

**INVESTIGATING, DESIGNING, AND VALIDATING  
A HAPTIC-AFFECT INTERACTION LOOP  
USING THREE EXPERIMENTAL METHODS**

by

THOMAS WILLIAM HAZELTON

B.C.S. (Hons), The University of Waterloo, 2008

B.Ed., Queen's University, 2008

Dip. East Asian Studies, Renison College, 2008

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

MASTER OF SCIENCE

in

THE FACULTY OF GRADUATE STUDIES

(Computer Science)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

August 2010

© Thomas William Hazelton, 2010

## **Abstract**

Computer interfaces commonly make large demands on our visual and auditory attention, which can make multi-tasking with multiple systems difficult. In cases where a primary task demands constant, unbroken attention from the user, it is often implausible for such a user to employ a system for a secondary task, even when desirable. The haptic modality has been suggested as a conduit for the appropriately-intrusive delivery of information from computer systems. Furthermore, physiological signals can be used to infer the affective state of a user without requiring attention. Combining these underexplored channels for implicit system command, control and display, we envision an automated, intelligent and emotionally aware interaction paradigm. We call this paradigm the Haptic-Affect Loop (HALO).

This work investigates the potential for the HALO paradigm in a specific use case (portable audio consumption). It uses three experimental techniques to gather requirements for the paradigm, validate its technological feasibility, and develop the feedback-supported language of interaction with a HALO-enabled portable audio system.

A focus group is first conducted to identify the perceived utility of the paradigm with a diverse – albeit technologically conservative – group of portable audio users, and to narrow its scope. Results of this focus group indicate that participants are sceptical of its technological feasibility (in particular, context resolution) and are unwilling to relinquish control over their players. This scepticism was alleviated somewhat by the conclusion of the sessions.

Next, technological validation of online affect classification is undertaken via an exploratory, but formally controlled, experiment. Galvanic skin response measures provided a means to make introductory measures of interruption and, in some cases,

musical engagement. A richer signal array is necessary to make the full array of required affect identifications for this paradigm, and is under development.

The final phase of work involves an iterative participatory design process with a single participant who was enthusiastic but practical about technology to better define system requirements and to evaluate input and output mechanisms using a variety of devices and signals.

The outcome of this design effort was a functioning prototype, a set of initial system requirements and an exemplar interaction language for HALO.

## Preface

Studies, experiments, focus groups and participatory design activities described in this thesis were performed under the approval of the Behavioural Research Ethics Board at the University of British Columbia. Relevant ethics applications were:

- **#H01-80470** “Low-Attention and Affective Communication Using Haptic Interfaces”
- **#H10-00783** “HALO Participatory Design”

Approval certificates for these applications are attached in Appendix E.

Summarization, planning and analysis of focus group sessions were assisted by Natalie Forssman, a then-undergraduate student working in the laboratory as a summer research assistant.

Detailed, model-based analysis for the data collected in the study described in Chapter 4 was undertaken by Susana Zoghbi, Ph.D. candidate in the Department of Mechanical Engineering at UBC. Susana also wrote Appendix D describing her results and produced Figure 12 in this thesis (used with permission).

The author of this thesis was involved in the following related publications:

Hazelton, T. W., Karuei, I., MacLean, K. E., Baumann, M. A., Pan, M. 2010. **Presenting a biometrically driven haptic interaction loop**. At Whole Body Interaction 2010, a SIGCHI 2010 Workshop. (*Workshop position paper summarizing the candidate's thesis research, presented throughout the thesis*)

Baumann, M. A., MacLean, K. E., Hazelton, T. W. and McKay, A. 2010. **Emulating human attention-getting practices with wearable haptics**. In Proceedings of Haptics Symposium 2010 (Waltham, Massachusetts, USA, March 25 - 26, 2010). (*Describes early investigatory work in human touch emulation, referenced in chapter 2, prototype used in chapter 3*)

Swerdfeger, B. A., Fernquist, J., Hazelton, T. W., and MacLean, K. E. 2009. **Exploring melodic variance in rhythmic haptic stimulus design**. In Proceedings of Graphics interface 2009 (Kelowna, British Columbia, Canada, May 25 - 27, 2009). ACM International Conference Proceeding Series, vol. 324. Canadian Information Processing Society, Toronto, Ont., Canada, 133-140. (*Human perception and inter-stimulus distinguishability research using haptic icons, referenced in Chapters 1 – 2*)

## Table of Contents

<b>Abstract.....</b>	<b>ii</b>
<b>Preface.....</b>	<b>iv</b>
<b>Table of Contents .....</b>	<b>v</b>
<b>List of Tables.....</b>	<b>x</b>
<b>List of Figures.....</b>	<b>xi</b>
<b>List of Abbreviations.....</b>	<b>xii</b>
<b>Acknowledgements .....</b>	<b>xiii</b>
<b>Dedication .....</b>	<b>xiv</b>
<b>1 Introduction.....</b>	<b>1</b>
1.1 An Anecdote.....	1
1.2 Emotionally Intelligent and Expressive Systems.....	3
1.3 Presenting HALO, the Haptic-Affect Loop .....	4
1.4 Overview of Thesis .....	6
<b>2 Related Work.....</b>	<b>8</b>
2.1 Natural and Low-Attention Human-Computer Interaction.....	8
2.2 Affect Sensing.....	9
2.2.1 Affect Sensing in Musical Contexts.....	10
2.3 Focus Groups for Requirements Gathering.....	11
2.4 Participatory Design.....	12

<b>3</b>	<b>Requirements Gathering – Focus Group .....</b>	<b>13</b>
3.1	Motivation, Overview and Research Questions.....	13
3.2	Methodology .....	14
3.2.1	Three-Session Design .....	14
3.2.2	Participant Recruitment.....	15
3.2.3	Scheduling and Remuneration .....	17
3.3	Session-by-Session Summaries.....	17
3.3.1	Session 1 – July 30, 2009.....	17
3.3.2	Session 2 – August 20, 2009 .....	21
3.3.3	Session 3 – September 18, 2009 .....	26
3.4	Results.....	32
3.4.1	Quantitative.....	32
3.4.2	Qualitative.....	34
3.5	Discussion and Open Questions.....	38
<b>4</b>	<b>Technological Validation – Music-Affect Experiment .....</b>	<b>42</b>
4.1	Motivation, Overview and Research Questions.....	42
4.2	Methodology .....	44
4.2.1	Participants.....	44
4.2.2	Overview of Experimental Process.....	44
4.2.3	Pre-Study Activities .....	45

4.2.4	Location of Study and Consent.....	46
4.2.5	Experimental Setup and Sensor Equipment.....	46
4.2.6	Trials .....	47
4.2.7	Measures for Self-Reported Affect Measurement.....	52
4.3	Results.....	53
4.3.1	Assessment of Song Selection .....	54
4.3.2	Visual Data Inspection and Observational Analysis .....	54
4.3.3	Further Observational and Behavioural Analysis .....	59
4.3.4	<i>k</i> -Nearest Neighbours Analysis.....	60
4.4	Discussion and Open Questions.....	61
<b>5</b>	<b>System Integration – Participatory Design.....</b>	<b>66</b>
5.1	Motivation and Overview .....	66
5.2	Methodology .....	67
5.2.1	Participant Recruitment.....	67
5.2.2	Scheduling and Remuneration .....	69
5.3	Session-by-Session Summaries.....	70
5.3.1	Session 1 – Focus Group Activities Revisited .....	71
5.3.2	Session 2 – Music-Affect Experiment Revisited .....	76
5.3.3	Session 3 – Exploring Interruptions and Distractions.....	80
5.3.4	Session 4 – Exploring the Haptic Feedback Design Space.....	86

5.3.5	Session 5 – Exploring Explicit Control and Haptic Feedback .....	94
5.3.6	Session 6 – Wizard of Oz Simulation of a Closed System.....	102
5.3.7	Session 7 – Follow-up Physiological Signal Measurement .....	110
5.4	Toward Design Guidelines for HALO-Enabled Systems .....	111
5.4.1	Loop Input Mechanisms.....	111
5.4.2	Loop Output Mechanisms .....	113
5.4.3	System Behaviour and Interaction Language .....	115
5.5	Discussion and Open Questions.....	116
5.5.1	Projected Customization Process .....	117
<b>6</b>	<b>Conclusions .....</b>	<b>119</b>
	<b>Bibliography .....</b>	<b>121</b>
	<b>Appendix A: Focus Group Materials .....</b>	<b>126</b>
A1	“As is” Scenarios .....	126
A2	“To be” Scenarios.....	128
A3	Sample Messages from Audio Player to User.....	132
A4	Inter-Session Survey .....	132
A5	Session 3 Scenario Set A.....	134
A6	Session 3 Scenario Set B.....	135
A7	Quantitative Survey Results – Intersession.....	138
A8	Quantitative Survey Results – Final.....	139



A9	Consent Form (Modified Formatting).....	140
<b>Appendix B: Affect Study Materials .....</b>		<b>142</b>
B1	Post-Trial Questionnaire (No Word Search Trials) .....	142
B2	Post-Trial Questionnaire (Word Search Trials).....	143
B3	Post-Trial Questionnaire (User Control Trial) .....	144
B4	Consent Form (Modified Formatting).....	146
<b>Appendix C: Participatory Design Materials.....</b>		<b>148</b>
C1	Recruitment E-mail.....	148
C2	Screening Interview Questions .....	148
C3	Consent Form (Modified Formatting).....	150
<b>Appendix D: <i>k</i>-Nearest Neighbours Analysis (written by Susana Zoghbi).....</b>		<b>152</b>
<b>Appendix E: BREB Approval Certificates .....</b>		<b>161</b>

## List of Tables

Table 1: Summary description of focus group participants .....	16
Table 2: Outcome of pain point scenario exercise .....	23
Table 3: Factors preventing HALO adoption.....	26
Table 4: Lingering apprehensions after final focus group session.....	31
Table 5: Inter-session and final survey results .....	33
Table 6: Overview of participatory design sessions.....	71
Table 7: Preliminary tap-based tactor input language.....	97
Table 8: User-defined tactor input language .....	99

## List of Figures

Figure 1: The proposed Haptic-Affect Loop (HALO) .....	5
Figure 2: THMB tactile display, wrist squeezer, and temperature glove prototypes .....	28
Figure 3: Overview of study protocol .....	45
Figure 4: Process for non-word search trials (diagram illustrates a single trial).....	48
Figure 5: Process for word search trials (diagram illustrates a single trial).....	50
Figure 6: User control trial study process .....	52
Figure 7: Affect circumplex diagram as illustrated by Posner <i>et al.</i> .....	53
Figure 8: SCR vs. trial time (no word search, projected “liked” song) .....	56
Figure 9: SCR vs. trial time (no word search, projected “disliked” song).....	56
Figure 10: All “no word search” GSR readings for one participant .....	57
Figure 11: SCR vs. trial time (word search, projected “liked” song).....	58
Figure 12: Recognition rate vs. time window for a subset of physiological features .....	61
Figure 13: SCR vs. trial time (liked, no word search) for participatory design subject....	78
Figure 14: SCR vs. trial time (liked, word search) for participatory design subject.....	79
Figure 15: C2 Tactor .....	87
Figure 16: The Twiddler.....	87
Figure 17: “Plucking” (a), “square” (b) and “rising” (c) tactor waveforms.....	89
Figure 18: Force profile of Twiddler scroll wheel with bookmarking.....	90
Figure 19: Tactor input visualizer and analyzer. ....	96
Figure 20: “Constant tone” (a), “fast pitches” (b), “faster pitches” (c), “slow-fast-fast” (d) tactor waveforms.....	98
Figure 21: Physical setup for Wizard of Oz sessions.....	104

## **List of Abbreviations**

<b><u>Abbreviation</u></b>	<b><u>Full Name</u></b>
BVP	Blood volume pulse
DJ	Disc jockey
EKG	Electrocardiography
GSR	Galvanic skin response (same as SCR)
HALO	Haptic-Affect Loop
ICICS	Institute for Computing, Information and Cognitive Systems
P1 – P8	Participant 1 – 8 (in focus groups)
SEMG	Surface electromyography
SCR	Skin conductance response (same as GSR)
$\mu\text{S}$	Micro-Siemens
$\mu\text{V}$	Microvolt(s)

## Acknowledgements

The author wishes to gratefully acknowledge the contributions of the following people in the creation of this thesis.

- Supervisors **Dr. Karon MacLean** and **Dr. Joanna McGrenere** for their unwavering support, confidence and advice throughout the research and writing process;
- **Natalie Forssman**, whose insights and contributions to the focus group sessions were invaluable to furthering and focusing the present research efforts as well as those of other students in the laboratory;
- **Susana Zoghbi**, whose detailed analysis of experiment data and model creation for real-time affect classification greatly strengthened evidence for the technological validity of the proposed interaction techniques and eased their evaluation;
- **Dr. Cristina Conati** for her time, patience and advice as second reader.

Additional thanks are warranted for **Matthew Pan**, **Gordon Jih-Shiang Chang**, **Gökhan Himmetoğlu**, and **AJung Moon**, who completed a relevant and insightful class project on a similar topic to this thesis and with whom many fruitful discussions and brainstorming sessions were held. **Bradley Swerdfeger**, **Jennifer Fernquist**, **Ricardo Pedrosa**, **Jeff Hendy** and **Steve Yohanan** offered invaluable support, advice and friendship throughout the tenure of my Master's studies.

## **Dedication**

This thesis is dedicated to Victoria Kim, for her ongoing support during the thesis-writing process, and to my parents, who will hang my diploma with pride in their dining room.

# **1 Introduction**

## **1.1 An Anecdote**

While enjoying a sabbatical in New Zealand in 2008, 40-something overworked technologist “Jane” finally got her first modern portable music player, an Apple iPod, and found time to fully explore its use. Using it as a supplement to her morning jogging routine, she discovered many improvements over the sequential-access cassette player she’d given up on 15 years earlier. She loaded music and audio books that spanned several genres and subject areas onto her player, and crafted personalized playlists for different activities and even stages of activities, like easy and intense parts of a workout. The iPod was customized as much as currently possible for Jane; it contained only the content of her choice to be played back in the order of her choice – an order changeable at a personal whim – and it would respond to Jane’s explicit functional demands based on her preference settings.

Jane’s self-reported user experience with her iPod, however, was far from optimal. Audio content often became inappropriate for her changing tastes and contexts throughout the runs, and interruptions to her exercise would require her to stop, remove the player from the arm band in which it was housed, and interact with it using her fingers. If Jane was ready to enter the “cool down” phase of her jog earlier than expected, for example, she would have to manually advance her music to an appropriate accompanying song. Likewise, seeking to stay at her target heart rate for longer than usual would require similar interaction with the device. A passing train would introduce unexpected noise into the listening environment, requiring Jane to manually adjust the volume of her iPod. As the train turned a corner, the volume would need to be manually

returned to its original setting. Stopping to study a map or concentrate on traffic while crossing a busy street, she would miss an interesting bit of a podcast; rewinding to the right point was nearly impossible, usually under- or over-shooting.

Oddly, these were never problems with the old Sony Walkman – it presented no choices beyond “pause” or “play”. What had happened? Increased richness of the media required greater control; but the controls haven’t changed in any essential way beyond nicer abilities to manually scroll through a list or a track. The result is *increased* cognitive demand on the part of the user, exactly when the opposite is desired, and a much greater sense of the user having to serve and focus on the interface, rather than the system doing its job quietly in the background.

This is all too common an outcome with today’s technological devices.

Why are devices unable to “just know” what users want them to do? The technology exists to automate many of the tasks that Jane found so frustrating in her running experiences. Context-aware devices, such as the pioneering Cyberguide [1] and SoundSense [30], are equipped with microphones, global positioning sensors, accelerometers, gyroscopes, video cameras and other technologies to make sense of the world around them and infer desired behaviour. Human physiological signals have been studied for decades, and are beginning to be used in real-time computer systems to help make behavioural decisions based on the affective (or otherwise physiological) state of the user. Given the relatively mature state of sensor and context-aware technologies, Jane (incidentally, a supervisor of this research) guessed that behavioural inference might be a logical solution to her audio player problem, and indeed, a feasible solution for many applications.



## 1.2 Emotionally Intelligent and Expressive Systems

A notable feature of Jane’s iPod anecdote is the existence of multiple conflicting tasks – fiddling with an audio player, focusing on a run, and perhaps even attending to a tertiary interruption – that competed for her attention and caused aggravation. The goal of a technological solution to Jane’s attentional dilemma would be to reduce or remove the effort required to tend to her audio player while ensuring that her perceived level of control over the device remains intact. In this manner, her focus could be maintained on a target task (i.e., running) without experiencing negative impacts on her music-listening experience. Transcending the sabbatical anecdote, for certain safety-critical scenarios that involve multitasking, the benefits of an interaction paradigm that allowed unbroken focus on a primary task are clear.

Providing high visibility of system status to the user is a well-established principle in interface design [37]. As the proposed interaction paradigm shifts system behaviour from the explicit to the implicit, the need for immediate and continuous feedback from the system becomes apparent. Bombarding the user with visual or auditory feedback undermines the goal of reducing the attentional requirements of the user when interacting with the system, leading us to conclude that the haptic channel may be best suited for this purpose: in addition to being underutilized, its proximity and the ways it is used in the natural world make it potentially well suited for background display [32].

Humans are generally experts at using touch to capture attention and deliver messages in subtle and nuanced ways; minor variations in intensity, pressure, frequency, and locus of contact impact the perceived meaning of a touch [19][16]. Based on the context of the recipient of a touch-based “message”, humans possess the capability to

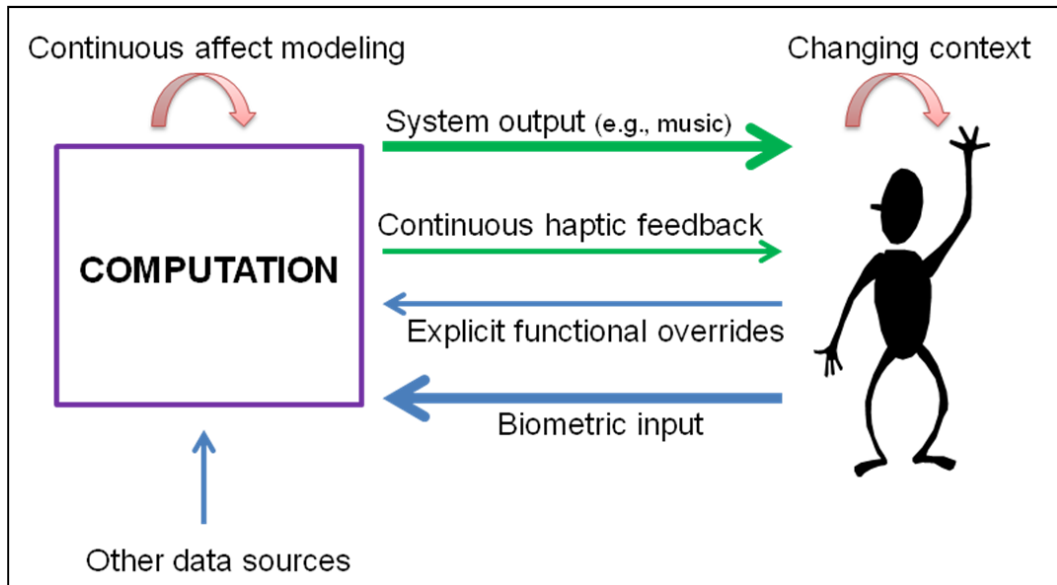
modulate the intensity and intrusiveness of their touches. Previous work shows that humans possess the ability to discriminate and classify haptic signals, and naturally associate these signals with emotions even in the absence of emotional intention [48]. Leveraging this knowledge, we endeavour to use human touch as inspiration for developing a natural, nonintrusive means to display system status in line with the attentional goals of the current work, and expect that users will be able to disambiguate feedback messages with low attentional requirements.

Pairing a system that provides contextually aware haptic feedback with a model of emotional intelligence is a natural step in the evolution of this work. Humans modulate their touches to one another based largely on the perceived emotional state of the recipient; we therefore require an autonomous means of performing affect classification in an online setting to meet the goals of natural touch-based communication. Users experiencing frustration rather than pleasure while using a computer device are likely to interpret, enjoy, tolerate and/or dismiss system notifications in vastly different ways. Returning to Jane’s running scenario, a robust emotional classification system would be able to identify her frustrations and catalyze appropriate behavioural assistance to her tasks. Additionally, even simple, raw physiological data (such as heart rate) in the absence of a robust model could provide useful information for selecting the correct course of action.

### **1.3 Presenting HALO, the Haptic-Affect Loop**

This thesis proposes an interaction paradigm that involves continuous affective state capture and classification from human users, autonomous behavioural decisions to affected applications based on the resultant affective model, and continuous, low-

intrusive feedback of system status via the haptic modality. This paradigm is termed the Haptic-Affect Loop (HALO), and is visualized in Figure 1. Blue arrows indicate input and streams to a system, green arrows output, and red cyclical arrows indicate continuous modeling and contextual changes, both in the software applications integrated with the HALO paradigm as well as the human environment. The thickness of the arrows in the diagram indicate the prominence of the input or output stream in the paradigm.



**Figure 1: The proposed Haptic-Affect Loop (HALO)**

Data sources complementary to physiological sensors can disambiguate desired system behaviour (explicit commands, GPS position, environmental sounds, etc.). Disambiguating channels provide an obvious benefit to the proposed interaction paradigm, and this thesis aims to focus on what can be accomplished using biometric data (as collected from physiological sensors) alone, and what must should (or necessarily) be deferred to secondary channels.

## 1.4 Overview of Thesis

The research presented in this thesis aims to uncover the specific requirements of the proposed HALO interaction paradigm, primarily with respect to a chosen use case (portable audio listening), and to evaluate its utility while honing its scope. A variety of experimental and evaluative methods are employed for these purposes.

A review of **related work** (chapter 2) is done to contextualize and validate the goals of the current efforts with respect to related research in human-computer interaction, engineering and the social sciences. The contributions and limitations of this related work are analyzed with respect to current goals to map out a space for our investigation.

Commencing the research efforts, **focus groups** (chapter 3) are used to first define and narrow the scope in which a HALO-enabled system could be of use. The purpose of the focus groups is to understand where frustrating and distracting interactions plague conventional devices (in particular, portable audio devices), what workarounds are currently being used to mitigate these issues, and gather insights as to how these interactions could be improved using the HALO paradigm.

An **exploratory experiment** (chapter 4) is next used to address questions of technological feasibility with regards to the affect sensing portion of the HALO paradigm. The potential for HALO-enabled devices to identify and classify affective states, startle responses and other physiological phenomena in an online setting are investigated.

Finally, an **iterative participatory design cycle** (chapter 5) is used to hone and evaluate prototypes of the HALO paradigm developed with, and specifically for, a single

user. Using Wizard of Oz (WoZ) testing, various languages of interaction, involving haptic-driven feedback and both implicit and explicit input channels, are evaluated.

**Conclusions** (chapter 6) are drawn at the end of this thesis based on the findings of this three-pronged approach to requirements gathering and interaction design. They involve recommendations for implementing a HALO-style interaction loop in a portable audio system, conjecture generalizability to other use cases, and suggest relevant areas for future work.

## **2 Related Work**

### **2.1 Natural and Low-Attention Human-Computer Interaction**

Previous research has revealed that traditional (WIMP) human-computer interfaces require tremendous amounts of – typically visual – attention. Devices often require validation or correction of behaviour to navigate program states; with intrusive visual and auditory alerts, this produces annoyances that are potentially embarrassing, especially in public contexts [32]. As a mitigating force, Weiser and Brown propose the notion of “calm technology” [52], under which devices “move easily from the periphery of our attention, to the center, and back”. Hundreds of research projects that fall under the umbrella of ubiquitous computing have used calm technology as a model in interaction design (e.g., [18] [53] [36]) and the current work falls in line with the same usability goals. Interestingly, despite this research very little carry-over of the calm technology ideology has been made in the design of modern consumer products, which are often anything but calm.

Haptic feedback has been proposed and implemented as a means to provide low-resolution, low-attention feedback in a continuous manner, allowing for parallel processing across tasks and modalities [31] [46] [49]. Studies have shown the effectiveness of haptic information delivery under workload (e.g., [49] [33]). Humans are able to learn and distinguish large volumes of abstract haptic signals [48] and there is evidence that signals motivated by human gestures are immediately and naturally associated with real-life counterparts [4].

## 2.2 Affect Sensing

Much work has been performed in non-musical contexts to autonomously classify human affective states. Significant work (e.g., [10] [3] [8] [29]) has focused on facial and speech recognition as well as eye tracking, leveraging machine learning techniques for this purpose. In addition, Conati and Maclaren [9] as well as others have built probabilistic models of user affect based on usage patterns of desktop systems which have been validated in testing with self-reported states.

A review of relevant literature reveals that recognition rates for affective or emotional states have been promising using visual and auditory input streams. For example, Yoshitomi *et al.* investigated the feasibility of modeling human emotional expressions using visual and auditory measures [56]. Voices were modeled using a hidden Markov model on various sonic attributes, while facial expressions were modeled using thermal and standard images on trained neural networks. Used together, total recognition rates over five emotional states amounted to 85%. In related work, Zeng *et al.* modeled emotional state using audio-visual input sources, and achieved 91.67% accuracy in detecting positive emotions from males, and 86.67% accuracy for negative emotions using the Adaboost multi-stream hidden Markov model framework [57].

Much work has been done to study physiological markers of human affect for the purposes of augmenting human-computer interactions and work to date has been promising and significant [11]. Using four physiological sensors (electromyography, blood volume pressure, skin conductance, and respiration), Picard *et al.* [41] achieved a recognition rate of 81% on eight classes of emotion (no emotion, anger, hate, grief, platonic love, romantic love, joy and reverence). This work indicates promise for affective classification in musical settings, but does not address contextual issues that

may apply to this particular use case. Similarly, Kulic and Croft [26] utilized a hidden Markov model (HMM) calibrated per-user in an endeavour to classify affective states based on physiological data in real time for human-robot interaction contexts.

This thesis focuses mainly on physiological techniques for affect detection, keeping in mind the potential for model augmentation via other explored means. In particular, classification frameworks that support our goal to require minimal attention from the user while being minimally invasive could be investigated in complementary work.

### **2.2.1 Affect Sensing in Musical Contexts**

Chung and Vercoe [6] developed a real-time music arranging system that selects music on the basis of physical and physiological cues. The goal of the system was to continuously transition the listener to a goal (enjoyable) state based on foot tapping (as recorded by a microphone), GSR, and subjective evaluation data. The disambiguating audio channel being leveraged in this work restricts its efficacy to particular contexts where recordings of foot taps are feasible (e.g., sitting at a desk), but this work gives promise to the efforts of this thesis.

Orienting responses have been associated with GSR in multiple studies [5] [15], however considerably less work to date has involved GSR-affect correlations in musical settings. Since music heavily influences emotions [55] it is not possible to assume that previous results will necessarily apply here; physiological markers on which this previous work relies could be drastically affected by the presence of music. In our own research group, we have extended parallel efforts to examine the use of GSR-based startle responses to drive the interaction of an audio book bookmarking system, and achieved an 84% recognition rate for interruptions [40].



## 2.3 Focus Groups for Requirements Gathering

Focus groups are a common method for gathering high-level requirements from a group of people in the early stages of project conceptualization and prototype development [25] [38]. Their purpose is to understand the mental models, existing practices, opinions and desires of a group of potential users of a technology in an efficient, self-stimulating manner [25]. Using qualitative open-ended interview techniques, they offer the benefits of [24]:

- Rich, qualitative data gathering on the perceptions of proposed technologies from several participants at once;
- Allowing participants to build on each other's comments and suggestions while mitigating the impact of extreme opinions and errors that could arise in one-on-one interviews;
- Allowing participants of differing backgrounds, experiences and technological comfort levels to consider their perspectives in the context of other participants';
- Facilitating dynamic changes to the structure and content of the discussion topics as needed.

Drawbacks of the focus group methodology include:

- The difficulty of many participants to envision or understand radical technological shifts;
- The possibility of group composition to influence discussion in unexpected, unrepeatable and undesirable ways.

Balancing the benefits of the focus group methodology with the drawbacks, they appear appropriate for understanding pain points with portable audio players, but less reliable for

evaluating the acceptance of new approaches to solving technological issues with radical solutions.

## **2.4 Participatory Design**

Participatory design has been utilized as a technique for human-computer interaction design since the 1970s [22]. It heavily involves the end user at all stages of the development of a computer system, from requirements gathering to final evaluation [44]. The philosophized benefits of this approach to system development are that user needs will remain paramount at all times, the benefits of end user knowledge will be directly and easily accessible, and prototypes can be evaluated by important stakeholders iteratively and rapidly [34]. Misinterpretations of requirements on the part of the designer should be rapidly uncovered through discussion or in prototype evaluation, which indicates that this approach is effective for defining the highly nuanced behavioural requirements of HALO.

### **3 Requirements Gathering – Focus Group**

#### **3.1 Motivation, Overview and Research Questions**

Motivated by the need to validate the conjectured benefits of the HALO paradigm over multiple users and scenarios, as well as to inform the direction of further research with respect to the paradigm, a focus group series that was comprised of three sessions was held. The specific research questions that directed the focus group sessions were:

- Are there identifiable frustrations (“pain points”) for users of portable audio devices that the HALO paradigm can mitigate?
- Is the audio use case appropriate for continued research on the HALO paradigm? If not, is another use case that these users would find more appropriate?
- Is the haptic modality appropriate for providing feedback and collecting input from users, and are there more appropriate (potentially mixed-modality) alternatives?
- What obstacles exist that would prevent people from adopting a HALO-enabled portable audio player?

The focus group targeted the portable audio use case in part due to the existence of easily segmented primary and secondary (background) tasks<sup>1</sup>. As the task of consuming content from the portable audio player could be considered either primary or secondary depending on user context and level of attention paid, an attractive feature of choosing this use case allowed for both options to be explored within the sessions.

---

<sup>1</sup> For example, cleaning (primary) while listening to music (secondary), or listening intently to an audio book (primary) while eating (secondary).

The qualitative design of the focus group sessions centred on the goal of identifying participants' "pain points" with respect to their portable audio players. Pain points are defined for current purposes as attributes of portable audio devices that cause the user's experience to be frustrating, ineffective, non-pleasurable or tedious. The aim was to identify features of audio players as well as usage scenarios that focus group participants felt were not well supported by their current audio players, present and discuss potential solutions that involved (and did not involve) the HALO paradigm, and based on participant feedback, present early prototypes of an envisioned HALO-based audio system.

## **3.2 Methodology**

### **3.2.1 Three-Session Design**

We aimed to address our research questions for early requirements gathering using the following three-phase plan:

- **Phase 1** would focus on conducting wide-ranging conversations and exercises with participants in order to uncover and understand their pain points with respect to current portable audio player usage habits and technology in general.
- **Phase 2** would focus on confirming our understanding of participants' pain points via tailored and personalized scenarios and on gathering early feedback on proposed technological solutions.
- **Phase 3** would give us an opportunity to evaluate early prototypes of our proposed technological solutions (created and revised based on the results of Phase 2) and provide direction for further design and evaluation efforts.

On the basis of this three-phase approach, we determined that three focus group sessions, one per phase, would be optimal for our purposes. The sessions were video and audio recorded for subsequent transcription and analysis.

### **3.2.2 Participant Recruitment**

The recruitment strategy utilized for the focus groups was conceived on the desire to bring a wide variety of perspectives – defined over several measures – to the table. Specifically, we wished to achieve representation from diverse demographic groups (age, gender, country of origin), different levels of comfort with information technology and portable devices and variable likelihood to adopt technologies soon after they are made available (be “early adopters”).

Approximately 50 – 75 advertisements were posted in various academic buildings, businesses and public advertising spaces on the UBC Vancouver campus publicizing the opportunity to participate in the focus group sessions. Aiming for eight participants in total, we required all interested parties to complete an online survey for consideration and evaluation. The survey asked prospective participants for:

- Demographic information;
- Feature lists of their current and past portable audio players;
- Common usage scenarios of their players (e.g., exercise, bus riding);
- Qualitative and quantitative measures of self-reported player satisfaction;
- Qualitative and quantitative measures of general comfort levels with technology.

The online survey was made available for approximately one week (from July 14 – July 22, 2009), during which time a total of 29 adequately completed surveys were submitted. From these submissions, eight participants were chosen whose responses fell in a

spectrum across our evaluative metrics. These eight participants were offered a position in the focus group via e-mail and asked to confirm their interest and ability to attend; five accepted our offer. Three replacement respondents were selected from the group of submitted surveys in an endeavour to maintain balance in the group, two of which accepted our offer. One final replacement was made to complete the recruitment process. In the end, 5 females and 3 males were recruited for participation, 3 of which were 18 – 25 years old, and 5 of which were 26 – 40 years old. Participants are coded as P1 – P8 throughout this chapter. Table 1 contains a summary description of each selected focus group participant.

<b>Code</b>	<b>Gender</b>	<b>Age</b>	<b>Occupation</b>	<b>Country of Origin</b>	<b># hours of player usage per week<sup>2</sup></b>	<b>Overall satisfaction with player</b>
P1	F	26 – 40	Journalist	Germany	[6, 10]	Moderately satisfied
P2	F	26 – 40	Nurse/Grad Student	Canada	[0, 1]	Extremely unsatisfied
P3	F	18 – 25	Engineering Undergrad Student	China	[6, 10]	Moderately satisfied
P4	F	18 – 25	Microbiology Grad Student	Canada	[3, 6]	Extremely satisfied
P5	F	18 – 25	Medical Laboratory Technologist	Canada	[6, 10]	Moderately satisfied
P6	M	26 – 40	Arts Undergrad Student	Singapore	[10, 20]	Extremely unsatisfied
P7	M	26 – 40	Consultant	Canada	[6, 10]	Neutral
P8	M	26 – 40	Post-doctoral Fellow	Portugal	[6, 10]	Moderately satisfied

**Table 1: Summary description of focus group participants**

---

<sup>2</sup> [x, y] indicates a lower bound of x hours per week and an upper bound of y hours per week.

### **3.2.3 Scheduling and Remuneration**

Three focus group sessions were held in the Observation Studio of the Institute for Computing, Information and Cognitive Systems (ICICS) building (room X725) on the UBC Vancouver campus, each with the same set of eight participants. The three sessions were held on July 30, August 20, and September 18, 2009. All sessions began at 11:30 A.M. and lasted 90 minutes. Lunch was provided at approximately 12:00 P.M. and participants were compensated \$15 at the conclusion of each session. As an incentive to participate in all focus group three sessions, participants were offered a \$30 bonus for their full attendance (all participants received this bonus). Each participant was thus remunerated a total of \$75 (plus three lunches) for their participation.

In addition to the three focus group sessions, two email surveys were administered: one between the second and third sessions (the “inter-session survey”), and one after the final session (the “final survey”).

## **3.3 Session-by-Session Summaries**

### **3.3.1 Session 1 – July 30, 2009**

#### ***3.3.1.1 Goals and Process***

The goal of the first session was to identify any and all major pain points, as defined above, that plague participants’ current portable audio listening habits.

After introducing the research team and briefing the participants on their ethical rights, discussion began by asking the group how they use their portable audio players as a background entity to some other task (e.g., cleaning, exercising, commuting). Discussion focused on the identification of reasons that audio helped or hindered the completion of suggested tasks. Participants were asked for specific situations in which

the experience of enjoying music on their player was negatively affected, and about the effects of “over-engagement” with their media (such that their focus on other tasks was diminished).

The next thread of discussion focused on participants’ general levels of comfort with technology. Questions focused on technological features that make devices perceptually “easy” or “hard” to use.

A survey was then issued asking participants to rank the importance of specific features offered by their players (e.g., pause, play, repeat, and shuffle) by means of a 5-point Likert scale. Subsequent discussion focused on the features that participants found either very important or very unimportant, as well as features that had fallen out of favour with the evolution of audio player interfaces over the years.

The means by which participants carry or hold their players in mobile and home environments (e.g., in their pockets, mounted on the arm) were then discussed.

Finally, two pieces of novel technology that use non-standard input mechanisms (physiological states and environmental noises) to make semi-automated behavioural decisions (Yamaha BODiBEAT [54], a portable audio player which autonomously aligns detected heart-rate with musical selections, and Dartmouth College’s SoundSense [30], a framework for modeling sounds for context resolution on portable devices) were introduced and evaluated, and the participants were given a blue sky exercise wherein they brainstormed on the following question: “Not thinking in practical terms, what would you really like a portable audio player to do that none that you know of currently can?”



### ***3.3.1.2 Summary of Session***

Most participants indicated they were physically active and use their audio players for commuting relatively frequently. Most of the issues they had with their players were related to form factor and battery constraints, but they were otherwise fairly satisfied with their players. Two female participants, P1 and P2, expressed general frustration with their players upfront and indicated that they use them relatively infrequently, seeing them as “cumbersome”, whereas other participants viewed their players as “essential” to their daily routines. P1 did not identify with the manner in which others use their players, especially in exercise contexts. She indicated that she dislikes the idea of going into her own world and “blocking out” her present surroundings. Most other participants identified with the “blocking out” experience that P1 described, but unlike P1, often welcomed the experience. P3 and P8 agreed that turning off their player is sometimes required when a primary task requires concentration (such as parking a car or working). Several participants suggested the idea of additional categorization or organizational utilities – without prompting – for their players. P2 doesn’t use her player very much and her reasons for this were mostly mechanical (battery life, headphones, low grade player).

Participants varied in their comfort levels with technology and many indicated that they are often convinced by others to try new devices. In particular, only one participant (P3) indicated that she actively learned about new technologies on her own accord, and two participants indicated that they use technology solely as a means to an end.

Participants rated volume, play, pause, seek and playlist-related features as most important, whereas the Genius feature, hold, and fast forward/rewind were explicitly

raised as ineffective by several participants. P3 and P6 indicated that volume controls were cumbersome to use as they require users to retrieve their audio players from where they are stored. P5 agreed with these assessments.

To this point, discussions were largely dispassionate, but participants were interested and engaged. Participants were happy to share their experiences with the facilitators, and had no occasion to react strongly to any line of questioning. However, when the discussion shifted to discussing exemplar technologies (i.e., when BODiBEAT was introduced to the participants) the atmosphere of the discussion changed dramatically. P5 indicated distrust in its abilities and labelled it as “useless”. P1 scoffed at the notion and found the idea “over-controlling” and “dominating”. P3 questioned the necessity of pulse detection in a musical context and called it “horrifying”, noting that “it’s trying to control [her], rather than [allowing her to] control it”. P2 considered the scope of the player’s abilities to be too narrow (only for exercise). P7 opined that the BODiBEAT was “almost a waste of technology” and “[could not] imagine people using it”. P4 indicated that the tool promoted laziness, exclaiming “we can’t even pick our own music?”

Shifting away from discussion of existing technologies, the mood of discussion lightened and participants were less passionate about their responses. The blue sky session primarily uncovered desires for storage capacity upgrades and extended battery life. P8 imagined a “touchless, AI kind of [interaction]” with his player, wherein control would be provided with thought, not touch. P4 imagined voice recognition being utilized to this end. P5 indicated a desire for contextual awareness that would adjust the player to her surroundings, providing safety and informational alerts where appropriate.

### 3.3.2 Session 2 – August 20, 2009

#### 3.3.2.1 Goals and Process

The goal of the second session was to present and validate scenarios that were written to confirm the experimenters' understanding of participants' pain points (gathered from the first session and incorporated into "as is" scenarios), and to present and validate corresponding "to be" scenarios that offer mitigation of these pain points.

Eight "as is" scenarios designed to portray the pain points identified by participants in the previous session were first presented and discussed (see Appendix A1 and Table 2 on page 23 for general themes). An example scenario is given below:

***Scenario 2.** Theresa is listening to her portable audio player while waiting for a bus on a serene corner of her neighbourhood. Her player is in her purse. Once on the bus, she can no longer hear her music due to a raucous group of passengers. Frustrated, she reaches for her player in her purse to adjust the volume, which involves unlocking her player using its touch screen interface. The raucous passengers exit the bus a few stops later, and Theresa wants to reduce the volume of her player, again requiring her reach for it and unlock it.*

Participants were asked to identify those scenarios that reflected their difficulties most adequately. Scenarios that were not chosen at all in this exercise were then discussed in an endeavour to identify reasons that these scenarios were not germane, and participants were given an opportunity to suggest scenarios that were not captured in the initial set of eight. Participants were also given an opportunity to brainstorm on possible solutions to the issues plaguing a scenario of their choice. Following this, a set of 16 "to be" scenarios (eight pairs) were presented (see Appendix A2). Each of the eight "to be" scenarios represented a possible solution to the pain point captured in the corresponding "as is" scenario. One solution in each pair involved the HALO paradigm in some form,

wherein sensed user parameters affected the portable audio player's behaviour at some level. An example scenario involving HALO is given below, which corresponds to the previous example:

***Scenario 2a.** Theresa is listening to her portable audio player while waiting for a bus on a serene corner of her neighbourhood. Her player is in her purse. Once on the bus, she can no longer hear her music due to a raucous group of passengers. Detecting her frustration, her player automatically increases its volume to compensate. The raucous passengers exit the bus a few stops later, and the high volume is no longer necessary; detecting her frustration again, the player returns to its previous volume setting.*

Participants were again asked to choose scenarios that “worked well” for them and scenarios that did not. Discussion on the impact of the HALO-based scenarios on the participants' perceived level of control over their portable players ensued. Discussion of techniques to overcome context issues associated with the HALO solutions followed.

Participants were provided with a list of sample messages and prompts that could be communicated to users of the proposed music player via the haptic modality and asked for their comments (see Appendix A3).

To gauge participants' comfort and preferences in terms of physiological sensor placement (as would be required by the HALO-based scenarios), participants were asked to demonstrate their vision of the HALO system in action using props from a supplied collection. Props included clothes (pants, shirts, etc.), accessories (head bands, bracelets, etc.) and other miscellaneous items (wood block, coins, etc.).

To conclude the session, participants were asked directly what features of their audio player they would like to see automated, what features they would not like to see automated, and for a specific benefit that they saw the HALO interaction paradigm offering.

### 3.3.2.2 Summary of Session

The outcomes of the first exercise are tabulated below:

Scenario #	Theme	# of participants that identified with scenario (scenario reflected their difficulties)
1	Interruption during immersion	5
2	Volume adjustment	2
3	Inappropriate content	3
4	Bookmarking content of interest	4
5	Inadvertent playback	2
6	Bookmarking song of interest	0
7	Content for exercise	3
8	Distraction while driving	1

**Table 2: Outcome of pain point scenario exercise**

Interruption was a common pain point for five participants; they indicated that they did not enjoy having to interact with their player while tending to the source of an interruption. Playback of inappropriate content was also a major theme in this exercise; content that was inappropriate for the participants' contexts or moods would be selected for playback, requiring interaction to adjust the player. No participants identified with scenario 6, which involved mentally noting the name of a song that the user was listening to in order to return to it in the future. Participants indicated that they would already be familiar with the content on their player, and that the playlist feature already sufficiently mitigated this issue. To address the pain points that participants identified with (as tabulated above), they created the following solution space in their brainstorming effort:

- Additional, more convenient controls for the player mounted on headphones, wires, or the device itself;
- Additional organization functionality;
- A tagging feature that would allow users to tag content for future consumption;

- Voice control.

When the “to be” scenarios were distributed to the group, reactions were strong and immediate. P3 was immediately suspicious that a device could “detect frustration”, and both P3 and P6 indicated that scenarios involving auditory feedback would only serve to deepen frustration. P1 said that frustration detection was “scary” and “hoped it would be difficult” to facilitate technologically. P2 pointed out that not all situations involving increased frustration warranted a change in behaviour by the player, and P8 indicated that incorrect decisions could easily be made by a HALO-enabled player due to a lack of context. P8 summarized his feelings about the HALO-enabled scenarios as follows:

*Through most of them, I don't like the idea of the iPod trying to be smart for me and detecting what I am doing and automatically doing something. But there was one that I really liked...the heartbeat one...If I decide to pre-set it to something, I feel like I have control. It's doing something that I actually programmed it to do. You know you're going to go run, and you know your heartbeat is going to increase, so you say, “when I reach this point, do that.” I thought that was really good, I liked that.*

P8 indicated that users’ anxiety levels could increase with a perceived lack of control over a portable device, a sentiment echoed by P5. P6 suggested that a perceived lack of control could be mitigated by allowing toggling between “manual” and “automatic” modes of the player. P4 suggested an “undo” feature to allow erroneous decisions to be reversed quickly, and a similar “panic button” that would stop the player was proposed by P8. P2 suggested that adding contextual cues could partially quell control issues.

When feedback (messages or questions) was proposed to the group as a means to address their concerns of trust and control with the HALO solutions, P5 indicated that a soothing notification would be welcomed. P3 imagined “pop-up” messages of some

form (which she imagined as a short beep) that would do nothing in the absence of a response, where P5 preferred a kneading motion that would not be “insistent” or delivered via audio. She suggested that “tapping” sensations would probably be annoying. P1 restated her discomfort for any solution that involved detection, likening it to a form of “manipulation”.

When the sample haptic messages were presented to participants, P1 expressed that she did not like the idea of “a little gadget telling her long stories” and that any message delivered to her should be important. She said that any message delivered by the device would distract her and demand her attention, and that machines touching her would, over time, reduce her situational awareness and numb her senses. P5 and P8 indicated that they would prefer warning messages to behavioural changes that would require no input from the user. P5 was interested in the notion of the device “apologizing” in the case of an error, as this would increase her happiness despite the “relationship” that this involved being “irrational”. P8 was concerned that “social niceties” would interfere with the interaction experience and could be potentially distracting.

Several participants indicated that any wearable sensor machinery should be minimally interfering and not cause harm to social relationships (for reasons of embarrassment, etc.). P1 rebutted that the pervasive use of even currently available audio players portrays a lack of social interest and integration on the part of the user.

Participants were finally asked to name the primary factor that would prevent them from adopting HALO-based technologies. Responses are tabulated in Table 3.

<b>Participant</b>	<b>Factor</b>
P1	“Over-manipulating”
P2	(no comment)
P3	Form factor and appearance
P4	(no comment)
P5	Price
P6	Accuracy
P7	Accuracy
P8	“Overstimulation and over-annoyance”

**Table 3: Factors preventing HALO adoption**

When asked what augmented functionality the participants would enjoy in their players, none involved the HALO paradigm except (potentially) automated sleep detection.

### **3.3.2.3 Inter-Session Survey**

As participants reacted strongly and negatively to the prospect of having their affect detected and to having their audio devices controlled in this manner, the goals of session 3 were reworked in an effort to uncover the specific reasons for these negative reactions. An inter-session survey was administered via email allowing participants to articulate their concerns, scepticism, and preferences with regards to sensing, user modelling, control and haptic messaging (see Appendix A4). All participants completed and returned this survey. Questions and results are compiled in Table 5.

## **3.3.3 Session 3 – September 18, 2009**

### **3.3.3.1 Goals and Process**

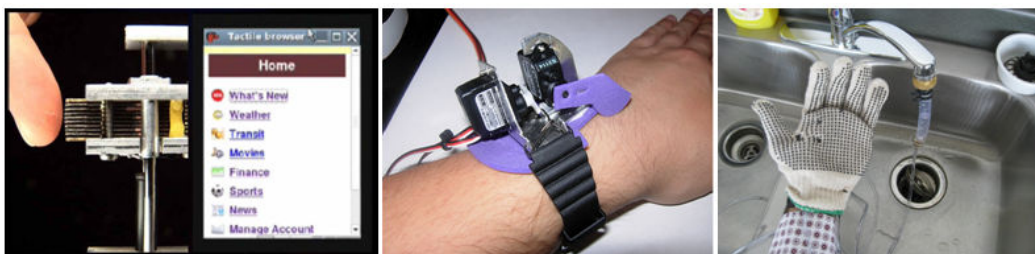
The goal of the third and final focus group session was originally to present and validate early prototypes of a HALO-based audio player, allowing further honing of interaction requirements for the HALO paradigm. The outcome of the previous session caused this focus to change to identify the specific areas where participants’ scepticism lay with regards to affect sensing, haptic feedback, and other aspects of the HALO



paradigm. By demonstrating the design space for haptic technologies and showing examples of working affect-enabled systems, the aim was to quell uncertainty and promote imagination in final discussions. In particular, we aimed to shift our focus away from the reliability of the technology to its potential benefits.

Participants were first informed of the research goals of the entire focus group series and were informed that their strong reactions to the proposed HALO-based scenarios had been noted. Participants were then informed that the focus of the final session would be to identify the specific sources of their scepticism and discomfort with the different aspects of the propositions. Participants were shown a short video demonstrating the capabilities of the Emotiv® EPOC headset [12] and were then split into two groups of four for technological demonstrations of research-stage and production-quality haptic displays and affect-enabled technologies (Figure 2). In total, these groups visited four stations in two rooms (with one group occupying a room at a time).

A haptic arm band with magnetic tappers and haptic squeezing bracelet were demonstrated in one room. The THMB tactile display [27] (which was also demonstrated using an ordinary comb) and Star Wars “Force Trainer” mind control toy [50] were demonstrated via slide show in the other room. Pictures of clunky affect sensors, with large form factors, connecting wires and electrodes were juxtaposed with the Force Trainer, consisting only of a helmet unit, to illustrate the vast differences between a prototype and a commercial product. A low-fidelity prototype of a variable-temperature glove (which operated by piping hot and cold water from a faucet around wearers’ fingers) was also demonstrated in the second room.



**Figure 2: THMB tactile display, wrist squeezer, and temperature glove prototypes**

Following these demonstrations, a set of scenarios (see Appendix A5) that involved automatic adaptation of systems (not portable audio players) based on the perceived needs of the user were presented and discussed. A discussion on real world disc jockeys (DJs) vs. an autonomous HALO-based DJ followed to uncover potential reasons for human-to-machine distrust in the face of human-to-human trust. An adapted set of scenarios from the second session (designed to avoid some contextual ambiguity in the original set) was presented and evaluated by participants (see Appendix A6). An example scenario from this set follows:

***Scenario 1b.** Theresa is listening to her portable audio player while waiting for a bus on a serene corner of her neighbourhood. Her player is in her purse. Once on the bus, she can no longer hear her music due to a raucous group of passengers. Detecting the increased ambient volume, and noting her preference to be deeply immersed in her audio, her player automatically increases its volume to compensate.*

A final discussion to expose the reasons for participants' remaining apprehensions, which would be at this point impacted by a better understanding of technological possibilities, closed the focus group. Participants were asked to imagine various aspects of the HALO paradigm working perfectly and to articulate any reasons why the technology would still not be desirable.

### **3.3.3.2 *Summary of Session***

Participants gave generally positive comments on the haptic display mechanisms that were presented to them. P8 said that, aside from the temperature-controlled glove, the devices were “informationally rich”. P2 indicated that the comb was effective at providing subtle indicators to users.

P1 had positive impressions of the haptic displays but was concerned that her sense of touch would be marginalized if she was “always wired” to a haptic device. P4 suggested that haptic-enabled devices could be distracting and potentially dangerous (when driving, for example) but both she and P1 agreed that they would offer a greater measure of safety than visual interfaces.

Returning to affective sensing, P8 restated his distrust in the device being able to resolve context to make appropriate decisions. “Even if [a HALO-enabled system] detects your emotions right, it may not ... [do] what you want.” P2 suggested that restricting or segmenting the contexts of use (e.g., to exercise only) would help resolve these contextual issues affecting behaviour, and P3 suggested an on-device preference management system to facilitate this.

After presenting the modified session 2 scenarios to the participants, P5 commented that they were “much more favourable” offering “nothing inherently objectionable”. P3 commented that she liked the proposed automatic features but questioned the need for feedback in every case, especially in light of changes to the audio channel. “I personally prefer that it just does it, I don’t need to know.” P8 said that the revised scenarios eased his control-related apprehensions, as they gave him the ability to determine when the portable audio player would make decisions on its own, and when it would not. He suggested that implementing a sensitivity level control for activating

autonomous behaviour – i.e., one that would act only if a certain threshold of certainty of action was reached – would increase his level of satisfaction. P3 indicated that there were still contextual issues that were not addressed by the revised scenarios, and gave an example where a user's annoyance could be directed at some entity outside of the system's modelling capabilities. P4 and P2 suggested an easily accessible input mechanism (P2 proposed a wristband) that would allow users to “confirm” annoyance in this case.

P8 said that he would trust a radio DJ that made a good song selection, and expect that he or she would make good decisions in the future. P1 and P5 indicated that human DJs can learn and grow, “surprising” listeners with content that they might like as new songs are released. P8 likened DJing to writing a story, something that a human would be far better at doing than a machine. He said that to him, “music is much more than just the instruments that are being played” and that based on his understanding, musical detection systems would be unable to replicate the emotional projections of a human DJ. He said that if he did not enjoy the music being played by a particular DJ, he would consider himself “incompatible” with him or her, and try a different radio station. At the same time, P8 saw utility in the Amazon recommender system, and said that his trust in the system continued to increase as he trained it with his preferences. He said that “it prompts [him] to try new things”, but conceded that friends could do that for him and it would be “much more interesting”. P5 summarized that people are far more forgiving to each other when they make mistakes than they are to computer devices.

The final exercise asked participants for any new or existing apprehensions with the technology, first with regards to their understanding of the capabilities of sensors and

feedback mechanisms, and then under the assumption that the technology could work perfectly. The outcomes of this discussion are tabulated below:

Participant #	Apprehensions	Apprehensions given “perfect” technology
P1	Potential misuse of the technology (privacy), risk of sensory degradation	No potential for inclusion of “new content”, “too easy”
P2	(no comment)	(no comment)
P3	Potential misuse of the technology (privacy)	No practical benefit, encourages laziness
P4	Potential annoyance from haptic feedback	(no comment)
P5	Weight, appearance	Style, privacy, battery life
P6	Accuracy	(no comment <sup>3</sup> )
P7	(no comment)	(no comment)
P8	Humans are too complex to model	(no comment)

**Table 4: Lingering apprehensions after final focus group session**

As a final point of discussion, participants suggested that the HALO paradigm would likely be well suited to navigation or other assistive tasks for disabled persons.

### **3.3.3.3 Final Survey**

Following the final focus group session, a final survey (containing a subset of questions from the inter-session survey described above) was issued allowing participants to indicate any shifts in their opinions on sensing, user modeling, control and haptic messaging after the final session. Seven of eight participants completed and returned this survey; questions and results are compiled in Table 5.

---

<sup>3</sup> In this table, “no comment” implies that participants volunteered no opinion.

Part of the motivation of the final survey was to verify the possibility that participants were “being nice” in their more positive feedback at the end of the final session (or that simply, they were not comfortable continuously repeating the same comments); we wanted to give them a chance to express their true opinions with less pressure using a familiar survey. Participants had already been fully compensated when the survey was administered, and were simply invited to respond.

### 3.4 Results

#### 3.4.1 Quantitative

The following questions were asked of participants in both the inter-session and final surveys. Participants were asked to specify how strongly they agreed with the following statements. Possible options were *Strongly Disagree (1)*, *Disagree (2)*, *Neutral (3)*, *Agree (4)*, *Strongly Agree (5)* and *I’m Not Sure (0)*. The values beside each question represent the average response (represented numerically according to the values given above) for the 8 participants. The raw responses are summarized in Appendixes A7 and A8.

Question	Inter-Session Mean (Stdev)	Final Survey Mean (Stdev)	$\Delta$ Mean (Stdev)
I dislike the idea of having my body signals detected.	3.8 (1.6)	3.4 (1.7)	-0.4 (+0.1)
I dislike the idea of having my body signals used to control a device.	4.3 (0.7)	3.7 (1.3)	-0.6 (+0.6)
I dislike the idea of having my body signals correlated to my “emotional state”.	4.0 (1.3)	3.9 (1.3)	-0.1 (0)

(continued on next page)

(continued from previous page)

Question	Inter-Session Mean (Stdev)	Final Survey Mean (Stdev)	$\Delta$ Mean (Stdev)
I am sceptical that my body signals can be correlated to my “emotional state”.	4.0 (1.0)	3.4 (1.1)	-0.6 (+0.1)
I am sceptical that the computer can reliably correlate my body signals to my “emotional state”.	4.9 (0.4)	4.1 (0.7)	-0.8 (+0.3)
I am sceptical that the computer can reliably use the information from my “emotional state” to do something useful for me.	3.9 (0.8)	4.3 (1.1)	+0.4 (+0.3)
I am sceptical that I would feel sufficiently “in control” of my audio player when I am controlling it partly through my body signals.	4.5 (1.1)	4.1 (1.2)	-0.4 (+0.1)
I am sceptical that I would be able to understand what a device is telling me through my sense of touch.	3.1 (1.5)	2.7 (1.4)	-0.4 (-0.1)
I am concerned that the device will need constant input from me to confirm decisions.	3.5 (1.3)	3.1 (1.2)	-0.4 (-0.1)
I am concerned that the wrong decisions would annoy me.	4.8 (0.5)	4.7 (0.5)	-0.1 (0)
I am concerned that I would have to stop what I'm doing to fix the player.	4.0 (0.8)	3.9 (1.1)	-0.1 (+0.3)
I imagine that system feedback that uses the sense of touch would be annoying.	3.6 (1.3)	3.0 (1.1)	-0.6 (-0.2)
I imagine that system feedback that uses the sense of touch would be distracting.	3.5 (1.2)	3.0 (1.0)	-0.5 (-0.2)
I imagine that system feedback that uses the sense of touch would be invasive.	3.3 (1.5)	3.0 (1.4)	-0.3 (-0.1)
I am unwilling to give the device time to learn about my body signals to better understand what I want to do.	3.1 (1.5)	2.7 (1.3)	-0.4 (-0.2)
To me, the proposed technology is of little value.	3.4 (2.1)	3.1 (1.1)	-0.3 (-1)
I wouldn't want to wear any extra peripherals to let this technology work.	3.3 (1.0)	3.1 (1.2)	-0.2 (+0.2)

Table 5: Inter-session and final survey results

Non-bracketed values highlighted in green indicate decreases in numerical response value to final survey statements (as compared to inter-session values), which corresponds to a larger degree of disagreement with the statements on average. Bracketed values indicate calculated standard deviation. Only one statement saw an increase in average response value (highlighted in red). As all statements were phrased in the negative, these results suggest an overall positive shift in scepticism, concern and fondness towards HALO after the third focus group session. In many cases,

however, the shifts were quite small, and only two statements saw a shift past neutral (i.e., values less than 3) into “positive” territory. Notable also is an increase in standard deviation for eight of the questions, with a decrease or zero change for the remaining eight. Fondness of haptic system feedback can be inferred from the reduction in both average and standard deviation for all related questions. Despite a reduction in average for most scepticism-related questions, participants tended to vary more substantially in their responses to these questions (as suggested by an increased standard deviation), indicating a lack of consistency.

### **3.4.2 Qualitative**

#### ***3.4.2.1 “Pain Point” Framing and Group Dynamics***

The aim of the recruitment strategy for the focus groups was to bring a broad spectrum of opinions forward to offer input on the audio player use case of HALO. The group dynamics of the focus group sessions, however, brings this general approach into question. Analysis indicates that framing the audio use case for HALO in terms of “pain points” was largely an unsuccessful strategy to convince participants of the utility of the proposed device. The suggestion that it would “fix” some problem they currently experience was met with strong resistance, with some participants claiming that this implied “laziness”.

The pain points experienced by the participants were varied and less severe than the anecdotal example given in the introduction; a 40-something technologist with no time or desire to administer her player may have given different responses; the demographics of our focus group may have been the root cause for the lack of success with our focus on pain points.



*“We can’t even pick our own music?” – P4*

*“To me, it doesn’t take that much effort to just take the thing out and change the song. We’re reducing our exercise as much as possible. We don’t have to be that lazy.” – P3*

Not wanting to appear “lazy” appeared to be a significant motivation for many participants to side with strong opponents to the proposed technology. Some participants’ strong distrust in the system to “make the right decisions” could have swayed those who saw potential benefits in the technology. P1 had particularly strong objections to the technology on the bases of privacy, invasiveness and unnaturalness, which appeared to spread to other participants.

It was made clear during the sessions that few “early adopters” of technology were in fact recruited for the focus group sessions, making it difficult to frame the discussion in terms of the novel or interesting attributes of the proposed affective technology. As a result, the experimenters faced trouble getting them to take a “leap of faith” and imagine with us the kinds of interactions this technology could make possible. Out of necessity sessions centred on discussing the technology in terms of practical utility which appeared to be a non-starter for the participants.

HALO-enabled devices appeared to be impractical for the focus group participants because they had no trouble imagining situations where it would interrupt, alter, or complicate their current style of interaction with their audio player. They required a “real” problem with the current interaction in order to be willing to try an alternative, and a sufficiently strong issue that fit this requirement could not be identified. It appeared to them that there was much more to lose with HALO than to be gained.

Participants in general had a more positive response to the proposed haptic communication aspect of the HALO paradigm. The signals were “informationally rich” (P8) and many participants indicated that they wouldn't be bothersome.

*“I love the ways of communicating, and I don't think they're startling or any of that. My problem wasn't the thing communicating with me, but more about the interpreting and making a decision for me. Some of these ways of communicating feel great.” – P8*

*“They don't bother me or disturb me. I think that if I was working or concentrating on something, if I wanted to ignore it, it's subtle enough that I think I could keep doing what I'm doing. I'm converted.” – P8*

This opinion was not, however, shared universally among participants:

*“I like how it automatically does some things, but I'm not sure about it always giving me a feedback. I personally prefer that it just does it, I don't need to know.” – P5*

*“I'm still very sceptical about the whole idea of using our senses and making them probably a little less sensitive...I can imagine that when you always have those little touches, you are not sensitive to other touches anymore.” – P1*

#### **3.4.2.2 Affect Detection, Modeling and Trust**

Participants were sceptical about exactly what was being detected and how it was being detected. They were unconvinced that these signals would be sufficient to “infer” emotion, but were also willing to defer authority on this matter to experts.

*“My biggest thing is that I think humans are complex. Machines can be elaborate, but humans are complex and subtle. That's just it. Even if it perceives my emotions right. Unless I decide to hand control to the machine, I would like to keep control and make my decisions.” – P8*

Participants worried that the affective model that transferred their physiological signals into emotional states would lack the context to make accurate and appropriate diagnosis of their emotional state. Particularly, they worried that their emotional state could not in contextual isolation prescribe action on the part of the device. They

questioned the reliability of the device to make the “right choices” and gracefully recover from errors.

*“What happens when you’re listening to a song you like and you’re annoyed about something else? It’s not really interpreted incorrectly. You are annoyed, just not at the song. I don’t think it’s possible to solve unless you track someone else’s thoughts.” – P3*

(in response) *“I’m not OK with that.” – P5*

Participants also worried about how reliability would be assured without over-distracted from informational prompts. Specifically there were concerns that in order to make the correct decisions, the device would constantly be irritating them with questions or “messages”, ultimately requiring more attention than the interface of their present audio player.

#### **3.4.2.3 Device Control and Privacy Concerns**

Participants worried that, since the model for action by the device was unknown to them, they would not feel “in control” of their device. Beyond this practical level of control, there were fears and suspicions about uses and abuses of the private information being collected, especially in the absence of knowledge about what was being used to drive system models. While the utility of the haptic-affect loop concept was comprehended conceptually, their instinctual reaction to the idea of the device having affective information and executing actions unbeknownst to them was often viewed as disturbing and undesirable. Issues around security, privacy, and historical allusions to so-called “controlling” technologies all came up in discussion.

*“I think any technology can be misused. I’m also sceptical about the degradation of our senses and our ability to sense. It might be a process where machines are changing us, we are adapting too.” – P1*

*“I wouldn't like that at all, I think the whole detecting thing is just scary. I don't want to be wired to something that detects my heart rate or anxiety. Maybe it's a [cultural] thing...You don't want to be manipulated.” – P1*

It is notable again that the extent to which users would “put up with” the lack of control over their device and worry about issues of privacy could vary by demographic.

#### **3.4.2.4 Alternative Use Cases**

In prompting the focus group participants to consider alternative applications for the device, there were suggestions that the audio use case was off base. Particularly, they cited an enjoyment in the experience of choosing music, and thought this choice to be fundamentally “human” and inappropriate for a computer to take over. They questioned the motivation for over-complicating their existing audio interface given that it works “fine” for their purposes.

### **3.5 Discussion and Open Questions**

Before the final focus group session, the experimenters saw cause to reformat the methodology to respond to a number of difficult questions and concerns. This process involved an in-depth look at criticisms that were likely to be – at least in part – shared by any future HALO user or experimental subject. By thoroughly exploring the specific sources of scepticism, discomfort, and weariness in our participant pool, insights to address these concerns in the iterative development of prototypes and further experiments were made. Although it was determined that the group that resulted from our selection process was composed of few early adopters (only 1 of 8 agreed with the statement “I am usually among the first to acquire new technology”) we cannot simply assume that, in the case the focus groups were re-run with different participants, outcomes would have significantly differed.

Previous work by Kamar *et al.* [20] shows that users of a computer system are less likely to accept an interruption (when given a choice) from what is perceived to be an automated agent than what is perceived to be a human agent when the benefits of the interruption is not clear. At the same time, users are equally likely to accept an interruption from either when the benefit is clear (whether low or high). Framing present focus group results in this context suggests that participants did not attach enough value to potential interruptions from their audio players to make the interaction worthwhile. The contexts in which interruptions should be made should be more thoroughly investigated in future work.

Quantitative results show a positive shift – although not drastic – in attitudes between surveys, which suggests that exposing participants to technologies related to the proposed HALO interaction was important and persuasive. A lack of exposure to related technologies prior to the final session meant that discussions were largely hypothetical. This, combined with the usage of words like “detect” and “control” to describe the action of the device, made participants uncomfortable with the technological propositions. At the very least, the participants had trouble imagining the device being useful or effective, but exposure to tangible instances of the technologies involved gave the participants an appropriate context in which to consider the HALO project. The results of the focus group sessions suggest that next steps in the project should involve less “telling” and more “showing”: by exposing people to the tools being proposed, discussions that involve imagining what is possible with HALO will become more relevant and grounded, and as a result, scepticism should be reduced.

Each of the research questions posed at the outset of this chapter is addressed in turn:

- Are there identifiable “pain points” for users of portable audio devices that the HALO paradigm can mitigate?
- Is the audio use case appropriate for continued research on the HALO paradigm, or is there something more appropriate?

Pain points were identified in the context of current audio player usage, but these were largely insufficient for participants to immediately adopt the proposed technology due to lingering privacy, control, context and trust issues. The pain points uncovered were largely not significant enough for many participants to ignore these concerns altogether, but there was evidence that users would adopt HALO in a highly configurable and compartmentalized form if it was reliable. Such a system would need to support manual overrides of automated decisions, minimize attentional demands and be transparent about its decisions via continuous feedback. Classification of the appropriateness of musical content being played as well as intelligent responses to interruptions were the clearest pain points captured from participants (suggested by Table 2, strengthened by subsequent discussion) and hence the strongest candidates for automation.

- Is the haptic modality appropriate for providing feedback and collecting input from users, and are there more appropriate (potentially mixed-modality) alternatives?

Focus group participants supported low-attention informational displays and reacted positively to the prototype haptic devices presented to them. They did not want to be distracted by their device while using it, which would preclude audio prompts from being practical in audio settings. Participants wanted to avoid retrieving their device from their pockets or purses for interaction, which suggests that (at least primary) visual displays would also be ineffective.

- What obstacles exist that would prevent people from adopting a HALO-enabled portable audio player?

Contextual issues, distrust in sensing technology, a desire to maintain full control over a system, form factor and cost issues were obstacles for our participants. This indicates that future work should focus on determining the technological feasibility of real-time affect detection and classification, defining an interaction language that supports a high level of perceived control and value, and uncovering ways to disambiguate contexts. This interaction language should be developed in pursuit of resolving the most significant pain points identified for these and other users examined in the future, as well as exploring potentially new features for audio players that participants could not imagine. In addition, the goal of producing prototypes (with as much autonomous functionality as possible) should be kept in mind to ensure that potential users are exposed to the technological possibilities HALO offers. These next steps are addressed in the following two chapters.

## **4 Technological Validation – Music-Affect Experiment**

### **4.1 Motivation, Overview and Research Questions**

The utility of the proposed HALO interaction paradigm relies heavily on the technological feasibility of real-time affect detection and classification. Focus group participants indicated that robust identification of continuously changing musical preferences (over the course of a music listening session) is required for the HALO paradigm to be effective in the portable music use case. In some scenarios (such as content bookmarking), startle responses or mind wandering must also be expediently and accurately identified. Scepticism of the proposed system to accurately make these detections and use them in a non-intrusive manner was common in the focus group sessions; the natural questions to be addressed in continuing research on the portable audio use case were thus whether or not the system can indeed make these detections with sufficient reliability, and which physiological channels provide the most reliable data for these detections.

While previous work has correlated physiological signals with orienting responses<sup>4</sup> and affective states (e.g., [5] [41]), no such work has been done in musical contexts without additional disambiguating system inputs (such as microphones, etc.). Since music heavily influences emotions [55] it is not necessarily reasonable to assume that previous results will apply; physiological markers on which this previous work relies could be drastically affected by the presence of music.

The primary goal of the technological validation experiment described in this chapter was to determine whether a prototype affect classification tool could be rapidly

---

<sup>4</sup> Orienting responses are defined for current purposes as a human's autonomous response to a stimulus that is insufficient to cause a startle response.



generated for subsequent use in the design, development, and evaluation of the HALO paradigm with respect to portable audio. Assuming such a tool could be developed, tremendous evidence for the utility and practicality of the paradigm would be afforded. Since focus group members' scepticism of the paradigm was relieved somewhat after demonstrations of biometric and haptic technology (session 3), further relaxation could be expected with a working – albeit preliminary – HALO model.

A secondary goal of the experiment was to collect further qualitative information to direct the interaction design of HALO prototypes for portable audio. The experimental design was performed to allow participants many opportunities to describe their feelings about music being played, and indicate situations in which they would like the player's behaviour to change. As the focus group sessions offered no opportunities for participants to actually experience frustration or any other emotion while listening to music, the qualitative data collected in this experiment provides a wider perspective on potential uses of HALO.

In summary, the specific research questions that the technological validation study aimed to address were:

- Can the level of enjoyment of a song be inferred using physiological data? How reliably?
- What physiological measure(s) are most promising for inferring correct system behaviour?
- To what extent is the reliability of the physiological measures affected by a secondary task (such as completing a word search or reading a newspaper)?
- For what reasons do music listeners wish to change a track they are listening to, even if they like what is playing?

## **4.2 Methodology**

A formally controlled but exploratory experiment was undertaken to address the research questions of this work. It is worth noting explicitly that the purpose of this experiment was not to test a pre-existing hypothesis about the feasibility of real-time affect classification; instead, it was undertaken to better understand the impact of chosen stimuli (music, secondary task) on the dependent measures (physiological metrics). We aimed to use a variety of analysis techniques, including eyeball observation as well as more formal approaches to analysis, to this end.

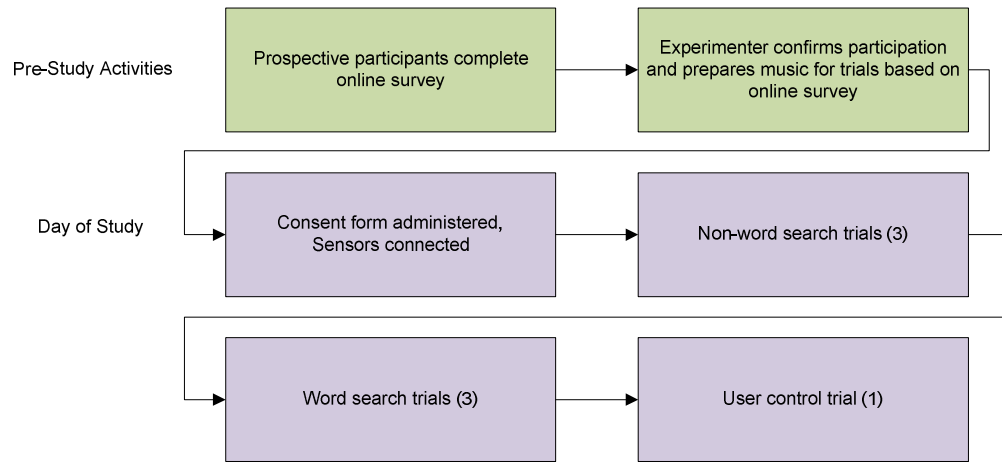
### **4.2.1 Participants**

12 participants (7 male) were recruited via e-mail invitations for a single one hour session and compensated \$10 for their efforts. The invitations were sent to academic UBC mailing lists and to former prospective participants of the focus group that were not offered a seat. Eight participants were aged 18 – 25, three were aged 26 – 40, and one was over 41. Most subjects were researchers or undergraduate students at the university.

### **4.2.2 Overview of Experimental Process**

The experimental process involved first capturing participants' musical preferences by means of an online survey, then collecting physiological data during exposure to music that the participants' were expected to like, dislike and feel neutral about (based on the survey). Data was collected with and without a secondary cognitive task (a word search). Following this, a trial was run wherein the user was given partial control over the music player (insofar as he or she could advance tracks at will). The

high level experimental process is illustrated in Figure 3; each stage of the experiment is discussed and illustrated in more detail in the sections that follow.



**Figure 3: Overview of study protocol**

#### 4.2.3 Pre-Study Activities

A preliminary online survey was issued to prospective participants that collected:

- Demographic information (including age and country of origin);
- A subjective ranking of musical genres (as identified on a 6-level Likert scale: hate, dislike, neutral, like, love, never heard);
- The names of participants' three "favourite" songs and of three that "annoy" them;
- The name of a the participants' "favourite" artist/group and of one whose music "annoys" them;
- A freeform description of participants' imagined reaction if subjected to their least favourite genre of music and were unable to control it.

The data collected in this survey was used to personalize the experiment for each participant. The songs that were hypothesized to be liked and disliked were selected on the basis of the genre question and particular songs and artists articulated by the participants.

#### **4.2.4 Location of Study and Consent**

The one hour sessions were run in a noise- and interruption-controlled environment in the Usability Lab (X727) of the ICICS building on the UBC Vancouver campus. Participants were briefed on their ethical rights upon arriving and provided consent at the outset of the session. All participants consented to having their sessions videotaped for subsequent analysis.

#### **4.2.5 Experimental Setup and Sensor Equipment**

Following ethical briefing, the following six physiological sensors from Thought Technology [51] were connected to the participants:

- Respiration amplitude, measured on a relative scale;
- Electrocardiography (EKG), measured in  $\mu V$ ;
- Surface electromyography (SEMG), measured in  $\mu V$ ;
- Skin conductance (SCR), measured in  $\mu S$ ;
- Blood volume pulse (BVP), measured on a relative scale;
- Peripheral temperature, measured in  $^{\circ}C$ .

Data from these sensors were collected using Thought Technology's FlexComp hardware system, using the FlexComp encoder connected via USB to a laptop. All signals were recorded at 256 Hz. Finally, the Emotiv® EPOC headset [12], a neuro-signal acquisition unit with 14 saline sensors, was placed on the participants' head and aligned with target

points on the scalp and forehead. The complete array of available sensors was selected to collect data from as many channels as possible for subsequent analysis.

A second laptop was used to play music for experimental trials through a pair of headphones, using YouTube videos and Apple iTunes® software as the source of media content. This laptop was hidden from participants' view.

#### **4.2.6 Trials**

##### ***4.2.6.1 Non-Word Search Trials***

Three repetitions of a basic trial were then executed on each participant that involved the measurement of physiological signals during exposure to music while the participant was not engaged by any other stimulus. The trials were distinguished by the extent to which users were expected to enjoy the music, based on preliminary survey results. Three levels of this variable were used – like, neutral, and dislike.

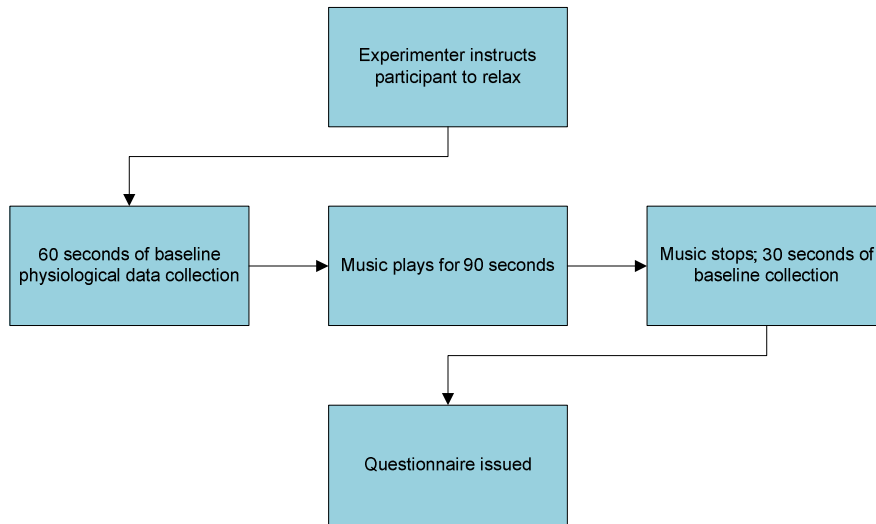
Each trial began with approximately 60 seconds of baseline data collection from the sensors, during which time the participants were instructed to relax and let their minds wander. This 60 seconds of data collection allowed for physiological data to stabilize and for participants to relax. After this baseline period, music began to play (from the secondary laptop) through a pair of headphones that the user was wearing. Music would continue to play for approximately 90 seconds before being stopped. Physiological data collection would then continue for approximately 30 seconds to allow data levels to return to baseline values. Following each trial, a questionnaire (Appendix B1) would be issued to participants that allowed them to:

- Self-report their affective state during the trial, both visually and descriptively (see section 4.2.7 for the selected visual measure);

- Indicate whether they recognized the song played during the trial;
- Indicate their level of enjoyment of the song played (using a continuous scale).

In total, these trials took approximately 3 minutes to complete, plus the time required by the participant to complete the questionnaire.

The described process for these trials is illustrated in Figure 4.



**Figure 4: Process for non-word search trials (diagram illustrates a single trial)**

#### **4.2.6.2 Word Search Trials**

The word search trials followed the same protocol as the non-word search trials, except that the participant was asked to continuously complete word searches during the trials. Word searches were selected specifically for these trials because they would dominate the participants' visual attention and introduce a moderate cognitive load. This was intended to mimic real-life attentional demands in common audio listening scenarios. Participants were given print-outs of word searches from a free online source [28] and asked to complete them with a provided pen using their dominant hand.

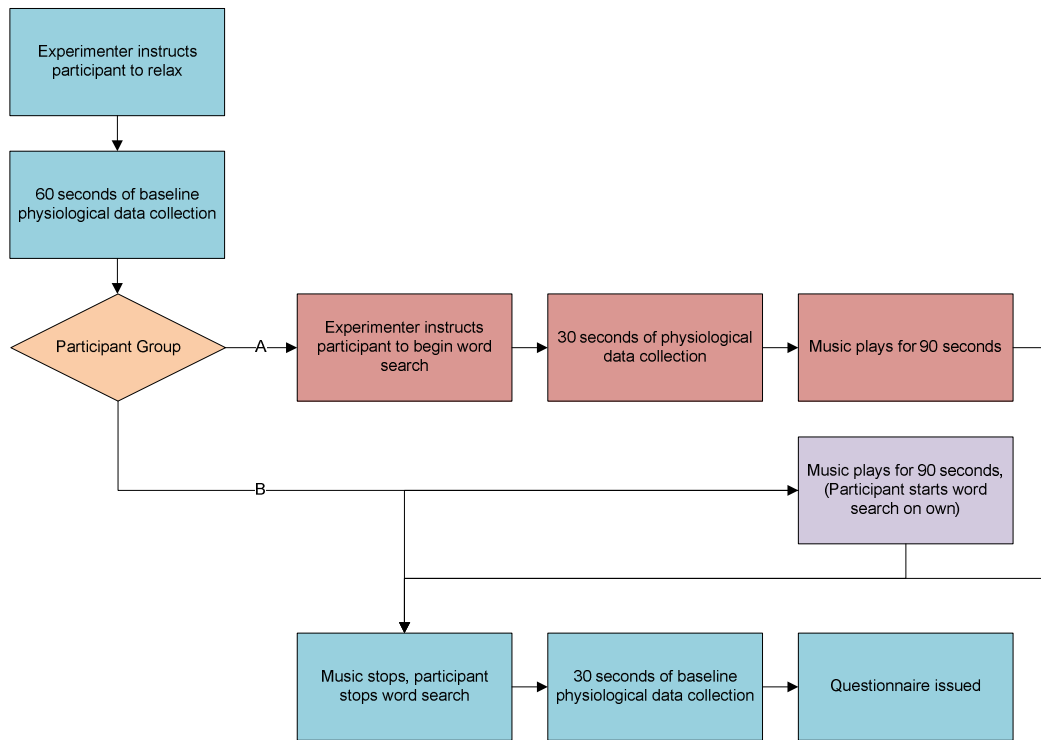
Half of the participants were instructed to begin the word search task as soon as the music began (after the 60 second baseline collection process). The other half was instructed to begin the word search after the baseline collection process in the absence of music; music commenced approximately 30 seconds later. Participants were assigned to these groups evenly. A between-subjects paradigm was selected for this purpose to ensure the study would not exceed one hour, and to keep the participants' tasks consistent throughout the study. We were interested in determining whether music and a cerebral task produced different indications in the collected data, so data from both groups was necessary. Both groups were instructed to cease work on the word search when the music stopped.

At the conclusion of each word search trial, a questionnaire (appendix B2) was administered to the participants. This questionnaire was identical to the non-word search trials questionnaire, with the exception of the following additions:

- A Likert scale to record the level of physical and mental effort required to complete the word search;
- A Likert scale to record the levels of enjoyment, engagement, frustration experienced while completing the word search;
- A Likert scale to measure the extent to which participants were distracted by the music while completing the word search.

In total, the word search trials took approximately 3 minutes or 3 ½ minutes to complete (depending on the experimental group) plus the amount of time required by the participants to complete the questionnaire.

The process described for these trials is illustrated in Figure 5.



**Figure 5: Process for word search trials (diagram illustrates a single trial)**

#### **4.2.6.3 Final “User Control” Trial**

Following the three non-word search and three word search trials, a final trial was run wherein each participant was instructed that they would have partial control over the music being played. In particular, they would be able to “advance” tracks in a predefined playlist at will by lightly tapping on a table using the hand not connected to physiological sensors.

Upon noticing the tap, the experimenter would manually advance tracks on the music player feeding the participants’ headphones, acting as a facilitator of a Wizard of Oz-style exercise.



Like the first six trials, approximately 60 seconds of baseline data collection was done, during which time the participants were instructed to relax and let their minds wander. Music began to play through participants' headphones following this process.

The predefined list of songs chosen for the participants (consistent for all participants, and containing no songs explicitly mentioned in the web surveys) would be played for approximately 5 minutes (300 seconds), with songs advancing based on participants' taps (or when they naturally ended).

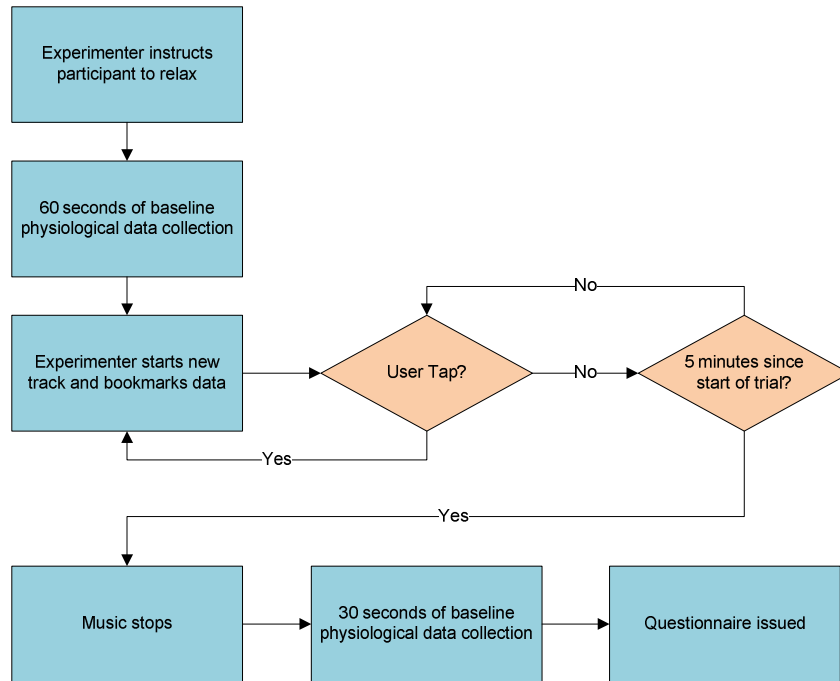
Following this final trial, a questionnaire (Appendix B3) was issued to participants. The questionnaire collected circumplex-based and freeform affect reporting the same way as the questionnaires for the first three (non-word search) trials. It also gave the participants an opportunity to:

- Describe what caused them to skip or linger on songs during the trial;
- Indicate their level of enjoyment of each of the songs played (using a continuous scale).

Songs used during the trial were played back upon request while the participants completed the questionnaire, as they may not have been familiar with their titles.

This trial took approximately 6 minutes to complete, plus the time required by the participant to complete the questionnaire.

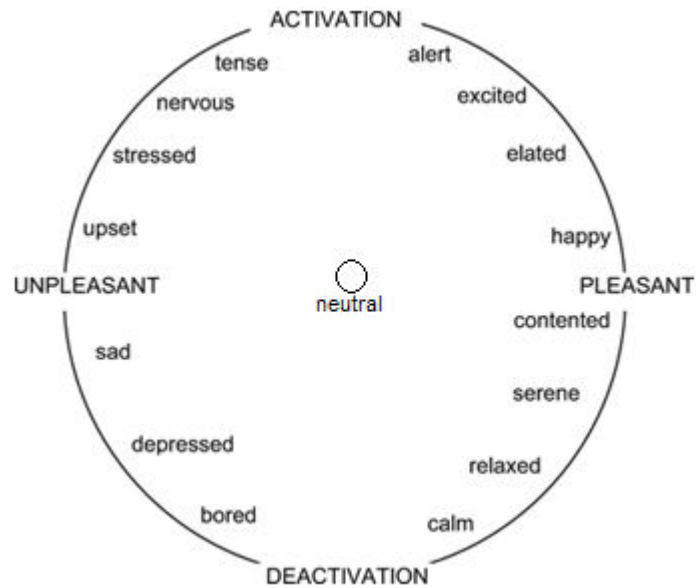
The process described for this trial is illustrated in Figure 6.



**Figure 6: User control trial study process**

#### **4.2.7 Measures for Self-Reported Affect Measurement**

Russell’s dimensional circumplex model of affect [43] (Figure 7, diagram from [42]) suggests that all emotions are linear combinations of valence and arousal dimensions. Strong, positive emotions appear in the upper right of the visualization, whereas weak, negative emotions appear in the lower left of the visualization. This intuitive approach to self-reporting affective state has been widely used in the social sciences [7], and was thusly used in the post-trial questionnaires. The model and diagram were explained to each participant prior to issuing the first questionnaire. The terms “valence” and “arousal” were explained to participants as “quality/type of emotion” and “intensity”, respectively, when necessary.



**Figure 7: Affect circumplex diagram as illustrated by Posner *et al.* [42]** This diagram was given to participants on post-trial questionnaires, where they were asked to indicate their valence and arousal graphically using an X.

### 4.3 Results

Several analyses of the data collected in this study were undertaken. This section describes these analyses and their findings in detail.

First, an informal visual inspection process of the collected data was undertaken by the experimenter in an attempt to identify clear physiological indicators of enjoyment, engagement, orientation and frustration. This process involved plotting and comparing collected data in line charts and correlating findings with observations of participants. The observations that backed up this analysis were made over the course of all trials in order to capture visual indicators of affective state. These observations were made during the study itself, as well as during review of video recordings.

Finally, a rigorous analytical process based on a  $k$ -nearest neighbours algorithm was undertaken by an associated research laboratory on campus, attempting to more robustly correlate physiological data with self-reported enjoyment.

#### **4.3.1 Assessment of Song Selection**

Before commencing analysis on the collected data, we aimed to validate our song selections for the first six experimental trials with respect to participants' self-reported levels of enjoyment. One third of the trials (total 24 over 12 participants) were executed using songs that participants were estimated to like, based on their questionnaire data. Likewise, 24 trials were executed using "neutral" music, and 24 with music that participants were estimated to dislike. Participants self-ranked songs in the "like" trials an average score of 9.06/10 ( $n = 12$ , stdev 0.79). "Neutral" songs were ranked on average 5.04/10 ( $n = 12$ , stdev 2.16) and "disliked" songs were ranked on average 1.42/10 ( $n = 12$ , stdev 1.10). A one-way analysis of variance test revealed that mean rankings were statistically significantly different between all groups ( $p < 0.0001$ ), suggesting that songs were properly selected for each trial.

#### **4.3.2 Visual Data Inspection and Observational Analysis**

Visual inspection of the collected data involved examining line charts of collected data across sensors and participants, comparing levels between and across subjects, with special attention paid to data at the start and finish of songs. We aimed to correlate these findings with visual observations of the participants and feedback from questionnaires.

Our analysis revealed that skin conductance response appeared to be most sensitive to major event changes during experimental trials (the onset and cessation of

song playback, for example) in the experiment. For most participants, a spike in SCR was recorded when music began to play, and either a second spike or “trailing off” effect when music was halted. These spikes also often occurred when participants requested song changes in the final experimental trial (although not universally).

Some participants experienced regular skin conductance spikes during the 90 seconds in which music was played; in many of these instances, participants described musical and emotional occurrences in post-trial questionnaires that appeared to correspond with these spikes (without knowing of their existence). For example, Figure 8 shows a plot of SCR vs. time for one participant, as collected during a trial involving no word search. The participant commented that she “really got into some parts of the music, especially the climax (near the end)” and rated the song 9.2/10 on the enjoyment scale on the post-trial questionnaire.

The same participant, when exposed to a song she disliked, commented that she was “at first really annoyed” but then became “resigned that [she] would have to listen” to the song. The participant ranked the song 1.1/10 on the enjoyment scale. Figure 9 illustrates her SCR measurements over that trial; spikes early after the music started could be associated with self-reported annoyance, while the smooth decline in the measurement could be associated with self-reported resignation.

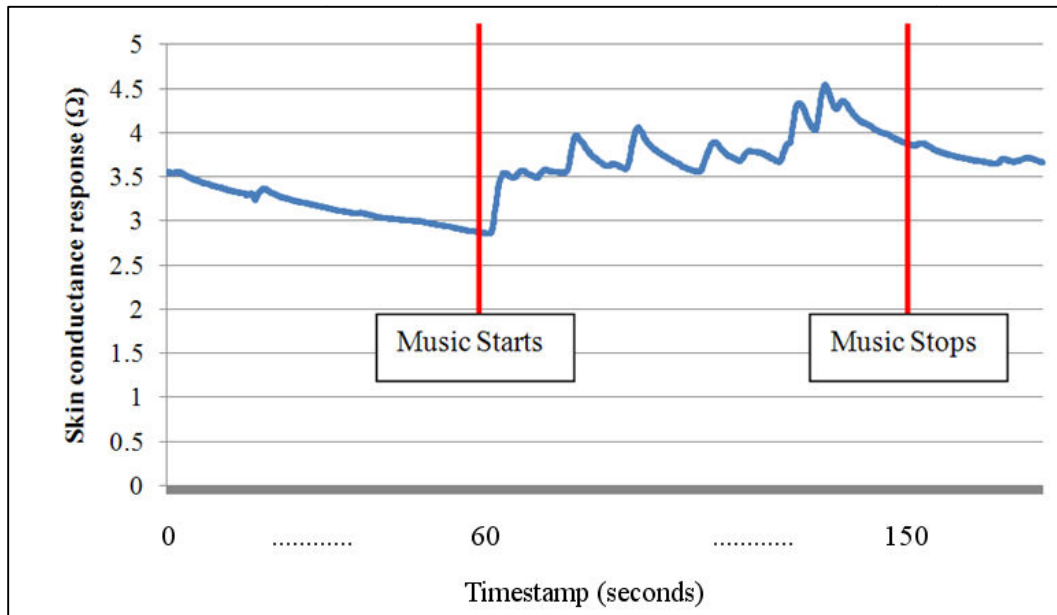


Figure 8: Skin conductance response vs. trial time (no word search, projected “liked” song) for one of the participants

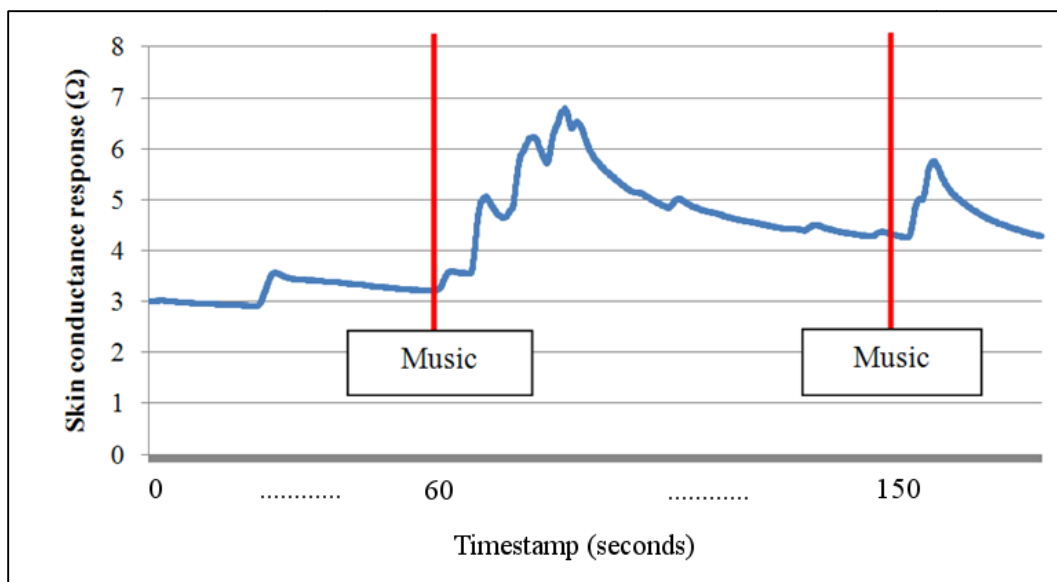
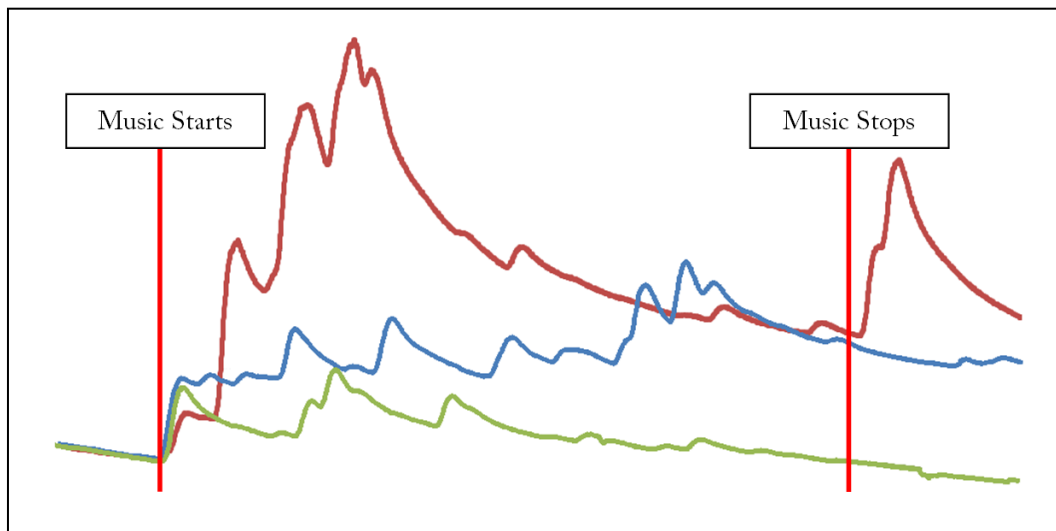


Figure 9: Skin conductance response vs. trial time (no word search, projected “disliked” song) for one of the participants (same as in Figure 8)

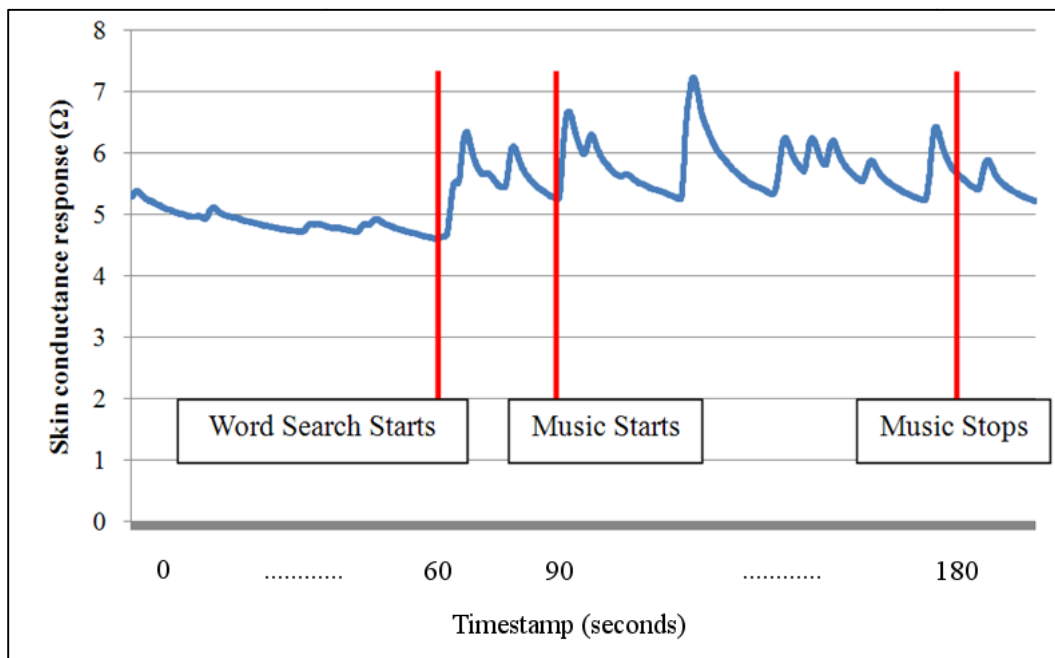
These observational results are consistent with results from previous work that associate orienting responses with GSR [5] [15] in non-musical contexts, and indicate that GSR could be a useful metric in detecting when users are surprised, startled or otherwise direct their attention to a source.

GSR in isolation did not appear to give a strong indication as to whether or not a song was enjoyed by a participant; rather, it indicated the strength at which the participant physically reacted to events in a song. For some participants, such as the one whose readings appear in Figure 10, the strength of their feelings in “dislike” trials appeared to be stronger than those in “like” trials. Other participants appeared to have the reverse reaction (based on visual analysis). The overlaid trials (all “no word search” condition) in the following diagram (Figure 10) indicate a clear difference in GSR magnitude across trials, but without context, these cannot be easily associated with levels of enjoyment.



**Figure 10:** All “no word search” GSR readings for one participant, overlaid for comparison. Data was shifted vertically to ensure baseline levels were calibrated. The “disliked” trial is indicated with the red line, the “liked” trial with the blue line, and the “neutral” trial with the green line.

Just as it was difficult to capture participants' level of enjoyment of a song using GSR alone, analysis revealed that disambiguating reactions made on the basis of music from other causes of GSR fluctuations was impractical. In Figure 11, we note similar peaks in the GSR signal when the participant begins the word search and when the music begins to play. As physical actions of a human can influence GSR readings [40], this finding also suggests that the peaks encountered in the final experimental trial that corresponded with requested song changes could potentially be explained by the physical action of the participant tapping on the table.



**Figure 11: Skin conductance response vs. trial time (word search, projected “liked” song) for one of the participants. In this trial, the participant was instructed to start the word search 30 seconds prior to the music commencing.**

No other source provided clear data that could be manually investigated for qualitative analysis. A comparison of GSR and other metrics revealed no apparent



associations. As the observational GSR analysis alone was not sufficient for our purposes, we rely on a more robust, model-based approach to make stronger conclusions.

#### **4.3.3 Further Observational and Behavioural Analysis**

A qualitative analysis of questionnaire results and video recordings of the experimental sessions uncovered a deeper understanding of the desired behavioural patterns of a HALO-enabled system.

Almost all participants (8/12) cited boredom as a reason for advancing tracks in the final experimental trial. Annoyance was also a highly cited reason (5/12) and curiosity about future songs in the playlist was also mentioned by some participants (3/12). Some participants indicated that if they recognized the song being played they were much more likely to continue listening to it, while others stated that this piqued their curiosity about songs to come and would cause them to advance.

Participants differed in their reactions to music they disliked. Some were frustrated throughout entire trials (those in which participants had no control over the playing music) and reported strong negative affect on the circumplex model. One participant in particular squirmed uncomfortably while music she disliked was playing, and asked the experimenter midway through a trial if the song could be changed. Other participants had much smaller reactions to songs they disliked, and devised coping strategies to ignore the music that was playing. Language used in the questionnaires often matched the severity of participants' visual reactions; words like "frazzled", "uncomfortable" and "painful" were used to describe the experience.

Similarly, reactions to enjoyable music varied drastically between participants. Some participants unconsciously tapped their hands or feet, hummed or sang along with the music, while others sat in an apparently calm silence.

#### 4.3.4 *k*-Nearest Neighbours Analysis

A machine learning approach (based on a *k*-nearest neighbour algorithm) to determine whether songs can be autonomously classified into the three targeted groups (like, dislike, neutral) was undertaken by an associated research laboratory on campus<sup>5</sup>. Data from 60 of the 72 trials were deemed valid for analysis, with other data containing unfilterable noise. Data was randomly separated into training and testing sets (50% each). Filtering techniques were used to smooth signals, extract features and normalize values. Several physiological features were analyzed over a range of time windows (1 – 20 seconds). A recognition rate of 76.67% was achieved with a time window of 9.02 seconds. The features that produced this recognition rate were the mean of normalized EMG [mean(nEMG)], standard deviation of normalized heart rate [std(nHR)], maximum normalized heart rate [max(HR)], and the differences between the maximum and minimum values of the following three signals: normalized derivative of skin conductance [diff(ndSCR)], normalized heart rate [diff(nHR)] and normalized heart rate acceleration [diff(nHRAccel)]. In spite of this promising recognition rate, high recognition rate fluctuations over time windows (see Figure 12) suggest unreliability of detection in real-time. It was hypothesized that abnormal variations in data, caused by unexpected noise and otherwise implausible data from sensors, were largely responsible for the analysis fluctuations, as well as significant physiological differences between users.

---

<sup>5</sup> All other analysis described in this chapter was performed by the experimenter. For more details on this analysis, please refer to Appendix D, Experiment 1.

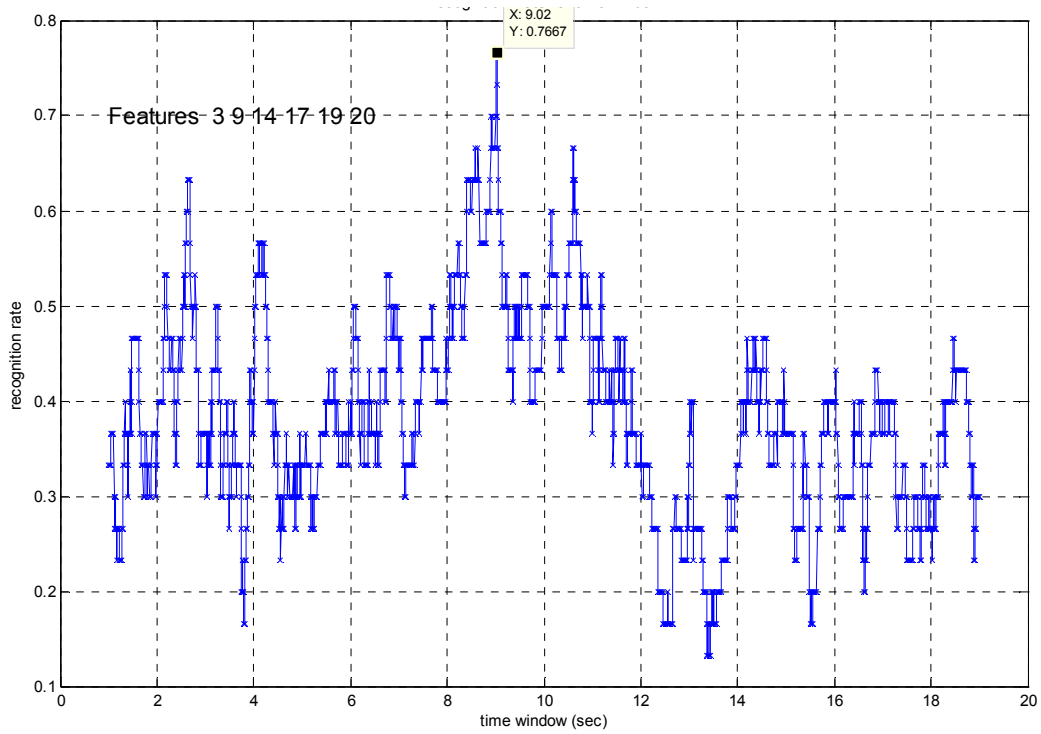


Figure 12: Recognition rate vs. time window for a subset of physiological features

#### 4.4 Discussion and Open Questions

A result from observational analysis is that the level of relative arousal (or activation) of a participant when exposed to songs of different genres (enjoyed to different degrees) was generally easily identifiable based on GSR readings. Correlating these measurements with polarity of appreciation (like and dislike) could not be done observationally, due to large between-subject differences in reactions to playing music, and a lack of other disambiguating, observationally clear measures. Despite this limitation, strong evidence that startle responses could be detected in audio use cases was collected. Subsequent work by Pan *et al.* [40] (which involved the author of this thesis) documents the development of an autonomous bookmarking system for audio books

based on orienting responses as detected via GSR. Using a threshold-based algorithm, an 84% identification rate of startle responses was achieved and validated via a user study. This provides promise that a similar GSR-based system could be developed for a music listening use case, but its utility would be limited to responding to startle effects, which were not the only physiological phenomena deemed important by focus group members.

Using manual, observational techniques, disambiguating the polarity of GSR-derived orienting responses is infeasible. Strong physiological signals could indicate either extreme pleasure or extreme annoyance, just as weak signals could indicate either displeased resignation or serenity. Autonomous classification is likely not possible on the basis of GSR alone; additional channels need to be incorporated into the system. Notwithstanding this result, large between-subject differences indicate that producing a highly tailored affect classification system for a specific subset of music listeners could be an appropriate next step for the current research. For participants where, for example, strong reactions were highly correlated with displeasure while pleasure elicited more subdued responses, a naïve but functional HALO prototype could be rapidly produced on the basis of GSR alone for further investigation.

It is also noteworthy that participants' own music collections could elicit largely different affect responses than the libraries used for the current study. It is unreasonable to assume that participants would own a large amount of music that they were not fond of, and in the case of streaming media (from an online source, for example) user-selected channels would not likely contain music that would be wholeheartedly rejected as unpleasant by the listener. Furthermore, far more music in participants' personal collections would be recognizable to them than the music used for this study. From a feasibility standpoint, extremes in musical selection were made for this study to

maximize the potential of identifying physiological markers in analysis; much subtler affective cues may need to be identified and used in a deployment system. At the same time, the strength at which users react with pleasure and displeasure toward a particular song can be mood or context specific, and large affective cues may still be possible to collect.

The development of a full-featured real-time affect classification model is outside the scope of this thesis, but confirming the technological feasibility of such a model was a significant milestone. We return to the research questions in turn to address this goal:

- Can the level of enjoyment of a song be inferred using physiological data? How reliably?

Visual inspection of physiological data revealed that GSR was strongly correlated with activation, but this alone could not classify level of enjoyment without additional disambiguating data. This fact notwithstanding, for single participants whose reactions to music vary in strength along an enjoyment scale, GSR could likely be used for rapid prototypes of a HALO-enabled audio player (i.e., no further disambiguation would be required).

Using a  $k$ -nearest neighbours approach to affect classification, a recognition rate of 76.67% for song enjoyment was achieved with a 9.02 second analysis time window. The recognition rate spectrum revealed high variability in recognition rates in adjacent time windows, however, inferring the potential for significant difficulty in the implementation of an online classification system. To address whether this limitation is universal or specific to the data that was collected, follow-up analysis with data collected for a single participant in a controlled environment without a secondary cognitive task

should be performed. We undertake this in the third phase of our work, which is reported in Chapter 5.

- What physiological measure(s) are most promising for inferring correct system behaviour?

GSR is an excellent measure for identifying level of activation, and can be done with little complex numerical analysis. While this was previously well-known, this work confirms the result for musical contexts. For scenarios involving startle responses, this measure could be used, potentially in isolation, to correctly drive system behaviour. Additionally, thorough investigation revealed that numerical analyses on GSR, heart rate and EMG data in combination facilitate affect classification with 76.67% accuracy, but large variances in classification scores across time windows and features weaken this result.

- To what extent is the level of reliability of the physiological measures affected by a secondary task (such as completing a word search or reading a newspaper)?

Using GSR alone, it was not possible to disambiguate orienting responses that were caused by music from those that were caused by the secondary task (word search). *k*-nearest neighbours analysis did not disambiguate word search from non-word search cases.

- For what reasons do music listeners wish to change a track they are listening to, even if they like what is playing?

Participants reported boredom, curiosity and repetitiveness as the primary reasons for choosing to advance over tracks, even if they enjoyed the music being played. If the music was not enjoyed by the listener, annoyance or general dislike of the genre were reported. It is especially important to consider curiosity in the development of future

HALO prototypes; a well-tuned system making excellent choices could pique the curiosity of some subjects, and without proper calibration, could be “misinformed” by users advancing over well-selected tracks.

Users portrayed their enjoyment and displeasure for musical selections in vastly different ways. These often-physical cues could be measured via additional sensors to aid in affect classification processes, as has been done in previous work. These cues were often observationally correlated with GSR data, which would permit fine-tuning of a naïve HALO implementation (that would use magnitude of orienting responses for affect classification) for particular user groups. In the production of a highly specialized interaction for a particular subset of users, physical observations of users, and their correlation to measurable signals, thus play an important role.

A generalized real-time classification system for a HALO-enabled music player does not appear immediately feasible given vast differences between users’ physiological responses to music and the ineffectiveness of a simple model (i.e., GSR-based). Even in the case that it was possible, the specific desired behaviour of an automated music player appears to differ between participants, and no concrete requirements have yet been established. The next step, therefore, focuses on a restricted subset of potential users to maximize the potential for real-time affect classification, and the likelihood of developing a highly tuned, specific interaction language for a HALO-enabled player. The following chapter discusses participatory design work that was undertaken to this end.

## 5 System Integration – Participatory Design

### 5.1 Motivation and Overview

Individual differences between users have hampered the generalizability of the work described in this thesis thus far. There is evidence from both the focus group sessions (chapter 3) as well as the music-affect study (chapter 4) that the HALO paradigm can provide utility to users – although perceived value varies – and that it is likely immediately technologically feasible to do so if it is tailored to a particular stream of physiological input (i.e., between-user disambiguation is not required). We therefore investigate, with a single user, the extent to which this evidence holds up in design, implementation and testing of a medium-fidelity HALO prototype.

In the work described in this chapter, we aimed to develop a preliminary interaction language for a HALO-enabled audio player and investigate its utility using a variety of approaches to scientific inquiry in human-computer interaction. These techniques were employed in a process of participatory design with a single user, who helped dictate the optimum system behaviour and feedback mechanisms of the device. While this interaction language proves to be highly tailored and potentially ungeneralizable, significant insights on the attentional demands, utilities, and general feasibility of the HALO paradigm were uncovered in the participatory process that can be generalized later. In particular, with a functional HALO prototype in place that works optimally for a *single user*, work focused on the generalizability of the paradigm, both for multiple users in the portable audio use case and across multiple use cases, can commence. This work aims to make primary identifications of potential areas for



interaction customization that would make HALO appropriate for other users, and potentially other use cases.

## **5.2 Methodology**

### **5.2.1 Participant Recruitment**

An outcome of our focus group sessions was that participants appeared to respond much more positively to proposed technologies when they could test them out and/or imagine their capabilities based on previous knowledge. By showing even potentially unrelated sensor and haptic technologies to the focus group members, scepticism and distrust in the paradigm was reduced. As an optimally working HALO prototype would require further investigation, we desired to recruit a single participant who was open to and excited about new technologies to help guide the interaction design process. We desired a participant that could be imaginative about the potential uses of HALO-enabled technologies while still offering honest, balanced feedback on potential designs. This desire stemmed from our own interest in finding areas of contrast between the technologically averse focus group members and the alternative perspective. We wanted the participant to be passionate about his or her music-listening experiences, and be open to evaluating potential enhancements to his or her experiences against current practices.

Our goal was to avoid recruiting a participant who would immediately reject the technology on the grounds of its imperfections or potential social and privacy concerns. While these issues certainly warrant attention, they are outside of the scope of the current work. Instead, our goal was to foster a design environment with direct, end-user feedback that would be exciting, productive and useful. With a working prototype in

place to test with and customize for other classes of users, we can better understand where apprehensions lie and future work should focus.

In pursuit of a participant that fit our criteria, we e-mailed invitations (see Appendix C1) to undergraduate research assistants to set up screening interviews. The students were targeted for their flexible availability over the timeframe of the participatory design process.

The screening interviews (see Appendix C2) focused on:

- Identifying usage scenarios of portable audio players;
- Identifying comfort levels with audio players and technology in general;
- Identifying likelihood of participants being early adopters of technology;
- Identifying genres of music enjoyed and disliked;
- Identifying the likelihood that the participant will remain committed and honest.

Interviews were conducted with three prospective participants. The selected participant, a female student in the Department of Applied Science aged 18 – 25, was chosen for the following reasons:

- **She uses her portable music player in a variety of contexts, and experiences different pain points in each.** In particular, the participant indicated that she uses her player while exercising, while commuting, at home for enjoyment, and as an aid for choreographing dance routines. This broad series of use cases gives her a broad perspective for participatory design activities to come, and allows for nuances in desired system behaviour between scenarios to be uncovered based on her first-hand experience.
- **She enjoys trying new technologies and learning how to use them via experimentation (such as scouring through settings menus), but adopts them**

**only if she deems them useful. She can “live with” her current pain points but welcomes potential solutions to them.** We sought a participant that would be open to trying new technologies but that was not likely to adopt technologies that did not serve a purpose (as someone with a fondness for “gadgets” would). In this manner, the participatory design sessions would focus on defining system behaviour for a purpose, rather than just for interest’s sake.

- **She has strong feelings about music she likes and employs a variety of coping techniques to endure music she dislikes.** The participant indicated that she undergoes significant physiological changes when listening to music she likes (increased general excitement), and feels the desire to dance when particular songs are played. She also indicated that she was able to calm herself down fairly effectively to bear music she disliked, as not to become too frustrated. She listens to a variety of music, both from North America and the Middle East.

Hereafter, the participatory design participant is anonymized as Beril.

### **5.2.2 Scheduling and Remuneration**

Participatory design sessions were scheduled in an impromptu manner based on the availability of the researcher and participant, and also based on the analysis process of previous sessions and emerging research questions in the research group. Each session was conducted with specific research questions in mind, and involved a combination of:

- Semi-structured interviews;
- Physiological data collection;
- Evaluation of low- and medium-fidelity prototypes;
- Wizard of Oz exercises.

All sessions were videotaped for subsequent analysis. In total, seven participatory design sessions were held that varied in length from one hour to two hours. In total, 9 hours were spent with Beril in the participatory design process. She was compensated at a rate of \$10/hour, and was paid \$90 in total over the sessions. Payments were made at the conclusion of each session.

### **5.3 Session-by-Session Summaries**

This section describes the research goals and questions that drove each of the seven sessions with Beril, the processes that were undertaken to address those goals and questions, and the resulting implications for design that were uncovered by the researcher.

At a high level, the stages of the participatory design process were:

- Calibrating the technological opinions and physiological responses of Beril with those of previous focus group and experimental subjects (sessions 1 and 2);
- Understanding the role of distractions in Beril's music listening experiences (session 3);
- Developing and evaluating a haptic feedback language for the paradigm (sessions 4 – 6);
- Developing and evaluating a supplementary input language for the paradigm (sessions 5 – 6);
- Simulating and evaluating a complete interaction loop (session 6);
- Collecting additional physiological data for future analysis (session 7).

More detail for each of the stages of this process is given in the subsequent sections of this chapter, on a session-by-session basis. For ease of consumption, Table 6 correlates

each session with the descriptive section name used in this thesis. These names were chosen to identify the areas of research, goals and questions we aimed to address in each section.

Session #	Section Title
1	Focus Group Activities Revisited
2	Music-Affect Experiment Revisited
3	Exploring Interruptions and Distractions
4	Exploring the Haptic Feedback Design Space
5	Exploring Explicit Control and Haptic Feedback
6	Wizard of Oz Simulation of a Closed System
7	Follow-up Physiological Signal Measurement

**Table 6: Overview of participatory design sessions**

### **5.3.1 Session 1 – Focus Group Activities Revisited**

#### **5.3.1.1 Goals and Research Questions**

We concluded that our focus group participants were largely unwilling to adopt HALO-enabled technologies due to scepticism in its abilities to properly classify human affect, as well as a discomfort with privacy, security and social issues related to the interaction paradigm. The recruitment strategy for the participatory design sessions was centred on finding a participant that would, in many ways, look past these practical concerns to help drive the development of the technology forward. With this in mind, the first participatory design session was run with the goal of addressing the following calibrating research questions:

- In light of the participatory design recruitment strategy employed, how will Beril respond to a subset of the same activities and questions that we posed in the first and second focus groups?

- How will Beril's responses compare with those of focus group participants?

#### **5.3.1.2 *Process***

The questions and activities that guided the first focus group session, as well as the scenario-related activities from the second focus group session, were rerun with the single participant following the same process originally designed. The semi-structured interview and exercises that were performed took approximately 60 minutes to complete.

#### **5.3.1.3 *Summary of Session***

Many of the subject's responses to interview questions were focused on her hobby of dance and how she uses her audio player (an iPod Classic) in related scenarios (such as teaching and rehearsing choreography). Aside from dance, the subject indicated that she uses her player while cooking, relaxing, attempting to sleep, exercising, and commuting by bus. She uses her player for functional purposes (e.g., to tune out noises around her, to make tedious tasks more enjoyable, and to teach dances) as well as for pure enjoyment. She indicated that her player sometimes distracts her from a primary task (such as studying) or is too cumbersome due to its form factor to make it effective in supplementing her tasks.

The subject's grievances with her player were largely related to its inability to completely block noise distractions, or that the player itself was distracting her from other tasks or preventing general situational awareness. In addition, she indicated difficulties scrolling through menus on her portable phone and iPod due to perceived button oversensitivity and small button sizes, and projected mild annoyance at this problem. She also indicated that it was frustrating to work with the "lock/hold" feature, as she would often forget to activate it and songs would inadvertently advance when the player was rubbed against with another object in her backpack (where she typically stores it).

She also mentioned that fast forward and rewind features were not sufficiently accurate for her purposes, as she would often wish to revisit a particular segment of a song repeatedly, both for purposes of choreography and general enjoyment. She also felt that adding content to playlists was not well supported by her iPod, as she would often forget to add music to both her music library and specific playlist. She said that she often finds herself advancing through music to find tracks that fit her current mood, since her playlists are not adequately structured.

When Beril described the circumstances of an ideal musical experience with her music player, she indicated that when a song she very much enjoyed began to play, she would lose focus on everything else and become completely immersed in the experience:

*I usually have an immediate reaction to [a song I like] ...I...stop doing whatever I was doing at that time, and immediately focus on the song, immediately think about the song, and am completely focused on the song – studying, cooking, whatever...I kind of start dancing, even if it wasn't an [appropriate] situation, I would do some sort of movement. After...I would be totally relaxed and try to listen to the song like 5 more times.*

Beril indicated that she would simply skip songs that she was “not in the mood for”, after giving them a few seconds to see if she could “get in the mood”. She indicated that in some cases, she would have the urge to return to a song that she had skipped earlier, after giving the situation more thought.

With regards to technological comfort, Beril indicated that she would rate herself a 9 out of 10. Beril indicated that she enjoys spending time going through features and experimenting with them when presented with new devices. She indicated that her comfort level with her audio player is slightly above her general level of technological comfort, due to her familiarity and previous experiences with the device.

Beril indicated that she preferred to store her iPod in accessible places for a given context, but often found this cumbersome or impossible and would resort to storing it less than ideal locations (i.e., in her backpack).

When Beril was given a description of the Yamaha BODiBEAT player (first introduced in Chapter 3.3.1.1), her initial reaction indicated intrigue with the technological propositions. She indicated that she would want to try the player to see if it was effective at selecting appropriate music, and questioned its abilities to do so given many potential contexts that could, for example, raise her heartbeat. She indicated that she saw utility in the device and would adopt it if it worked well.

When the subject was introduced to SoundSense technology, she indicated that it could “definitely be useful” and suggested some “neat” applications (like auto-turn off of a device). She indicated that there would definitely be times that she would want to turn off the “intelligence” of the system, especially in cases that she did not want to be distracted. Without prompting, she suggested that she could imagine the technology (embedded in an audio player) determining instances when she was interrupted by other people, and taking some appropriate action (depending on whether she wanted to be interrupted or not).

In the blue sky exercise, Beril suggested that she would enjoy it if her player could provide some form of guidance for choreography instruction. Her initial suggestions involved analysis of the teaching context (through an embedded video recorder) and the provision of verbal corrections to students’ errors.

When the scenarios were presented to Beril, she identified with interruption scenarios quite strongly (although not always in the specific contexts presented in the scenarios). She indicated that she does not listen to audio books, but indicated that



autonomously captured bookmarks would be useful for returning to portions of songs for which she was teaching dance choreography. She also identified strongly with the desire to advance through media based on perceived inappropriateness for the situation or her current tastes.

As a relevant side note, Beril indicated that she was a very visual person and likes to be provided with a large amount of information from her devices.

#### ***5.3.1.4 Implications for Design***

Beril's responses to the focus group activities were highly consistent with those she gave in her screening interview. She uses her player in a wide variety of contexts for which interruptions and musical selections cause different perceived physiological effects. She feels very strongly about music she enjoys, often losing focus on her surroundings and beginning to dance when particular songs begin. Songs she dislikes or feels neutral about are not likely to cause her significant annoyance or pleasure; the arousal of her affect in these contexts is likely to be low. Her grievances with her current audio player generally focus on the manual steps required to pause, advance or otherwise adjust playback settings. A particular issue for her was fast forwarding or rewinding to particular positions in songs in order to listen to content repeatedly.

Beril is open minded to new technologies and enjoys experimenting with new devices. At the same time, Beril strongly indicated that she would not adopt new technologies that had no purpose. For her, technology must serve some purpose, but she is willing to give "neat" devices a try before dismissing them.

Early indications for design requirements, as captured from this session are:

- Explore automatic song selection based on affective responses. Affective responses large in magnitude are likely to be correlated with enjoyment for this subject.
- Explore content bookmarking, both automatically placed in areas of the song with large affective measurements (to allow repeated playback of well-loved selections), as well as for manually placed bookmarks (to facilitate choreography tasks).
- Provide transparent feedback to the participant about the choices of the audio player. The subject indicated that she enjoys consuming large amounts of information from devices, and given her curiosity for technology would likely be interested in receiving information about a subset of the internal processes of a HALO-enabled system. This feedback should be provided in a manner that does not require her to interact directly with the device, as retrieving it from storage was one of her pain points.

### **5.3.2 Session 2 – Music-Affect Experiment Revisited**

#### **5.3.2.1 Research Questions**

The second session of the participatory design process involved rerunning the music-affect experiment with the chosen participant in order to compare findings with the original group of 12 participants. In particular, we aimed to find physiological markers and trends that could be leveraged in the production of a partially functioning HALO audio player for further testing. The research questions for this session, therefore, were:

- How will Beril’s physiological markers compare with those of the participants of the music-affect study?

- Does the hypothesis that liked songs will be correlated with high-magnitude affective measurements for this subject (specifically in GSR) appear to be correct?

#### **5.3.2.2 Process**

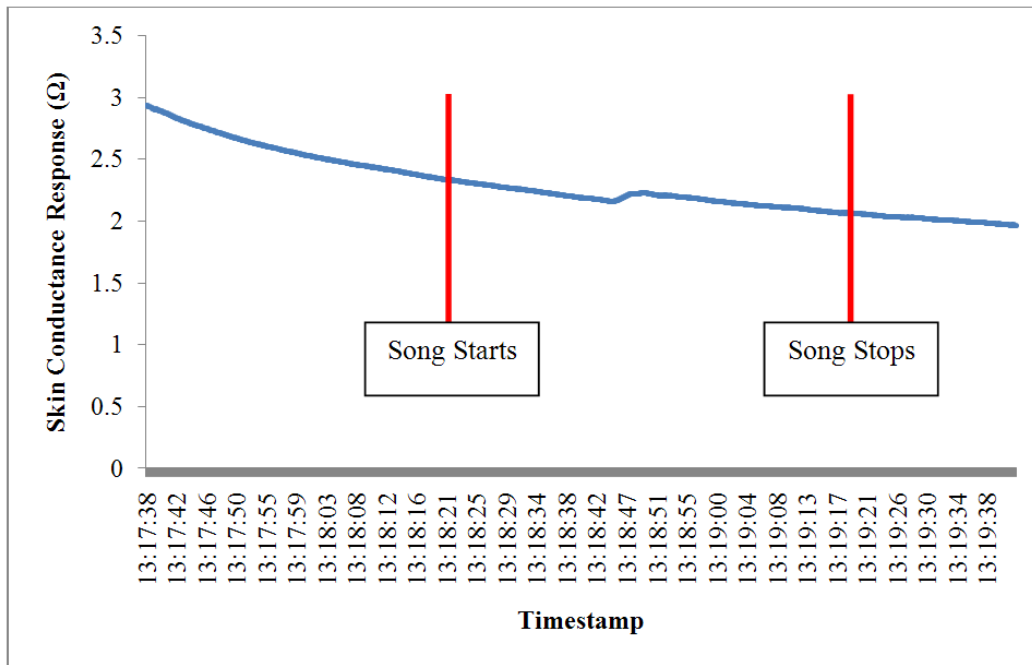
The second participatory design session took one hour to complete. The same experimental procedure as the music-affect study was followed with the single participant, except the Emotiv® helmet was not used. This was largely motivated by the thickness of the subject's hair, which based on experiences from the music-affect study would have drastically reduce the feasibility of making sufficiently close connections between the scalp and saline sensors. Additionally, previous results indicated that no usable data could be collected from the helmet for musical affect identification.

#### **5.3.2.3 Summary of Session**

The first song played during the session was hypothesized to be enjoyed by Beril based on her survey responses. She ranked it a 10/10 on the enjoyment scale and commented:

*When the song came on I felt very calm and felt like I was in my own world. After [the trial] I felt sad that the song was over because I was just really getting into the mood of being relaxed.*

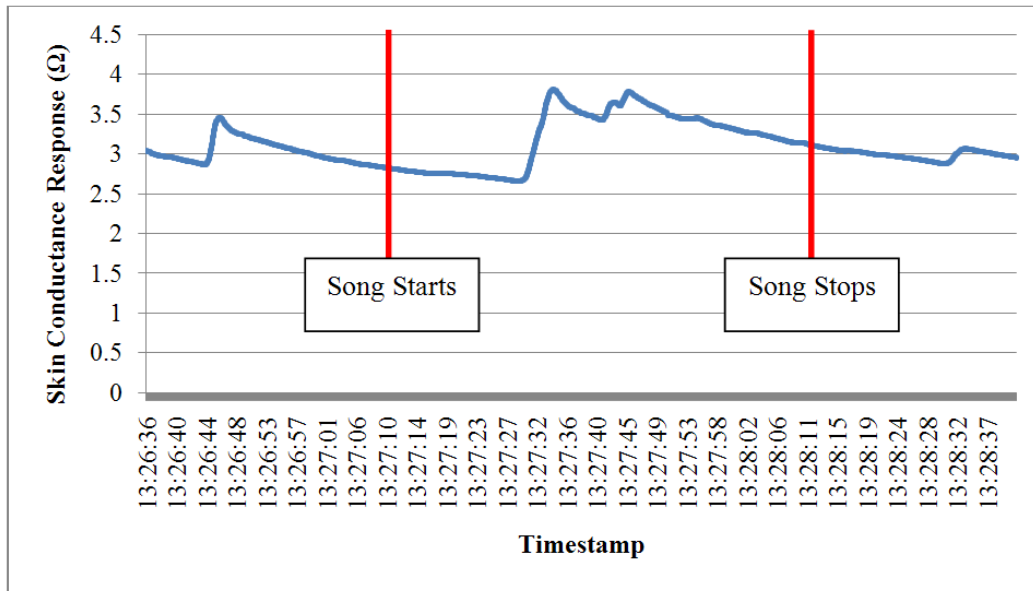
A plot of GSR vs. time for this trial appears in Figure 13.



**Figure 13: SCR vs. trial time (liked, no word search) for participatory design subject**

The smooth, continuous GSR readings pictured in Figure 13 could indicate the “calm” that Beril reported. However, further verbal comments from the subject that she “could ‘feel’ the music and ... even started moving to it” toward the end of the song do not appear to be correlated with the readings at all. Generally, physical movements and excitement have shown to cause significant spikes in GSR readings; Beril’s remained flat throughout the trial with the exception of a minor upward fluctuation approximately 45 seconds in.

In the following trial, where a second song projected to be liked by the subject was played and she was required to complete a word search, the subject indicated that “at specific parts of the song (which I liked) my focus went to the song”. She also said that “at the beginning I felt happy about the song.” She ranked the song 9/10 on the enjoyment scale. A plot of GSR vs. time for this trial appears in Figure 14.



**Figure 14: SCR vs. trial time (liked, word search) for participatory design subject**

The “happy” feelings reported by the subject could be visually correlated to early spikes in the GSR readings, but the first spike of visual significance did not appear until approximately 30 seconds of the song had been played. This GSR activity could perhaps instead be correlated with her concentration difficulties (or with neither state), leaving no obvious means to infer happiness.

GSR readings for songs that were disliked by Beril (those that were ranked 3/10 and 1/10) contained continuous spikes that were difficult to disambiguate. Beril indicated in both cases that she was “curious” about the songs and gave them a chance, and also that had a hard time directing focus away from the songs. This indicates a high level of engagement with the music, which was not expected based on discussions in the previous session (i.e., Beril was expected to be able to ignore these songs, and become engaged only in music she liked).

#### **5.3.2.4 *Implications for Design***

Based on earlier interviews, we expected Beril's GSR responses to disliked music to be minimal as compared to those elicited by music that caused Beril continuous enjoyment, but there was no clear means of disambiguating these cases using GSR alone. However, trials involving music that relaxed the participant rendered smooth, consistent GSR readings, and as Beril identified relaxation as a pleasant and enjoyable state, we consider this to be a positive finding.

For scenarios that involve bringing Beril to a state of serenity, such as helping her to fall asleep, a naïve prototype of the HALO affect classification system could likely be employed. However, as this would not address the subject's largest pain points and would narrow the focus of this research too greatly, this is not deemed an appropriate avenue of inquiry for the sessions to follow.

Based on the findings of this session, an implication for the design of subsequent participatory design sessions is that a HALO-enabled system that addresses the pain points of the subject cannot be prototyped using GSR alone. In the future, more robust data collection from the subject should be performed to facilitate a more robust affect modeling system. Subsequent work in the participatory design sessions should focus on designing and evaluating the interaction loop with low-fidelity methods, such as Wizard of Oz testing.

### **5.3.3 Session 3 – Exploring Interruptions and Distractions**

#### **5.3.3.1 *Research Questions***

We aim to address the effect of interruptions to music listening experiences on the subject in this participatory design session. The goal was to uncover sources of

distraction, the responses the subject generally makes to these distractions, and any form of support her portable audio player could lend to mitigate these distractions and rapidly return her to a state of contentment with her audio player.

The specific research questions we aimed to address were:

- In what situations is the participant comfortable with distractions to her music?
- In what situations does the participant dislike being distracted from her music?
- What behaviour does the participant elicit in distracted states?
- What form of system behaviour would mitigate the negative effects of an unwanted distraction?

#### **5.3.3.2 *Process***

This participatory design session was conducted for approximately 90 minutes. This session began with an open-ended interview that attempted to address each of the research questions with subjective responses from the subject. Following this, an informal experiment was undertaken to correlate these initial subjective responses with observations of responses to distraction and interruption, at first when just listening to music, and then when listening to music during a task.

The array of sensors used in the study described in Chapter 4 was attached to the subject to measure the physiological effects of distractions and interruptions. The subject listened to music from her personal iPod during the experiment. In the first trial, the subject was simply asked to listen to the music being played and focus on it at her leisure. In the second trial, the subject was also provided with a task (reading a three-page history of bubble gum retrieved from the internet [39]) and was told that there would be a follow-up content quiz on the material. The subject was interrupted by the experimenter on a regular basis (approximately every 2 minutes) during both trials. These interruptions

involved asking the subject both simple and complex questions about herself, her leisure activities, and her academic plans, as well as noise-based distractions, such as clapping.

Following the informal experiment, a follow-up, open-ended interview was conducted to subjectively measure levels of annoyance and frustration caused by the experimenter's interruptions, and to determine how realistic the interruptions were to real-life encounters.

In a follow-up study (which involved the author of this thesis), a similar but more formal interruption experiment was run with the participatory design subject (see [40] for details). This experiment, whose sessions lasted 30 minutes, involved the analysis of user biometric responses to interruptions while listening to an audio book. This task was not reported to be a common portable audio listening habit for Beril. Some relevant results from this work are presented in the following summary section to introduce further data for designing a HALO system to respond appropriately to interruptions.

#### **5.3.3.3 *Summary of Session***

In the initial interview, the subject reiterated that she would welcome interruptions to her music listening experience when required by a secondary task (such as studying). She also indicated that she would be open to distractions when portions of songs that she was disinterested in were playing. If the subject was "listening to music with the intention of listening to music" she indicated that she would not like to be distracted. Scenarios she gave as examples of this involved listening to music on the bus, when relaxing at home, and in all cases where focus on the music is required (such as when attempting to memorize lyrics).



Beril described a scenario where she was exercising at the gym and was highly distracted by noises being emitted by exercise equipment within her vicinity. She claimed that she was so distracted that she found it difficult to exercise. Raising the volume of the music was insufficient to completely block out the distracting sounds, and her attempts to focus on the music were unsuccessful. She stopped exercising as a final response to the distraction.

The subject indicated that in general, she would first attempt to increase the volume of her iPod to mitigate the effects of an unwanted distraction. If the source of the distraction was not in her control, she would next attempt to tune out the distraction and focus intently on the music. In circumstances of extreme annoyance, the subject is unable to completely block out her distractions. In these cases, Beril attempts to change the currently playing song to something qualitatively different to encourage greater engagement in the music. Beril indicated that she generally does not fast forward or rewind within a particular song for this purpose.

Beril described a scenario wherein she repeatedly listened to a song that she thoroughly enjoyed, which caused her immense distraction from her primary task (studying). She indicated that the particular song causing this distraction contained lyrics that were occasionally “shocking” and “lyrical” and that these poetic instances often corresponded to dramatic musical turns of events. During approximately 20 seconds of extreme distraction by the music, the subject indicated that she would get up from her desk and jump or dance to release energy.

Beril next described a situation where a friend and she were listening to an enjoyable song and her friend’s dance movements distracted her from focusing on the music. This distraction was not unpleasant for Beril, and it instead intrigued her to learn

her friend's dance. Asked to describe additional scenarios where a person or object was distracting her to an annoying degree, Beril was only able to relate stories about acquaintances wherein direct control over the interruption was possible.

During the informal experiment, the subject appeared to be surprised by the experimenter's interruptions, and appeared to face difficulty maintaining concentration on the task (especially for reading trials) as a result of these interruptions.

In the follow-up interview to the experiment, the subject discussed her frustration at being distracted and confirmed her general difficulty returning to concentration after an interruption. The subject answered three of five questions for which she had read the related content (all answers were correct) on the quiz during the experiment.

Beril stated that she was most frustrated with interruptions during the experimental trial that involved the reading task. She said that she does not often listen to music while reading (her studying involves more problem solving – generally in the areas of vector calculus and thermodynamics), and this exacerbated her frustration. She indicated that she would not do the two activities together given a choice. She also mentioned that verbal interruptions were less intrusive to her than the clapping interruptions, which surprised her, and happened to coincide with “nice” parts of the music, producing a stronger reaction.

The subject said at the conclusion of the interview that in the event of a significant distraction that legitimately demanded her attention, she would be likely to turn off the music in favour of the interruption.

During the related interruption study performed by Pan *et al.* [40] and attended by the thesis author, Beril was consistent in her responses to experimenter interruptions. She was startled by the first interruption (a simple verbal question), began to expect

subsequent interruptions, and became distracted later in the experiment as her focus intensified on the audio book listening task. Beril in many cases was required to rewind the audio book to listen to content that had been lost in the interruption, but results from the participatory design session indicate that this activity would not be likely for Beril in musical contexts.

#### **5.3.3.4 *Implications for Design***

This participatory design session and related follow-up study shed light on desirable behaviour of a HALO-enabled portable audio player with respect to interruptions for this participant. Several classes of interruptions were uncovered during the interviews and activities that made up this session, each with different implications for design. Specifically,

- Some interruptions to music listening, such as friends dancing or difficult homework demanding full concentration were not seen as frustrating or annoying, but rather helpful. If a HALO-enabled music player could infer when the user was distracted by music from a more critical task, it might be able to, for example, disable itself.
- Beril responded to undesirable interruptions that did not require her attention in a number of ways. Her first line of action was to adjust the volume of her audio player to drown out noises, and then to focus intently on her music in an effort to cognitively exclude the interruption from her thought processes. Failing this, the song is changed, perhaps to a louder or more involved selection. A HALO-enabled system should leverage this granular problem solving process and correlate it with detected levels of frustration or annoyance in order to automate distraction resolution, perhaps semi-pre-emptively.

- Undesirable interruptions that required the subject's attention were not considered as frustrating as undesirable interruptions that had nothing to do with Beril. Since the former class of interruption necessarily requires Beril to discontinue use of her player, an autonomous pausing or volume adjustment mechanism may be appropriate based on detected orienting responses to the interruption.
- Beril's level of engagement in the music is important in dictating the appropriate behaviour of the HALO system. The subject indicated that she would be more welcoming of an interruption while she was listening to a musical selection (or portion thereof) in which she was not heavily invested. This indication was further evidenced in the informal experiment by a negative reaction to a clapping interruption that coincided with a highly enjoyable part of a song. This suggests that continuous measurement of engagement may be important for dictating appropriate system responses to interruptions in real-time.

### **5.3.4 Session 4 – Exploring the Haptic Feedback Design Space**

#### **5.3.4.1 Research Questions**

The goal of the fourth session was to perform introductory investigation of the role of informative haptic feedback in the interaction loop with the subject. In so doing, we did not wish to pre-define a haptic language of communication that would be immediately imposed on her. Instead, we aimed to give the subject the tools to brainstorm and imagine how haptically-delivered messages from a portable audio player would fit into the interaction loop. These tools took the form of simple haptic stimuli

with no pre-defined meanings. This approach was followed to ensure that Beril remained prominent in the iterative design of the paradigm.

The specific research questions we aimed to address in this session were:

- Can the subject associate a pre-composed array of haptic signals with potentially useful messages or status notifications from her portable audio player?
- What associations between haptic signals and audio player messages will be made by the subject?
- What should the form factor of a haptic display for a portable audio player be?

#### **5.3.4.2 Process**

Prior to conducting this session, a series of haptic renderings was produced for an Engineering Acoustics vibrotactile transducer (called the C2 Tactor, hereafter “tactor”) [13] driven by a standard 1/8” audio port (Figure 15) and for the proprietary “Twiddler” [45], an actuated knob fitted with an encoder for dual input and output purposes (Figure 16).



**Figure 15: C2 Tactor**



**Figure 16: The Twiddler**

These devices were chosen for primary evaluation with the user in this session because they differed in their intensity, method of force or tactile feedback, and form

factor. The Twiddler device mimics the scroll wheel interface of the subject's portable audio player, allowing the evaluation of haptic feedback that would be delivered in direct response to finger input. The tactor could be flexibly placed or mounted on the subject's body for evaluation in different placement scenarios, and as it delivers a buzz-like tactile sensation to the skin, the user could leverage previous experiences with cellular phone or pager motors in the current design process while being subconsciously reminded of their potential drawbacks.

The following qualitatively disparate haptic signals were produced for the tactor using the Audacity [2] application<sup>6</sup> on the basis of experimentation:

- 1) **Plucking** – A 50 ms “plucking” profile (with gradually but non-uniformly decreasing amplitude) at 300 Hz, repeated indefinitely.
- 2) **Square** – A 50 ms, 300 Hz square wave followed by 50 ms of silence, repeated indefinitely.
- 3) **Rising** – 4 300 Hz square waves with duration 20 ms with gradually increasing amplitude and fading on the edges, repeated indefinitely.

Screenshots of the tactile waveforms are seen in Figure 17. 300 Hz was selected as the base frequency for each waveform, as preliminary experimentation with the tactor revealed that this produced the most intense tactile sensations. The strength of the haptic feedback could be controlled by sound card volume in the case that smaller amplitudes are desirable.

---

<sup>6</sup> Audacity is a free, cross-platform digital audio editing tool. Its waveform generation tools were utilized for the efforts described in this chapter.

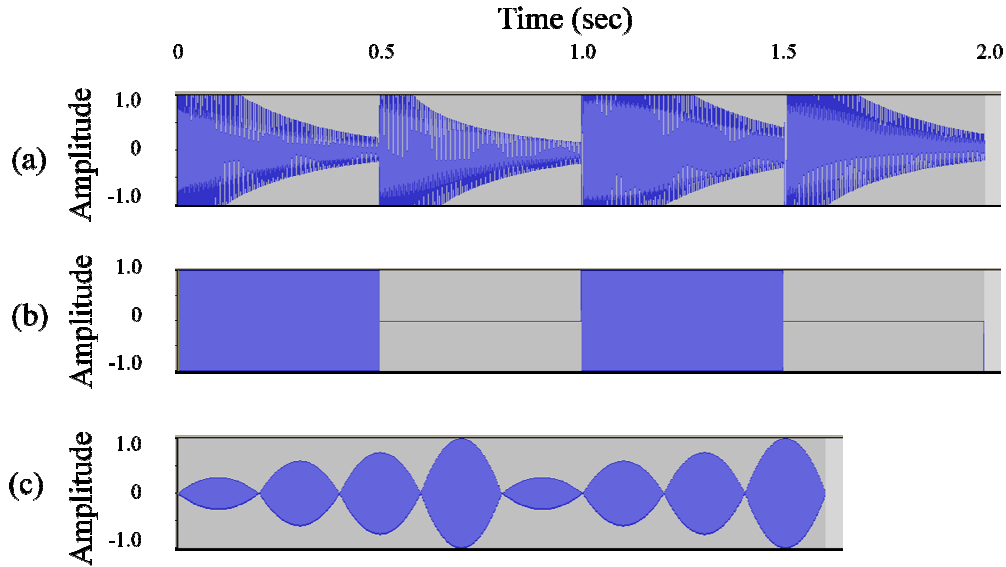
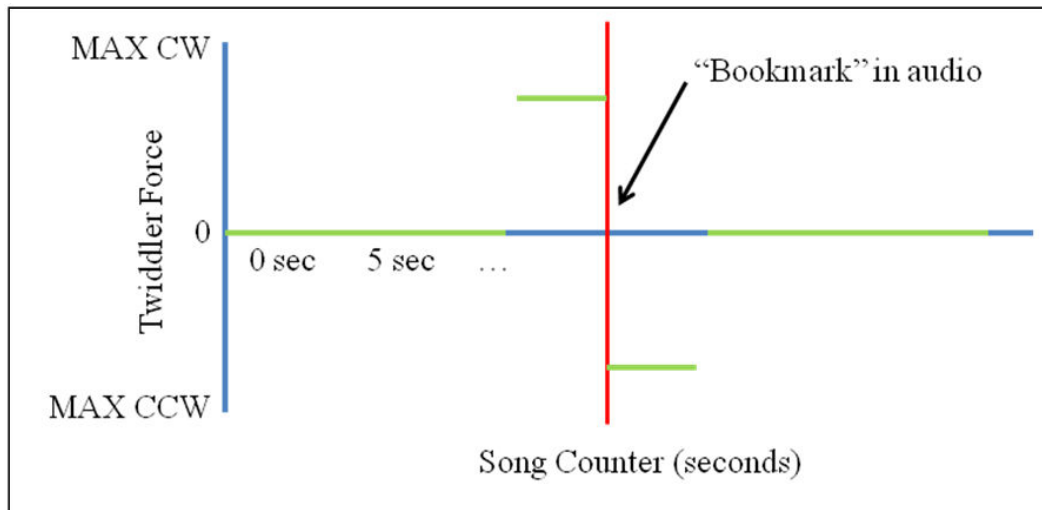


Figure 17: “Plucking” (a), “square” (b) and “rising” (c) factor waveforms

A Java application was developed to interface the Twiddler device with a customized MP3 player running on a PC. As the Twiddler’s knob was scrolled left or right, the song would be advanced backward or forward in real time to match. A haptic rendering environment was developed for the device to deliver forces that would pull the user’s finger towards designated “magnetic” areas of the song. These areas corresponded to several locations of interest in one of the subject’s favourite songs (as provided to the experimenter via e-mail). The following diagram (Figure 18) illustrates the force profile of the Twiddler designed for the current session. Green lines indicate forces in clockwise or counter-clockwise directions. Note that the forces attempt to hold the user to a position of interest, similar to the neutral point of a spring (indicated with a red line). The forces were not modeled specifically on the behaviour of a spring, but instead directed the knob with constant torque toward the target positions within a certain threshold.



**Figure 18: Force profile of Twiddler scroll wheel with bookmarking**

An additional sound card was installed into the PC used for this session to allow simultaneous haptic (tactor) and musical output from the system.

With technological preparations in place, the subject was introduced to the Twiddler and tactor devices, and was exposed to each of the haptic signals in succession. For the Twiddler case, Beril was asked to scroll through a song, imagining that her player was attempting to communicate information to her through the knob, and to indicate what the player was telling her. Likewise for the tactor cases, the subject was asked to assume that a portable music player was transmitting the haptic feedback to her fingers, and to assign meaning to the signals. Her reactions to the signals was also discussed, to determine the extent to which the signals were pleasant and informative, and where improvements could be made for the next iteration of haptic feedback development. During this final interview process, potential locations for wearing or mounting the devices were discussed, as well as potential alternatives to the sensations produced by the tactor and Twiddler.



#### **5.3.4.3 *Summary of Session***

When Beril encountered the “magnetic” areas of the Twiddler environment, she was initially surprised. She indicated that the knob was attempting to keep her at a particular point in the song and disallowing her control over the player. Beril was immediately aware that the magnetic areas were aligned with her self-reported favourite parts of the song being played, and said that during rapid scrolling, this form of feedback would indicate “special” parts of the song to which she should direct her focus. She said that she would prefer to have these special indicators approximately 30 seconds earlier than she experienced in this session, to allow the music time to build into the climaxes she enjoys.

Beril suggested that a similar feedback mechanism could be used to provide information to her about tracks currently being played, such as minute-by-minute “chimes” that would give her temporal awareness during blind scrolling (an idea previously explored in [47]), or “tick”-like forces that would match the rhythm of the song being played. She also suggested placing force indicators at musical landmarks of a song (such as the onset of the bridge or chorus) to help with blind navigation. She indicated that she preferred not to be restricted in her scrolling movements, but light guidance would be helpful.

When the “plucking” haptic signal was continuously displayed via the tactor to the user, she indicated that it raised her awareness immediately. She imagined it being annoying if it was displayed for long periods of time, but that it served as an alert of an important message. In the musical context, she indicated that this message would suggest a pending major change to the music, and could intensify the anticipation of arriving at musical resolution.

Beril identified the “square” haptic profile as “strong” and “boring” and as something that would not sustain her attention. She indicated that she would switch it off if possible. She indicated that she could imagine this signal being displayed when a song she does not listen to often begins to play as a signal that the song may not be “appropriate” for her. At the same time, she indicated that the signal could “calm her down” when she is nervous.

Beril noted that the “rising” profile was faster than the other two, and contained a number of “steps”. She labelled it as “upbeat” and “intriguing” and suggested that it would indicate that a part of a song that she would really like is fast approaching.

Comparing the tactor with the Twiddler, Beril indicated that the tactor felt more natural, and she felt that she remained in better control of the system when receiving haptic feedback using a separate mechanism from the scrolling device. She indicated that she would probably be confused with “messages” from the Twiddler with unfamiliar music, but that she would “get used to it” over time.

Beril indicated that she could “tune out” the tactor much more easily than the Twiddler (as the latter required explicit finger movements to activate), and anticipates that messages delivered using this device would fade into the background over time (a slight contrast to earlier responses, but made after greater exposure to the signals).

When asked to imagine alternative mechanisms for delivering feedback (aside from the Twiddler and tactor), Beril indicated that she would welcome some form of visual augmentation to the haptic mechanisms – temperature and colour changes to the device were specifically suggested. Beril suggested that she could imagine wearing the tactor display on her triceps so that message could be easily detected but would leave her hands free.

#### **5.3.4.4 *Implications for Design***

Passively delivered haptic feedback appears to be the most promising avenue for future research. Beril associated the three tactor haptic signals with different messages from the player. Many of these imagined messages had qualities that would be ordinarily associated with human companionship, such as warnings that upcoming content is likely to be enjoyed or disliked. It appears that Beril is comfortable fostering a relationship with her player that transcends a master-slave scenario, and is open to receiving suggestions from her player that would ordinarily only be communicated between humans.

Receiving direct information regarding the status of musical playback appeared to be preferred to be delivered by the Twiddler during blind scrolling activities. This information should be communicated directly on the scroll wheel of her music player in a manner that does not inhibit her finger movement.

The implications of the results of this session on the HALO interaction design process are that:

- Continuous, low-attention feedback via a wearable tactor would be appreciated by the subject and could be ignored if the subject became disinterested.
- Human-like suggestions or indications about the musical experience should be portrayed via the tactor.
- Song and scrolling information should be delivered using feedback directly on the iPod scroll wheel.
- Haptic signals that involve animated “rising” magnitude excite Beril and heighten her anticipation.

- Haptic signals that involve animated “lowering” magnitude capture Beril’s attention and act as “wake up” calls to direct her to important information.
- Consistent haptic signals (i.e., “square”) appear to bore the user, but the sensation of the signal can act as a calming mechanism in situations of anxiety.

### **5.3.5 Session 5 – Exploring Explicit Control and Haptic Feedback**

#### **5.3.5.1 Research Questions**

The fifth participatory design session aimed to deepen our understanding of an appropriate haptic output space for a HALO audio player as well as its control space: i.e. to investigate the augmentation of (implicit) physiological input signals with explicit input mechanisms. We hypothesized that Beril would welcome some form of override control over the player, and aimed to evaluate an integrated tactor input system that we conjectured would require minimal attention to interact with.

We aimed to determine how Beril would communicate certain ideas or commands to the player, and when she felt these communications would be necessary. We also aimed to evaluate several more haptic signals using the tactor to broaden the communication language initially established in the previous session.

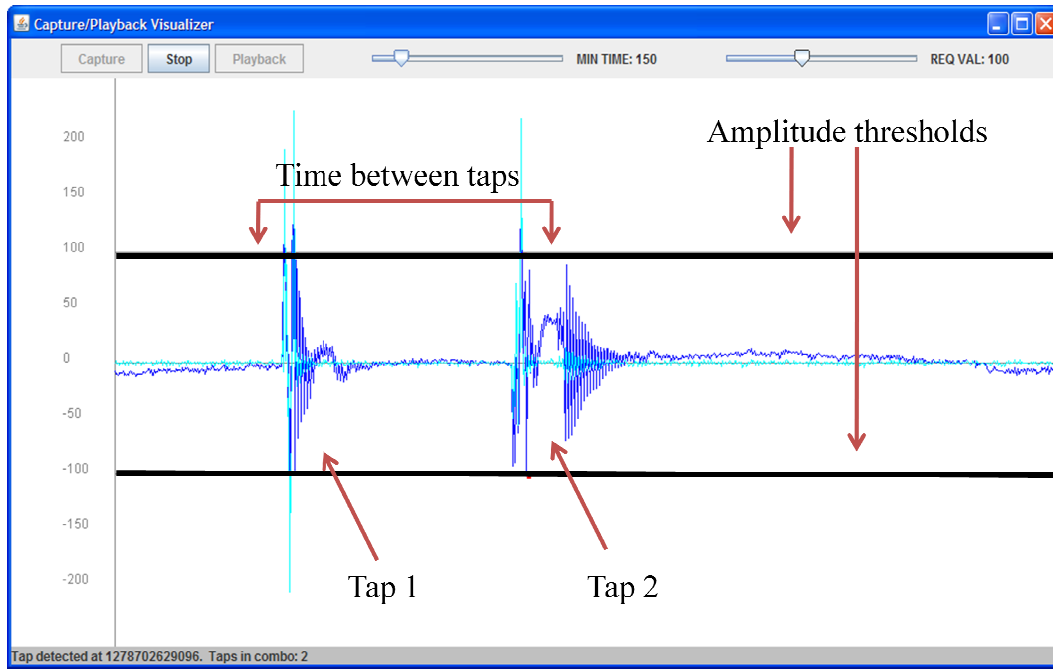
The summarized research questions of this session were:

- Using a broader set of haptic signals, what associations between these signals and audio player messages will be made by the subject? Will these be consistent with her previous associations?
- When will the user feel it necessary to communicate messages to the player, and what will the messages be?
- How would the user communicate these messages to the player using gestures?

#### **5.3.5.2 *Pre-Session Tactor Modification***

Prior to conducting this session, the tactor display introduced to Beril in session four was augmented to support input. As the tactor behaves like a speaker when connected to an audio source, connecting the device to an input audio jack causes allows it to function as a low-fidelity microphone, which in turn allows for low-resolution touch-based input. Experimentation with the modified device revealed that taps on its surface could be easily captured by recording software, but anything more nuanced could not be easily captured.

An application was authored using Java that visualizes and processes input from the tactor in real-time. This application was integrated with the music player used in the previous (and current) session to enable control over audio playback using the tactor. Due to the low resolution of the tactor's microphone capabilities, the input system was designed to identify "taps" on the device and group immediately consecutive taps together (allowing, for example double- and triple-tapping). A screenshot of the input analysis software is shown in Figure 19 (modified for legibility). In the figure, a double-tap combination has been identified by the software, as the magnitude of the input signal crossed the configured threshold twice; the time between the taps was greater than the minimum required time to ensure disparity (in this case, 150 ms) and less than the maximum time specified for multi-tap correlation (in this case, 500 ms).



**Figure 19: Tactor input visualizer and analyzer.**

### 5.3.5.3 Process

At the outset of this session, Beril was supplied with the tactor and asked to communicate a variety of control messages to the player. These varied from informational (“I don’t like what you’re playing”) to direct commands (“Pause here”). The full set of messages is presented in Table 8. The subject was asked to demonstrate and explain any other messages she could imagine communicating to the player at the conclusion of this phase.

Following this, a preliminary input language for direct system commands based on tapping (the only possibility given technological constraints) was evaluated with Beril. The language was described to the user as follows (Table 7):

# of taps	Function
1	Pause or resume song.
2	Advance to the next song.
3	Start the current song again.
4	Advance the current song 20 seconds.

**Table 7: Preliminary tap-based tactor input language**

Beril was then asked to perform each of the functions approximately 20 times, on demand, in a random order. Following this, a discussion on the input language ensued to evaluate its naturalness (as subjectively perceived by the user) and its limitations.

The final activity of this session was to allow the user to associate a larger suite of haptic signals with messages from the player. Two of the three signals from the previous session (“pluck” and “rising”) were re-evaluated to check for consistency between sessions. Four additional tactile signals were developed using Audacity for use in this session. They are described as follows:

- 1) **Constant tone** – a constant 300 Hz square waveform, repeated indefinitely (Figure 20a).
- 2) **Fast pitches** – 20 ms 300 Hz square waveforms followed by 20 ms of silence, repeated indefinitely (Figure 20b).
- 3) **Faster pitches** – 10 ms 300 Hz square waveforms followed by 10 ms of silence, repeated indefinitely (Figure 20c)
- 4) **Slow-fast-fast** – 40 ms 300 Hz square waveform followed by 10 ms of silence, then a 15 ms 300 Hz square waveform followed by 10 ms of silence, and another 15 ms 300 Hz square waveform followed by 10 ms of silence (Figure 20d).

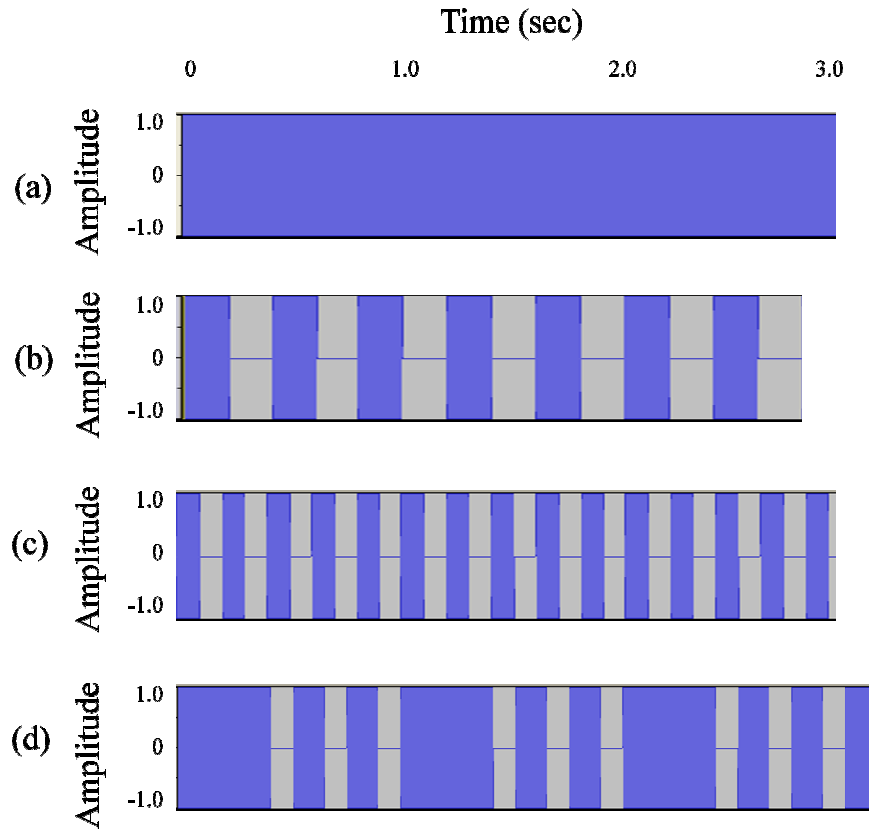


Figure 20: “Constant tone” (a), “fast pitches” (b), “faster pitches” (c), “slow-fast-fast” (d) tactor waveforms

#### 5.3.5.4 Summary of Session

Beril developed the following gesture-based input language for communicating with her audio player via the tactor:

Message	Gesture
I don't like what you're playing.	Tap (~ ½ second).
I like what you're playing.	Nothing (upbeat songs) or tap along with the beat; Rub (slow/melodic songs).
Change the song immediately.	Very quick rub (away from cable).
Move ahead to a part I will like.	Very quick rub (toward cable).

(continued on next page)



(continued from previous page)

Message	Gesture
Is there a part of this song I will like? Move there.	Shake device (question). Double click (confirm).
Pause here.	Single long press.
Put a bookmark here.	“Checkmark” symbol over surface.
Turn up the volume.	Diagonal rub, lower left to upper right.
Turn down the volume.	Diagonal rub, upper right to lower left.
(User suggestion) Change speed of song.	Shake at target speed continuously.
(User suggestion) That was wrong.	Squeeze

**Table 8: User-defined factor input language**

The subject was presented with the tap-based input language and tested on her ability to use the language to perform tasks. Over the course of her test, she made no mistakes, asked no questions, and performed all tasks with 100% accuracy. Analysis of her tapping patterns indicated that delays between grouped taps were an average of 234 ms in length, with a minimum pause time of 125 ms and a maximum time of 282 ms. Delays tended to increase between consecutive tasks further in the group (i.e., there were larger average pauses between the second and third taps in a sequence than between the first and second). The default configuration of the system to group taps between 100 and 500 ms apart appeared successful for this participant.

Beril said that the tap-based language easily became “automatic” to her, especially for pausing/resuming (1 tap) and advancing to the next song (2 taps). She said that she was reminded of single- and double-clicking a mouse in these cases. Despite making no mistakes, Beril indicated that it required much more mental effort to make three- or four-tap series, and said that such a language “would take some getting used to”. She also indicated that during slow pieces she would like find it unnatural to tap quickly, and may require the system to allow a larger delay between related taps.

When re-exposed to the “pluck” and “rising” haptic signals, Beril’s associations were not entirely consistent with the previous session. She indicated that the “pluck” signal indicated that “something is coming up” that requires no immediate response, a similar but slightly weaker importance judgment than before. “Rising” was described in this case as an indication that “something bad is coming” – for example, that the next song in the playlist may not be enjoyable. Previously, Beril associated this signal with a message that was upbeat and intriguing – much more positive labels.

Beril likened the “constant tone” signal to a “ring of energy...like slamming down a staff” (as in fantasy movies). She indicated that it was highly authoritative, and could be used to indicate the onset or conclusion of a player process of some sort. She indicated that she could ignore it if desired, but would feel compelled to look at her player to determine the cause of the signal provided it was used sparingly.

Beril imagined “fast pitches” being used to confirm a requested tempo change, but that the speed of the pitches would match the altered tempo rather than remaining constant. “Faster pitches” elicited a similar response, but she imagined exercise and walking scenarios requiring this signal. She said that this signal was more “shocking” and “attention-getting” than “fast pitches” and that it would hold her attention indefinitely until she dealt with the source of the interruption.

Beril said that “slow-fast-fast” gave her the feeling of “jump[ing] and skid[ding]”, and imagined the movements involved in snowboarding and children playing. She indicated that the signal would be appropriate when the player was warning of an impending automatic change to the music, and described this warning as indicating the music is “destined to be changed”. She suggested that she should be able to cancel the impending change using a squeeze (as suggested in the first exercise), and expanded

this comment to stopping the device from any autonomous action or continuous haptic feedback cycle.

#### **5.3.5.5 *Implications for Design***

Beril supplied us with a rich gesture-based input language for her audio player. None of the gestures directly conflicted with each other, despite strong similarities (especially in directional rubs). In contrast to this rich user-provided input language, our limited, naïve tap-based system proved effective in investigative evaluation with Beril. This suggests that it would be more beneficial to support gestures for this participant that can be autonomously disambiguated in an autonomous fashion than those that she “naturally” defined. In defining an input language for the player, listening for more than two grouped taps at a time should be avoided due to the mental load this requires. The gesture language should transcend taps for this reason, and Beril’s inclination to make spatial gestures (such as checkmarks and directional lines) should be considered. The modified tactor display does not have the input resolution to support such a language, so other technological alternatives will need to be considered. Input devices using only sound would likely be infeasible for detecting directional input.

There were inconsistencies between sessions in the haptic signal classification process. This indicates that while natural associations with haptic messages should be used as a basis for output design, easily distinguishable and identifiable signals should be the main focus. Signals chosen for haptic output must be largely attention-neutral and ignorable. Many of the waveforms displayed to the user were classified as “attention grabbing” and difficult to ignore, and in some cases Beril indicated that these signals would cause her to investigate the causes on the main player console. Unless extremely important messages are required to be conveyed to Beril that demand her continued

attention, these signals should be avoided in practice. One can imagine, for example, using highly distracting and attention-holding signals to compel Beril to focus on an important task (such as studying). Signals such as “constant tone” can be used to capture attention for less mission-critical messages, as Beril indicated that these could be easily ignored.

### **5.3.6 Session 6 – Wizard of Oz Simulation of a Closed System**

#### **5.3.6.1 Research Questions**

The purpose of the sixth participatory design session was to simulate and evaluate a closed haptic-affect interaction loop, whose behaviour was dictated based on results from previous sessions. Due to the present inability to richly and automatically model affect to the extent required for robust testing, a Wizard of Oz [21] process was used to simulate the affect classifier and to make changes to system behaviour.

The specific research questions we aimed to address were:

- What will Beril’s immediate reactions to the closed system be when the system is successful in its estimations and actions, and what will they be when it fails?
- What inconsistencies in Beril’s behaviour will make future implementation of an autonomous system difficult?
- What frustrations do the simulated HALO loop cause that Beril has not already identified?
- What changes to the interaction loop, as tested, would Beril like to make?

#### **5.3.6.2 Process**

Two playlists of music were prepared from the researcher’s personal collection, each containing five songs. In the first playlist, one song was conjectured to be liked

very much by Beril (based on previous surveys and discussions), one song was conjectured to be highly disliked, and three conjectured to be fall between neutral and liked. In the second playlist, two songs were chosen with the intention of being disliked, two with the intention of being liked, and one was chosen to be neutral.

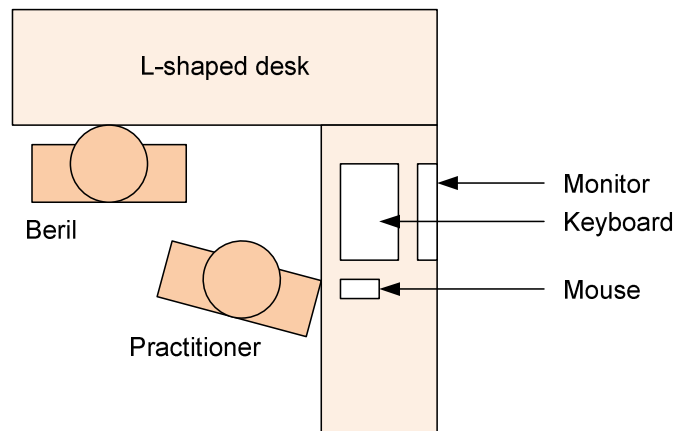
Each song in the playlist was associated with haptic signals that had been evaluated and classified by Beril in the previous sessions. The signals were designed to communicate that the player was aware that the song was being enjoyed, that a good part of a song was coming up, or that the player intends to change the currently playing track. The signal association was made on the basis of Beril's feedback in the previous sessions, and were either chosen to be helpful or purposely incorrect to gauge reactions.

Beril was asked to establish a single, unambiguous gesture to communicate to the Wizard of Oz practitioner that she would like to explicitly change the song. This was done to ensure that the experiment did not go completely off track at any time and require the practitioner to break out of his "computer persona". Beril was informed that she could use this signal anytime during the process.

Two Wizard of Oz sessions were run, each with one of the playlists. Beril was informed that she would be unable to communicate with the practitioner in natural language, and that he would act like a computer, responding to input on the tactor device and also responding to perceived emotional state. At the conclusion of each session, a discussion about correct and incorrect system behaviour was conducted and Beril was given a chance to rank each of the songs she was exposed to on a scale of enjoyment (0 – 10). Feedback on the first Wizard of Oz session was used where possible to improve the realism and helpfulness of the simulated system in the second Wizard of Oz session.

The practitioner inferred Beril's affective state by observing her body language (i.e., facial expressions, posture, etc.) throughout the Wizard of Oz exercises. The practitioner came to become very familiar with Beril's body language as it relates to musical contexts over the course of the previous five participatory design sessions, and as such this method of measurement was deemed satisfactorily reliable for current purposes. Discussions at the conclusion of each Wizard of Oz session were made to confirm observations and resolve ambiguities.

Beril and the practitioner were seated at an L-shaped desk for the Wizard of Oz sessions. Beril faced a wall behind the desk, and the practitioner sat behind her at an angle that would facilitate observing her facial expressions. The practitioner sat at such an angle as to minimize the extent he and the computer equipment controlling the simulation were visible to Beril (i.e., in her peripheral vision). Figure 21 illustrates the physical setup used for these sessions. Beril's headphones, which were connected to the computer equipment on the right side of the figure, are not pictured.



**Figure 21: Physical setup for Wizard of Oz sessions**

#### **5.3.6.3 *Summary of Session***

Beril selected a quick tap as a gesture to immediately change the current song during the experiment.

During the first song of the first Wizard of Oz session, the “plucking” waveform was displayed to the user in an attempt to communicate that a portion of the song that she would enjoy was coming up. Beril gestured up to move forward in the song, which the Wizard of Oz practitioner interpreted as a request to skip to the next song. This error was identified in the follow-up discussion. Beril indicated that the haptic signal was useful, and that her trust in the system would increase knowing that correct information had been conveyed and captured.

The second song, which was rated a 0/10 on the enjoyment scale by Beril, was skipped over almost immediately (using the established “quick tap” gesture). The intention of the Wizard of Oz practitioner had been to display the “slow-fast-fast” signal to indicate that the song was about to be automatically change, but he was unable to do so given the immediate change.

The third song in the first session was selected to be highly liked by Beril. Again, the “plucking” waveform was displayed to indicate an upcoming favourable portion of the song. Beril squeezed the display to silence this display, and the practitioner halted haptic playback. Beril made a clockwise motion with her finger on the tactor (a gesture never before witnessed by the practitioner) and this was interpreted as a request for a volume change. In subsequent discussion, Beril indicated that she intended to scroll forward in the song to the more desirable location inferred by the haptic feedback. Due to this misinterpretation and an inability by the practitioner to determine Beril’s true intentions, Beril advanced to the final song using the established “quick tap” gesture.

During the final song Beril tapped along to the beat. “Constant tone” was displayed to determine Beril’s reaction to an unanticipated haptic signal. She indicated in the follow-up interview that she felt the player was “listening” to her taps and tapping along with her. During this discussion, it was determined that whenever she would make contact with the surface of the tactor, she would perceive an intensity shift in the constant vibration pattern being displayed, producing the co-operative tapping effect.

Interview feedback was considered by the practitioner when conducting the second Wizard of Oz session. Fewer unplanned misinterpretations occurred in the second session, and the instances of error were again due to misinterpreting scrolling gestures. The second session involved two songs that were hypothesized to be disliked by Beril, two that were hypothesized to be liked, and one that the practitioner was unsure of.

The “plucking” waveform was again played during the first song, which Beril later confirmed as being helpful for identifying an enjoyable upcoming change to the song. “Faster pitches” was the assigned haptic message for the second song, an upbeat jazz fusion piece. Beril tapped along with this piece and, about 30 seconds in, repeated the clockwise gesture indicating a desired advance in time. The practitioner interpreted this correctly, but Beril disliked the part of the song that was advanced to, and ultimately changed the song completely using a “quick tap”.

During playback of the third song in the second session, an associate interrupted Beril. The Wizard of Oz practitioner paused playback to allow Beril to attend to the interruption. When the interruption ceased, Beril tapped to resume playback, and then tapped again to skip to the next song. Beril later related that she regretted skipping over the song and attempted to return to it with a left scroll gesture. This was not properly interpreted by the practitioner; instead, he reduced the volume of the device, which Beril



reversed with the opposite (right scroll) gesture. Ultimately, the song was skipped by Beril using the “quick tap” gesture.

The final song – a heavy metal selection – was selected to be disliked by Beril. The practitioner halted playback of the song on her behalf, hypothesizing that she would dislike the song. Beril stated that she was still trying to “give the song a chance” when it was terminated, but admitted that it was the right choice and she likely would have stopped the song herself shortly after.

Beril related that constant haptic feedback provided a sense of comfort, and enjoyed the expressive capabilities of the player over her current player. She mentioned that the feedback provided by the player was “like having a friend there” and that she was willing to forgive mistakes due to the relationship she felt was developing with her player. She also said that haptic feedback was ergonomic and private, and she would not be uncomfortable using such a device in public.

Beril indicated that she would prefer to remain in explicit control of most system functions, and simply receive “suggestions” from the device that she could ignore, silence or act upon. She indicated that the tactor gave her an additional layer of control over the player than her iPod does.

In subsequent discussions, Beril dictated the appropriate behaviour of the system when an interruption was noticed. She indicated that depending on the importance and requirements of the interruption, she would want the player to pause playback or play ambient, relaxing music. Depending on her level of enjoyment of the song she was listening to at the time of the interruption, she would want the player to either resume playback where the song was paused (highly enjoyable songs), or skip to the beginning of a completely new song (songs that she was not deeply involved in).

Beril said that she would definitely prefer to mount the tactor on her upper arm rather than holding it, despite making it less accessible for input. She reiterated that colour or temperature feedback on some piece of the device would provide her with more natural and non-intrusive sources of information.

Discussion in the session concluded with the following quote:

*“I did not think that ... I would like the way it interacts with me. But now I like it. If it's able to sense my physiological [state], you know, and it tries to do something with me, even if it makes mistakes I don't mind it. From my perspective, I understand it can't be something perfect, but I still like the fact that it gives me feedback and does stuff with me.”*

#### **5.3.6.4 Implications for Design**

Beril's trust in the system to make appropriate decisions was reinforced with a history of correct decisions (despite a number of incorrect responses). Her threshold for error tolerance was increased over what she expected because of the human-like relationship that she appeared to foster with the device. As a highly active music listener with strong feelings towards songs she enjoys, the added haptic signals heightened her enjoyment and simulated the feelings she has when sharing music with friends. Messages communicated with a purpose were either followed or ignored with a brief squeeze; Beril indicated that these interruptions were not drastic enough to diminish her experience.

The results of this session have many implications for design. First, inappropriate haptic interruptions made by the system and other system behaviour impact Beril according to the extent to which she is engrossed in her music. Regular tapping on the tactor appeared to represent enjoyment in every case during the Wizard of Oz sessions, and Beril confirmed these findings.

Many haptic messages were not found to be distracting by Beril, and she indicated many were easily ignorable (although other channels may be, in fact, easier to ignore). If a HALO device is unsure if the user is enjoying an audio track due to a lack of tapping, a low-attention haptic message could be displayed to the user. Based on the response to these messages, the player may be able to make better behavioural decisions for a given context.

It appears to be beneficial for this participant to receive frequent haptic signals as music is played. Beril likened the sensations caused by the tactor to the presence of a friend, and this made her more likely to tolerate mistakes. The benefit of this finding is it allows subtle differences in haptic output to communicate large volumes of information via a medium that this participant finds easy to ignore if needed. With a continuous stream of haptic feedback, the tactile sensations become more ambient and attention can be engaged at will. In the case that a highly important message needs to be communicated, there is evidence that strong, repetitive and “uninteresting” signals can be used for this purpose.

It is important to consider the effects that the human Wizard of Oz practitioner may have had in the outcomes of this session. Despite the practitioner’s endeavours to simulate the HALO paradigm realistically and his steadfast failure to respond to Beril’s natural language, the friend-like bond that Beril felt with the “system” may have, unconsciously, referred to the human practitioner. As this was the sixth participatory session in the series, a relationship between the researcher and subject had already been fostered, and simply asking the participant to pretend she was interacting with a computer system may have been insufficient to completely mitigate the effects of this bond.

Regardless of the extent of this issue, we can still conclude that if the HALO system *can* foster a human-like relationship with the user, the effects are positive.

Engagement with music is also an important factor to consider when determining appropriate behaviour upon detecting an interruption to the user. System inferences in this regard can be confirmed or negated based on the user's behaviour after an interruption takes place. If the system, for example, fails to select a new track after resuming the music and this function is quickly performed by the user, the system model can potentially learn from this mistake.

### **5.3.7 Session 7 – Follow-up Physiological Signal Measurement**

Due to our previous lack of success with real-time affect classification and model production, we endeavoured to collect more data for further analysis. A limitation of our music-affect study (Chapter 4) was that data was collected from a number of participants who had disparate physiological reactions to music. Additionally, approximately half of the data collected was done so as participants were exposed to both music and a secondary task (word search), which could have a significant impact on the categorization abilities of the system – the secondary task could have introduced significant irregularities in the data.

We determined that collecting data from a single subject, exposed only to music of different genres, would maximize consistency in collected data. Due to the participatory design subject's familiarity with the sensor technology, the author and an associated research group selected her for participation in this data collection process.

A follow-up data collection process was therefore undertaken with the participatory design subject. 22 trials were conducted using the same experimental protocol as the music-affect study, skipping word search and “user control” trials.

9/22 trials involved songs that were self-rated as “liked”, 6/22 involved “neutral” songs, and 7 involved “disliked” songs. At the time of this writing, analysis of this follow-up work (using  $k$ -nearest neighbours) as well as continued data collection is ongoing by the associated research group. Details appear in Appendix D (experiments 2 and 3).

## **5.4 Toward Design Guidelines for HALO-Enabled Systems**

On the basis of these participatory design sessions, with consideration of the results from the focus group and music-affect study, a set of design guidelines for a HALO-enabled audio player for this participant are compiled below, and are categorized according to their function in the interaction loop. A summary, per function, is given at the conclusion of each sub-section.

The guidelines presented in this section are based largely on the preferences for system interaction that were expressed by Beril during the participatory design sessions. These preferences were preliminarily validated using Wizard of Oz testing as well as exercises to check for consistency between sessions, but have not yet been tested or verified with a completely working prototype.

### **5.4.1 Loop Input Mechanisms**

Sensors that detect user state are essential to the correct operation of the loop. Whatever sensors are used, physiological or not, they need to be able to rapidly and effectively measure and identify:

- The user’s level of engagement in her music.
- The user’s level of enjoyment in her music.
- Extraneous interruptions to the user’s audio listening experience.

In addition to this list, measures of anxiety, as well as other physiological measures, could facilitate additional system functionality and help to disambiguate uncertain contexts. For example, a heightened heart rate could be associated with a desirable (e.g., excitement) or undesirable (e.g., anxiety) state. In this case, a measure of anxiety could influence whether the system should sustain or attempt to lower the arousal of the user. Additionally, non-physiological models and sensors (e.g., vision, eye tracking) have been previously investigated to determine users' interruptability in a given context and could be integrated into the HALO paradigm for the same or related purposes [14].

In addition to affect sensors, an override mechanism should be developed that allows the user the ability to rapidly correct erroneous system behaviour. Usage of this override mechanism should result in changes to the system model. A tactile input system has been investigated in the current research, but alternatives, such as voice control, can also be considered.

The social and practical implications of the particular mechanisms used to facilitate explicit control of the system need to be considered before implementation. For example, voice control may not be effective in environments that are too noisy (such as a crowded streetcar) or too quiet (such as a library). Recognition accuracy plays an important role in ensuring that attentional requirements are minimized; this metric is tied directly to the difficulty level for the user to articulate his or her control demands in various environments.

A HALO-augmented audio player should maintain its existing control framework (manual audio controls) as well as settings to control the extent to which action is taken in the absence of explicit commands. Users should perceive a high level of control and be able to demand it fully when they desire, or else the system will not likely be adopted.

This could be further enhanced by developing a rich gesture-based language for explicit interaction and an associated accessible input device, likely wearable.

**Summary of Guidelines:**

- Sensors that measure engagement, enjoyment and interruptions are the main inputs to the system. These need not be physiological, but should be autonomous and require no attention from the user.
- Other physiological or non-physiological sensors can potentially help with disambiguation.
- An override mechanism is helpful for rapid correction of errors or overriding control. Tactile systems for this purpose are promising, but alternatives, such as voice control, can be considered.
- Existing device interfaces should remain intact, but can be supplemented with more easily accessible input systems.

#### **5.4.2 Loop Output Mechanisms**

Haptic feedback appears to be a highly effective mechanism for displaying information to the user that augments existing portable audio player capabilities; tactile and continuously variable temperature signals could both be used for this purpose. A wearable tactile device, likely to be placed on the upper arm, is preferred to force devices or vibrators embedded directly in the player, as this facilitates easy access and ensures consistent contact with the wearer's skin. In the case that the user manipulates the player console itself, feedback in the interactive surfaces can provide information, which is especially helpful for blind scrolling.

Continuous feedback of system status is important to build a human-machine relationship with the player. This relationship mitigates the impact of system failures and provides a calming, enjoyable experience.

Haptic signals that are used frequently should not dominate the user's attention and should be easily ignorable. It is best to avoid extremely rapid or monotonous signals for these purposes, and reserve these for conveying messages that are intended to interrupt and demand the user's attention. More appropriate signals for continuous feedback involve amplitude variations and soft edges.

Informational content can be delivered to the user directly on the audio player's body to enable detection during blind scrolling. This category of largely system status-related messages can also be conveyed using non-haptic means. Ambient changes to the colour or temperature of the device or a wearable extension are appropriate for these purposes, but it is important to note that colour changes require the user's visual attention to acknowledge.

**Summary of Guidelines:**

- Use a wearable haptic device to communicate continuous system feedback.
- Choose intrusive signals only when a message is important, and non-intrusive for ongoing communication.
- Augment existing audio interfaces with haptic displays to allow delivery of information during blind scrolling.
- Other ambient changes (e.g., temperature) can be investigated as alternatives.



### 5.4.3 System Behaviour and Interaction Language

System behaviour depends heavily on the user's affective state. A high level of engagement indicates that the system should not interfere with current operation, except to warn users of impending negative events.

System behaviour in response to an external, detected interruption should vary based on level of engagement as well. If the user is heavily engaged and must attend to an interruption, the system should preserve its current settings to allow a return to this engagement level. If the user is heavily engaged and does not have to attend to an interruption, a granular series of adjustments should be made to preserve engagement; volume should first increase, and failing the effectiveness of this step, an alternate song should be selected for playback. Low levels of engagement allow the system to make more mistakes and also to take more drastic steps with external interruption resolution; songs can be changed or paused with minimal effect to the user experience.

Level of enjoyment of a song should be used to determine whether the system should select and play back an alternative. If the level of enjoyment is uncertain, low-attention haptic messages should be used to indicate to the user that there are other options available. If the level of enjoyment is certain and low, the system should immediately change it, first giving a pre-emptive warning of the impending change. Based on a historical affect model, haptic messages should be delivered to the user when segments of upcoming songs are projected to be enjoyable or disliked.

Explicit input to the system can be used to effect system behaviour changes by the user and maintain his or her feeling of control over the player. Explicit commands (such as “stop” or “pause”), if defined, should be respected immediately, and should be used to tailor the system model for future decision-making. Non-command input, such as

rhythmic tapping, should be used for similar model development, and should be responded to with a low-attention, ignorable haptic signal.

As a general rule, the system should behave in a tentative but helpful manner. Confident, autonomous changes of songs should only occur when the system is certain of the user's displeasure based on affect measurement and an internal model, or when the level of engagement is low enough that large mistakes are tolerable. Like a new friend, the player should take few bold risks early in the "relationship" with the user. This design requirement falls in line with the findings of previous work in adaptive interfaces (e.g., [17]).

#### **Summary of Guidelines:**

- Use measured level of engagement as a means to vary the intrusiveness of a system interruption and to determine how best to mitigate an external interruption.
- Use level of enjoyment to further modulate reactions to an interruption – the higher, the less likely the system should be changed from its current state.
- The system should be tentative but helpful. Avoid large, intrusive system interruptions and large variations in behaviour unless the system has very high confidence in their utility.

## **5.5 Discussion and Open Questions**

The participatory design sessions allowed us to gather rich, immediate feedback from a single participant with regards to the design and implementation of the HALO

paradigm. This allowed us direct access to feedback on Beril's ideas and their impact on her uncovered pain points as the interaction paradigm was developed. Rather than soliciting iterative feedback from a variety of participants, we endeavoured to produce the core interaction loop on the basis of a single participant's needs, assuming that at least a base set of needs spanning a spectrum of use cases will be addressed in this manner.

The limitations of this approach, however, are clear and were understood from the start. Between-user differences in player usage patterns, media types, and the desired behaviour of the player remain largely unknown; hence the utility of the system for users other than the participatory design subject is still unclear. In the focus group efforts described in Section 3, we found that when multiple focus group members were assembled, some of these differences began to appear, but the overwhelming collective scepticism of the player's abilities dominated discussions; this single-participant exercise was intended as a triangulation that would avoid that issue. In the next step, future work will need to test the generalizability of our resultant HALO interaction design across multiple subjects, and to determine what aspects of the paradigm need to be customized on a per-user basis. This work will likely need to be performed in follow-up design and testing sessions with stakeholders. The present preparatory work provides insight on the basics of HALO-enabled design, and provides a list of important design guidelines for performing this follow-up work.

#### **5.5.1 Projected Customization Process**

It is hypothesized that an effective HALO implementation can be rapidly realized for additional users, using the current design principles as a basis. However, as indicated above, optimal system behaviour may differ, perhaps substantially, on a per-subject basis. The following aspects of the interaction loop should be investigated via one-on-one or

dyad interviews and experimental processes, to mitigate the scepticism encountered in the focus group sessions:

- Situations when the HALO audio player should pre-empt its more “aggressive” actions (such as changing a song) with warning messages, versus those when actions should simply occur, potentially with post-hoc notification.
- The extent to which HALO behaviour should be guided implicitly rather than explicitly.
- The extent to which continuous feedback would be helpful or a detriment to the low-attention goals of the interaction experience, and what specific signals are most appropriate for each user.
- Which biometrically-accessible factors will be most relevant and informative to facilitate other users’ desired musical experiences.
- The specific steps that the system takes in response to an interruption may differ due to factors other than user’s engagement in music. What should be considered on a per-user basis?

It is hypothesized that by investigating these three main categories of system design that an appropriate system could be made for other specific users. If this process is performed with a number of users, “groups” of users with similar customizations may emerge. From these groups, a series of generalized design principles for universal HALO design may become clearer.

## 6 Conclusions

This preliminary work, which investigates the utility and behaviour of the proposed Haptic-Affect Loop in musical settings, attempts to rigorously gather requirements for implementation using a variety of experimental methodologies. It aims to address the specific technological requirements for affect detection and classification, and performs preliminary validation work for producing these classifications in real-time.

Focus groups and participatory design sessions were undertaken to primarily address the former goal. Early results indicated that users were unlikely to adopt the technology for portable audio scenarios due to scepticism in its technological abilities, perceived control domination, and privacy and form factor concerns. At the same time, demonstrations of haptic and physiology-driven technologies to these users appeared to quell these concerns to a degree. Focusing on a single participant to design and evaluate the interaction loop in an iterative process, several design principles emerged. These design principles were, in summary:

- Focus on user engagement and level of enjoyment as a first indicator for suggesting or adapting system behaviour.
- Provide continuous, low-attention haptic output to the user to foster a new human-system relationship, foster trust, and mitigate the effects of erroneous decisions.
- Facilitate a “way out” for the user – a highly accessible override mechanism, at a minimum, improves the utility of the interaction paradigm substantially and instils confidence and a perception of control in the user.

Specific guidelines for producing an interactive HALO-based system were given with implications for required sensor technologies.

Technological validation work was performed to confirm the results of previous work, specifically with regards to GSR sensing, in musical contexts. Orienting responses appear to persist in this setting and can be used to drive autonomous systems. GSR signals are best suited for indicating unexpected interruptions to users' music consumption experiences, and such systems can be rapidly prototyped to test proposed interaction paradigms in computer software. Real-time affect classification using a richer array of signals in musical contexts shows promise, but to date is not ready for implementation.

Future work must be performed to generalize the results presented here, and to evaluate a fully implemented HALO interaction paradigm in context. A series of follow-up studies collecting and evaluating necessary customizations to the currently highly personalized system must be undertaken.

Finally, sensor technologies must be sufficiently portable and minimally intrusive in order to function effectively as the backbone of this technology. Wireless sensors with minimal noise effects would be optimal for the implementation of a HALO system that can be effectively evaluated. With the continued "miniaturization" of computer technologies, finding such sensors in the future seems more than plausible.

## Bibliography

- [1] G.D. Abowd et al., "Cyberguide: a mobile context-aware tour guide," *Wirel. Netw.*, vol. 3, no. 5, pp. 421-433, 1997.
- [2] Audacity. Audacity: Free Audio Editor and Recorder. [Online].  
<http://audacity.sourceforge.net/>
- [3] M.S. Bartlett et al., "Recognizing Facial Expression: Machine Learning and Application to Spontaneous Behavior," *IEEE International Conference on Computer Vision and Pattern Recognition*, pp. 568-573, 2005.
- [4] M.A. Baumann, K.E. MacLean, T.W. Hazelton, and A. McKay, "Emulating human attention-getting practices with wearable haptics," in *Haptics Symposium 2010*, Waltham, MA, 2010.
- [5] J.T. Cacioppo and L.G. Tassinary, "Inferring psychological significance from physiological signals," *American Psychologist*, vol. 45, no. 1, pp. 16-28, 1990.
- [6] J. Chung and G.S. Vercoe, "The affective remixer: Personalized music arranging," in *CHI'06 extended abstracts on Human factors in computing systems*, 2006, p. 398.
- [7] J.A. Coan and J.J.B. Allen, *Handbook of emotion elicitation and assessment*. New York, NY: Oxford University Press, 2007.
- [8] J.F. Cohn, "Foundations of Human Computing: Facial Expression and Emotion," *Int. Conf. on Multimodal Interfaces*, pp. 233-238, 2006.
- [9] C. Conati and H. Maclaren, "Data-driven refinement of a probabilistic model of user affect," *User Modeling 2005*, pp. 40-49, 2005.
- [10] R. Cowie et al., "Emotion recognition in human-computer interaction," *IEEE Signal Processing Magazine*, vol. 18, no. 1, pp. 32-80, 2001.
- [11] S.K. D'Mello and A. Graesser, "Multimodal semi-automated affect detection from conversational cues, gross body language, and facial features," *User Modeling and User-Adapted Interaction*, pp. 1-41, 2010.
- [12] Emotiv Systems. (2009) Emotiv EPOC Neuroheadset. [Online].  
<http://www.emotiv.com/apps/epoc/299/>

- [13] Engineering Acoustics, Inc. C2 Tactor (Data Sheet). [Online].  
<http://www.eaiinfo.com/EAI/PDF%20Documents/C-2%20tactor.pdf>
- [14] J Fogarty et al., "Predicting human interruptibility with sensors," *ACM Transactions on Computer-Human Interaction*, vol. 12, no. 1, pp. 119-146, 2005.
- [15] C.D. Frith and H.A. Allen, "The skin conductance orienting response as an index of attention," *Biological Psychology*, vol. 17, no. 1, pp. 27-39, 1983.
- [16] A. Haans and W. IJsselsteijn, "Mediated social touch: a review of current research and future directions," *Virtual Reality*, vol. 9, no. 2, pp. 149-159, 2006.
- [17] K. Hook, "Steps to Take before Intelligent User Interfaces Become Real," *Interacting with Computers*, vol. 12, pp. 409-426, 2000.
- [18] H. Ishii and B. Ullmer, "Tangible bits: towards seamless interfaces between people, bits and atoms," in *Proceedings of the SIGCHI conference on Human factors in computing systems*, 1997, p. 241.
- [19] S.E. Jones and A.E. Yarbrough, "A naturalistic study of the meanings of touch," *Communication Monographs*, vol. 52, no. 1, pp. 19-56, 1985.
- [20] E. Kamar, Y. Gal, and B.J. Grosz, "Modeling User Perception of Interaction Opportunities for Effective Teamwork," in *CSE '09: Proceedings of the 2009 International Conference on Computational Science and Engineering*, Washington, DC, USA, 2009, pp. 271-277.
- [21] J.F. Kelley, "An empirical methodology for writing user-friendly natural language computer applications," in *ACM SIG-CHI '83 Human Factors in Computing systems*, New York, 1983, pp. 193-196.
- [22] F. Kensing and J. Blomberg, "Participatory design: Issues and concerns," *Computer Supported Cooperative Work (CSCW)*, vol. 7, no. 3, pp. 167-185, 1998.
- [23] P.S. Kidd, "Getting the Focus and the Group: Enhancing Analytical Rigor in Focus Group Research," *Qualitative Health Research*, vol. 10, no. 3, pp. 293-308, 2000.
- [24] R.A. Kruger and M.A. Casey, *Focus groups: A practical guide for applied research*, 4th ed.: Pine Forge Press, 2008.
- [25] K. Kuhn, "Problems and benefits of requirements gathering with focus groups: a case study," *Ease and Joy of Use for Complex Systems at Siemens: A Special Double Issue of the International Journal of Human-computer Interaction*, vol. 2, no. 3-4, pp. 309-325, 2001.



- [26] D. Kulic and E.A. Croft, "Affective state estimation for human-robot interaction," *IEEE Transactions on Robotics*, vol. 23, no. 5, pp. 991-1000, 2007.
- [27] V. Levesque and J. Pasquero. (2007) Laterotactile.com - Devices - THMB. [Online]. <http://www.cim.mcgill.ca/~haptic/laterotactile/dev/thmb/>
- [28] Livewire Puzzles. FREE Printable Word Search Puzzles. [Online]. <http://www.puzzles.ca/wordsearch.html>
- [29] S. Lucey, A.B. Ashraf, and J.F. Cohn, "Investigating Spontaneous Facial Action Recognition through AAM Representations of the Face," *Face Recognition*, pp. 275-286, 2007.
- [30] H. Lu, W. Pan, N.D. Lane, T. Choudhury, and A.T. Campbell, "SoundSense: scalable sound sensing for people-centric applications on mobile phones," in *Proceedings of the 7th international conference on Mobile systems, applications, and services*, New York, NY, USA, 2009, pp. 165-178.
- [31] K.E. MacLean, "Designing with haptic feedback," in *IEEE International Conference on Robotics and Automation*, vol. 1, 2000, pp. 783-788.
- [32] K.E. MacLean, "Putting haptics into the ambience," *IEEE Transactions on Haptics*, vol. 2, no. 3, pp. 123-135, 2009.
- [33] J. McGrenere, A. Chan, and K.E. MacLean, "Learning and Identifying Haptic Icons under Workload," in *First Joint Eurohaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, Pisa, Italy, 2005.
- [34] M.J. Muller, "Participatory design: the third space in HCI," in *Human-Computer Interaction: Development Process*, A., Jacko, J.A. Sears, Ed., 2003, pp. 1051-1068.
- [35] H. Neilson, M. Enriquez, and K.E. MacLean, *Haptic Signals for Communication Under Workload*, 2006.
- [36] L. Nelson, S. Ichimura, E.R. Pedersen, and L. Adams, "Palette: a paper interface for giving presentations," in *CHI '99: Proceedings of the SIGCHI conference on Human factors in computing systems*, 1999, pp. 354-361.
- [37] J. Nielsen. (2005) useit.com. [Online]. [http://www.useit.com/papers/heuristic/heuristic\\_list.html](http://www.useit.com/papers/heuristic/heuristic_list.html)
- [38] P.J. O'Donnell, G. Scobie, and I. Baxter, "The use of focus groups as an evaluation technique in HCI," *People and Computers VI*, pp. 211-224, 1991.

- [39] Pagewise. (2002) The history of chewing gum. [Online].  
[http://www.essortment.com/all/chewinggumhis\\_rdjz.htm](http://www.essortment.com/all/chewinggumhis_rdjz.htm)
- [40] M. Pan et al., "Now, where was I? Physiologically triggered bookmarks for audio books," The University of British Columbia Department of Computer Science, Vancouver, Technical Report TR-2010-09, 2010.
- [41] R.W. Picard, E. Vyaz, and J. Healey, "Toward machine emotional intelligence: Analysis of affective physiological state," *IEEE transactions on pattern analysis and machine intelligence*, pp. 1175-1191, 2001.
- [42] J. Posner, J.A. Russell, and B.S. Peterson, "The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology," *Development and Psychopathology*, vol. 17, no. 3, pp. 715-734, 2005.
- [43] J.A. Russell, "A circumplex model of affect," *Journal of personality and social psychology*, vol. 39, no. 6, pp. 1161-1178, 1980.
- [44] D. Schuler and A. Namioka, *Participatory design: Principles and practices.*: CRC, 1993.
- [45] M.J. Shaver and K.E. MacLean, "The Twiddler: A Haptic Teaching Tool: Low-Cost Communication and Mechanical Design," Vancouver, 2003.
- [46] A.E. Sklar and N.B. Sarter, "Good vibrations: Tactile feedback in support of attention allocation and human-automation coordination in event-driven domains," *Human factors*, vol. 41, no. 4, p. 543, 1999.
- [47] S.S. Snibbe et al., "Haptic Metaphors for Digital Media," in *ACM Symposium on User Interface Software & Technology*, Orlando, FL, 2001.
- [48] B.A. Swerdfeger, J. Fernquist, T.W. Hazelton, and K.E. MacLean, "Exploring melodic variance in rhythmic haptic stimulus design," in *Proceedings of Graphics Interface 2009*, Kelowna, BC, 2009, pp. 133-140.
- [49] A. Tang, P. McLachlan, K. Lowe, C.R. Saka, and K.E. MacLean, "Perceiving ordinal data haptically under workload," , Toronto, Italy, 2005, pp. 317-324.
- [50] ThinkGeek. ThinkGeek : Star Wars Force Trainer. [Online].  
<http://www.thinkgeek.com/interests/giftsforkids/bf1b/>
- [51] Thought Technology. Biofeedback Equipment: Thought Technology Ltd. [Online].  
<http://www.thoughttechnology.com/>

- [52] M. Weiser and J.S. Brown, "Designing calm technology," *PowerGrid Journal*, vol. 1, no. 1, pp. 75-85, 1996.
- [53] C. Wisneski et al., "Ambient displays: Turning architectural space into an interface between people and digital information," *Cooperative buildings: Integrating information, organization, and architecture*, pp. 22-32, 1998.
- [54] Yamaha Corporation of America. BODiBEAT mp3 player. [Online].  
<http://www.yamaha.com/bodibeat/>
- [55] D. Yang and W.S. Lee, "Disambiguating music emotion using software agents," in *Proceedings of the 5th International Conference on Music Information Retrieval (ISMIR04)*, 2004, pp. 52-58.
- [56] Y. Yoshitomi, S. Kim, T. Kawano, and T. Kitazoe, "Effect of sensor fusion for recognition of emotional states using voice, face image and thermal image of face," in *Proc. ROMAN*, 2000, pp. 178-183.
- [57] Z. Zeng et al., "Audio-visual emotion recognition in adult attachment interview," in *Proceedings of the 8th international conference on Multimodal interfaces*, 2006, p. 145.

## **Appendix A: Focus Group Materials**

### **A1 “As is” Scenarios**

1. John is listening to a Podcast of a talk show while he works in the garden. He is deeply immersed in the show. His neighbour suddenly interrupts him to borrow hedge trimmers, causing John to become startled and remove his headphones. After retrieving the hedge trimmers for his neighbour, John puts his headphones back on to find that he has missed an important part of the show. He iteratively rewinds and plays back the Podcast in order to find his place. Eventually he recognizes some of the content and begins listening to his media again from that point. Due to his neighbour’s interruption, his level of immersion in the show is reduced to almost nothing.
2. Theresa is listening to her portable audio player while waiting for a bus on a serene corner of her neighbourhood. Her player is in her purse. Once on the bus, she can no longer hear her music due to a raucous group of passengers. Frustrated, she reaches for her player in her purse to adjust the volume, which involves unlocking her player using its touch screen interface. The raucous passengers exit the bus a few stops later, and Theresa wants to reduce the volume of her player, again requiring her reach for it and unlock it.
3. Susie is riding her bicycle and listening to music using her portable player. Her player is mounted on her arm and set to shuffle mode. After an upbeat song that she was enjoying ends, an economics lecture that her professor had put online for the class unexpectedly begins. Susie becomes annoyed at this change and wants to return to listening to music. She thus stops her bicycle, removes her player from the

- arm mount, and presses the “forward” button on the player until she finds a song she likes. She then resumes cycling.
4. Mario is cleaning his kitchen, listening to a Podcast on an engrossing but complex lecture. A certain fact that he hears excites him, and he decides that he would like to return to that part of the Podcast in the future to jot down some notes in preparation for an essay. He pauses his player to preserve its current playback location, and searches around for a pen and paper. After finding them, he writes down the name of the Podcast and the time at which he paused it.
  5. Monique is resting in bed, listening to calm, ambient music on her iPod to block external distractions. She drifts off to sleep, and the iPod continues to play. After waking up refreshed 4 hours later, she discovers that her player is out of batteries.
  6. Steven is walking around town on a beautiful day listening to his portable audio player. He is in a good mood and the song that he is playing is matching this mood perfectly. He retrieves his player from his pocket in order to mentally note the name and artist of the song in order to return to it another time, but after returning home, he can't for the life of him remember these details.
  7. Brian is going for his daily morning run. As he warms up, he prefers to listen to relatively slow-paced, happy pop. After his warm-up, however, he much prefers driving, intense, Euro-infused electronica. Knowing his preferences, Brian makes an appropriate playlist of music ahead of time, but on his run, he discovers that the lengths of the songs in this list do not match up with the schedule of his exercise routine, requiring him to manually advance through the playlist after his warm-up and before his cool-down.

8. Mark is listening to music in his car (using his portable player connected through the auxiliary jack) as he drives to work. He comes upon a messy construction zone that requires him to manoeuvre his car through a series of tight lanes marked off by metal pylons. Knowing he will need his full attention to avoid the pylons, he looks down for his player to turn it off. He needs to unlock it and press the pause button in its touch screen, requiring him to momentarily shift his attention away from the road.

## **A2    “To be” Scenarios**

1. (a) John is listening to a Podcast of a talk show while he works in the garden. He is deeply immersed in the show. His neighbour suddenly interrupts him to borrow hedge trimmers, causing John to become startled and remove his headphones. After retrieving the hedge trimmers for his neighbour, John puts his headphones back on to find that his show has been paused. He presses play and the show begins playing 10 seconds before the point at which he was interrupted.  
  
(b) John is listening to a Podcast of a talk show while he works in the garden. He is deeply immersed in the show. His neighbour interrupts him to borrow hedge trimmers. Knowing that removing the headphones from the headphones jack automatically pauses the player, John pulls them out. When he returns to the Podcast, he manually skips back 15 seconds using rewind and then resumes listening.
2. (a) Theresa is listening to her portable audio player while waiting for a bus on a serene corner of her neighbourhood. Her player is in her purse. Once on the bus, she can no longer hear her music due to a raucous group of passengers. Detecting her frustration, her player automatically increases its volume to compensate. The raucous passengers exit the bus a few stops later, and the high volume is no longer

necessary; detecting her frustration again, the player returns to its previous volume setting.

(b) Theresa is listening to her portable audio player while waiting for a bus on a serene corner of her neighbourhood. Her player is in her purse. Once on the bus, she can no longer hear her music due to a raucous group of passengers. Detecting the increased ambient volume, her player automatically increases its volume to compensate. The raucous passengers exit the bus a few stops later, and the high volume is no longer necessary; detecting the shift in ambient volume again, the player returns to its previous setting.

3. (a) Susie is riding her bicycle and listening to music using her portable player. Her player is mounted on her arm and set to shuffle mode. After an upbeat song that she was enjoying ends, an economics lecture that her professor had put online for the class unexpectedly begins. Susie becomes annoyed at this change and wants to return to listening to music. Detecting her annoyance, the player switches to a song it knows Susie will like.

(b) Susie is riding her bicycle and listening to music using her portable player. Her player is mounted on her arm and set to shuffle mode. After an upbeat song that she was enjoying ends, an economics lecture that her professor had put online for the class unexpectedly begins. Susie becomes annoyed at this change and presses the skip button. The player buzzes as if to shrug “oops” and goes back to playing some favourites. It will not make that mistake again.

4. (a) Mario is cleaning his kitchen, listening to a Podcast on an engrossing but complex lecture. A certain fact that he hears excites him, and he decides that he would like to return to that part of the Podcast in the future to jot down some notes in preparation

for an essay. His player detects this excitement and automatically bookmarks the current playback location for future reference. His player confirms the bookmarking action with a gentle kneading motion.

(b) Mario is cleaning his kitchen, listening to a Podcast on an engrossing but complex lecture. A certain fact that he hears excites him, and he decides that he would like to return to that part of the Podcast in the future to jot down some notes in preparation for an essay. Mario pulls out his player and holds down a push button to tag the current playback location for future reference.

5. (a) Monique is resting in bed, listening to calm, ambient music on her iPod to block external distractions. Prior to her rest, she programs in a sleep timer to shut off the iPod after an hour. After an hour passes, the player shuts off.

(b) Monique is resting in bed, listening to calm, ambient music on her iPod to block external distractions. She drifts off to sleep, which the iPod detects. The player then switches off to preserve battery life.

6. (a) Steven is walking around town on a beautiful day listening to his portable audio player. He is in a good mood and the song that he is playing is matching this mood perfectly. The player, noticing that Steven's good mood was preserved throughout the song, catalogues it as a potential favourite for Steven's future reference. Steven, who wants to make sure the song was catalogued, retrieves his player and notices that there is a heart symbol next to the title of the song, indicating that it has indeed been marked as one of his favourites.

(b) Steven is walking around town on a beautiful day listening to his portable audio player. He is in a good mood and the song that he is playing is matching this mood



perfectly. Steven retrieves his player from his pocket and adds the current song to his “On the Go” playlist by holding down the centre button.

7. (a) Brian is going for his daily morning run. As he warms up, he prefers to listen to relatively slow-paced, happy pop. After his warm-up, however, he much prefers driving, intense, Euro-infused electronica. Knowing his preferences, Brian makes an appropriate playlist of music ahead of time. On his run, he completes his warm-up and his heart rate increases to the appropriate level for a sustained workout. Despite the warm-up pop tune not yet being finished, his player moves to electronica in response to his bodily changes.  
  
(b) Brian is going for his daily run. As he warms up, he prefers to listen to relatively slow-paced, happy pop. After his warm-up, however, he much prefers driving, intense, Euro-infused electronica. He sets his player’s target heart rate to 150. As he begins his run, his player, sensing his resting rate of 60, starts his warm-up track. As Brian runs, his heart rate climbs to his target of 150, a cue to the player to begin playing electronica. When the shift occurs, the player taps his leg rapidly to indicate that he is ready to begin the bulk of his run.
8. (a) Mark is listening to music in his car (using his portable player connected through the auxiliary jack) as he drives to work. He comes upon a messy construction zone that requires him to manoeuvre his car through a series of tight lanes marked off by metal pylons. Detecting his increasing anxiety, his player turns itself off.  
  
(b) Mark is listening to music in his car (using his portable player connected through the auxiliary jack) as he drives to work. He comes upon a messy construction zone that requires him to manoeuvre his car through a series of tight lanes marked off by

metal pylons. Detecting his increasing anxiety, his player warns him that he might want to turn the music off by beeping rapidly through the stereo.

### **A3 Sample Messages from Audio Player to User**

- "I'm going to do something"
- "I have an option for you, and require input. I am waiting."
- "I have an option for you, but am continuing as I was."
- "I'm sorry Dave, but I can't do that." (error)
- "I think I understood you like this."
- "Is this how you feel?"
- "You are mad about what I did and I understand that. (Sorry!)"
- "This next song is rated 5 stars; you're going to love it!"
- "Wi-fi is strong here."

### **A4 Inter-Session Survey**

In the second focus group, we presented two sets of scenarios involving audio players. With the first set, we attempted to identify your frustrations with your audio player as it stands, and with the second set, we aimed to introduce some ways to reduce these frustrations.

We observed some resistance to the proposed technology, which detects signals from your body to inform changes to an audio player, and communicates messages with you via the sense of touch. This survey is intended to help us make sure we've heard you correctly on some of the issues you were most concerned with, and to identify more exactly where your concerns lie. Since this survey is designed to get a more complete

picture of your feelings on these topics, please answer each question as honestly and completely as possible. If you have any questions about how to respond to this survey, please don't hesitate to email me (*email address removed*).

Specify how strongly you agree with each of the following statements by placing an “X” in the appropriate box.	<i>Strongly Disagree</i>	<i>Disagree</i>	<i>Neutral</i>	<i>Agree</i>	<i>Strongly Agree</i>	<i>I'm Not Sure</i>
I dislike the idea of having my body signals detected.						
I dislike the idea of having my body signals used to control a device.						
I dislike the idea of having my body signals correlated to my “emotional state”.						
I am skeptical that my body signals can be correlated to my “emotional state”.						
I am skeptical that the computer can reliably correlate my body signals to my “emotional state”.						
I am skeptical that the computer can reliably use the information from my “emotional state” to do something useful for me.						
I am skeptical that I would feel sufficiently “in control” of my audio player when I am controlling it partly through my body signals.						
I am skeptical that I would be able to understand what a device is telling me through my sense of touch.						
I am concerned that the device will need constant input from me to confirm decisions.						
I am concerned that the wrong decisions would annoy me.						
I am concerned that I would have to stop what I'm doing to fix the player.						
I imagine that system feedback that uses the sense of touch would be annoying.						
I imagine that system feedback that uses the sense of touch would be distracting.						
I imagine that system feedback that uses the sense of touch would be invasive.						
I am unwilling to give the device time to learn about my body signals to better understand what I want to do.						
To me, the proposed technology is of little value.						
I wouldn't want to wear any extra peripherals to let this technology work.						

1. As you see it, what is the major problem with using detected body signals to control an audio player?
2. Can you think of an application/device/system/object that you use in your daily life that communicates useful information using the sense of touch? For example, the feeling you get from the road when driving a car.
3. Can you think of an application/device/system that would be more suitable for control by detection of body signals than an audio player?
4. Please think back to the prospective scenarios we discussed in focus group 2, and bring to mind your reaction to the one(s) that seemed most interesting and potentially useful to you. On a scale of *extremely low* – *extremely high*, how would you rate the value of the proposed feature? Give the first answer that comes to mind.
5. Do you have any additional comments about the proposed technology or any of your answers?

## **A5 Session 3 Scenario Set A**

1. Jen wishes to purchase a book that her friend recommended for her, so she logs into her Amazon account. After placing the book in her shopping basket, she notices that the web page has some suggestions about other books that might interest her.
2. Jon is late for work and driving quickly - his car has an automatic transmission. Detecting (via his foot's pressure on the pedal) that he wishes to accelerate rapidly, his car drops down a gear to give him better torque. Jon comes upon a stalled car in his lane and must brake fast to avoid a collision. Once again

- detecting that his foot pressure on the brake is high, his ABS kicks in to bring him to a safer, faster stop.
3. Anita is also late for work and driving quickly - her car has a manual transmission. To accelerate rapidly around a car in her path, Anita drops a gear using the stick-shift to achieve better torque. Upon encountering the same stalled car as Jon, Anita also must brake quickly; she pumps the brakes to avoid skidding and hitting the car.
  4. Michael is scanning through radio stations trying to find one he likes. He lands upon a station that is playing one of his favourite songs, and stays tuned to the station, confident that the disc jockey will play more songs that he likes.
  5. Stacey is writing a letter to her mother in her word processor. After writing the salutation (Dear Mom,) the word processor pops up a message informing her that it noticed she is writing a letter, and would like to help format it for her.
  6. Jason's desktop contains over 50 assorted shortcuts, documents and downloaded images, ordered by date copied. Noticing that he hasn't used some of these items in a while, his operating system pops up a message offering to archive these items. He accepts the offer, and the items on his desktop are reduced to those that he uses frequently.

## **A6    Session 3 Scenario Set B**

### **Scenario 1**

- (a) Theresa is listening to her portable audio player while waiting for a bus on a serene corner of her neighbourhood. Her player is in her purse. Once on the bus, she can no longer hear her music due to a raucous group of passengers. Detecting the increased

ambient volume, and noting her preference to be deeply immersed in her audio, her player automatically increases its volume to compensate.

(b) Theresa is listening to her portable audio player while waiting for a bus on a serene corner of her neighbourhood. Her player is in her purse. Once on the bus, she can no longer hear her music due to a raucous group of passengers. Noting her preference to remain aware of her surroundings, her player takes no action.

### **Scenario 2**

(a) Brian is going for his daily run. Beginning with a slow jog, Brian is listening to a Podcast. He selects the “begin workout” setting on his player, which plays high-intensity music to bring his heart rate up. Upon detecting his target heart rate, the music becomes more moderate. 20 minutes later, Brian taps the “end workout” button, which causes his player to play soothing music until a resting heart rate is detected.

### **Scenario 3**

(a) Steven is walking around town on a beautiful day listening to his portable audio player on shuffle mode. He really enjoys some of the songs that come up, and is lukewarm about the others. The next day, Steven takes another walk around town. He selects the “play detected favourites” option on the player, and the songs he really enjoyed the day before begin to play.

### **Scenario 4**

(a) Susie is riding her bicycle and listening to music using her portable player. Her player is mounted on her arm and set to shuffle mode. After an upbeat song that she was

enjoying ends, an economics lecture that her professor had put online for the class unexpectedly begins. Susie becomes annoyed at this change. Detecting her annoyance and noting that Susie has enabled auto-skip mode, the player switches to a different song. Susie feels a light tap on her wrist informing her that the change in media can be undone if desired.

(b) Susie is riding her bicycle and listening to music using her portable player. Her player is mounted on her arm and set to shuffle mode. After an upbeat song that she was enjoying ends, an economics lecture that her professor had put online for the class unexpectedly begins. Susie becomes annoyed at this change. Detecting her annoyance but noting that Susie has not enabled auto-skip mode, the player catalogues the sequence of tracks that caused her annoyance quietly for future reference. Susie's wrist-watch squeezes her arm to inform her that her annoyance was detected.

## A7 Quantitative Survey Results – Intersession

	<i>Strongly Disagree (1)</i>	<i>Disagree (2)</i>	<i>Neutral (3)</i>	<i>Agree (4)</i>	<i>Strongly Agree (5)</i>	<i>I'm Not Sure (0)</i>
I dislike the idea of having my body signals detected.	1	1	1	1	4	0
I dislike the idea of having my body signals used to control a device.	0	0	1	4	3	0
I dislike the idea of having my body signals correlated to my “emotional state”.	0	2	0	2	4	0
I am skeptical that my body signals can be correlated to my “emotional state”.	0	1	0	4	2	1
I am skeptical that the computer can reliably correlate my body signals to my “emotional state”.	0	0	0	1	7	0
I am skeptical that the computer can reliably use the information from my “emotional state” to do something useful for me.	0	0	1	2	4	1
I am skeptical that I would feel sufficiently “in control” of my audio player when I am controlling it partly through my body signals.	0	1	0	1	6	0
I am skeptical that I would be able to understand what a device is telling me through my sense of touch.	1	2	2	1	2	0
I am concerned that the device will need constant input from me to confirm decisions.	1	1	0	5	1	0
I am concerned that the wrong decisions would annoy me.	0	0	0	2	6	0
I am concerned that I would have to stop what I'm doing to fix the player.	0	0	2	4	2	0
I imagine that system feedback that uses the sense of touch would be annoying.	0	2	2	1	3	0
I imagine that system feedback that uses the sense of touch would be distracting.	0	2	2	2	2	0
I imagine that system feedback that uses the sense of touch would be invasive.	0	4	1	0	3	0
I am unwilling to give the device time to learn about my body signals to better understand what I want to do.	0	3	2	2	1	0
To me, the proposed technology is of little value.	0	1	3	2	1	1
I wouldn't want to wear any extra peripherals to let this technology work.	0	3	2	1	2	0



## A8 Quantitative Survey Results – Final

	<i>Strongly Disagree (1)</i>	<i>Disagree (2)</i>	<i>Neutral (3)</i>	<i>Agree (4)</i>	<i>Strongly Agree (5)</i>	<i>I'm Not Sure (0)</i>
I dislike the idea of having my body signals detected.	1	2	0	1	3	0
I dislike the idea of having my body signals used to control a device.	0	2	0	3	2	0
I dislike the idea of having my body signals correlated to my “emotional state”.	0	2	0	2	3	0
I am skeptical that my body signals can be correlated to my “emotional state”.	0	2	1	3	1	0
I am skeptical that the computer can reliably correlate my body signals to my “emotional state”.	0	0	1	4	2	0
I am skeptical that the computer can reliably use the information from my “emotional state” to do something useful for me.	0	1	0	2	4	0
I am skeptical that I would feel sufficiently “in control” of my audio player when I am controlling it partly through my body signals.	0	1	1	1	4	0
I am skeptical that I would be able to understand what a device is telling me through my sense of touch.	1	3	1	1	1	0
I am concerned that the device will need constant input from me to confirm decisions.	1	1	1	4	0	0
I am concerned that the wrong decisions would annoy me.	0	0	0	2	5	0
I am concerned that I would have to stop what I'm doing to fix the player.	0	0	1	3	2	0
I imagine that system feedback that uses the sense of touch would be annoying.	0	3	2	1	1	0
I imagine that system feedback that uses the sense of touch would be distracting.	0	3	1	3	0	0
I imagine that system feedback that uses the sense of touch would be invasive.	0	4	1	0	2	0
I am unwilling to give the device time to learn about my body signals to better understand what I want to do.	0	5	0	1	1	0
To me, the proposed technology is of little value.	0	2	3	1	1	0
I wouldn't want to wear any extra peripherals to let this technology work.	0	3	1	2	1	0

## A9 Consent Form (Modified Formatting)

### PARTICIPANT'S COPY CONSENT FORM

Department of Computer Science  
2366 Main Mall  
Vancouver, B.C. Canada V6T 1Z4  
tel: (604) 822-3061  
fax: (604) 822-4231

**Project Title:** Portable Audio Player Focus Group Sessions  
(UBC Ethics #H01-80470)

**Principal Investigators:**

Dr. Karon MacLean, Department of Computer Science, 604-822-8169

Dr. Joanna McGrenere, Department of Computer Science, tel. 604-827-5201

**Student Investigator:**

Thomas Hazelton, Department of Computer Science, tel. 604-827-3982

The purpose of this series of focus groups is to examine how people use portable audio players to listen to media.

In each of the focus groups, you will be asked to share your thoughts with other participants about the utility of portable audio players. Discussions will centre on the specific contexts in which you currently use portable audio players and/or situations in which you find them ineffective or cumbersome. You will be asked to rank and otherwise evaluate current and potential features for portable audio players. You will be asked to evaluate a series of prototype portable audio player designs. Data will be collected by video and/or audio recordings, and by questionnaires and surveys.

**REIMBURSEMENT:** \$15 per focus group + \$30 bonus for completing all three (\$75 total)

**TIME COMMITMENT:** 3 × 1 ½ hour sessions

**CONFIDENTIALITY:** *You will not be identified by name in any study reports. Data gathered in the focus group will be stored in a secure Computer Science account accessible only to the experimenters. We encourage all participants to refrain from disclosing the contents of the discussion outside of the focus group; however, we cannot control what other participants do with the information discussed.*

You understand that the experimenter will ANSWER ANY QUESTIONS you have about the instructions or the procedures of these focus groups. After participating, the experimenter will answer any questions you have about the focus groups.

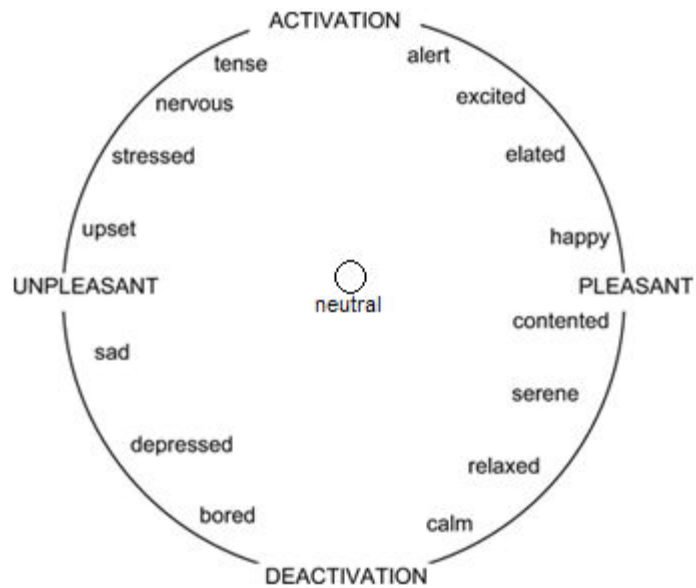
Your participation in these focus groups is entirely voluntary and you may refuse to participate or withdraw at any time without jeopardy. Your signature below indicates that you have received a copy of this consent form for your own records, and consent to participate in these focus groups.

If you have any concerns about your treatment or rights as a research subject, you may contact the Research Subject Info Line in the UBC Office of Research Services at 604-822-8598.

## Appendix B: Affect Study Materials

### B1 Post-Trial Questionnaire (No Word Search Trials)

Indicate with an X on the following grid how you felt overall during this trial.



Did you recognize the song that was playing? Yes No

How much did you like the song (mark an X)? Not at all |-----| I loved it

Would you have changed the song if you had the choice? Yes No

Did your feelings change throughout the trial? If so, describe. Yes No

---

---

Did your feelings peak anywhere during the trial? If so, describe.      Yes      No

---

---

## **B2      Post-Trial Questionnaire (Word Search Trials)**

(same as above with the following additions)

How *physically* taxing was working on the word search (mark an X)?

Not at all   |-----|   Extremely taxing

How *mentally* taxing was working on the word search (mark an X)?

Not at all   |-----|   Extremely taxing

How much did you like working on the word search (mark an X)?

Not at all   |-----|   Couldn't love it more

How *engaged* in the word search were you (mark an X)?

Not at all   |-----|   Fully engaged

How *frustrated* by the word search were you (mark an X)?

Not at all   |-----|   Fully frustrated

To what extent were you distracted by the music while trying to complete the word search?

Not at all |-----| Extremely distracted

### **B3 Post-Trial Questionnaire (User Control Trial)**

(affect grid and first two questions from above two questionnaires, plus the following additions)

Describe what caused you to skip songs during the trial.

---

---

Describe what made you listen to a song (i.e., not skip it) during the trial.

---

---

How much did you like each of the songs that were played during the trial? If you do not recognize a song title, you may skip it [participant was instructed to ask the experimenter to replay it during the experiment]. The songs are presented here in the order they were played.

<b>Skater Boy</b>	Not at all  -----  I loved it
<b>Hounds of Spring</b>	Not at all  -----  I loved it
<b>What is Hip?</b>	Not at all  -----  I loved it

<b>Chopin Op. 28, #13</b>	Not at all	-----	I loved it
<b>Dark Horse</b>	Not at all	-----	I loved it
<b><i>Tian Hei Hei</i></b>	Not at all	-----	I loved it
<b>You Know My Name</b>	Not at all	-----	I loved it
<b>Turn the Beat Around</b>	Not at all	-----	I loved it
<b>All Eyes on Me</b>	Not at all	-----	I loved it
<b>Sing Sang Sung</b>	Not at all	-----	I loved it
<b>Anomoly</b>	Not at all	-----	I loved it
<b>Like a Prayer</b>	Not at all	-----	I loved it
<b>Nothing Without You</b>	Not at all	-----	I loved it

## **B4    Consent Form (Modified Formatting)**

### **PARTICIPANT'S COPY CONSENT FORM**

Department of Computer Science  
2366 Main Mall  
Vancouver, B.C. Canada V6T 1Z4  
tel: (604) 822-3061  
fax: (604) 822-4231

**Project Title:** Affective State Measurement in Audio Contexts  
(UBC Ethics #H01-80470)

#### **Principal Investigators:**

Dr. Karon MacLean, Department of Computer Science, 604-822-8169  
Dr. Joanna McGrenere, Department of Computer Science, tel. 604-827-5201

#### **Student Investigator:**

Thomas Hazelton, Department of Computer Science, tel. 604-827-3982

The purpose of this experiment is to examine and measure the effect of music on users' affective states as they complete a series of tasks.

In this experiment, you will be asked to listen to music through headphones and, for some parts of the experiment, complete word searches. For some parts of the experiment, you will have control over the music that is playing, while in others, you will not. You will be asked to wear external (i.e. non-invasive) sensors that collect some basic physiological information such as heart rate, respiration rate, some muscle activity, and perspiration. Please tell the experimenter if you find the sensors uncomfortable and adjustments will be made. You will be asked to answer questions in a questionnaire as part of the experiment.

This session will be videotaped. The contents of these videotapes will be used for analysis purposes. No parts of the videotapes will be publically presented without your further consent. You have the option not to be videotaped.

REIMBURSEMENT:	\$10
TIME COMMITMENT:	1 × 60 minute session
CONFIDENTIALITY:	<i>You will not be identified by name in any study reports. Data gathered from this experiment will be stored in a secure Computer Science account accessible only to the experimenters.</i>



You understand that the experimenter will ANSWER ANY QUESTIONS you have about the instructions or the procedures of this study. After participating, the experimenter will answer any other questions you have about this study.

Your participation in this study is entirely voluntary and you may refuse to participate or withdraw from the study at any time without jeopardy. Your signature below indicates that you have received a copy of this consent form for your own records, and consent to participate in this study.

If you have any concerns about your treatment or rights as a research subject, you may contact the Research Subject Info Line in the UBC Office of Research Services at 604-822-8598.

## **Appendix C: Participatory Design Materials**

### **C1 Recruitment E-mail**

My name is Tom Hazelton. I'm a MSc student working under Dr. Karon MacLean. Over the next couple of months, I will be engaged in a participatory design process for a piece of novel interactive technology that uses physiological signals to gather data and haptic signals to communicate messages. I am currently in the process of searching for a subject with whom I will work closely on this participatory design, and hope that you may be interested in chatting a bit more about this project to see if you might be a good fit.

I was given your contact information by Karon MacLean directly, as you have been hired to work with her in the summer and therefore may take an interest in this project and will likely be around at the right time.

Please contact me if you are interested in learning more.

### **C2 Screening Interview Questions**

1. Do you have a portable media player such as an iPod?
2. How long have you had it? What kind of features does it have? What kind of features do you wish it had?
3. Not counting your current mobile audio player, how many mobile audio devices have you owned throughout your life?
4. Tell me what sorts of situations you use your audio player in.
5. Are there times your player doesn't meet your needs?
6. What sorts of things do you listen to on your player?
7. How often do you use your player?
8. How satisfied are you with your player?

9. On a scale of 1 – 10, how comfortable are you with technology?
10. How do you stay current?
11. Would you say you're usually the first to acquire new technologies, sometimes the first, or rarely the first?
12. How do you evaluate a new piece of technology when it becomes available on the market? How do you decide whether you need it or not?
13. Would you say you follow trends with respect to technology, or make up your own path?
14. What's your favourite genre of music? How do you react when a song you really like comes on?
15. What's your least favourite genre of music? How do you react when you can't stand the music that's playing?
16. Give the name of a few songs you really like. Give the names of a few songs you really dislike.
17. On a scale of 1 – 10, how passionate do you think you are about music? Explain why it's not higher and why it's not lower.
18. What do you think makes a really good experimental subject for research experiments? What role do you think subjects play in good research practices?

### C3 Consent Form (Modified Formatting)



## PARTICIPANT'S COPY CONSENT FORM

Department of Computer Science  
2366 Main Mall  
Vancouver, B.C. Canada V6T 1Z4  
tel: (604) 822-3061  
fax: (604) 822-4231

**Project Title:** Portable Audio Player Participatory Design Sessions  
(UBC Ethics # H10-00783)

**Principal Investigators:**

Dr. Karon MacLean, Department of Computer Science, 604-822-8169  
Dr. Joanna McGrenere, Department of Computer Science, tel. 604-827-5201

**Student Investigator:**

Thomas Hazelton, Department of Computer Science, tel. 604-827-3982

The purpose of this participatory design session is to examine how people use portable audio players to listen to media and collaboratively develop novel interaction techniques for portable audio players using haptic signals and human affect models.

In the participatory design sessions, you will be asked to share your thoughts with the experimenter about the utility of portable audio players. Discussions will centre on the specific contexts in which you currently use portable audio players and/or situations in which you find them ineffective or cumbersome. You will be asked to rank and otherwise evaluate current and potential features for portable audio players. You will be asked to evaluate a series of prototype portable audio player designs. You will be asked to wear external (i.e. non-invasive) sensors that collect some basic physiological information such as heart rate, respiration rate, some muscle activity, and perspiration. Please tell the experimenter if you find the sensors uncomfortable and adjustments will be made. Data will be collected by video and/or audio recordings, and by questionnaires and surveys.

REIMBURSEMENT:	\$10 per hour (\$50 total)
TIME COMMITMENT:	5 × 1 hour sessions
CONFIDENTIALITY:	<i>You will not be identified by name in any study reports. Data gathered in the sessions will be stored in a secure Computer Science account accessible only to the experimenters.</i>

You understand that the experimenter will ANSWER ANY QUESTIONS you have about the instructions or the procedures of these sessions. After participating, the experimenter will answer any questions you have about the sessions.

Your participation in these sessions is entirely voluntary and you may refuse to participate or withdraw at any time without jeopardy. Your signature below indicates that you have received a copy of this consent form for your own records, and consent to participate in these sessions.

If you have any concerns about your treatment or rights as a research subject, you may contact the Research Subject Info Line in the UBC Office of Research Services at 604-822-8598.

## **Appendix D: $k$ -Nearest Neighbours Analysis (written by Susana Zoghbi)**

---

### **Estimating Affect Using Physiological Responses to Music Technical Report**

**by Susana Zoghbi**

*Abstract— This technical report was written as a supporting document to present the experimental analysis in a joint project between CARIS Lab and SPIN Lab at UBC, the Haptic-Affect Interaction Loop (HALO). It investigates the use of physiological signals to estimate human affective states while interacting with a music player.*

### **Introduction**

In daily life, when humans interact among each other, we use explicit and implicit cues to communicate both the actions we intend to take and our affective states towards the interaction. A portion of interpersonal communication relies on implicit cues. Communication and recognition of affective states are important and expected in a human-human interaction. If media devices are to interact with users in a less intrusive way requiring low effort and cognitive loads, they should be able to perceive the user's affective states in both explicit and implicit modes. Several explicit and implicit cues can be used to estimate affective states, *e.g.*: characteristics of speech, facial expressions, gestures, postures, and physiological signals. This research will focus on the last of these cues. Physiological signals provide quantifiable measures that tend to be involuntary as

well as age and culture independent. *The goal of this study is to investigate the use of physiological responses to various songs or genres of music to estimate the affective state of users.* Specifically, the research questions aimed to address are: How reliably can users' affective states (e.g., level of enjoyment) be estimated using physiological responses? What physiological features are most promising for inferring users' affect? What is an appropriate time window for analyzing physiological responses?

## **Methodology**

An experimental approach was undertaken to address the research questions. The general process consisted of subjecting participants to a series of songs while measuring their physiological responses. After each song, participants were asked to self report their valence (i.e., level of enjoyment). Six physiological sensors were connected to the participants: Respiration, electrocardiography (EKG), electromyography (EMG), skin conductance (SC), blood volume pulse (BVP) and skin temperature (ST). Data from these sensors were collected using Thought Technology's FlexComp hardware system, using the FlexComp encoder connected via USB to a laptop. All signals were recorded at 256 Hz.

These physiological signals are filtered and analyzed using machine learning techniques.

## **Study Implementation**

Three data-collection experiments have been performed as of August 2010. This section describes the implementation of each.

In Experiment 1, 12 participants (7 male) were subjected to 6 songs for 90 seconds each, which were preselected based on a questionnaire to be variably enjoyable.

In Experiment 2 and Experiment 3, only one subject (female) was recruited (same subject for both in order to remove variability between subjects). In Experiment 2, the subject was exposed to 22 songs of different genres of 90 seconds each. In Experiment 3, 36 songs belonging to one genre of music were used for 45 seconds each. Only one genre of music was used in this case to reduce potential variability caused by various intrinsic characteristics of different genres. The playing time was reduced because the subject's initial reaction to the song is the main interest. This way, both the data volume and experiment time are reduced without compromising the validity of the data. In all experiments, participants were asked to self-report their affective states and ratings were translated from a Likert scale into like/neutral/dislike scale.

Experiment 1 contained an additional final trial where users were instructed that they would have partial control over the music being played. In particular, they would be able to “advance” tracks in a predefined playlist at will by lightly tapping on a table using the hand not connected to physiological sensors. Upon noticing the tap, the experimenter would manually advance tracks on the music player.

## **Data Analysis**

The aim of the data analysis is to create a model to predict users' reported affective states using a set of physiological features. To this end, each physiological signal was smoothed using filtering techniques and features were extracted. Each feature was normalized to both account for day-dependent baseline variations and allow feature comparisons across individuals. The features used were mean, standard deviation (std),



maximum value (max), minimum (min) and difference between max and min (diff) of the following signals: heart rate (hr), heart rate acceleration (hrA), skin conductance response (scr), derivative of skin conductance (dScr), electromyography from the corrugator muscle (emg), respiration (resp) and skin temperature (temp) .

A k-nearest neighbour algorithm (KNN) was implemented to estimate valence, given an input vector containing a set of physiological features. When a test input is presented, the Mahalanobis distance to its  $k$  nearest points is computed and a probabilistic estimate of the valence is returned. Each song played for each subject provided a data point. For Experiment 1, all data points were randomly separated into training and a testing set (50% of the data for each set).

In Experiment 2 and Experiment 3, one data point was tested each time and all other data points were used for the training set. This test was repeated sequentially for all data points so that each point was tested.

Relevant features, adequate time intervals and appropriate number of  $k$  neighbours were estimated using cross validation. All possible combinations of features were systematically tested as well as time intervals ranging from 1 second to 20 seconds and  $k$  (number of neighbours) values ranging from 1 to 15.

## **Results and Discussion**

### **Experiment 1**

A recognition rate of 76.67% was achieved with a time window of 9.02 seconds. The features that produced this recognition rate were the mean of normalized EMG [mean(nEMG)], standard deviation of normalized heart rate [std(nHR)], maximum normalized heart rate [max(HR)], and the differences between the maximum and

minimum values of the following three signals: normalized derivative of skin conductance [diff(ndSCR)], normalized heart rate [diff(nHR)] and normalized heart rate acceleration [diff(nHRAccel)]. In spite of this promising recognition rate, high recognition rate fluctuations over time windows (see Figure 1) suggest unreliability of detection in real-time. It was hypothesized that high-frequency variations in recognition rate may be due to high variability caused by intrinsic physiological differences between subjects. This result led to explore the physiological responses of one subject at a time, which was done in Experiment 2 and 3.

## Experiment 2

In this experiment all possible combinations of features were tested for a time window of 5 seconds and  $k = 1$ . The highest recognition rate for all the combinations of features tested was 72.73%. In particular, 22 combinations yielded this result. Figure 2 shows how often individual features were used to achieve such rate. The features most frequently used in decreasing order were: Std(emg), Diff(emg), Std(temp), Max(resp), Min(dScr) and Max(dScr).

In order to characterize the conditions under which the algorithm's performance is more accurate, the Mahalanobis distance to the closest neighbor for each data point was compared between correctly recognized songs and incorrectly ones. Figure 3 presents a histogram of the relative frequency of several ranges of distances. It shows that it is more likely to correctly recognize a data point when the distance to the closest neighbor is within a range from 0 to 5 units. A t-test was performed and significant differences were found between the two groups,  $t(131)$ ,  $p < 0.05$ .

### **Experiment 3**

In an attempt to validate results obtained in Experiment 2, a subsequent experiment was performed. All possible combinations of features were tested on 36 new songs (time window of 5 seconds and  $k = 1$ ). The highest recognition rate for all the combinations of features tested was 83.33%. In particular, 4 combinations yielded this result. Lower –yet acceptable- recognition rates were found for other combinations of features. In total 666 combinations yielded to a recognition rate higher than 70%. Figure 4 shows how often individual features were used to achieve such rate. The features most frequently used in decreasing order were Mean(hrA), Min(dScr), Mean(temp), Min(tem), Diff(hr), Max(temp) and Max(hrA). Unfortunately, only Min (dScr) was commonly used in both Experiment 2 and 3. All other the other features mostly used in Experiment 3 were not commonly used in Experiment 2.

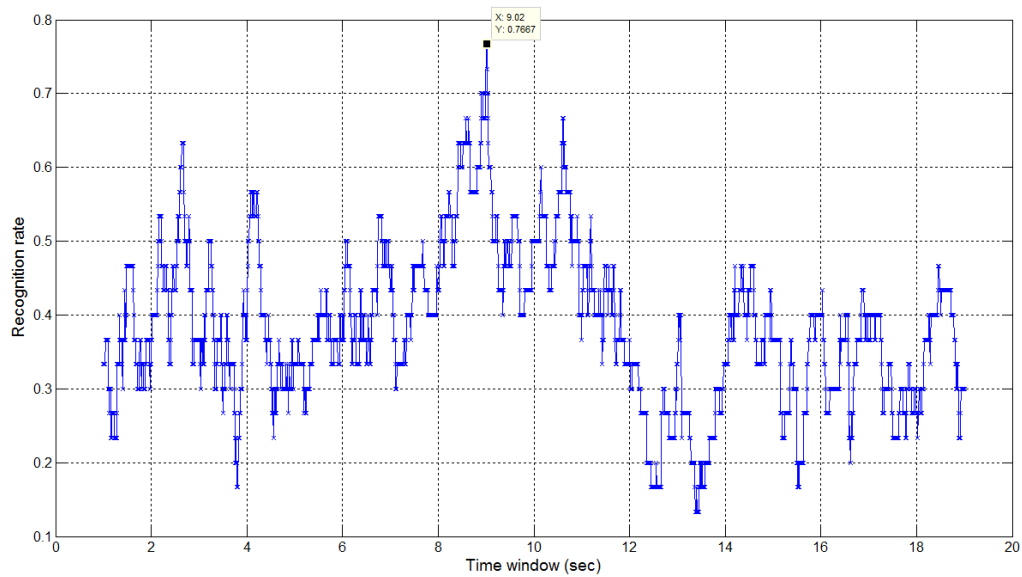
### **Combining Experiment 2 and Experiment 3**

The data collected from the last two experiments was combined and the features that yielded the highest recognition rates for both experiments were tested in the aggregated data set. The highest recognition rate was 58.60% using the features mean(hr), min(temp), mean(hrA), diff(dScr), max(hr) and min(hr).

### **Summary and Discussion**

In order to investigate the use of physiological signals to estimate users' affective states while listening to music, three experiments have been performed. The first experiment analyzed the physiological data collected from 12 subjects. The highest recognition rate was 76.67%. However, high-frequency variations were observed over

different time windows, indicating that this result would not be reliable for implementation in an audio player. The second and third experiment investigated the physiological responses of one subject. When the data from each experiment was analyzed separately, the highest recognition rates were 72.3% and 83.33%, for experiment 2 and experiment 3 respectively. However, aggregating the data collected from both experiments yielded to 58.60% rate. These results, while promising, still require further study and validation. A new data collection experiment has been planned in an attempt to achieve this.



*Figure 1. Recognition Rate vs. Time Windows in Experiment 1*

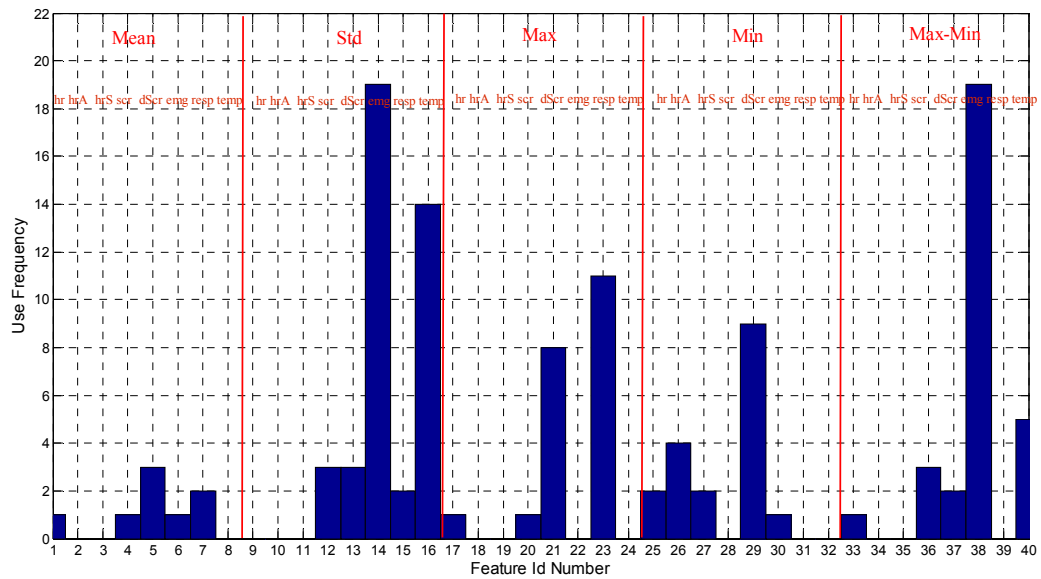


Figure 2. Frequency of Features- Experiment 2

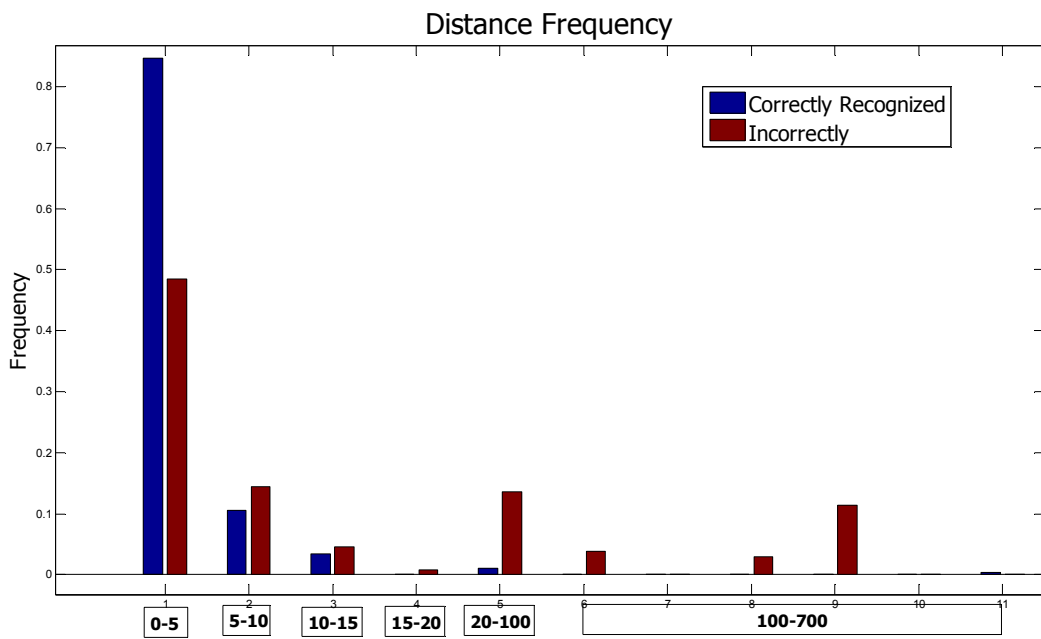


Figure 3. Distance to the Closest Neighbour -Experiment 2

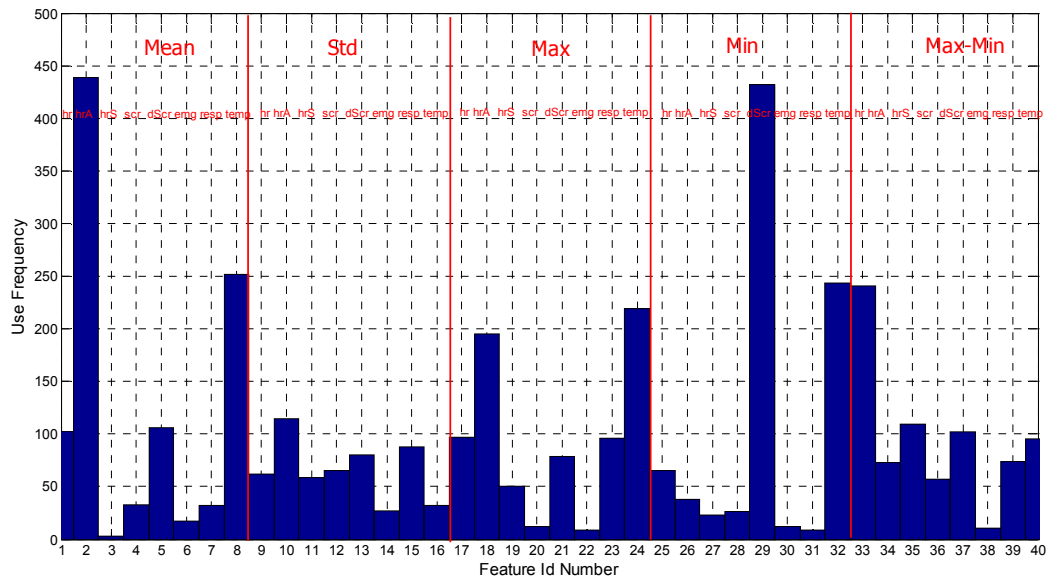


Figure 4. Frequency of Features -Experiment 3

## **Appendix E: BREB Approval Certificates**

Certificates are appended on the following three pages.

**Page 162** contains the approval certificate for focus group-related activities.

**Page 163** contains the approval certificate for the audio-affect experiment.

**Page 164** contains the approval certificate for the participatory design sessions.



The University of British Columbia  
Office of Research Services  
Behavioural Research Ethics Board  
Suite 102, 6190 Agronomy Road, Vancouver, B.C. V6T 1Z3

## CERTIFICATE OF APPROVAL - AMENDMENT & RENEWAL

<b>PRINCIPAL INVESTIGATOR:</b> Karon E. MacLean	<b>DEPARTMENT:</b> UBC/Science/Computer Science	<b>UBC BREB NUMBER:</b> H01-80470
<b>INSTITUTION(S) WHERE RESEARCH WILL BE CARRIED OUT:</b>		
<b>Institution</b> UBC Other locations where the research will be conducted: N/A		<b>Site</b> Vancouver (excludes UBC Hospital)
<b>CO-INVESTIGATOR(S):</b> Joanna McGrenere Ricardo Pedrosa Matthew Baumann Susan Gerofsky Joseph Hall Karin Krauei Thomas William Hazelton Elizabeth A. Croft Steve Yohanan Jonathan Chang		
<b>SPONSORING AGENCIES:</b> Natural Sciences and Engineering Research Council of Canada (NSERC) Nokia Vancouver Spark Robotics Technology Various Sources		
<b>PROJECT TITLE:</b> Orsil title - Low-Attention and Affective Communication Using Haptic Interfaces		

**CERTIFICATE EXPIRY DATE:** July 10, 2010

<b>AMENDMENT(S):</b>	<b>RENEWAL AND AMENDMENT APPROVAL DATE:</b> July 10, 2009	
<b>Document Name</b>	<b>Version</b>	<b>Date</b>
<b>Consent Forms:</b>		
Audio Practices Survey consent form	1.0	June 18, 2009
Audio Practices Focus Group Consent Form	1.0	June 18, 2009
<b>Advertisements:</b>		
Audio Practices Recruitment Poster	1.0	June 18, 2009
<b>Questionnaire, Questionnaire Cover Letter, Tests:</b>		
Audio Practices Recruitment Survey	1.0	June 18, 2009
<b>Other:</b>		
All web-based aspects of this study will be accessed via <a href="http://www.cs.ubc.ca/labs/spin/">http://www.cs.ubc.ca/labs/spin/</a> All content used online is attached above.		
The application for continuing ethical review and the amendment(s) for the above-named project have been reviewed and the procedures were found to be acceptable on ethical grounds for research involving human subjects.		
Approval is issued on behalf of the Behavioural Research Ethics Board and signed electronically by one of the following:  ..... Dr. M. Judith Lynam, Chair Dr. Ken Craig, Chair Dr. Jim Rupert, Associate Chair Dr. Laurie Ford, Associate Chair Dr. Anita Ho, Associate Chair		





The University of British Columbia  
Office of Research Services  
**Behavioural Research Ethics Board**  
Suite 102, 6190 Agronomy Road, Vancouver, B.C. V6T 1Z3

## CERTIFICATE OF APPROVAL - MINIMAL RISK AMENDMENT

<b>PRINCIPAL INVESTIGATOR:</b> Karon E. MacLean	<b>DEPARTMENT:</b> UBC/Science/Computer Science	<b>UBC BREB NUMBER:</b> H01-80470						
<b>INSTITUTION(S) WHERE RESEARCH WILL BE CARRIED OUT:</b>								
<table border="1"><thead><tr><th>Institution</th><th>Site</th></tr></thead><tbody><tr><td>UBC</td><td>Vancouver (excludes UBC Hospital)</td></tr><tr><td colspan="2">Other locations where the research will be conducted: N/A</td></tr></tbody></table>			Institution	Site	UBC	Vancouver (excludes UBC Hospital)	Other locations where the research will be conducted: N/A	
Institution	Site							
UBC	Vancouver (excludes UBC Hospital)							
Other locations where the research will be conducted: N/A								
<b>CO-INVESTIGATOR(S):</b> Joanna McGrenere Ricardo Pedrosa Matthew Baumann Susan Gerotsky Joseph Hall Idin Karuei Thomas William Hazelton Elizabeth A. Croft Steve Yohanan Jonathan Chang								
<b>SPONSORING AGENCIES:</b> Natural Sciences and Engineering Research Council of Canada (NSERC) Nokia Vancouver Spark Robotics Technology Various Sources								
<b>PROJECT TITLE:</b> Orsil title - Low-Attention and Affective Communication Using Haptic Interfaces								

**Expiry Date** - Approval of an amendment does not change the expiry date on the current UBC BREB approval of this study. An application for renewal is required on or before: **July 10, 2010**

<b>AMENDMENT(S):</b>	<b>AMENDMENT APPROVAL DATE:</b> December 10, 2009																														
<table border="1"><thead><tr><th>Document Name</th><th>Version</th><th>Date</th></tr></thead><tbody><tr><td colspan="3"><b>Consent Forms:</b></td></tr><tr><td>Affective State Measurement in Audio Contexts Consent Form</td><td>1.0</td><td>November 27, 2009</td></tr><tr><td colspan="3"><b>Questionnaire, Questionnaire Cover Letter, Tests:</b></td></tr><tr><td>Affective State Measurement in Audio Contexts - Post-Trial (Baseline) Questionnaire</td><td>1.0</td><td>November 27, 2009</td></tr><tr><td>Affective State Measurement in Audio Contexts Recruitment Survey</td><td>1.0</td><td>November 27, 2009</td></tr><tr><td>Affective State Measurement in Audio Contexts - Post-Trial (User Control) Questionnaire</td><td>1.0</td><td>November 27, 2009</td></tr><tr><td>Affective State Measurement in Audio Contexts - Post-Trial (Word Search) Questionnaire</td><td>1.0</td><td>November 27, 2009</td></tr><tr><td colspan="3"><b>Other:</b></td></tr><tr><td colspan="3">All web-based aspects of this study will be accessed via <a href="http://www.cs.ubc.ca/abs/spin/">http://www.cs.ubc.ca/abs/spin/</a> All content used online is attached above.</td></tr></tbody></table>	Document Name	Version	Date	<b>Consent Forms:</b>			Affective State Measurement in Audio Contexts Consent Form	1.0	November 27, 2009	<b>Questionnaire, Questionnaire Cover Letter, Tests:</b>			Affective State Measurement in Audio Contexts - Post-Trial (Baseline) Questionnaire	1.0	November 27, 2009	Affective State Measurement in Audio Contexts Recruitment Survey	1.0	November 27, 2009	Affective State Measurement in Audio Contexts - Post-Trial (User Control) Questionnaire	1.0	November 27, 2009	Affective State Measurement in Audio Contexts - Post-Trial (Word Search) Questionnaire	1.0	November 27, 2009	<b>Other:</b>			All web-based aspects of this study will be accessed via <a href="http://www.cs.ubc.ca/abs/spin/">http://www.cs.ubc.ca/abs/spin/</a> All content used online is attached above.			
Document Name	Version	Date																													
<b>Consent Forms:</b>																															
Affective State Measurement in Audio Contexts Consent Form	1.0	November 27, 2009																													
<b>Questionnaire, Questionnaire Cover Letter, Tests:</b>																															
Affective State Measurement in Audio Contexts - Post-Trial (Baseline) Questionnaire	1.0	November 27, 2009																													
Affective State Measurement in Audio Contexts Recruitment Survey	1.0	November 27, 2009																													
Affective State Measurement in Audio Contexts - Post-Trial (User Control) Questionnaire	1.0	November 27, 2009																													
Affective State Measurement in Audio Contexts - Post-Trial (Word Search) Questionnaire	1.0	November 27, 2009																													
<b>Other:</b>																															
All web-based aspects of this study will be accessed via <a href="http://www.cs.ubc.ca/abs/spin/">http://www.cs.ubc.ca/abs/spin/</a> All content used online is attached above.																															
The amendment(s) and the document(s) listed above have been reviewed and the procedures were found to be acceptable on ethical grounds for research involving human subjects.																															
<p style="text-align: center;">Approval is issued on behalf of the Behavioural Research Ethics Board and signed electronically by one of the following:</p> <div style="text-align: center;">..... Dr. M. Judith Lynam, Chair Dr. Ken Craig, Chair Dr. Jim Rupert, Associate Chair Dr. Laurie Ford, Associate Chair Dr. Anita Ho, Associate Chair</div>																															



The University of British Columbia  
Office of Research Services  
**Behavioural Research Ethics Board**  
Suite 102, 6190 Agronomy Road, Vancouver, B.C. V6T 1Z3

## CERTIFICATE OF APPROVAL - MINIMAL RISK

<b>PRINCIPAL INVESTIGATOR:</b> Karon E. MacLean	<b>INSTITUTION / DEPARTMENT:</b> UBC/Science/Computer Science	<b>UBC BREB NUMBER:</b> H10-00783						
<b>INSTITUTION(S) WHERE RESEARCH WILL BE CARRIED OUT:</b>								
<table border="1"><thead><tr><th>Institution</th><th>Site</th></tr></thead><tbody><tr><td>UBC</td><td>Vancouver (excludes UBC Hospital)</td></tr><tr><td colspan="2">Other locations where the research will be conducted: N/A</td></tr></tbody></table>			Institution	Site	UBC	Vancouver (excludes UBC Hospital)	Other locations where the research will be conducted: N/A	
Institution	Site							
UBC	Vancouver (excludes UBC Hospital)							
Other locations where the research will be conducted: N/A								
<b>CO-INVESTIGATOR(S):</b> Joanna McGrenere Thomas William Hazelton Elizabeth A. Croft								
<b>SPONSORING AGENCIES:</b> Spark Robotics Technology								
<b>PROJECT TITLE:</b> HALO: Haptic Affect Loop Participatory Design								

**CERTIFICATE EXPIRY DATE:** April 26, 2011

<b>DOCUMENTS INCLUDED IN THIS APPROVAL:</b>	<b>DATE APPROVED:</b> April 26, 2010												
<table border="1"><thead><tr><th>Document Name</th><th>Version</th><th>Date</th></tr></thead><tbody><tr><td><b>Consent Forms:</b> Portable Audio Player Participatory Design Sessions</td><td>1.1</td><td>April 19, 2010</td></tr><tr><td><b>Advertisements:</b> Recruitment Email</td><td>1.0</td><td>March 18, 2010</td></tr><tr><td><b>Questionnaire, Questionnaire Cover Letter, Tests:</b> Sample Interview Questions</td><td>1.0</td><td>April 12, 2010</td></tr></tbody></table>	Document Name	Version	Date	<b>Consent Forms:</b> Portable Audio Player Participatory Design Sessions	1.1	April 19, 2010	<b>Advertisements:</b> Recruitment Email	1.0	March 18, 2010	<b>Questionnaire, Questionnaire Cover Letter, Tests:</b> Sample Interview Questions	1.0	April 12, 2010	
Document Name	Version	Date											
<b>Consent Forms:</b> Portable Audio Player Participatory Design Sessions	1.1	April 19, 2010											
<b>Advertisements:</b> Recruitment Email	1.0	March 18, 2010											
<b>Questionnaire, Questionnaire Cover Letter, Tests:</b> Sample Interview Questions	1.0	April 12, 2010											

The application for ethical review and the document(s) listed above have been reviewed and the procedures were found to be acceptable on ethical grounds for research involving human subjects.

Approval is issued on behalf of the Behavioural Research Ethics Board  
and signed electronically by one of the following:

Dr. M. Judith Lynam, Chair  
Dr. Ken Craig, Chair  
Dr. Jim Rupert, Associate Chair  
Dr. Laurie Ford, Associate Chair  
Dr. Anita Ho, Associate Chair