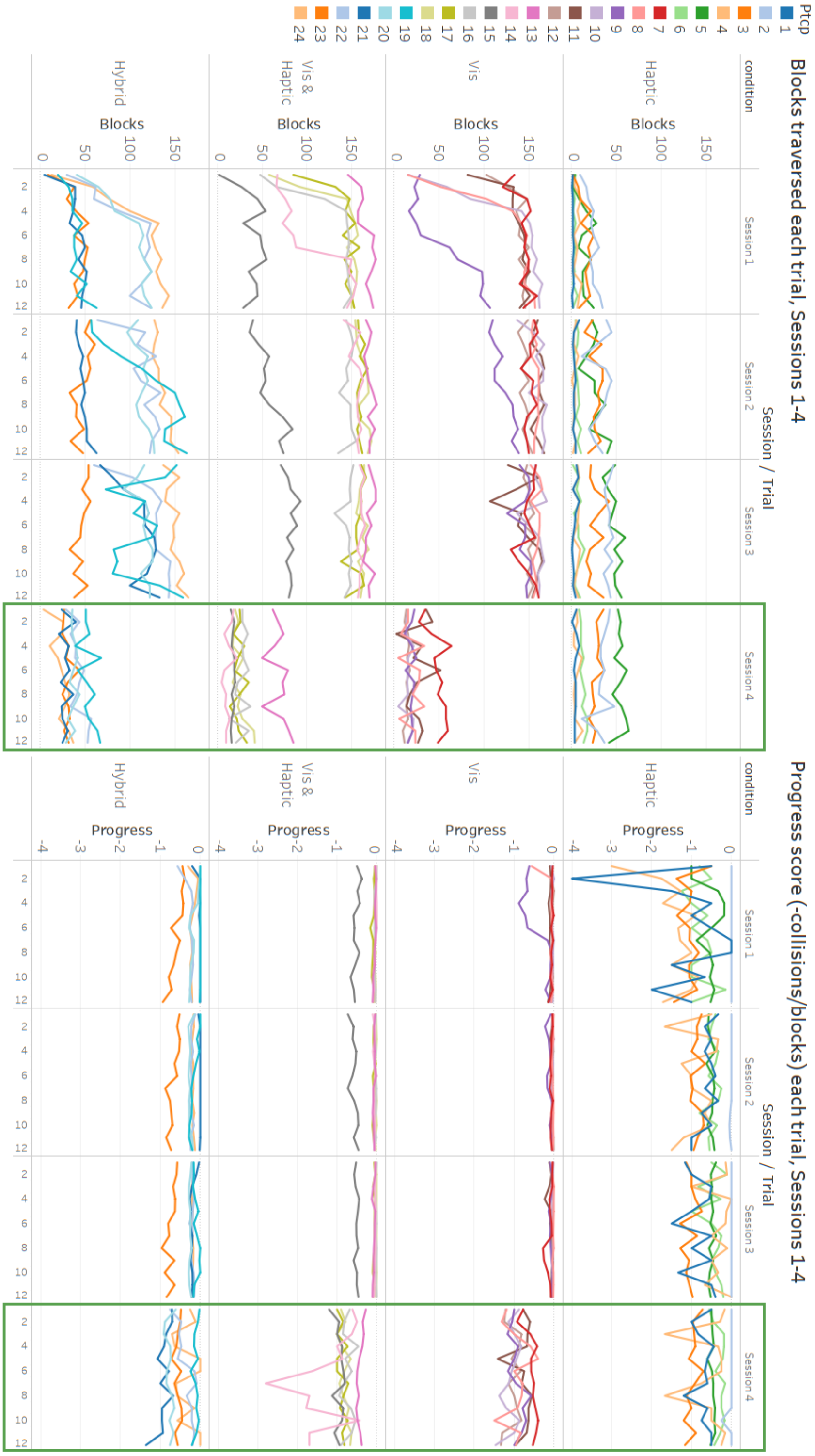


Figure 6.20: Chart showing each participant's number of blocks traversed (left) and progress score (right) during each trial, across all four sessions. Data from sessions 1-3 not used in further analysis as conditions render metrics incomparable. In session 4 (highlighted green), all participants performed the task under identical conditions (haptic feedback only), allowing comparative analysis.

CHAPTER SEVEN

INSIGHTS ON SENSORIMOTOR INTERFACES - DISCUSSION AND FUTURE WORK

To conclude this thesis, I re-visit and reflect on the insights and contributions of my PhD research, alongside limitations and questions my work raises to be explored in the future.

7.1 Introduction

The work in this thesis originated from a desire to create novel interfaces which draw on the principles of SSADs to augment perceptual experience and allow human-computer interaction in a natural, intuitive fashion. My initial goal was to build and/or modify a portable vibrotactile device for communication of task-specific information across a wide range of applications. Such a device would support communication of a variety of input data types, allowing extension of human consciousness and sensorimotor skills into the digital realm. The prospect of such enactive engagement is explored through a ‘day in the life’ story, featured in Appendix A, where I describe a flexible haptic glove capable of communicating a variety of contextually-relevant spatiotemporal signals through discreet and diverse vibrotactile patterns on the back of the hand.

It was initially my impression that there would be consistent (if not already well-established) principles and approaches to building, testing and developing devices in the existing literature, which I could draw on to guide the creation of encodings, the communication of sensory information and the learning and training of participants on interaction principles. To some degree, I assumed that these devices could simply be built and evaluated in line with the literature base. However, it quickly became apparent that there are major gaps in the literature on SSADs and related enactive and sensorimotor interfaces. Some of these gaps were as elementary as their definition and the principles of interaction, but some revealed more complicated limitations, regarding the optimisation of signal design and the provision of a suitable training and testing environments for device evaluation. The work presented in this thesis, can therefore be considered an exploration of the overlap between what I knew, as an HCI researcher, and where I wanted to be as a designer and creator of systems that augment human perception through non-invasive, comfortable and intuitive interfaces.

This PhD research sought to understand and define the principles of what makes an SSAD feasible (and effective), and to extract these principles for application in generalised interaction techniques. In starting this work, it became clear that there is little consensus (though many opinions) on exactly what an SSAD is. Thus, the factors which they have in common (i.e. how certain SSADs are related) are unclear, manifesting as loose functional definitions on their assistive capacities [163] or alluding to ‘sensory’ access to otherwise inaccessible environmental information [112], neither of which satisfy scientific scrutiny with application and engineering in mind. An exploration into the specific subset of devices which have promise for giving rise to externalisation (the phenomenon by which the device or tool becomes invisible during use, briefly introduced in Section 2.5.2) led to the identification of enactive interaction principles as the common factor to be understood and applied. At the centre of the enactive school of cognition (see Fig. 3.1) lies the rejection of representationalist (symbolic) signals, and the idea of action being critical to mastery. Both of these stances reflect properties of the SSADs which have exhibited the potential for externalisation [100].

While devices are being built in line with the presented principles in order to allow for dynamic, enactive interactions with the goal of augmenting perception, labelling remains inconsistent, and the associated terminology has not found its way into mainstream HCI. When building devices with the goal of enabling new interactive techniques, the device or application is understandably of greater concern than the psychological or perceptual insights gained through its evaluation. As such, detailing the factors important to the fabrication and presentation of information in such interfaces is of secondary importance, leaving a gap between the disciplines of psychology and HCI, which is where this thesis is situated.

Addressing sensorimotor interfaces from an enactive theoretical foundation (rather than the modification and derivation from SSADs) clarifies their potential role in HCI—sensorimotor interfaces are systems designed to improve task performance, not through the presentation of new information for cognitive decision-making, but through the augmentation of perception, and the revelation of new affordances.

To this end, there is a clear distinction between the literature review of devices and research exploring enactivity in interface design, through SSADs and associated history (Chapter 2) as a survey of the current state-of-the-art, and the introduction to Sensorimotor Theory and related cognitive models as a foundation to ground future research directions (Chapter 3).

At this point, rather than further complicating the field through offering yet another definition for the SSAD, it seemed more appropriate to create a fresh, clearly-derived term for future consideration. This labelling is a term of convenience used to describe the herein discussed devices and interaction techniques, and should not be mistaken for a contribution to the philosophy of Sensorimotor Theory or enactivism itself.

Enactivity already principally exists within HCI, too. However, searching through most computer science-related publishing sources returns mixed results representing robotics, interaction design, cognitive modelling, social engagement, teaching and learning. This may indicate the ambiguity surrounding the terminology, which is utilised in different ways at interactive, cognitive and social levels. Through this thesis I offer the formalisation and consolidation of the principles of the enactivist psychological school of cognition, specifically for the design of exciting new interfaces.

7.2 Chapter Recap and Overview

Recapping the chapters of this thesis, my research into enactivity in HCI constitutes the following:

Chapter 2 comprises a literature review exploring the history of enactivity in HCI through the history and description of SSADs and their influence on enactive interaction design. The need for enactive models of cognition is outlined.

Chapter 3 introduces and summarises Sensorimotor Theory as an enactive model of cognition, describing its components and utility in the field of enactive interface design.

Chapter 4 synthesises the fields of HCI and Sensorimotor Theory to derive the schematic of the sensorimotor interface (Fig. 4.1). Principles and paradigms of design, interaction and learning are proposed and summarised in the design document (Fig. 4.3).

Chapter 5 outlines the importance of calibration of vibrotactile systems for use in enactive applications, from theoretical and experimental perspectives. The Xbox One controller is utilised as an exemplar and prepared for future experimentation.

Chapter 6 reflects the published work [31] undertaken to explore the effects of sensory concurrency on the learning of a haptic spatiotemporal signal. The study directly addresses and confirms learning mechanisms and predictions of direct learning theory, and reinforces the validity and robustness of the model and wider theoretical contributions of Chapter 4, and characterisation/calibration work described in Chapter 5.

I now specifically re-address the research contributions of the thesis, and how the theoretical and experimental exploration detailed thus far has clarified the approach to device and application design, improved accessibility of the field, and opened new pathways and paradigms of interface design.

7.3 Research Contributions

This section revisits the research questions posed in Chapter 1 in detail, and addresses them in light of the contributions made.

7.3.1 How can we define and characterise enactive interfaces?

In this thesis, I have focused specifically on identifying and defining the enactive techniques (utilised by some SSADs) that enable task improvement through perceptual augmentation: devices which give rise to externalisation and the attribution of sensory stimuli to objects 'out there' at a perceptual level. Many SSADs exemplify these principles, though other definitions cloud the field, ranging from the functional claims of sensory assistance to simple stipulation of which sense or modality a stimulus is presented to. In order to cleanly work towards a coherent understanding of enactivity's relevance to HCI, in my work I adopt Sensorimotor Theory (SMT) [133] as a base line cognitive model, due to its association with SSADs and the emphasis it places on enactivity at the foundations of perception. By extracting the key components of this model and extending them beyond natural perception to enactive interfaces, I thus defined the sensorimotor interface: *"Sensorimotor interfaces are devices which represent and apply the principles of sensorimotor theory in interaction design"*. The schematic and associated descriptions further outline its construction, specification and the principles of its application as an interaction technique. As SMT is an enactive theory of mind, providing such clear labelling therefore contributes directly to the definition and description of enactive interfaces.

The key theoretical contribution of this work lies in the explanation and review efforts leading to this definition, as described in Chapter 4. As discussed, alongside the framework presented and the introduction to Sensorimotor Theory outlined in this thesis, the definition is generally applicable, and serves as a firm definition for a range of enactive interfaces gaining traction in HCI. However, I wish to highlight further impact and implications of accepting and adopting this definition and framework.

Impact. The framework presented has proven consistent with many observations, such as Auvray & Myin who argue that: *"perception with SSDs occurs neither in the substituting nor in the substituted modality [6, p1052]"*.

I agree, that such enactive engagements using such devices result in the formation of new sensorimotor contingencies, each one of which give rise to new perceptual experiences. By this, the oft-posed philosophical question which asks: 'In which modality is the substituted sense experienced?' can be answered—it is a new modality. This is further useful in that such insight can be used to *define* what a modality is: the unique perceptual experience arising through the exercise of a sensorimotor contingency. Thus, sensorimotor interfaces are systems designed to create entirely new modalities, offering associated benefits in revealing previously inaccessible environmental affordances.

While the sensorimotor interface shares much in common with definitions of the SSAD (such as Visell's 'perceptual claims' [163]) or enactive interfaces (as defined by Froese [62]), I argue that the detailed derivation through SMT and its direct applicability to interface design beyond assistive technologies justifies the new labelling. The wider impact and further advantages of this specific derivation and labelling are further explored.

While similar in derivation and nature, there is nuance in the differentiation between my definition of the *sensorimotor* interface and Froese's *enactive* interface (*"an enactive interface is a technological interface that is designed for the purpose of augmented sense-making [62, p368]"*). While the enactive interface refers generally to the manner in which the interface is used (that is to say, enactively), the sensorimotor interface defined here appeals specifically and directly to O'Regan and Noë's Sensorimotor Theory (SMT) as a framework for its construction and application [133]. SMT is an enactive model of cognition and is therefore entirely compatible with enactivist principles. However, the detailed description of the sensorimotor interface from a perspective of interaction without specific reference to biology also allows for generalisation which more conventional accounts of enactivism may not. For example, the concept of ‘embodiment’ in enactivism typically reflects biological presence as a requirement for agency. However, as Sensorimotor Theory does not describe the motor or sensory biology of the acting agent, it allows for SMCs to be entirely virtual, and overall more abstract than those under the physical limitations imposed by enactive frameworks. The effectiveness of this abstraction is exemplified by my work described in Chapter 6, which demonstrates the potential development of SMCs in a completely virtual, abstracted environment. Abstracting this application further (e.g., by replacing the controller by a brain-computer interface) may test the limits of enactivism, due to the lack of biological motor movement in acting upon the ‘world’. However, such activities are entirely compatible with sensorimotor approaches, provided there remain lawful SMCs which can be learned through exploration. Essentially, the sensorimotor interface and associated principles promises creation of new affordances which do not even have real world equivalents—a complete extension of perceptual experience into the digital realm.

Limitations & Open Questions. The advantages of applying non-representationalist theories of cognition to the definition and specification of enactive interfaces are numerous. It could be argued that the primary benefit to the adoption is that critical observations surrounding the importance of active exploration and user control in learning SSAD signals [5] are inherently accounted for in a robust and consistent fashion. Indeed, if we are hoping to build devices that enable new perceptual experiences through interactivity, then starting from a theory of mind which claims that natural perception arises through enactive engagement seems an appropriate place to start. However, there are many related non-representationalist philosophies of mind which fulfil the requirement of holding action and perception as ‘two sides of the same coin’, such as environmental psychology (explored by Lobo through literature survey [105] and experimentation [108, 125]). Yet, SMT lures and convinces through the simplicity of its claims, its direct applicability (leading to the formation of my taxonomy and model), and its proven history in the testing and development of SSADs through the feelSpace experiments [85, 96, 124]. Furthermore, as demonstrated through the use of direct learning theory, combining elements from the range of post-cognitivist theories (the 4E approaches [75]) offers ample material for experimentation and application.

The 4E approaches share many principles in their rejection of the representationalist perspective, but also contribute individually and uniquely to models of cognition. While I have shown that the principles of enactivity may be quite effectively applied to interface design, there remains much room for an exploration of the principles of *embodiment, embedded* and *extended* cognition to the creation of engaging new interaction techniques.

7.3.2 How can we construct, apply and evaluate enactive interfaces in interface design?

Beyond the theoretical value of defining and characterising enactive principles for interaction design, the model of the sensorimotor interface introduced and described in Chapter 4 is detailed enough to guide construction and comparison of associated devices.

The enactive philosophy driving the construction and evaluation of sensorimotor interfaces reflects the key insight derived through Chapters 2 to 4, best summarised with reference to affordances: Most devices in 'cognitive' HCI present affordances, communicating information that is somehow represented or symbolised through signal output. When we evaluate these devices, we are interested in measuring how well these affordances are communicated, which is connected with how readily the information is available and how comprehensible it is. By contrast, with sensorimotor interfaces, the goal is to communicate new environmental affordances. For example, much as a child may be able to identify an object, they may not truly perceive that it has the affordance of 'carryable' until they learn their own strength, and have interacted with the object in a relevant capacity. With this knowledge it is clear how environmental affordances, and, consequently, the perception of objects and features in terms of available actions, may change through mastery of new sensorimotor skills.

The results of the experimental work presented in Chapters 5 and 6 demonstrates the ease of application (and importance of following) theoretically-derived frameworks for both device construction and learning.

In summary, the practical construction of suitable sensorimotor devices, as demonstrated by the numerous devices outlined in Tables 4.1 & 4.2, along with the appropriation of the off-the-shelf Xbox controller described in Chapter 5 illustrates the simplicity of meeting the hardware requirements of an enactive interface. However, through this thesis I have showed that the hardware is only a small component of enactive interface implementation. As such, noting the paradigms of device construction ('Make it *valuable, practical* and *sensory*' - see Section 4.2.3), introducing systems through principles derived from direct learning theory (*'intention, attention, calibration'* - see Section 4.2.8) and following the design document guidelines (Fig. 4.3, shown in Section 4.3 and re-visited below in Section 7.3.4) outline the steps required for interface conformation to sensorimotor and wider enactive principles.

Regarding evaluation, in Chapter 5, the metrics utilised in the evaluation of the Xbox One controller are in-line with cognitivist principles valuing the communication of explicit values in human-computer interfaces, and highlight the unsuitability of continuous encodings for this communication (in accordance with Novich [130]). However, in the context of this thesis, it should now be clear that such metrics provide little insight into the suitability of the device itself for sensorimotor interactions. The conclusions therefore also serve to highlight the importance of carefully-designed interactivity, implicit feedback, application choice, and observation of behaviour as providing stronger insight into the strengths and limitations of a sensorimotor interface. This is further reinforced through the wide range of quantitative measurements and qualitative evaluations conducted as described in Chapter 6, where the guidance derived from the model of the sensorimotor interface reinforces the importance of analysing behaviour, performance and sentiment in evaluating sensorimotor skill mastery.

The learning principles derived from direct learning theory (Section 4.2.8), and their demonstration in Chapter 6 also contribute to the comparative evaluation of SSADs and sensorimotor interfaces under lab conditions. With well-trained participants, greater assurance can be that behavioural results reflect device potential and the utility of the presented environmental information, rather than the initial intuitiveness of the device, or worse, the cognitive use of the presented signal to achieve task goals in the lab. The timescales demonstrated in the study outlined in Chapter 6 reinforce the importance of longer-term (multiple sessions over more than one day) use of enactive devices. With specific reference to Figure 6.20, learning continued well into the third session for many participants across all conditions. Ensuring sufficient learning time under carefully-developed training conditions is a critical part of the fair comparison of enactive and sensorimotor interfaces.

To summarise, the evaluation process of sensorimotor interfaces should be centred around forming an understanding of how users perceive the environment and affordances *through* the device, and how this changes as mastery is acquired, rather than on their ability to memorise, recognise or interpret specific patterns. While such design [62] and evaluation methods [84] have been proposed before, this work further reinforces the importance of such an approach, not least in improving comparability between studies (a well-noted issue [19, 84]), but further in improving on objective utility of research outcomes.

Impact. While the thorough derivation of principles for sensorimotor interfaces from theory renders my model robust, my research efforts described in Chapters 5 & 6 to test its implications reinforce its utility. The findings presented in these chapters indicate that the model is suitable for the design of flexible interfaces with perception-altering potential. Furthermore, the application of direct learning theory to training conditions opens up new avenues for self-guided training in both sensorimotor interfaces and SSADs.

My experimental work also highlights the importance of measured calibration of (vibrotactile) sensorimotor interfaces, highlighting the importance of considering existing SMCs and associated perceptual biases when designing new interfaces.

Limitations & Open Questions. The key limitations to the experimental work in this thesis reflect its generalisability and robustness.

Regarding the work outlined in Chapter 5, the demonstration of the characterisation and calibration of vibrotactile system is useful as the vibrotactile stimulus (haptic) represents a large portion of sensorimotor interfaces (see Figures. 4.1 & 4.2). However, the work can only indicate principles that are applicable across other modalities: the results should not be implicitly generalised to all modalities without further experimentation. By contrast, the implications of this process do highlight new opportunities for research: as *existing* SMCs are shown to have an impact on the perception of stimuli designated for *new* SMCs, this points towards longer-term future work towards understanding the limitations of brain plasticity in fully mastering new sensorimotor contingencies.

Regarding Chapter 6, in an effort to create highly-controlled conditions in an accessible, practical environment, the maze navigation tasks are somewhat abstract, to some degree limiting generalisability to wider sensorimotor tasks. This is a typical trade-off; the reliability of lab conditions against the practicality and applicability of more contextual experimentation. Given the fledgling status of theoretically-driven comparative experimentation in enactivity and sensorimotor interfaces, the decision to lean towards gathering robust evidence from controlled studies is well-justified.

While the controlled, comparative methodology and the analysis through generalised linear mixed modelling yields statistically sound evidence for the effectiveness of the hybridised condition, it cannot be denied that participant numbers are limited. This is especially obvious when reflecting on the NASA TLX and Likert scale questionnaire results, which are not statistically significant, only warranting attention as they align with the hypotheses and predictions of DLT and wider sensorimotor learning principles.

Further experimentation on several fronts would be valuable. A longer term, closer look at how these skills are learned, across studies with more participants, with different modalities and potentially in more generalised environments, may be fruitful. Consideration for closer measurement of *improvement* during training under varying conditions may also yield interesting results, and perhaps allow observation and highlight of individual participant differences in the learning of new sensorimotor skills.

7.3.3 How can we improve the effectiveness of learning to use enactive interfaces?

In keeping consistent with enactive psychology and Sensorimotor Theory, I have summarised and applied the principles of direct learning theory (discussed in Chapters 4 & 6) to show how exercise of sensorimotor skills in a goal-oriented task is fundamental to the mastery of sensorimotor interfaces.

Direct learning theory introduces the ideas of *intention*, *attention* and *calibration* as components of self-guided learning of sensorimotor skills, and the general consideration and realisation of these components formed the basis for the experimental design of the investigative studies conducted in Chapter 6. These learning components were carefully balanced through provision of easily-interpreted sensory information in a goal-oriented environment which depended on the formation of new sensorimotor skills. Carefully-controlled and designed sensory concurrency was shown to have a significant effect on the effectiveness of a fixed period of training time on the integration of new sensorimotor contingencies, from a quantitative, behavioural and experiential perspective. This suggests that control of such variables may lead to promising results in the design of self-guided learning environments for sensorimotor interfaces.

Notably, the absence of expert or explicit feedback in the experiments represents a paradigm-shifting approach to the idea of what constitutes ‘training’ of such devices. Clearly, we need to provide enough task information for user actions to be goal-oriented (rather than randomly explorative), yet there is no need to provide explicit task performance and dictate user strategy at a granular level.

Impact. As discussed in the introduction to Chapter 6, approaches to studying the learning of enactive and sensorimotor interfaces are often sporadic [19], and the sensory information available during training is often designed with reasonable logical or empirical justification but limited theoretical grounding [106]. By contrast, my research takes a unique, systematic approach to learning and training of procedural, sensorimotor knowledge, derived from a cognitive model and associated learning theories. As such, the study represents a novel approach to conducting research on abstract sensorimotor skill learning to-date.

I suggest that further understanding and exploration of direct learning theory, and the effects that dynamically-adjusted sensory concurrency can have on learning, would prove invaluable to the creation of environments and training mechanisms that promote fast, self-guided learning of sensorimotor interfaces.

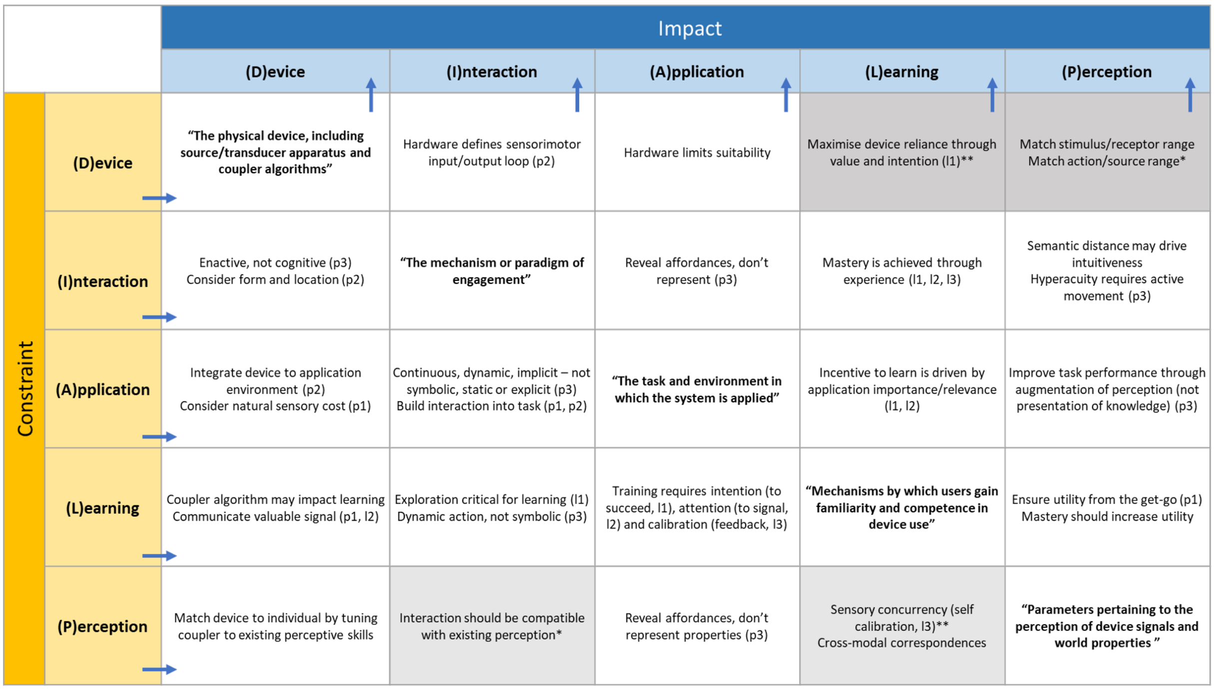
Limitations & Open Questions. There is immediate scope for future optimisation of several parameters here. Is there a specific degree of concurrency (overlap of information between senses) optimal for this learning trajectory? Or, as mentioned before, can dynamic adjustment of the information available in peripheral senses at a trial-by-trial level improve learning beyond static sensory concurrency? Direct learning theory and wider dynamical systems approaches would suggest that both of these are valid and exciting questions with promising answers, yet they remain to be tested through experimentation. If conducted robustly within the framework of non-representationalist psychology, such experiments would contribute both to the practical uptake of sensorimotor interfaces (through provision of training insights), and as an evaluation of direct learning theory itself.

Lobo [125] also suggests that training of Sensorimotor Interfaces should be based on the theory of direct learning, and quotes: *"we might design programs in which the usefulness of variables to perceive a property is manipulated"*. Such suggestion aligns closely with the arguments presented in preparation for experimentation in Chapter 6, where the utility of the visual was manipulated specifically between conditions to observe the effect on self-guided learning, isolated (as far as possible) from explicit task feedback. An insightful next step would be to dynamically adjust the visual feedback provided, perhaps on a trial-by-trial basis in response to user performance, in order to optimise and maximise the user dependence on the vibrotactile signal and demonstrate an even faster path to learning. Critically, post-experimental analysis of the dynamics of the parameters used to define the ever-changing degree of visual feedback [75] would then provide a highly informative measurement of individual and overall learning effectiveness, certainly more useful than the traditional evaluation of task performance alone [125].

In the longer-term, further understanding the effects of explicit vs implicit training, and optimising user mastery of SSADs and sensorimotor interfaces through structured training could also be re-applied to wider sensorimotor skills. Directly, understanding how humans acquire practical knowledge could impact the way we approach teaching and developing sensorimotor skills in general, and observation of the use of sensorimotor interfaces may provide a pathway to such understanding. Further, if sensorimotor interfaces themselves provide implicit information, perhaps they may be used more extensively as implicit feedback sources themselves, aiding the acquisition of wider sensorimotor skills in fields such as music or sports [66].

7.3.4 The Sensorimotor Interface Design Document

As a contribution that spans the research questions and contributions outlined above, the insight arising from my PhD research has also resulted in a detailed design document (see Fig. 4.3, reproduced in miniature in Figure 7.1). This design document encapsulates the value of the



		Impact				
		(D)evice	**(I)nteraction**	**(A)pplication**	**(L)earning**	**(P)erception**
Constraint	**(D)evice**	**“The physical device, including source/transducer apparatus and coupler algorithms”**	Hardware defines sensorimotor input/output loop (p2)	Hardware limits suitability	Maximise device reliance through value and intention (l1)**	Match stimulus/receptor range Match action/source range*
	(I)nteraction	Enactive, not cognitive (p3) Consider form and location (p2)	**“The mechanism or paradigm of engagement”**	Reveal affordances, don’t represent (p3)	Mastery is achieved through experience (l1, l2, l3)	Semantic distance may drive intuitiveness Hyperacuity requires active movement (p3)
	(A)pplication	Integrate device to application environment (p2) Consider natural sensory cost (p1)	Continuous, dynamic, implicit – not symbolic, static or explicit (p3) Build interaction into task (p1, p2)	**“The task and environment in which the system is applied”**	Incentive to learn is driven by application importance/relevance (l1, l2)	Improve task performance through augmentation of perception (not presentation of knowledge) (p3)
	(L)earning	Coupler algorithm may impact learning Communicate valuable signal (p1, l2)	Exploration critical for learning (l1) Dynamic action, not symbolic (p3)	Training requires intention (to succeed, l1), attention (to signal, l2) and calibration (feedback, l3)	**“Mechanisms by which users gain familiarity and competence in device use”**	Ensure utility from the get-go (p1) Mastery should increase utility
	(P)erception	Match device to individual by tuning coupler to existing perceptive skills	Interaction should be compatible with existing perception*	Reveal affordances, don’t represent properties (p3)	Sensory concurrency (self calibration, l3)** Cross-modal correspondences	**“Parameters pertaining to the perception of device signals and world properties ”**

Figure 7.1: Table of Design Dimensions—revisited. Reading across then up allows identification of design influences and constraints for the effective evaluation or creation of a sensorimotor interface. Grey cells indicate insights reflecting experimental contributions from: * = Chapter 5, ** = Chapter 6

theoretical contributions made in this thesis and highlights existing and potential pathways of experimentation. It is a structured reflection of the principles and guidelines detailed in Chapter 4, offering a scaffolding for the design of new sensorimotor interfaces and associated experimentation.

In Section 4.4, this potential was illustrated through detailed reference to the Lateral Line [93] and associated experimentation. This retrospective application of the design document to an existing study is an effective illustration of its utility. However, as noted when introduced, the document may also be used in combination with the principles of interaction and device design (as revisited above in Section 7.3) to address opportunities for future research, to aid in the building of new commercial interfaces, or to iterate on existing systems to improve uptake or apply to new areas.

In brief recap, five design dimensions and their interacting effects were highlighted for consideration in the design of sensorimotor interfaces (Figure 4.3): (D)evice, (I)nteraction, (A)pplication, (L)earning and (P)erception.

To provide a brief example of how the document can illuminate potential for future work, suppose we begin with access to an existing device, such as the feelSpace belt [1], fixing the parameter of

Dimension D. The output component of interaction (Dimension I) and applications (Dimension A) is limited by the vibrotactile array and torso-mount design of the belt. These hardware features therefore lend the system well to navigational applications (considering the arrangement of motors aligns well with the natural sense of orientation). The chosen information space should represent a parameter of high task relevance, which may to some degree be sensed or deduced through existing perceptual experience but which a user will derive greater enjoyment or ease from when appropriated through the device (Dimension L). Output intensity should be mapped to this input information space and calibrated to allow intuitive perception (Dimension P) in accordance with the sensing of vibrotactile intensity (see Chapter 5). Within these constraints, the feelSpace belt is suited to the communication of environmentally-sourced (virtual or real) topographical information relevant to whole-body orientation. Such information may for example represent previously invisible parameters (e.g. the 'happiness' of a route [140]) or deducible yet pertinent upcoming information, e.g. dead-ends or traffic-heavy areas. This set of design decisions guides towards the implementation of a system which aligns well with the feelSpace's original intended uses (to provide augmentative navigation information) while also strongly reflecting enactive principles.

Limitations. There is, of course, much room for improvement in this document and associated principles. Primarily, some degree of experience is still required to make the insightful leaps from simply meeting the document requirements to making design decisions. Of additional consideration is the verbose reading of the prior literature required to truly understand the document's summarised guidance. While this work is structured in such a way as to stand alone and address this in minimum fashion, the reading involved is still not insignificant, and a move from cognitive to enactive interaction design still requires dedication, representing a considerable investment.

A further important criticism of the document, and the dimensions more generally, is that it is difficult to know whether the design dimensions are complete, and offer a full description of all facets of design.

I must firstly address this criticism by openly accepting that there are undoubtedly missing dimensions. For example, ethical considerations are omitted entirely, mainly because such ethical considerations are beyond the scope of this research. On this note, it is not clear how ethical considerations (aside from general comfort/ergonomics) may influence design and implementation, or if they would simply inform cultural suitability once the interface had been designed. A further iteration of the document may seek to address this 'known unknown'.

However, of more threat to the document's validity are the 'unknown unknowns': the important dimensions influencing device and interaction design which may have been omitted unintentionally, or those we may simply not yet be aware of due to limitations in scientific knowledge. In this respect, some elements may be more directly addressable than others. For example, one may accept that some properties of the user are not represented well by the document, such as *individuality* (perhaps requiring device customisation or adaptation), while also realising that research supporting such customisations is thin (but not non-existent [123,148]). In creating this guide, I have erred on the side of justifiable, actionable dimensions supported by consensual literature evidence, and the insights gained through experimentation in this thesis.

In summary, while I acknowledge the difficulty in arguing the document's completeness, I reiterate that I have fairly described its origins and reasoning, and demonstrated its utility through both retrospective and hypothetical application.

7.4 Future Applications of Sensorimotor Theory

In this thesis, the combination of theory-derived modelling, tested and reinforced through experimental work, constitutes a strong argument for the application of sensorimotor theory to enactive interface design, and offers an exciting scaffolding for future work.

While use of sensorimotor theory has been justified throughout, its claims have largely been assumed without challenge, and its insights into the nature of human cognition have simply been used as tools for application to interaction design. Indeed, while its *utility* should be clear by now, it is not my intention (nor within the scope of the thesis) to reinforce or refute sensorimotor theory's *accuracy*.

With that said, in researching SSADs specifically, I have encountered numerous observations pertaining to the overlap between interaction design and theory of mind, which may be interesting to consider from the perspective of sensorimotor theory. These insights represent further opportunities for research on enactive systems, and the development of sensorimotor theory itself.

Cross-modal Correspondences and Plasticity. Sensory information provided to one modality can affect, influence or inform user perception typically associated with a different modality.

Interestingly, Visual to Audio SSADs have shown activation of the visual cortex in multiple contexts [4, 141], hinting at the existence of a cross-modal link between audio and vision. While cross-modal correspondences [155] (the somewhat cognitive associations between modalities) are real phenomena that have been linked to observations of SSADs [69], it is difficult to argue that they fully account for the vision-like *experience* claimed by users of the vOICe [166].

Instead, we must find a mechanism for explaining cross-modal plasticity (the re-purposing or flexible application of cortical areas usually associated with other modalities [133]), which may then give rise to this experience. In this case, how should we account for the re-utilisation of neural circuitry for the processing of audio information?

While this may initially seem a difficult question to answer, such observations can interestingly be used as an argument *in support of* Sensorimotor Theory. Approaching from the 'what it is like' qualities of Sensory Augmentation, Sainz Martinez [145] introduces the concept of 'perceptual distance' as an indication of how similar the sensorimotor laws and comprehensibility of various perceptual experiences are, touching on the idea that some sensorimotor engagements are more *experientially* similar than others. To illustrate, if users of the vOICe are exposed to 'vision-like' contingencies (through a device which creates stimuli in a manner comparable to the retinal response to eye and body movement), then the resulting experience may be 'vision-like' despite being auditorily stimulated, resulting in a close contingency distance to natural vision. By this argument, and in agreement with the original claims of Sensorimotor Theory [133] it may be the nature of the interaction and the sensorimotor contingencies (rather than the visual receptor itself) which gives rise to the vision-like experience. At this stage, we may argue that the visual cortex is optimised for processing of vision-like information, and that this information is dependent on the contingencies of sensorimotor engagement (rather than the output of the biological receptor), then it stands to reason that the brain will engage the visual cortex in response to the perceptually close 'vision-like' contingencies of the vOICe.

Hyperacuity, Apparatus- and Object-SMCs. Sensorimotor interactions often allow sensory hyperacuity: where agent-driven movement enables sensory localisation beyond natural static sensory resolution. This is most commonly observed in vision [55]. Participants utilising exploratory oscillations to achieve hyperacuity have been observed in SSAD studies [100, 108]. Lobo [108] notes that such oscillations appear to arise naturally from active exploration during the use of spatio-temporal SSADs. Such observations reinforce the notion that information may be extracted from dynamic and comparative stimulatory sensations (as explored in Chapter 5), and its observation during SSAD use further reinforces the role of action in perception, and the relevance of sensorimotor models to the design of enactive interfaces.

More interestingly, deliberate experimentation surrounding active information detection could play an important role in understanding how information is communicated through sensorimotor interfaces: the observation of hyperacuity may potentially lead to reliable, comparable methods for the identification of new sensorimotor couplings [100].

Specifically investigating active exploration and how it is linked to information extraction in sensorimotor interactions has also been approached by Bermejo [17]. In this experiment, it is claimed that the exploratory, oscillatory movements manifesting between different SSADs and in identification of different shapes correspond to apparatus-SMCs and object-SMCs respectively.

Here, I note that it is possible that some of the oscillatory patterns observed in the participant trials reported in Chapter 6 arose as a direct result of active exploration. However, whether these patterns reflect hyperacuity, object- or apparatus-SMCs, or even some other potentially interesting emergent behaviour, cannot be confirmed without specific further tests. There is much scope for further work investigating the role active exploration plays in information transfer, the creation of new perceptual experiences, and the mastery of sensorimotor interfaces.

Illusions. Illusions are usually defined as a discrepancy between physical stimulus and the corresponding percept [97]. Illusions indicate a surprising difference between the way the world is, and the way it is perceived, caused by a combination of biological and neurological factors. While visual illusions are most commonly known and researched, there also exist auditory illusions (Shepard Tone illusion [151]), tactile illusions (subcutaneous rabbit [21]), and many multi-modal illusions that depend on integrated perception (such as the rubber hand illusion [53]). It is conceivable that new modalities, when integrated into perception, may allow perception of illusions comparable in nature or effect to existing illusions, or to entirely new illusions with hitherto unknown qualities. Such potential and its relation to existing illusions is touched on briefly through an exploration of pseudo-haptics by Collins and Kapralos [**?**]. In any case, illusions arise by definition through perceptual experiences, and are therefore indicative of a sensation being attributed (albeit incorrectly) to a phenomenon ‘out there’ in the world. Therefore, the measurement of new forms of illusion arising from the engagement with sensorimotor interfaces, may, if observed, represent a confirmation of the method’s potential for true, perceptual alteration.

7.5 The Future of Enactivity in HCI

Enactivity remains a niche paradigm with limited literature (and even more limited examples in everyday technology) guiding and inspiring its application to interface design.

Prior to my work, as outlined in Section 7.1, applying enactive principles to interface design was primarily reserved to groups with prior experience or interests in assistive technologies, specifically SSADs, or to researchers of perception science collaborating with engineers to implement specialist tools for exploration of the principles of cognitive psychology. With no clear starting point, the design and development of enactive interfaces presented numerous challenges for researchers outside of these fields.

In response, this thesis has provided a broad, cross-disciplinary literature review of the field (Chapter 2), a detailed and derived model for interaction illustrated by a schematic taxonomy (Chapter 4, Section 4.2.1), a comprehensive set of design guidelines (Chapter 4, Section 4.3), and a series of user studies reinforcing the value of such guidelines to the application of enactivity to interface design (Chapters 5 & 6.

In presenting this work, it is intended that the application of enactivity to user interface design be more accessible through provided explanation and guidelines, the motivations be outlined through demonstration of potential, and the direction of research be clearer through presentation of the limitations and opportunities for future studies.

7.6 Final Thoughts

Despite nearly 60 years passing since Paul Bach-y-Rita constructed the TVSS, the field of Sensory Substitution has in many ways not changed. There remains questions surrounding how to build, test and deploy enactive, sensorimotor systems, and under which conditions they may be learned. At a psychological level, the potential for how human perception may be affected, manipulated and augmented through non-invasive technologies such as those outlined here remains unclear. Furthermore, it remains to be seen the degree to which rehabilitation or sensory compensation may occur through such means; these being applications of high relevance to such technologies. Beyond this, the pursuit of manufacturing and deployment, and guiding consumer interest, towards augmentative approaches and technologies as ubiquitous and comprehensive as the mouse, keyboard and monitor, seems a task for the distant future. The dream, as described in Appendix A, is a way off yet. However, this thesis represents another small stepping stone in the journey.

The work herein presented is intended to form a foundation to new researchers. It aims to highlight, through review and experimentation, several promising pathways towards the application of enactive principles in HCI, the pitfalls that may lead to failure along the way, and the how and why of user studies and evaluation of their results. In illustrating and applying enactive cognitive models suitable to the building of devices that engage with the way we perceive the world, I aim to have provided a self-consistent and helpful entry point to researchers within the HCI community for new types of interaction.

What started as a simple, but exciting, desire to create new interfaces in HCI from the principles of Sensory Augmentation quickly evolved into a long and diligent process of understanding, organising, formalising and communicating the structure of perception itself. This task has been exceptionally rewarding. This thesis brings together the fields of cognitive psychology, Sensory Substitution and Sensory Augmentation, and Human Computer Interaction, and I'm proud of the research I have conducted in contributing to these fields.

APPENDIX A

APPENDIX-A: LIFE WITH A MULTI-CONTEXT VIBROTACTILE SENSORY AUGMENTATION SYSTEM

I am wearing a multi-function vibrotactile interface: fingerless gloves featuring multiple tactors on the back. An extension of my senses and muscles, they work as both input and output for a lot of my life, interfacing wirelessly with the people, devices and systems that I engage with, reacting intelligently depending on the context just as my natural senses do.

I'm part of a team that controls complex systems in the electrical energy industry. I work with a large terminal covered in charts and buttons, reporting the overall state of the local grid, supporting generators. The terminal is useful when we need explicit details, but the overall flow and state of the system is usually best "felt". I visually scan the terminals with a soft focus as I gesture subtly to change the gloves' context, each refresh giving me a tactile overview of the systems with a level of detail that isn't subject to the change blindness my eyes are fallible to. Something not quite right about the sensation in one context prompts me to investigate the figures more closely. Upon inspection, the figures suggest an unwanted oscillation starting to develop. It's well within safety, so the automated system didn't pick it up. But if we stabilise it now, it'll save us an emergency later.

After work, I head with some colleagues to play football against another local team. I feel the sensation in the gloves change, indicating an awareness of the sport I'm playing, and a connection to the others around me. As I play, my eyes are on the ball, but I sense nearby players' position and trajectory on the field – it's not precise, like numbers or paths drawn in my mind, it's more of a feeling, the knowledge that if I deflect this next pass backwards and to my left… yes, it's been picked up by my teammate and we've scored! I feel a brief shimmer of excitement as the gloves confirm the goal, and a certain smugness as I sense the opposing team trudging back towards the halfway line behind me.

After the game, I head to my car to drive home. The vehicle unlocks with a twist of my fingers as I approach. Getting in, the gloves vibrate and warm briefly as they connect seamlessly with the vehicle's wireless charging system, and interface with the navigation and proximity sensors. I haven't been to this field for months, and can't remember the exact route home, but I have an innate sense of which direction I should be going… it crosses my mind briefly that this might be how birds feel as they migrate. I rarely check the mirrors as I change lanes or overtake, I trust the knowledge of nearby vehicles as I would my eyes. Perhaps more so – my eyes can't tell me what is around the corner, but the gloves can.

Arriving at home, I'm excited to try a new, immersive video game I've purchased. As I put on my VR headset and select the game, the familiar pattern of vibration indicating a change of context rushes through my hands. As I embody the character in the opening scene, I'm instantly excited and itching to explore my new environment, and the senses associated with my virtual embodiment. The haptic sensation in my hands is new in some ways, as I've never played this game before and can't be sure of what it all means, yet it's familiar enough for me to feel confident in my exploration of the environment, much like my old console controllers were familiar and under my thumbs even though the controls changed between games. As I move, I have a sense of a potential or gradient around me, like a pressure, or a proximity. It seems this has something to do with my abilities in this new world - an alien or perhaps magical potential, something primal in sensation, but beyond human in capabilities. I gesture to test my new affordances and feel the rush as the sensation in my hands confirms the invisible nuances of the environment around me, and I move forward, eager to discover more. The immersivity of the game I'm playing is magnified immeasurably by the connection to the environment. I'm experiencing a world that is beyond human experience.

At some point, an unexpected yet familiar pattern cuts through the game's sensations like a doorbell during a movie - someone in the real world is trying to call me. I disengage from the game and the gloves change context to give me the warming sensation I associated with the caller. It's my mum, I guess I'd better pick up…

APPENDIX-B: PARTICIPANT INFORMATION SHEETS AND QUESTIONNAIRES FOR EXPERIMENTAL WORK:

University of St Andrews

Participant Information

Sensory Augmentation through Nonvisual Interfaces in Video Games

Iain Carson, Uta Hinrichs, Aaron Quigley

What is the study about?

We invite you to participate in a research project investigating nonvisual interfaces in video games.

Do I have to take part?

This information sheet has been written to help you decide if you would like to take part. It is up to you and you alone whether you wish to take part. If you do decide to take part you will be free to withdraw at any time without providing a reason.

What would I be required to do?

You will be invited to complete a short questionnaire.

Following this, you are invited to engage with a task much like a video game, where you will be using the provided controller to control a character in order to avoid hazards in a virtual environment. Vibrational or “rumble” feedback will be used to provide additional information throughout the task, in various forms. Your visual feedback and the rumble feedback will change throughout the task, and you will be asked to provide feedbck on the perceived difficulty of the task at various stages by means of a short questionnaire.

This session should take no more than 60 minutes. Upon completion of the session, you may be invited to return to the study at a later date.

Are there any risks associated with taking part?

Taking part in this experiment will not involve risks any greater than engaging with a video game. If at any point you feel nauseous, fatigued or otherwise impacted by engagement with the task, please do let us know, and do not feel obliged to continue.

Informed consent

It is important that you are able to give your informed consent before taking part in my project and you will have the opportunity to ask any questions in relation to the research before you provide your consent (oral or written).

Reward

Participation in this experiment will be rewarded by a £20 book voucher.

What information about me or recordings of me (‘my data’) will you be collecting?

You will be asked to fill out a short questionnaire. All questions are optional.

Game performance measurements such as button presses and progress will be collected and recorded during the task.

Further short questionnaires evaluating your experience will be presented at the end of each session.

How will my data be stored, who will have access to it?

Your data will be stored in a PSEUDONYMISED form, which means that your data will be edited so that you are referred to by a unique reference such as a code number or different name, and the original data will remain accessible only to the above-named researchers. Your data will be stored in a locked cabinet in the SACHI lab, and only the above-named researchers will be able to access it. There will be a ‘key’ document, which will link your unique reference to your real identity. The key will be kept in a separate locked cabinet in the SACHI lab, and only the above-named researchers will have access to it and be able to reconnect your data to you at a later date. Audio and video recordings will be stored on an encrypted device.

How will my data be used, and in what form will it be shared further?

Collected data (such as game performance measurements, questionnaire results) will be analysed as part of the research study, and may be published in anonymized form.

It is expected that the project to which this research relates will be finalised by 05/2021 and written up as part of Iain Carson's PhD thesis.

When will my data be destroyed?

Your data will be shared as described above, and then the data held by the researcher will be destroyed 3 years following completion of my PhD.

Will my participation be confidential?

Yes, your participation will only be known to the above-named researchers.

Lawful basis for making use of personal data and data protection rights

The lawful basis that the University will rely on to make use of your personal data during the research and for related research projects in the future, as described to you is public task; where special category personal data are used the lawful basis is archiving purposes in the public interest, scientific or historical research purposes or statistical purposes.

The University of St Andrews is a Data Controller for the information you provide about you. You have a range of rights under the data protection legislation, including the right of complaint. However, some of those rights may not be available where you provide personal data for research purposes. For questions, comments or requests, consult the University website at https://www.st-andrews.ac.uk/terms/data-protection/rights/, or email dataprot@st-andrews.ac.uk.

You will be able to withdraw your data before 1st April 2019. If your data is anonymised, we will not be able to withdraw it after that point, because we will no longer know which data is yours.

Ethical Approvals

This research proposal has been scrutinised and subsequently granted ethical approval by the University of St Andrews Teaching and Research Ethics Committee.

What should I do if I have concerns about this study?

In the first instance you are encouraged to raise your concerns with the researcher and if you do not feel comfortable doing so, then you should contact my Supervisor. A full outline of the procedures governed by the University Teaching and Research Ethics Committee is available at www.st-andrews.ac.uk/utrec/guidelinespolicies/complaints/

Contact details

Researcher(s)	Iain Carson ic48	**Supervisor(s)**	Uta Hinrichs uh3 Aaron Quigley aquigley

Consent Form

Sensory Augmentation through Nonvisual Interfaces in Video Games
Iain Carson, Aaron Quigley, Uta Hinrichs

The University of St Andrews attaches high priority to the ethical conduct of research. We therefore ask you to consider the following points before signing this form. Your signature confirms that you are willing to participate in this study, however, signing this form does not commit you to anything you do not wish to do and you are free to withdraw your participation at any time.

Please initial box

- I understand the contents of the Participant Information Sheet (PIS_22/07/2019_2_NONVISUALGAMES1) ☐
- I have been given the opportunity to ask questions about the study and have had them answered satisfactorily. ☐
- I understand that my participation is entirely voluntary and that I can withdraw from the study at any time without giving an explanation. ☐
- I understand who will have access to my data, how it will be stored, in what form it will be shared, and what will happen to it at the end of the study. I understand that I will be able to withdraw my data up to 1 month following participation, and that if my data has been anonymised, it cannot be withdrawn after that point. ☐
- I agree to take part in the above study ☐

Signatures			
I confirm that I am willing to take part in this research			
	Print name	Date	**Signature**
Participant			
Researcher			

P

Nonvisual Videogames Experiment Screening Questionnaire

All questions are optional.

Q1. Gender?

____

Q2. Age?

____

Q3. Dominant hand?

a. Right
b. Left
c. Ambidextrous

Q4. Do you have any sensory/perceptual impairments that you feel the researchers should know about (e.g. corrected vision, colour blindness, manual touch/tactile impairments)?

a. No
b. Yes (please specify)

____

Nonvisual Videogames Experiment Post-session Questionnaire

All questions are optional.

Q1. Where would you say your focus of attention was during this session’s trials?

____

Q2. Did this change throughout the session?

____

Q3. If it changed, how would you describe the change?

____

Q4. Please rate the following statements:

	Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
I had to explicitly think about how to respond to the game signals					
I found today's trials intuitive					

Please now fill out the NASA TLX questionnaire on the reverse side.

Nonvisual Games Task NASA TLX

Mental Demand How mentally demanding was the task?

Very Low Very High

Physical Demand How physically demanding was the task?

Very Low Very High

Temporal Demand How hurried or rushed was the pace of the task?

Very Low Very High

Performance How successful were you in accomplishing what you were asked to do?

Perfect Failure

Effort How hard did you have to work to accomplish your level of performance?

Very Low Very High

Frustration How insecure, discouraged, irritated, stressed, and annoyed were you?

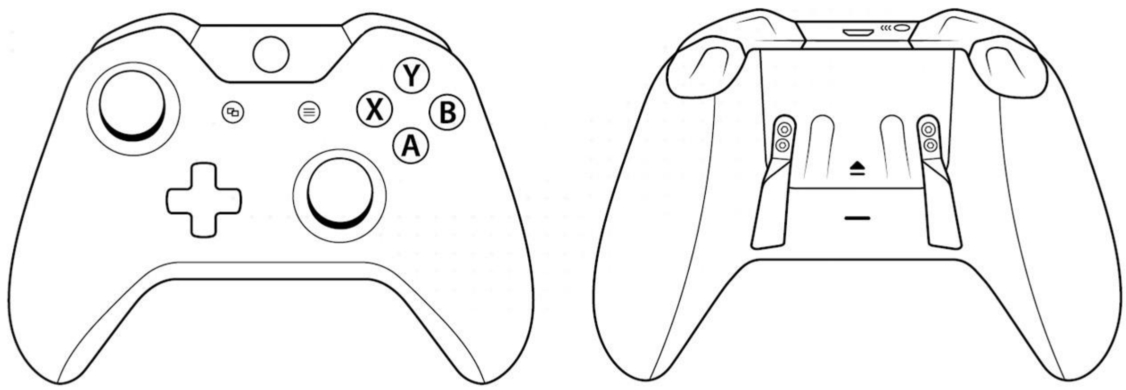
Very Low Very High

P Final

Nonvisual Videogames Experiment Post-study Questionnaire

All questions are optional.

Q1. What buttons on the controller have you used most across the sessions? Please describe in your own words how you have used these buttons to control the game. If your use of buttons has changed across the study sessions, please describe these changes as well.



Q2. Please rate the following statements:

	Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
I found the final session enjoyable					
The first three sessions helped me prepare for the final session					
I felt the final session required completely different skills to the previous sessions					
I felt my performance (control of and progress through the game) improved across the sessions.					
I experimented with the controls in every session					
I have a strong understanding of the maze and game rules					
I learned more about the game controls over the sessions					
I learned more about the game feedback over the sessions					

Q3. If you would like to comment on your answers to Q2. or provide further details, please do so here:

Q4. What did you find most enjoyable about the final session (if you found it enjoyable at all).

Q5. What did you find the most frustrating about the final session?

Q6. Any final/further comments, thoughts or ideas you would like to communicate?

Maze Navigation Task: Participant Pre-Session Brief

The controls, environment and rules of the game environment will be the same for every trial in every session you attend, however the visual and vibrational feedback you receive as you move around may vary. Some combinations of visual or haptic feedback may be easier to interpret than others. Don’t worry, keep trying.

For the first trial, you will find yourself at the beginning of a straight corridor. Your only goal is to reach the end of the corridor. You will be told when you reach the end of the corridor. There is no time limit and you are encouraged to explore the control sticks and buttons to figure out the controls and feedback you receive.

I will then leave the room. For the subsequent 12 trials, instead of a corridor, you will find yourself at the entrance to a "maze". You will be given 2 minutes exactly to make as much progress through each maze as you can. The game will tell you how many blocks you have progressed through the maze.

We will not tell you what the controls are nor what any feedback you receive means in any of the trials. It is up to you to experiment to figure out the controls, and to use your best judgement to interpret the environment and to make progress through the mazes.

The screen in front of you will provide further instructions and indicate your time remaining in the breaks between mazes. Good luck!

REFERENCES

[1] feelSpace - taktile Information. `https://www.feelspace.de/?lang=en`. [Online; accessed 2021-06-08].

[2] ABBOUD, S., HANASSY, S., LEVY-TZEDEK, S., MAIDENBAUM, S., AND AMEDI, S. EyeMusic: introducting a visual colorful experience for the blind using auditory sensory substitution. *Restor Neurol Neuroscience 32*, 2 (2014), 247–257.

[3] ALSBERG, B. K. Is sensing spatially distributed chemical information using sensory substitution with hyperspectral imaging possible? *Chemometrics and Intelligent Laboratory Systems 114* (2012), 24–29.

[4] AMEDI, A., STERN, W. M., CAMPRODON, J. A., BERMPOHL, F., MERABET, L., ROTMAN, S., HEMOND, C., MEIJER, P., AND PASCUAL-LEONE, A. Shape conveyed by visual-to-auditory sensory substitution activates the lateral occipital complex. *Nature Neuroscience 10*, 6 (2007), 687–689.

[5] AUVRAY, M., HANNETON, S., AND O'REGAN, J. K. Learning to perceive with a visuo-auditory substitution system: Localisation and object recognition with 'The vOICe'. *Perception 36*, 3 (2007), 416–430.

[6] AUVRAY, M., AND MYIN, E. Perception with compensatory devices: From sensory substitution to sensorimotor extension. *Cognitive Science 33*, 6 (2009), 1036–1058.

[7] AUVRAY, M., AND ROHDE, M. Perceptual crossing: The simplest online paradigm. *Frontiers in Human Neuroscience*, JUNE 2012 (2012).

[8] AVILÉS, C., NAVIA, J. A., RUIZ-PÉREZ, L.-M., AND ZAPATERO-AYUSO, J. A. How enaction and ecological approaches can contribute to sports and skill learning. *Frontiers in Psychology 11* (2020).

[9] BACH-Y RITA, P. *Brain mechanisms in sensory substitution.* Academic Press, New York, 1972.

[10] BACH-Y RITA, P. Sensory substitution in rehabilitation. *Rehabilitation of the neurological patient* (1982), 361–383.

[11] BACH-Y-RITA, P., COLLINS, C. C., SAUNDERS, F. A., WHITE, B., AND SCADDEN, L. Vision Substitution by Tactile Image Projection. *Nature 221*, 5184 (1969), 963–964.

[12] BACH-Y RITA, P., AND KERCEL, S. Sensori- “motor” coupling by observed and imagined movement. *Intellectica*, 608 (2002), 287–297.

[13] BACH-Y-RITA, P., TYLER, M. E., AND KACZMAREK, K. A. Seeing with the Brain. *International Journal of Human-Computer Interaction 15*, 2 (2003), 285–295.

[14] BACH-Y RITA, P., AND W. KERCEL, S. Sensory substitution and the human-machine interface. *Trends in Cognitive Sciences 7*, 12 (2003), 541–546.

[15] BACH-Y RITA, P., WEBSTER, J. G., AND THOMPKINS, W. J. Sensory Substitution for Space Gloves and for Space Robots. *Proceedings of the Workshop on Space Telerobotics JPL Public* (1987), 50–57.

[16] BAU, O., POUPYREV, I., ISRAR, A., AND HARRISON, C. Teslatouch: Electrovibration for touch surfaces. In *Proceedings of the 23Nd Annual ACM Symposium on User Interface Software and Technology* (New York, NY, USA, 2010), UIST ’10, ACM, pp. 283–292.

[17] BERMEJO, F., DI PAOLO, E., HÜG, M. X., AND ARIAS, C. Sensorimotor strategies for recognizing geometrical shapes: A comparative study with different sensory substitution devices. *Frontiers in Psychology 6*, MAY (2015).

[18] BERNING, M., BRAUN, F., RIEDEL, T., AND BEIGL, M. Proximityhat: A head-worn system for subtle sensory augmentation with tactile stimulation. In *Proceedings of the 2015 ACM International Symposium on Wearable Computers* (New York, NY, USA, 2015), ISWC ’15, Association for Computing Machinery, p. 31–38.

[19] BERTRAM, C., AND STAFFORD, T. Improving training for sensory augmentation using the science of expertise. *Neuroscience and Biobehavioral Reviews 68* (2016), 234–244.

[20] BIRD, J., MARSHALL, P., ROGERS, Y., AND ROGERS, V. Low-fi skin vision: a case study in rapid prototyping a sensory substitution system. *Proceedings of the 23rd British HCI Group Annual Conference on People and Computers: Celebrating People and Technology* (2009), 55–64.

[21] BLANKENBURG, F., RUFF, C. C., DEICHMANN, R., REES, G., AND DRIVER, J. The Cutaneous Rabbit Illusion Affects Human Primary Sensory Cortex Somatotopically. *PLOS Biology 4*, 3 (mar 2006), e69.

[22] BLUM, J. R., BOUCHARD, M., AND COOPERSTOCK, J. R. What's around me? Spatialized audio augmented reality for blind users with a smartphone. *Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering 104 LNICST* (2012), 49–62.

[23] BREWSTER, S., AND BROWN, L. M. Tactons: Structured tactile messages for non-visual information display. In *Proceedings of the Fifth Conference on Australasian User Interface - Volume 28* (AUS, 2004), AUIC '04, Australian Computer Society, Inc., p. 15–23.

[24] BREWSTER, S. A., WRIGHT, P. C., AND EDWARDS, A. D. N. A detailed investigation into the effectiveness of earcons. *Auditory display: sonification, audification, and auditory interfaces* (1994), 471–498.

[25] BROWN, D. J., SIMPSON, A. J., AND PROULX, M. J. Visual objects in the auditory system in sensory substitution: How much information do we need? *Multisensory Research 27*, 5-6 (2014), 337–357.

[26] BROWN, L. M., BREWSTER, S. A., AND PURCHASE, H. C. Multidimensional tactons for non-visual information presentation in mobile devices. In *Proceedings of the 8th Conference on Human-computer Interaction with Mobile Devices and Services* (New York, NY, USA, 2006), MobileHCI '06, ACM, pp. 231–238.

[27] BUCHS, G., HAIMLER, B., KEREM, M., MAIDENBAUM, S., BRAUN, L., AND AMEDI, A. A self-training program for sensory substitution devices. *PloS one 16* (2021), e0250281.

[28] CANCAR, L., DIAZ, A., BARRIENTOS, A., TRAVIESO, D., AND JACOBS, D. M. Tactile-Sight: A sensory substitution device based on distance-related vibrotactile flow regular paper. *International Journal of Advanced Robotic Systems 10* (2013).

[29] CAPELLE, C., TRULLEMANS, C., ARNO, P., AND VERAART, C. A real-time experimental prototype for enhancement of vision rehabilitation using auditory substitution. *IEEE Transactions on Biomedical Engineering 45*, 10 (1998), 1279–1293.

[30] CARDIN, S., THALMANN, D., AND VEXO, F. A wearable system for mobility improvement of visually impaired people. *Visual Computer 23*, 2 (2007), 109–118.

[31] CARSON, I., QUIGLEY, A., CLARKE, L., AND HINRICHS, U. Investigating the Effect of Sensory Concurrency on Learning Haptic Spatiotemporal Signals. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies 5*, 1 (2021), 1–30.

[32] CARTON, A., AND DUNNE, L. E. Tactile distance feedback for firefighters: Design and preliminary evaluation of a sensory augmentation glove. In *Proceedings of the 4th Augmented Human International Conference* (New York, NY, USA, 2013), AH '13, Association for Computing Machinery, p. 58–64.

[33] CASSINELLI, A., REYNOLDS, C., AND ISHIKAWA, M. Haptic radar/extended skin project. *ACM SIGGRAPH 2006: Sketches, SIGGRAPH '06 1* (2006), 1991.

[34] CAUCHARD, J. R., CHENG, J. L., PIETRZAK, T., AND LANDAY, J. A. Activibe: Design and evaluation of vibrations for progress monitoring. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (New York, NY, USA, 2016), CHI '16, ACM, pp. 3261–3271.

[35] CHALMERS, D. Facing up to the problem of consciousness. *Journal of Consciousness Studies 2*, 3 (1995), 200–19.

[36] CHANG, A., O'MODHRAIN, S., JACOB, R., GUNTHER, E., AND ISHII, H. ComTouch: Design of a vibrotactile communication device. *Proceedings of the Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques, DIS* (2002), 312–320.

[37] CHEBAT, D. R., MAIDENBAUM, S., AND AMEDI, A. Navigation using sensory substitution in real and virtual mazes. *PLoS ONE 10*, 6 (2015), 1–18.

[38] CHEN, J., CASTILLO, P., TURCOTT, R., ISRAR, A., AND LAU, F. Feeling speech on the arm. In *Conference on Human Factors in Computing Systems - Proceedings* (2018), vol. 2018-April.

[39] CHEN, T., WU, Y.-S., AND ZHU, K. Investigating different modalities of directional cues for multi-task visual-searching scenario in virtual reality. In *Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology* (New York, NY, USA, 2018), VRST '18, ACM, pp. 41:1–41:5.

[40] CHERNYSHOV, G., RAGOZIN, K., CHEN, J., AND KUNZE, K. DubHap: A sensory substitution based superhuman sport. *ACM International Conference Proceeding Series Part F1376* (2018).

[41] COLMAN, A. M. *A Dictionary of Psychology*. Oxford University Press, jan 2015.

[42] CRAIG, J. C. Difference threshold for intensity of tactile stimuli. *Perception & Psychophysics 11*, 2 (1972), 150–152.

[43] DANILOV, Y., AND TYLER, M. Brainport: An Alternative Input to the Brain. *Journal of Integrative Neuroscience 04*, 04 (dec 2005), 537–550.

[44] DE JESUS OLIVEIRA, V. A., NEDEL, L., MACIEL, A., AND BRAYDA, L. Spatial discrimination of vibrotactile stimuli around the head. *IEEE Haptics Symposium, HAPTICS 2016-April* (2016), 1–6.

[45] DE PAZ, C., TRAVIESO, D., IBÁÑEZ-GIJÓN, J., BRAVO, M., LOBO, L., AND JACOBS, D. M. Sensory substitution: The affordance of passability, body-scaled perception, and exploratory movements. *PLoS ONE 14*, 3 (2019), 1–18.

[46] DEROY, O., AND AUVRAY, M. Reading the world through the skin and ears: A new perspective on sensory substitution. *Frontiers in Psychology 3*, NOV (2012), 1–13.

[47] DEROY, O., AND AUVRAY, M. *Perception and its modalities*. Oxford University Press Oxford, UK, 198 Madison Avenue, New York, NY 10016, 2014, ch. A crossmodal perspective on sensory substitution, pp. 327–349.

[48] DI PAOLO, E., BUHRMANN, T., AND BARANDIARAN, X. *Sensorimotor Life*, vol. 1. Oxford University Press, jun 2017.

[49] DÍAZ, A., BARRIENTOS, A., JACOBS, D. M., AND TRAVIESO, D. Action-contingent vibrotactile flow facilitates the detection of ground level obstacles with a partly virtual sensory substitution device. *Human Movement Science 31*, 6 (2012), 1571–1584.

[50] DUBLON, G., AND PARADISO, J. A. Tongueduino: Hackable, high-bandwidth sensory augmentation. In *CHI '12 Extended Abstracts on Human Factors in Computing Systems* (New York, NY, USA, 2012), CHI EA '12, Association for Computing Machinery, p. 1453–1454.

[51] EAGLEMAN, D. Plenary talks: A vibrotactile sensory substitution device for the deaf and profoundly hearing impaired. In *2014 IEEE Haptics Symposium (HAPTICS)* (2014), pp. xvii–xvii.

[52] EBRAHIMI, D., AND KUNOV, H. Peripheral Vision Lipreading Aid. *IEEE Transactions on Biomedical Engineering 38*, 10 (1991), 944–952.

[53] EHRSSON, H. H., SPENCE, C., AND PASSINGHAM, R. E. That's my hand! activity in premotor cortex reflects feeling of ownership of a limb. *Science 305*, 5685 (2004), 875–877.

[54] ERNST, M. O., AND BÜLTHOFF, H. H. Merging the senses into a robust percept. *Trends in Cognitive Sciences 8*, 4 (2004), 162–169.

[55] FAHLE, M., AND POGGIO, T. Visual hyperacuity: spatiotemporal interpolation in human vision. *Proceedings of the Royal Society of London. Series B. Biological Sciences 213*, 1193 (nov 1981), 451–477.

[56] FAVELA, L. H., RILEY, M. A., SHOCKLEY, K., AND CHEMERO, A. Perceptually Equivalent Judgments Made Visually and via Haptic Sensory-Substitution Devices. *Ecological Psychology 30*, 4 (2018), 326–345.

[57] FECHNER, G. T. *Elemente der psychophysik*, vol. 2. Breitkopf u. Härtel, Leipzig, 1860.

[58] FORTA, N. G., MORIOKA, M., AND GRIFFIN, M. J. Difference thresholds for the perception of whole-body vertical vibration: Dependence on the frequency and magnitude of vibration. *Ergonomics 52*, 10 (2009), 1305–1310.

[59] FORTIN, P. E., SULMONT, E., AND COOPERSTOCK, J. Detecting perception of smartphone notifications using skin conductance responses. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (New York, NY, USA, 2019), CHI '19, ACM, pp. 190:1–190:9.

[60] FRANCISCO, E., TANNAN, V., ZHANG, Z., HOLDEN, J., AND TOMMERDAHL, M. Vibrotactile amplitude discrimination capacity parallels magnitude changes in somatosensory cortex and follows Weber's Law. *Experimental Brain Research 191*, 1 (2008), 49–56.

[61] FRISSON, C., DECAUDIN, J., PIETRZAK, T., NG, A., PONCET, P., CASSET, F., LATOUR, A., AND BREWSTER, S. A. Designing vibrotactile widgets with printed actuators and sensors. In *Adjunct Publication of the 30th Annual ACM Symposium on User Interface Software and Technology* (New York, NY, USA, 2017), UIST '17, ACM, pp. 11–13.

[62] FROESE, T., MCGANN, M., BIGGE, W., SPIERS, A., AND SETH, A. K. The enactive torch: A new tool for the science of perception. *IEEE Transactions on Haptics 5*, 4 (2012), 365–375.

[63] GELDARD, F. A. Adventures in tactile literacy. *American Psychologist 12*, 3 (1957), 115–124.

[64] GESCHEIDER, G. A. *Psychophysics: The fundamentals, 3rd ed.* Lawrence Erlbaum Associates Publishers, Mahwah, NJ, US, 1997.

[65] GORI, M., CAPPAGLI, G., TONELLI, A., BAUD-BOVY, G., AND FINOCCHIETTI, S. Devices for visually impaired people: High technological devices with low user acceptance and no adaptability for children. *Neuroscience and Biobehavioral Reviews 69* (2016), 79–88.

[66] GU, H. Exploring sensory augmentation through tactile interfaces. *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2015 ACM International Symposium on Wearable Computers - UbiComp '15* (2015), 1243–1249.

[67] GÜNTHER, S., MAKHIJA, M., MÜLLER, F., SCHÖN, D., MÜHLHÄUSER, M., AND FUNK, M. Pneumact: Pneumatic kinesthetic actuation of body joints in virtual reality environments. In *Proceedings of the 2019 on Designing Interactive Systems Conference* (New York, NY, USA, 2019), DIS '19, ACM, pp. 227–240.

[68] HACKATHORN, R., AND MARGOLIS, T. Immersive analytics: Building virtual data worlds for collaborative decision support. *2016 Workshop on Immersive Analytics, IA 2016* (2017), 44–47.

[69] HAMILTON-FLETCHER, G., WRIGHT, T. D., AND WARD, J. Cross-modal correspondences enhance performance on a colour-to-sound sensory substitution device. *Multisensory Research 29*, 4-5 (2016), 337–363.

[70] HANNETON, S., AUVRAY, M., AND DURETTE, B. The Vibe: A versatile vision-to-audition sensory substitution device. *Applied Bionics and Biomechanics 7*, 4 (2010), 269–276.

[71] HART, S. G. Nasa-task load index (nasa-tlx); 20 years later. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting 50*, 9 (2006), 904–908.

[72] HELD, R., AND HEIN, A. Movement-produced stimulations in the development of visually guided behavior. *Journal of Comparative and Physiological Psychology 56*, 5 (1963), 872–876.

[73] HENDERSON, J., AVERY, J., GRISONI, L., AND LANK, E. Leveraging distal vibrotactile feedback for target acquisition. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (New York, NY, USA, 2019), CHI '19, ACM, pp. 485:1–485:11.

[74] HEYES, T. Sonic pathfinder-a programmable guidance aid for the blind. *Electronics & Wireless World 90*, 1579 (1984), 26.

[75] HIGUERAS-HERBADA, A., DE PAZ, C., JACOBS, D. M., TRAVIESO, D., AND IBÁÑEZ-GIJÓN, J. The direct learning theory: a naturalistic approach to learning for the post-cognitivist era. *Adaptive Behavior 27*, 6 (2019), 389–403.

[76] HOFFMANN, R., VALGEIRSDÓTTIR, V. V., JÓHANNESSON, Ó. I., UNNTHORSSON, R., AND KRISTJÁNSSON, Á. Measuring relative vibrotactile spatial acuity: effects of tactor type, anchor points and tactile anisotropy. *Experimental Brain Research 236*, 12 (2018), 3405–3416.

[77] HTC. Vive cosmos overview. `https://www.vive.com/us/product/vive-cosmos/overview/`, 2020. [Online; accessed 2021-27-09].

[78] HUET, M., CAMACHON, C., FERNANDEZ, L., JACOBS, D. M., AND MONTAGNE, G. Self-controlled concurrent feedback and the education of attention towards perceptual invariants. *Human Movement Science 28*, 4 (2009), 450–467.

[79] HWANG, I., SEO, J., KIM, M., AND CHOI, S. Vibrotactile perceived intensity for mobile devices as a function of direction, amplitude, and frequency. *IEEE Transactions on Haptics 6*, 3 (2013), 352–362.

[80] JONES, L. A. Tactile communication systems. optimizing the display of information. *Progress in Brain Research 192* (2011), 113–128.

[81] JONES, L. A., KUNKEL, J., PIATESKI, E., AND JONES, L. A. Vibrotactile pattern recognition on the arm and torso. *Proceedings - 1st Joint Eurohaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems; World Haptics Conference, WHC 2005 38*, 1 (2005), 90–95.

[82] JONES, L. A., AND SARTER, N. B. Tactile displays: Guidance for their design and application. *Human Factors 50*, 1 (2008), 90–111.

[83] KACZMAREK, K. A., WEBSTER, J. G., BACH-Y RITA, P., AND TOMPKINS, W. J. Electrotactile and Vibrotactile Displays for Sensory Substitution Systems. *IEEE Transactions on Biomedical Engineering 38*, 1 (1991), 1–16.

[84] KAŁWAK, W., REUTER, M., ŁUKOWSKA, M., MAJCHROWICZ, B., AND WIERZCHOŃ, M. Guidelines for quantitative and qualitative studies of sensory substitution experience. *Adaptive Behavior 26*, 3 (2018), 111–127.

[85] KASPAR, K., KÖNIG, S., SCHWANDT, J., AND KÖNIG, P. The experience of new sensorimotor contingencies by sensory augmentation. *Consciousness and Cognition 28*, 1 (2014), 47–63.

[86] KENNA, K., AND RYAN, A. Superhearo. *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing Adjunct - UbiComp '16* (2016), 481–486.

[87] KERDEGARI, H., KIM, Y., AND PRESCOTT, T. J. Head-mounted sensory augmentation device: Designing a tactile language. *IEEE transactions on haptics 9*, 3 (2016), 376–386.

[88] KL, G., G, M., A, M., RS, C., PJ, B., AND GM, C. A comparison of Tactaid II+ and Tactaid 7 use by adults with a profound hearing impairment. *Ear and hearing 20*, 6 (1999), 471–482.

[89] KRAUS, M., DIEHL, A., AND BACH, B. Visualization in the VR-Canvas : How much Reality is Good for Immersive Analytics in Virtual Reality ?

[90] KRISHNA, S., BALA, S., MCDANIEL, T., MCGUIRE, S., AND PANCHANATHAN, S. VibroGlove: an assistive technology aid for conveying facial expressions. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI) Extended Abstracts* (2010), 3637–3642.

[91] KRISTJÁNSSON, Á., MOLDOVEANU, A., JÓHANNESSON, Ó. I., BALAN, O., SPAGNOL, S., VALGEIRSDÓTTIR, V. V., AND UNNTHORSSON, R. Designing sensory-substitution devices: Principles, pitfalls and potential1. *Restorative Neurology and Neuroscience 34*, 5 (2016), 769–787.

[92] KRUGER, M., DRIESSEN, T., WIEBEL-HERBOTH, C. B., DE WINTER, J. C., AND WERSING, H. Feeling uncertain—effects of a vibrotactile belt that communicates vehicle sensor uncertainty. *Information 11*, 7 (2020), 353.

[93] KRÜGER, M., WIEBEL-HERBOTH, C. B., AND WERSING, H. The lateral line: Augmenting spatiotemporal perception with a tactile interface. In *Proceedings of the Augmented Humans International Conference* (New York, NY, USA, 2020), AHs '20, Association for Computing Machinery.

[94] KUCHERA, B. How the nintendo switch's hd rumble makes tumbleseed feel real. `https://www.polygon.com/2017/5/1/15499328/tumbleseed-hd-rumble-nintendo-switch`, 2017. [Online; accessed 2021-27-09].

[95] KULL, K. Umwelt and Modeling, 2009.

[96] KÖNIG, S. U., SCHUMANN, F., KEYSER, J., GOEKE, C., KRAUSE, C., WACHE, S., LYTOCHKIN, A., EBERT, M., BRUNSCH, V., WAHN, B., KASPAR, K., NAGEL, S. K., MEILINGER, T., BÜLTHOFF, H., WOLBERS, T., BÜCHEL, C., AND KÖNIG, P. Learning new sensorimotor contingencies: Effects of long-term use of sensory augmentation on the brain and conscious perception. *PLOS ONE 11*, 12 (12 2016), 1–35.

[97] LEDERMAN, S. J., AND JONES, L. A. Tactile and Haptic Illusions. *IEEE Transactions on Haptics 4*, 4 (2011), 273–294.

[98] LEE, J., SINCLAIR, M., GONZALEZ-FRANCO, M., OFEK, E., AND HOLZ, C. Torc: A virtual reality controller for in-hand high-dexterity finger interaction. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (New York, NY, USA, 2019), CHI '19, ACM, pp. 71:1–71:13.

[99] LENAY, C., CANU, S., AND VILLON, P. Technology and perception: the contribution of sensory substitution systems. 44–53.

[100] LENAY, C., GAPENNE, O., HANNETON, S., MARQUE, C., AND GENOUËLLE, C. Sensory Substitution : Limits and Perspectives. *Touching for Knowing Cognitive psychology of haptic manual perception 19* (2003), 275–292.

[101] LEVINE, J. Materialism and qualia: The explanatory gap. *Pacific Philosophical Quarterly 64*, 4 (1983), 354–361.

[102] LEVY-TZEDEK, S., NOVICK, I., ARBEL, R., ABBOUD, S., MAIDENBAUM, S., VAADIA, E., AND AMEDI, A. Cross-sensory transfer of sensory-motor information: visuomotor learning affects performance on an audiomotor task, using sensory-substitution. *Scientific reports 2* (2012), 949.

[103] LINDEMAN, R. W., YANAGIDA, Y., NOMA, H., AND HOSAKA, K. Wearable vibrotactile systems for virtual contact and information display. *Virtual Reality 9*, 2-3 (2006), 203–213.

[104] LINVILL, J. G., AND BLIS, J. C. A Direct Translation Reading Aid for the Blind. *Proceedings of the IEEE 54*, 1 (1966), 40–51.

[105] LOBO, L. Current alternatives on perceptual learning: introduction to special issue on post-cognitivist approaches to perceptual learning. *Adaptive Behavior 27*, 6 (2019), 355–362.

[106] LOBO, L., NORDBECK, P. C., RAJA, V., CHEMERO, A., RILEY, M. A., JACOBS, D. M., AND TRAVIESO, D. Route selection and obstacle avoidance with a short-range haptic sensory substitution device. *International Journal of Human Computer Studies 132*, March (2019), 25–33.

[107] LOBO, L., TRAVIESO, D., BARRIENTOS, A., AND JACOBS, D. M. Stepping on obstacles with a sensory substitution device on the lower leg: Practice without vision is more beneficial than practice with vision. *Plos one 9*, 6 (2014), e98801.

[108] LOBO, L., TRAVIESO, D., JACOBS, D. M., RODGER, M., AND CRAIG, C. M. Sensory substitution: Using a vibrotactile device to orient and walk to targets. *Journal of Experimental Psychology: Applied 24*, 1 (2018), 108–124.

[109] LUZHNICA, G., AND VEAS, E. Background perception and comprehension of symbols conveyed through vibrotactile wearable displays. In *Proceedings of the 24th International Conference on Intelligent User Interfaces* (New York, NY, USA, 2019), IUI '19, ACM, pp. 57–64.

[110] LUZHNICA, G., AND VEAS, E. Boosting word recognition for vibrotactile skin reading. In *Proceedings of the 23rd International Symposium on Wearable Computers* (New York, NY, USA, 2019), ISWC '19, ACM, pp. 135–144.

[111] MACPHERSON, F. Taxonomising the senses. *Philosophical Studies 153*, 1 (2011), 123–142.

[112] MACPHERSON, F., AND MACPHERSON, F. Sensory Substitution and Augmentation: An Introduction. *Sensory Substitution and Augmentation* (2018), 1–42.

[113] MAEREG, A. T., NAGAR, A., REID, D., AND SECCO, E. L. Wearable Vibrotactile Haptic Device for Stiffness Discrimination during Virtual Interactions. *Frontiers Robotics AI 4*, SEP (2017), 1–9.

[114] MAIDENBAUM, S., ABBOUD, S., AND AMEDI, A. Sensory substitution: Closing the gap between basic research and widespread practical visual rehabilitation. *Neuroscience and Biobehavioral Reviews 41* (2014), 3–15.

[115] MAIDENBAUM, S., LEVY-TZEDEK, S., CHEBAT, D. R., AND AMEDI, A. Increasing accessibility to the blind of virtual environments, using a virtual mobility aid based on the "EyeCane": Feasibility study. *PLoS ONE 8*, 8 (2013), 1–7.

[116] MEIJER, P. B. L. The vOICe. `https://www.seeingwithsound.com/`. [Online; accessed 2021-09-06].

[117] MEIJER, P. B. L. An Experimental System for Auditory Image Representations. *IEEE Transactions on Biomedical Engineering 39*, 2 (1992), 112–121.

[118] MITTERMAYER, C. On the determination of dynamic errors for rise time measurement with an oscilloscope. *IEEE Transactions on Instrumentation and Measurement 48*, 6 (1999), 1103–1107.

[119] MOLONEY, J., SPEHAR, B., GLOBA, A., AND WANG, R. The affordance of virtual reality to enable the sensory representation of multi-dimensional data for immersive analytics: from experience to insight. *Journal of Big Data 5*, 1 (2018), 1–19.

[120] MOSCATELLI, A., MEZZETTI, M., AND LACQUANITI, F. Modeling psychophysical data at the population-level: The generalized linear mixed model. *Journal of Vision 12*, 11 (2012), 1–17.

[121] MOSER, E. I., ROUDI, Y., WITTER, M. P., KENTROS, C., BONHOEFFER, T., AND MOSER, M. B. Grid cells and cortical representation. *Nature Reviews Neuroscience 15*, 7 (2014), 466–481.

[122] MUNZNER, T. *Visualization analysis and design.* CRC press, 2014.

[123] MYERS, D. Tactile Situational Awareness System (TSAS), 2004.

[124] NAGEL, S. K., CARL, C., KRINGE, T., MÄRTIN, R., AND KÖNIG, P. Beyond sensory substitution - Learning the sixth sense. *Journal of Neural Engineering 2*, 4 (2005), 13–26.

[125] NAVAS, L. L. *An Ecological Approach to Sensory Substitution.* PhD thesis, Facultad de Psicologìa, Departamento de Psicologìa Bàsica, Universidad Autònoma de Madrid, 2017.

[126] NEGEN, J., WEN, L., THALER, L., AND NARDINI, M. Integration of A New Sensory Skill with Vision After Less Than 3 Hours of Training. *bioRxiv* (2017), 232579.

[127] NEOSENSORY. Neosensory – experience sound as touch and gain new awareness. `https://neosensory.com/`, 2020.

[128] NIELSEN, J. Ten usability heuristics, 2005.

[129] NORMAN, D. *The Design of Everyday Things*. 2013.

[130] NOVICH, S. D., AND EAGLEMAN, D. M. Using space and time to encode vibrotactile information: toward an estimate of the skin's achievable throughput. *Experimental Brain Research 233*, 10 (2015), 2777–2788.

[131] OCULUS. Oculus rift: Vr headset for vr ready pcs, 2020.

[132] OLWAL, A., AND KRESS, B. 1d eyewear: Peripheral, hidden leds and near-eye holographic displays for unobtrusive augmentation. In *Proceedings of the 2018 ACM International Symposium on Wearable Computers* (New York, NY, USA, 2018), ISWC '18, Association for Computing Machinery, p. 184–187.

[133] O'REGAN, J. K., AND NOË, A. A sensorimotor account of vision and visual consciousness. *Behavioral and brain sciences 24*, 5 (2001), 939.

[134] PAN, Y.-T., YOON, H. U., AND HUR, P. A Portable Sensory Augmentation Device for Balance Rehabilitation Using Fingertip Skin Stretch Feedback. *IEEE transactions on neural systems and rehabilitation engineering : a publication of the IEEE Engineering in Medicine and Biology Society 25*, 1 (2017), 28–36.

[135] PEZENT, E., ISRAR, A., SAMAD, M., ROBINSON, S., AGARWAL, P., BENKO, H., AND COLONNESE, N. Tasbi: Multisensory Squeeze and Vibrotactile Wrist Haptics for Augmented and Virtual Reality. *2019 IEEE World Haptics Conference, WHC 2019*, July (2019), 1–6.

[136] PONGRAC, H. Vibrotactile perception: Differential effects of frequency, amplitude, and acceleration. *Proceedings of the 2006 IEEE International Workshop on Haptic Audio Visual Environments and Their Applications, HAVE 2006*, November (2007), 54–59.

[137] PROULX, M. J., BROWN, D. J., PASQUALOTTO, A., AND MEIJER, P. Multisensory perceptual learning and sensory substitution. *Neuroscience and Biobehavioral Reviews 41* (2014), 16–25.

[138] PROULX, M. J., STOERIG, P., LUDOWIG, E., AND KNOLL, I. Seeing 'where'through the ears: effects of learning-by-doing and long-term sensory deprivation on localization based on image-to-sound substitution. *PloS one 3*, 3 (2008), e1840.

[139] PROUZEAU, A., CORDEIL, M., ROBIN, C., ENS, B., THOMAS, B. H., AND DWYER, T. Scaptics and highlight-planes: Immersive interaction techniques for finding occluded features in 3d scatterplots. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (New York, NY, USA, 2019), CHI '19, ACM, pp. 325:1–325:12.

[140] QUERCIA, D. Chatty, Happy, and Smelly Maps. pp. 741–741.

[141] RENIER, L., COLLIGNON, O., POIRIER, C., TRANDUY, D., VANLIERDE, A., BOL, A., VERAART, C., AND DE VOLDER, A. G. Cross-modal activation of visual cortex during depth perception using auditory substitution of vision. *NeuroImage 26*, 2 (2005), 573–580.

[142] REYNOLDS, Z., AND GLENNEY, B. Interactive Training for Sensory Substitution Devices. *AP-CAP* (2009).

[143] REYNOLDS, Z., GLENNEY, B., AND COLLEGE, G. When Sensory Substitution Devices Strike Back: An Interactive Training Paradigm. *Philosophy Study 2*, 6 (2012), 451–457.

[144] ROSSI, M., BIANCHI, M., BATTAGLIA, E., CATALANO, M. G., AND BICCHI, A. HapPro: A Wearable Haptic Device for Proprioceptive Feedback. *IEEE Transactions on Biomedical Engineering 66*, 1 (2019), 138–149.

[145] SAINZ MARTINEZ, C. *Sensory Integration in Substitution and Augmentation Devices*. PhD thesis, University of Reading, 2019.

[146] SAINZ MARTINEZ, C., AND HWANG, F. TESSA -Toolkit for experimentation with multimodal sensory substitution and augmentation. *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems - CHI EA '15 2* (2015), 259–262.

[147] SALMONI, A. W., SCHMIDT, R. A., AND WALTER, C. B. Knowledge of Results and Motor Learning: A Review and Critical Reappraisal. *Psychological Bulletin 95*, 3 (1984), 355–386.

[148] SCHOONMAKER, R. E., AND CAO, C. G. Vibrotactile force feedback system for minimally invasive surgical procedures. *Conference Proceedings - IEEE International Conference on Systems, Man and Cybernetics 3* (2007), 2464–2469.

[149] SCHUMANN, F., AND O'REGAN, J. K. Sensory augmentation: integration of an auditory compass signal into human perception of space. *Scientific Reports 7*, February (2017), 42197.

[150] SEO, J., MUN, S., LEE, J., AND CHOI, S. Substituting motion effects with vibrotactile effects for 4d experiences. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (New York, NY, USA, 2018), CHI '18, ACM, pp. 428:1–428:6.

[151] SHEPARD, R. N. Geometrical approximations to the structure of musical pitch. *Psychological review 89*, 4 (1982), 305.

[152] SHIN, S., AND CHOI, S. Effects of haptic texture rendering modalities on realism. In *Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology* (New York, NY, USA, 2018), VRST '18, ACM, pp. 42:1–42:5.

[153] SHULL, P. B., AND DAMIAN, D. D. Haptic wearables as sensory replacement, sensory augmentation and trainer - A review. *Journal of NeuroEngineering and Rehabilitation 12*, 1 (2015).

[154] SOHL-DICKSTEIN, J., TENG, S., GAUB, B. M., RODGERS, C. C., LI, C., DE WEESE, M. R., AND HARPER, N. S. A device for human ultrasonic echolocation. *IEEE Transactions on Biomedical Engineering 62*, 6 (2015), 1526–1534.

[155] SPENCE, C. Crossmodal correspondences: A tutorial review. *Attention, Perception, and Psychophysics 73*, 4 (2011), 971–995.

[156] STAFFORD, T., JAVAID, M., MITCHINSON, B., GALLOWAY, A., AND PRESCOTT, T. Integrating augmented senses into active perception: a framework. In *the Royal Society Meeting on Active Touch Sensing* (Kavli Royal Society International Centre, Buckinghamshire, January 2011).

[157] STRONKS, H. C., PARKER, D. J., AND BARNES, N. Vibrotactile Spatial Acuity and Intensity Discrimination on the Lower Back Using Coin Motors. *IEEE Transactions on Haptics 9*, 4 (oct 2016), 446–454.

[158] TECHNOLOGIES, B. Brainport balance plus | united states. `https://www.wicab.com/brainport-balance-plus`. BrainPort Technology | United States | Vision and Balance.

[159] TECHNOLOGIES, B. Brainport vision pro | united states. `https://www.wicab.com/brainport-vision-pro`. BrainPort Technology | United States | Vision and Balance.

[160] TREUTWEIN, B. Adaptive psychophysical procedures. *Vision Research 35*, 17 (sep 1995), 2503–2522.

[161] ULLMAN, M. T. The Declarative/Procedural Model. *Neurobiology of Language* (2016).

[162] VAN ERP JAN, B. F., VAN VEEN HENDRIK, A. H., AND MARK, R. More than a feeling: Bringing touch into astronauts' spatial orientation. *Microgravity Science and Technology 19*, 5-6 (2007), 108–112.

[163] VISELL, Y. Tactile sensory substitution: Models for enaction in HCI. *Interacting with Computers 21*, 1-2 (2009), 38–53.

[164] WAHNL, B., AND KÖNIG, P. Multimodal integration, attention and sensory augmentation? *5th International Winter Conference on Brain-Computer Interface, BCI 2017* (2017), 24–26.

[165] WANG, Y., MILLET, B., AND SMITH, J. L. Designing wearable vibrotactile notifications for information communication. *International Journal of Human Computer Studies 89* (2016), 24–34.

[166] WARD, J., AND MEIJER, P. Visual experiences in the blind induced by an auditory sensory substitution device. *Consciousness and Cognition 19*, 1 (2010), 492–500.

[167] WARD, J., AND WRIGHT, T. Sensory substitution as an artificially acquired synaesthesia. *Neuroscience and Biobehavioral Reviews 41* (2014), 26–35.

[168] WATSON, A. B., AND PELLI, D. G. Quest: A Bayesian adaptive psychometric method. *Perception & Psychophysics 33*, 2 (mar 1983), 113–120.

[169] WHITE, B. W., SAUNDERS, F. A., SCADDEN, L., BACH-Y-RITA, P., AND COLLINS, C. C. Seeing with the skin.

[170] YILDIRIM, D., FRAGUADA, L. E., AND BIGGER, E. E. Dualskin: Ambient electric field sensing wearable. In *Proceedings of the 23rd International Symposium on Wearable Computers* (New York, NY, USA, 2019), ISWC ’19, Association for Computing Machinery, p. 339–345.

[171] ZHAO, S., ISRAR, A., LAU, F., AND ABNOUSI, F. Coding tactile symbols for phonemic communication. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (New York, NY, USA, 2018), CHI ’18, Association for Computing Machinery, p. 1–13.