# Adaptive affective computing: countering user frustration


by

Behzad Aghaei


A thesis

presented to the University of Waterloo

in fulfillment of the

thesis requirement for the degree of

Master of Applied Science

in

Systems Design Engineering

Waterloo, Ontario, Canada, 2013

©Behzad Aghaei 2013

# AUTHOR'S DECLARATION

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Behzad Aghaei

# ABSTRACT

With the rise of mobile computing and an ever-growing variety of ubiquitous sensors, computers are becoming increasingly context-aware. A revolutionary step in this process that has seen much progress will be user-awareness: the ability of a computing device to infer its user's emotions. This research project attempts to study the effectiveness of enabling a computer to adapt its visual interface to counter user frustration.

A two-group experiment was designed to engage participants in a goal-oriented task disguised as a simple usability study with a performance incentive. Five frustrating stimuli were triggered throughout a single 15-minute task in the form of complete system unresponsiveness or delay. An algorithm was implemented to attempt to detect sudden rises in user arousal measured via a skin conductance sensor. Following a successful detection, or otherwise a maximum of a 10-second delay, the application resumed responsiveness. In the control condition, participants were exposed to a “please wait” pop-up near the end of the delay whereas those in the adaption condition were exposed to an additional visual transition to a user interface with calming colours and larger touch targets. This proposed adaptive condition was hypothesized to reduce the recovery time associated with the frustration response.

The experiment was successfully able to induce frustration (via measurable skin conductance responses) in the majority of trials. The mean recovery half-time of participants in the first trial adaptive condition was significantly longer than that of the control. This was attributed to a possibility of a large chromatic difference between the adaptive and control colour schemes, habituation and prediction, causal association of adaptation to the frustrating stimulus, as well as insufficient subtlety in the transition and visual look of the adaptive interface.

The study produced findings and guidelines that will be crucial in the future design of adaptive affective user interfaces.

# Acknowledgements

I would like to express my deepest gratitude to my supervisor, Professor Catherine Burns. She not only accepted me as her graduate student despite my unrelated academic background, but she supported and guided me in all of my academic and career endeavours. I owe my success in design, research, and human factors, present and future, to Professor Burns.

I am also very thankful of my readers, Professor Carolyn MacGregor and Professor Mark Hancock for their constructive comments and also for agreeing on such short notice to become my official thesis readers.

This rather ambitious project would not have been possible without the help of one talented and incredibly hard working friend and lab mate, Anson Ho. I am indebted to Anson for developing the entire software system, assisting in experimental design, assisting in running participants, working late nights, and listening to my ever-changing and last-minute demands.

My gratitude also goes to Dr. Jerome Pasquero of Research In Motion for providing us with the opportunity, flexibility, and insight we needed to carry this project through to completion.

I would also like to thank my Canadian family, AIDL. The warm, supportive, social, and funny people of AIDL took me in 2 years ago and instantly made me feel part of their team. They are, and will remain, among my closest and most cherished friends.

Not to be forgotten are the unconditional love of my family, their infinite patience, and the support shown to me by them and those closest to me, Hamed Shateri and Jessica Li.

Finally, I would like to thank NSERC and Research In Motion for making this project possible through their funding.

# TABLE OF CONTENTS

# List of Figures

# LIST OF ABBREVIATIONS

**AC** - Augmented Cognition

**ECG** - Electrocardiography

**EDA** – Electro-Dermal Activity

**EEG** - Electroencephalography

**EKG** – See ECG

**HCI** - Human Computer Interaction

**IAA** - Intelligent Adaptive Automation

**IAI** – Intelligent Adaptive Interfaces

**GSR** – Galvanic Skin Response

**SCR** – Skin Conductance Response

# 1. INTRODUCTION

With the rapid growth of mobile computing, smartphone devices are equipped with more and increasingly complex sensors. The ability of a device to collect information on its surroundings, as well as its user, allows it to provide more context-aware output and thereby improve the user's experience. Some examples of this include providing relevant search results, nearby food locations, and changing between landscape and portrait display modes depending on how the user is physically orienting the device. In a sense, the device can be said to adapt to the user's preferences, location, and behaviour.

Adaptive computing is not new. There exists decades of research studying the effects and feasibility of software programs that are able to partially map their outputs to the state of their users. An early example of adaptive technology in consumer software is the ability to view recently opened files on desktop computer programs. Several such examples exist where software programs are designed to customize their content to the behaviour of their users.

There is another dimension to computing that is far less studied, however. With the rise of neuroscience and emotion research in psychology, more emphasis was put on understanding the "human" component of human-computer interaction. The area of research that studies human emotion as it relates to computing was coined "affective computing" (Picard, 1995).

As progress is made in the development of computers that can adapt to user preferences, as well as their understanding of human emotion, it is easy to imagine that our everyday devices may someday understand how we feel and attempt to soothe or excite us through visual, auditory, or tactile stimuli. One significant application of this technology is its potential ability to address its user's frustration with the device itself. Frustration with computers is commonplace and is often sourced in back-end issues (e.g. non-interface problems such as bandwidth or connection) (Ceaparu et al., 2004).

This specific research project draws on previous research and attempts to break new ground into adaptive affective computing. With the use of a Galvanic Skin Response sensor and a tablet computer, it aims to take a first step at exploring a tablet application that can:

a. trigger frustrating events that intentionally frustrate the user,

b. attempt to detect if a user is frustrated via a rise of in-context arousal, and if so,

c. adapt its look and feel in an attempt to counter the frustration.

## 1.1 Objectives

The purpose of this project was to determine if an appropriate visual adaptation of the user interface could decrease the recovery half-time of a user's skin conductance response to a frustrating stimulus.

There were two hypotheses for the results of the experiment:

1. Frustration would be induced in the majority of trials and represented by a skin conductance response peak when the application was triggered to be unresponsive.
2. The recovery halftime for the skin conductance response peak to the frustrating stimulus would be significantly shorter for participants in the adaptive condition than those in the control condition, implying faster recovery from the frustrated state. This is because adaptive user interfaces have been known to decrease cognitive load, and visual aesthetics have been known to provide calming and pleasing experiences synonymous with low levels of psychophysiological arousal.

## 1.2 Structure of Thesis

The following sections of the thesis are briefly introduced below:

- Chapter 2 introduces the different fields and concepts associated with this project and provides a background review of each. This includes a study of adaptive interfaces, affective computing, the relationship between cognition and emotion, and finally the motivation behind the research.
- Chapter 3 discusses how the research experiment was designed and how the experiment was set up and executed.
- Chapter 4 describes how the raw data gathered from the experiment was analyzed for outliers and also the statistical tests that were applied.

- Chapter 5 presents the observations and findings of the experiment.
- Chapter 6 will discuss the findings and attempt to explain them based on existing research.
- Chapter 7 draws conclusions and lessons learned from the research project and lists suggestions for future research based and commercial devices on what was learned from the current study.

# 2. BACKGROUND

The following sections provide a literature review of various fields of study related to this thesis experiment, ranging from adaptive computing in human-computer interaction to the basics of the neuroscience of emotion. The section will conclude by providing the motivation behind the current study.

## 2.1 ADAPTIVE COMPUTING

Adaptive computing refers to the automatic personalization of a computer interface to its user(s) (Benyon and Murray, 1993) and has been a topic of research and experimentation at least as early as the 1970's (Chu and Rouse, 1979). As with many studies of human-centered automation, it had its beginnings in the fields of aviation and defense – which fell under the umbrella of human factors research in complex systems.

As with much of technological advancement developed for aerospace and military applications, adaptive computing eventually made its way into the consumer world – in this case, the rapidly growing consumer computing industry. Its applications were focused on "users" instead of "operators" and became its own field of research in Human-Computer Interaction (or HCI).

Although Human Factors and Human-Computer Interaction view adaptive computing in the same light, they vary in their approach, implementations, and terminology of the matter. The following two sub-sections will outline the approaches taken by Human Factors and Human-Computer Interaction, respectively.

### 2.1.1 HUMAN FACTORS APPROACH TO ADAPTIVE COMPUTING

In human factors, adaptive computing often takes on the label "Intelligent Adaptive Automation" (IAA) or "Augmented Cognition" (AC) and its applications are largely focused around aviation and defense (Morrison et al., 1992).

At its core, the purpose of intelligent adaptive automation is to improve an operator's judgment and decision-making ability in complex systems (Geddes and Shalin, 1997). This can be done at a high level through directing an operator's attention, regulating their arousal and workload, as well as providing lower-level decision support to an operator – based on various factors such as the nature of the task and the state of the operator (Prinzel et al., 1999; Young & Eggleston, 2002).

In general, IAA has been proven to be effective in improving operator task performance (Chu and Rouse, 1979; Freedy et al., 1985; Morris and Rouse, 1986). Although there is no shortage of successful experimentation with IAA, its real-world adoption has been scarce. Several aviation and defense organizations have existing IAA programs but have yet to produce feasible products due to limitations imposed by real-world complexity and technology (Joubert et al., 1995).

### 2.1.2 Human-Computer Interaction Approach to Adaptive Computing

In Human-Computer Interaction, adaptive computing is a subset of personalized computing, which also includes adaptable user interfaces. Adaptable interfaces are ones that allow users to manually alter, or personalize, parts of the interface to suit their needs. Adaptive interfaces, on the other hand, are ones that perform this personalization automatically (Benyon and Murray, 1993; Velsen et al., 2008). Although personalization in interfaces have been shown to be beneficial over traditional static graphical user interfaces (e.g. Gajos et al., 2006; McGrenere et al., 2002), each of these approaches offer distinct advantages. For example, Findlater and Mcgrenere (2004) found an advantage of an adaptable over an adaptive interface in the context of menu design, recommending a mixture of both to cater to a variety of users.

Adaptive computing in HCI has also been dubbed "intelligent adaptive interfaces" (IAI) and has been focused on less complex systems. Its studies and applications are user-centered and are more focused on the commercial consumer industry.

The purpose of IAI is to improve a user's interaction with a system by making it more "efficient, effective, and easy to use" (Banbury et al., 2005). As with its human factors counterpart, IAI is also focused on improving task performance by decreasing workload. It also employs very similar means of doing so: augmenting cognition by supporting the user's memory, perception, and decision-making processes.

Notable applications of adaptive computing in HCI include the ability to view recently opened files in software applications. This aids memory and reduces cognitive workload by providing quick access to files to which the user is likely seeking access. Another popular implementation was the Microsoft Office Assistant, which manifested itself as an animated paper clip, "Clippy". This example of IAI would offer assistance on the user's tasks – which it inferred from the user's actions and typed phrases. Despite the promising idea of an intelligent agent, Microsoft removed the Microsoft Office Assistant in their later products in response to overwhelming user frustration and academic research, which proved the agent to be obtrusive (Veletsianos, 2007).

All is not lost, however, since adaptive computing has seen much use since the introduction of smartphones. Developers of mobile applications commonly make use of location, weather, orientation, acceleration, and ambient light detection capabilities of modern smartphones to provide users with relevant content. Another type of implementation has been used by Google's Ads feature, which collects information from a user's emails and caters relevant ads. Similarly, a commonplace practice in e-commerce is to suggest products in which the user might be interested in, depending on his or her purchase or view history. Studies in HCI have also shown promise in systems which can infer a person's emotional state and make appropriate suggestions (Lee et al., 2011; Picard, 2001; Duric et al., 2002).

It becomes apparent, then, that most successful applications of affective computing in HCI occur at a relatively basic level: suggesting relevant information to the user. Despite the potential cognitive benefits of more complex adaptive interfaces (e.g. agents), their realization has been limited by challenges in their implementation – such as poor context awareness and aesthetics (Veletsianos, 2007). Both of these factors are essential to this project in its pursuit of a realistic and commercial-like application of affective computing. Specifically, the study involves the design of a mobile map application that is familiar to the user in terms of function and design, as well as unobtrusive, meaning without the added clutter of automated cognitive suggestions.

### 2.1.3 Summary

Regardless of the field of research, adaptive computing has yet to become a mainstream concept in existing complex systems or consumer products (Findlater and McGrenere, 2008), even in light of its benefits that have been shown by research in both fields. For example, Microsoft Office featured adaptive menus in its early-2000 versions, only to be removed in more recent versions due poor user experience. Few examples exist today, such as responsive web design and a small number of graphical user interface elements, such as self-resizing text boxes. It is also important to note most of these existing adaptive interfaces involve changing the layout of interfaces, whereas the user interface developed in this project involves no layout or functionality changes.

In HCI, adaptive computing has yet to mature beyond its application of providing the user with suggestions. With the existence of unprecedented amounts of user-specific data and device sensors, surely there must be a solution for more integrated and practical intelligent personalization without hindering user adoption. The study outlined in this report attempts to add an emotional dimension to adaptive graphical user interfaces, built upon existing research in the field of Affective Computing.

## 2.2 Affective Computing

Sensory perception, memory, information processing and decision-making cannot be sufficient in describing how humans think and behave. Affect, as well, plays a significant role in how we experience the world. It plays a significant role in how we experience *computers*: frustration with technology has become part of the human experience – so perhaps there is more to computer development and intelligence than building cognitive assistants. Perhaps adapting an application interface to a user's cognition – though beneficial – stands to gain by considering human affect in its pursuit of personalized computer intelligence.

The following sections will discuss the current state of affective computing, methods for the measurement of affect, and a comparison to adaptive interfaces.

### 2.2.1 Emotion in Human Computer Interaction

The realization of this apparent gap in computing research led to creation of the phrase, "Affective Computing" (Picard, 1995). By her definition, affective computing is one that "relates to, arises from, or influences emotions." Affective computing has since become a branch of research under HCI with a variety of approaches to the unlikely duo of emotions and machines. The approach which is of most interest to this project, and arguably the most practical given current technological limitations in artificial intelligence, is one in which computers are able to read and "understand" emotions.

The ability of computers to detect or infer user affect can have many practical implications. Thus far, the major practical focus of affective computing revolves around the detection of user frustration and attempts to make displays of sympathy or empathy to address a user's apparent emotional state (e.g. Picard & Klein, 2002; Picard, 2001; Nasoz et al., 2003). For example, Klein et al. (2002) found that even by simply using empathetic language in a post-study questionnaire, users chose to interact with a frustrating game significantly longer. This finding hinted at the likeliness that the display of emotion from the side of automation can potentially undo the negative emotional effect of failure or frustration by the automation itself. Of course, this success is not limited to the "treatment" of negative emotions caused by automation - it can also be used therapeutically (e.g. Krijn et. al, 2004). In the affective computing realm, empathetic agents have been shown to decrease user stress (Prendinger et al., 2003).

### 2.2.2 Measuring Affect

There is still one major hurdle – perhaps the most challenging of all – which has thus far limited affective computing to the research world and prevented it from becoming a mainstream commercial reality. That is the question of "how" – how to enable a machine to infer a user's affect? The answer is, by necessity, as complex as the study of human emotion.

Inferring human affect would, at the very least, require the simultaneous detection of user-specific and task-specific variables (Scheirer et al., 2002). Environmental variables also play a role, but reasonable inferences can be made without the latter – at least in the context of laboratory-controlled affective computing experiments. Once a piece of automation is able to detect these variables, it must make use of appropriate models to recognize patterns (e.g. Hidden Markov Models as used by Scheirer et al. (2002)) or even learn the user's behavior using neural networks. Access to a larger number of appropriate sensors with the use of appropriate models to consolidate the data in real-time has been shown to predict user frustration with a significantly better than chance probability (Kapoor et al., 2007). Kapoor et al. made use of Gaussian process classification to combine data from a pupil tracking camera, a pressure-sensitive mouse, a skin conductance sensor, and a pressure-sensitive chair, and self-reported frustration data to attempt to learn a user's non-verbal and physiological response to frustration. Patterns based purely on periodic self-reports of affect have also been used as an attempt to predict user affect with better-than-chance results (LiKamWa et al., 2011).

Several techniques are available in detecting user affect, based on proven physiological and behavioural methods. An example of a physiological measure is the use of Galvanic Skin Response sensors, which detect changes in skin conductance that vary with a user's level of physiological arousal. Other general physiological measures can be detected via several categories of sensors: temperature sensors, heart-rate sensors, facial recognition, blood pressure sensors, electroencephalograms (EEG) for monitoring brain waves, respiration rate sensors, and eye-trackers or similar technology for detecting eye movement or blink rate (Lisetti & Nasoz, 2004).

Various combinations of these physiological sensors have been used to reliably infer basic human emotions (Lisetti & Nasoz, 2004). Although simple self-ratings of emotion can be used, they are subjective ratings – they are not verifiable and ultimately not as vigorous as objective, quantifiable, physiological measures.

What is meant by basic human emotions? One of the most common methods to classify human emotion is through the use of dimensional models of emotion. These models were developed on the assumption that emotions could be considered as continuous elements which could vary along different dimensions. Although several attempts have been made to develop dimensional models of emotion, the most commonly used model in HCI and affective computing research has been the circumplex model (Russel, 1980). Russel proposed a two-dimensional model of affect. The first dimension was valence, which varied from negative to positive. The second was activation or arousal, which varied from low to high. Individual affective states could then be plotted along this 2D model. As an example, boredom and calmness/peacefulness are both low-arousal states, but boredom has a negative valence while calmness/peacefulness is attributed with positive valence. Figure 1 below shows an example plot of the circumplex model with different positions on the 2D plot representing distinct emotional states which were derived based on the arousal and valence axis.

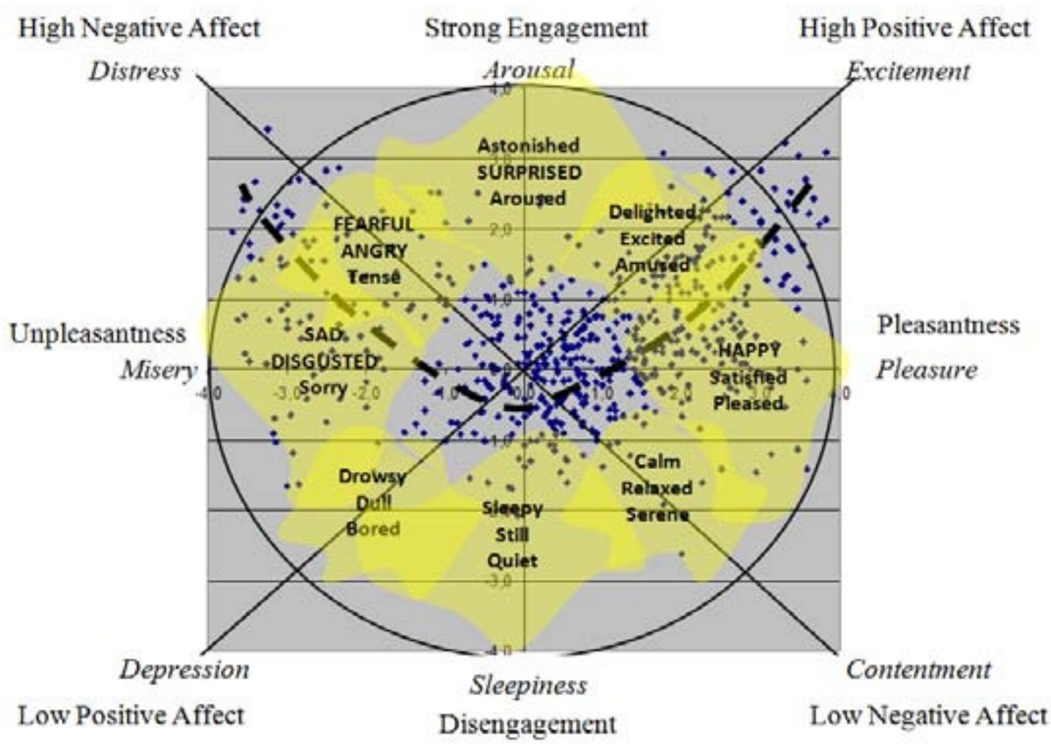


FIGURE 1: (ADAPTED FROM GOKCAY, 2011): CIRCUMPLEX MODEL OF EMOTION WITH VALENCE AS THE X-AXIS AND AROUSAL AS THE Y-AXIS

In order to plot the affective states of a user on the continuous dimensional circumplex model, separate measurements must be taken along each dimension. Many of the previously stated physiological sensors enable the measurement of physiological arousal – both combined as a model or separately. The most commonly used physiological measure of arousal is Galvanic Skin Response. Since a person's physiological signals vary throughout the day, any physiological sensor would need an initial calibration, as well as an algorithm to re-establish a baseline reading for quick successive time intervals (Picard et al., 2001). Other sensors such as heart rate and temperature may be combined to provide redundancy in an attempt to verify GSR measurements.

Detecting valence physiologically is more technically difficult and less reliable than that of arousal. The most common physiological method of sensing valence is via facial recognition – subtle changes in a person's facial expression can be used to infer if they are experiencing positive or negative emotions (Nasoz et al., 2003).

Measurements of valence and arousal can be combined to plot a user's affective state on the 2D circumplex model. Picard et al. (2001) were able to obtain 81% accuracy in distinguishing between a user's eight different affective states, including neutral. This is a promising level of reliability for machine affect detection. In other studies, computers were able to infer affect with 80-98% accuracy (Yacoob & Davis, 1996; Gardner & Essa, 1997) whereas humans obtained 70-98% accuracy (Bassili, 1979) for a similar type of study (facial expression recognition) (Picard et al., 2001).

Although physiological measures provide the major advantage of real-time, unobtrusive, and continuous monitoring, it has its own set of challenges. The technology is still regarded as immature for commercialization due to its relatively high potential of noise interference (Wilson, 2002; St john et al., 2003, Prinzel et al., 2003). In addition, there are currently no standards of real time physiological data analysis (Scheirer et al., 2002). The physiological fluctuations experienced by people in a day require any physiological sensor to constantly re-calibrate itself and calculate new baselines.

Behavioural measures of affect can be task-dependant and thus can vary. For example, Scheirer et al. (2002) recorded the number of mouse clicks the user made in combination with physiological measures in order to infer the user's arousal state. Further behavioural measures of arousal and confusion can also be gained by observing how fast a user moves the mouse, the force with which they click, and the path they use to navigate to a clickable object (Duric et al., 2002). Caution must be exercised when using behavioural measures, however, due to the risk of subjective biases when interpreting behavior, and the context under which it is taking place.

### 2.2.3 Adaptation in Affective Computing

Although some examples of affective computing imply adaptation of some kind, the majority do not involve any real-time emotional feedback. Several examples can be found in the literature review performed by Akgun et al. (2011). Past studies include positive enjoyment caused by a computer's use of flattery, as well as the use of empathy by Klein (2002). Also mentioned were several studies which investigated the use of sympathetic (e.g. apologetic) feedback to the user upon system failure (Nielsen, 1998; Tzeng, 2004; Tzeng, 2006). These interfaces simply provided an affective failure message instead of the usual mechanical failure message, and as such were not truly providing feedback to user frustration. Another branch of studies has explored the use of text-based and embodied agents (Akgun et al., 2011) with greater support for the latter. However, the commercial feasibility of agents may be questionable due to similar issues faced by the Microsoft Office Agents – aesthetics, obtrusiveness, and difficulty in inferring the user's current task or state.

It should also be noted that none of the studies reviewed thus far have exhibited a bi-directional adaptive affective nature where not only is user affect measured in real-time, but adaptive feedback is provided to the user in real-time as well.

### 2.2.4 Summary

The scientific understanding of emotion in psychology and neuroscience paved the way to a new way of thinking about human-computer interaction. The promise of affective computing was to imbue computers with the ability to infer, process, and express affect. Its benefits to user experience have been shown to provide improvements over the traditional emotionless and mechanical user-machine interactions.

The success of the studies in affective computing gives promise to its eventual commercialization and widespread use. As technology advances, the accuracy of affect-recognition systems will improve. The commercialization of affect-recognition devices is already under way (e.g. Affectiva, www.affectiva.com) and it may be a matter of time before computers are able to detect human affect more accurately than humans themselves.

The study outlined in this report heavily relied on existing groundwork in the field of affective computing. Specifically, the basis of measuring and inferring affect, inducing frustration in participants, and experimental design of this study were derived, as closely as possible, from existing literature in this research area.

## 2.3 Cognition and Emotion

Following the era of behavioural psychology, and with the onset of the technological revolution, cognitive psychology popularized a new way of thinking about the human mind – the information-processing model. In many ways it mirrored the technical innovations of that era: computing. Several analogies can be formed between the information processing model of the human mind and that of the computer. For example, permanent storage and temporary storage correspond to long-term memory and working memory, respectively. The same applied to the processor and its resources – even the concept of cognitive overload and its PC counterpart.

Overall, the early manifestation of cognition generally disregarded emotion (Phelps, 2006). It was merely involved with the mechanics of information processing: input (or stimulus), output (or response), memory, motor control, bandwidth, processing resources, and so on.

The following section will discuss the relationships found between cognition and emotion in the contexts of psychology, psychophysiology, and neuroscience.

### 2.3.1 Links between Cognition and Emotion

With the maturation of neuroscience and our understanding of the brain, research began to show links between the information processing view of the brain and emotions (Damasio, 1996). This paved the way for further research into the brain to discover how emotion manifests itself in the neural networks of our minds and what effects it has on how we process information. In order to develop adaptive systems, it is necessary to understand the links between cognition and emotion. These links can provide context to designing affective adaptive experiments and also interpreting their results. Topics such as coping strategies and arousal will be discussed. To that end, the following paragraphs will attempt to provide a brief summary of the links between cognition and emotion in the fields of psychology, neurophysiology and cognitive neuroscience.

The study of vigilance in the field of psychology has attempted to describe the decrease in human performance observed in sustained-attention tasks. The most recent and integrated model has been the attentional resources theory (e.g. Matthews et al., 1990, Matthews, 2001). A significant finding (relevant to cognition and emotion) of this field of research was that low physiological arousal states brought about by sustained attention tasks were decreasing the availability of information processing resources (Caggiano & Parasuraman, 2004). In addition, Hitchcock et al. (2003) found that cerebral blood flow in the right hemisphere decreased along with vigilance performance, which agrees with the attentional resource model due to the right hemisphere being responsible for the allocation of attention. This field of research provided a fundamental link between affect and the information-processing model by suggesting that physiological arousal (affect) and attentional resources (cognition) were closely related.

Perhaps the strongest psychological indicator of the affect-cognition relationship lies in our strategies of regulating emotion. Based on a thorough review of the matter by Gross (2002), humans have several ways of inhibiting or "down-regulating" emotion – both internally and externally. First of all, the emotional impact of an event can be limited by how the event is perceived – known as cognitive change. This refers to the ability to perceive an emotional event in a non-emotional way. The second strategy is cognitive reappraisal, through which one can choose to evaluate negative emotion-provoking thoughts and choose to replace them. The third strategy is to suppress emotions prior to making an emotional response.

A more thorough view of the relationships between cognition and emotion inherent in human nature was provided by neuroscience. This view starts with the limbic system. The limbic lobe consists of several brain components and is a distinct region of the brain in close proximity to the brain stem. Situated within the limbic lobe are a set of closely situated and thoroughly entwined networks known as the hypothalamus, hippocampus, and the amygdala, among others (Erdem & Karaismailoglu, 2011). It is the working relationship between these components that provides the neurological links between cognition and emotion.

The hippocampus plays a significant role in memory formation and information processing (Erdem & Karaismailoglu, 2011). It has been deemed responsible for the recollection of episodic memory, and is also in a close neural connection with the proposed link between cognition and emotion: the amygdala (Phelps, 2006). Phelps explained that the amygdala has been directly related to emotional changes in various studies, and that damage to it has caused a change in emotional behaviour. Due to this close proximity and vast connections to the cognition center of the brain (hippocampus and the prefrontal cortex – Young et al., 1994), it is thought to have direct influence on information processing. Phelps (2006) also pointed at the observed influence of emotions on the encoding and recall of episodic memory, which is described by the close interactions of the amygdala and the hippocampal complex. Richards and Gross (2000) also found a close interaction between memory and emotional processes.

Phelps (2006) also found similar links between high-level cognitive processes (such as perception and attention) as were described in the vigilance studies. In addition to those findings, Anderson and Phelps (2001) presented a direct link between the amygdala itself and the attentional resource theory. They found that participants with a damaged amygdala failed to show increased attentional processing resources as a result of induced arousal. This also provided a link between the effect of arousal brought about by the amygdala and information processing performance.

The relationship between the amygdala and cognitive processes of the brain is not unidirectional, however. Evidence also exists for the influence of cognition on the amygdala itself. For example in a study by Phelps et al. (2001), it was observed that the amygdala is influenced by cognitive or learned emotional properties of a stimulus (e.g. being instructed what a stimulus will feel like, as opposed to having experienced it).

The psychological coping strategies of emotion that were discussed previously were identified through psychological studies, but now a deeper explanation has become apparent: the cognitive centers of the brain and its decision making processes make frequent and significant regulations on the amygdala (Phelps, 2006). This, in part, is what makes us function in society. We constantly make decisions to regulate our emotional outbursts (Gross, 2002).

The hypothalamus is another essential component of the limbic lobe. It controls much of the human emotional behaviour (Erdem & Karaismailoglu, 2011). It too is in close proximity to the amygdala and the hippocampus: information is constantly passed between the hippocampal formation (cognition) and hypothalamus (emotion). This is the fundamental physiological link behind the design of adaptive affective interfaces. The notion that the human brain's processing of perceptual information is closely interconnected with what humans feel as emotion, gives designers of adaptive systems the ability to create perceptual elements that can invoke specific types of emotions in users.

### 2.3.2 Summary

The role of emotion in human cognitive processes can no longer be neglected. Studies in psychology have hinted at the interaction of the two for some time. In addition, neuroscientists have already found distinct relationships between the parts of the brain responsible for information processing and the experience of emotion. A close two-way interaction exists between these two phenomena. First, cognition can be used to regulate emotion at different stages and emotion – we do this on a regular basis in order to function in society. Second, emotion also makes a significant impact on our cognition: from modulating our attentional resources via arousal to determining how our memories are stored and recalled.

The relationship between cognition and emotion is a deep but essential foundation of the study outlined in this report. Specifically, it provides a basis for why visual perception of graphical user interface elements such as colour and animation can have direct effects on physiological emotion.

## 2.4 Affective Adaptive Computing

By the mid-90's it had become apparent that emotion played a significant role in how we interact with computers (Picard, 1995). The concept of Affective Computing was introduced and opened a new path for computing. Despite all this, nearly two decades later, computers still show little sign of affect, regardless of its commercialization potentials. The argument can be made that affect-inference technology has not yet matured for commercialization, but that could be attributed to a lack of interest by the computing industry.

The potential for affective computers is growing, however. With Apple's introduction of its intelligent vocal agent "Siri", it may have found a feasible solution to the Microsoft agent problem. "Siri" is not void of affect, either: it (or, rather, she) also displays sympathetic behaviour for its failures – surely one of the factors behind its quick adoption and use (Vascellaro, 2012).

Despite the advances in consumer products and research, there is still a gap – and it exists between the realms of affective and adaptive computing. Filling this gap with research and development would pave the way for *adaptive affective systems*: intelligent computers which adapt themselves in real-time to user affect. Of course, the goal is for them to do so with usability and consumer adoption in mind. The components of such a system already exist: the field of adaptive automation has made advances into adapting computers to user cognition for complex systems (e.g. Chu and Rouse, 1973; Morris and Rouse, 1986; Geddes and Shalin, 1997; Hitchcock et al., 2003, St John et al., 2003). Affective computing researchers have developed agents which make positive impacts on user performance and emotional states (e.g. Picard, 2001; Klein et al., 2003; Tseng, 2004). The significance of affect on how we think has already been shown through neuroscience and psychology.

### 2.4.1 Motivation of Research

Little work has been done on adapting a commercial UI to user affect. In one instance, researchers developed a system on the Android operating system (Lee et al., 2011) which displayed certain applications to users based on their measured mood states. Although interesting, no real user interface/interaction modifications were made to the system. A new, additional application was simply developed with suggested content. There are currently no published studies looking at real time automated adaptation of a user interface look-and-feel based on a user's real time physiological readings. This is despite the growing emotional attachment to mobile phones (Vincent, 2005), which necessitates at the least a consideration for affect in the design and function of these devices. Even more relevant to this study is demand for devices that automatically adapt their colour schemes according to their users' moods (LiKamWa et al., 2011).

The motivation behind this project was to explore the feasibility of adapting the visual look-and-feel of a software application to human emotion as measured in real-time - as opposed to existing research that attempts to adapt content and functionality to a user's cognitive state. This fills a gap that currently exists in the research on adaptive affective computing. The study also links together the three general areas covered in this background report: adaptive computing, affective computing, and emotion research. It is closely based on existing studies in affective computing for the measurement and induction of negative affect (frustration), while using emotion research coupled with adaptive computing guidelines to develop an adaptive graphical user interface aimed at reducing the effects of said frustration.

Although this work is exploratory, it is expected to lead to further research on visual adaptations to real-time human affect. Advancements in this area of research could potentially lead to an improved user experience of every day computing devices by making them emotionally aware and adaptive.

# 3. Experiment

This chapter first familiarizes the reader on certain characteristics of the electro-dermal response, and then provides the rationale and theoretical basis behind the experimental design via reviewing similar experiments in affective computing. The experimental protocol is then presented in detail. Finally, it presents the design and rationale of the affective adaptive graphical interface.

## 3.1 The Electro-Dermal Response

The galvanic skin response sensor used in this experiment measures the skin conductance response or electro-dermal activity of the user. The unit of the output for this sensor (further discussed in the Apparatus section) is a micro-Siemen.

When the user experiences a rise in emotional arousal such as frustration or excitement, the sweat glands on their skin activate and a proportional physiological response can be detected via a GSR sensor. These types of events generally show as local peaks in the EDA (Electro-Dermal Activity) data, also referred to as Skin Conductance Response (or SCR) (Andreassi, 2006). Although participants can simply be queried for self-reported frustration data, it can be disruptive to the task at hand and therefore not a practical solution for real-world commercial products. In addition, self-reported emotion data is notoriously variable (Scheirer et al., 2002).

Once an emotional trigger or stimulus has been presented, there is a latency of typically 1-2 seconds (Figner, 2010) before the occurrence of an SCR. In other words, it takes time for the body's sweat glands to activate in response to an emotional stimulus.

The SCR is characterized by certain parameters, shown below in Figure 2.

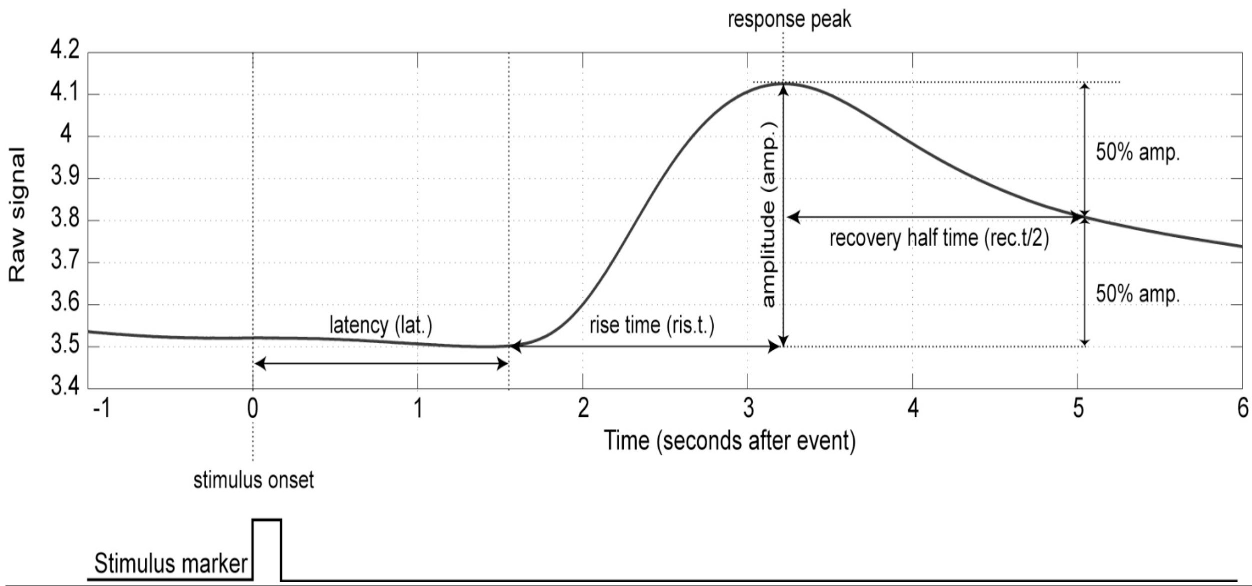


FIGURE 2: PARAMETERS OF A SKIN CONDUCTANCE RESPONSE PEAK (FIGNER, 2010).

Since this experiment is concerned with how users recover from their frustrated states, the parameter selected for analysis was the recovery half time, or the amount of time it takes for the SCR to reach half of its peak value, measured from the occurrence of the peak value. The recovery half-time has been shown to vary with stimulus and thus can be treated as an independent variable (Janes, 1982; Venables and Fletcher, 1981).

## 3.2 Experimental Design

In order to test for the effects of affective adaptation, a single factorial (two-group), between-subjects experiment was designed as follows:

Group 1: Control: Presentation of frustrating stimuli without affect-support adaptation

Group 2: Adaptive: Presentation of frustrating stimuli followed by affect-support adaptation

In a similar study by Klein et al. (2002), a 2x3 factorial design was used. The first factor was the existence of frustrating stimuli and the second included 3 degrees of affect-support in the form of questionnaires. This factorial design was feasible for their study due to the fact that the affect-support provided could also apply to, and have an effect on, the no-frustration conditions. In the current experiment, however, the affect-support adaptation depends on the existence of a frustrating stimulus. It was decided that little knowledge would be gained by applying an affect-support adaptation at a random point in the experiment where no frustrating stimulus is present.

Participants were equally divided into each experimental group and were run in individual sessions.

The study in both experimental conditions involved participants undergoing a timed map navigation challenge while wearing a GSR sensor. Participants were asked to use a customized map application to provide answers to a set of particular geographic questions. A $100 incentive was provided for the participant with the most correct answers provided. The compensation scheme and amounts were based closely on Scheirer et al. (2002).

During each experimental session, the customized application was manually triggered to freeze for a fixed amount of time while the participant was interacting with the application. Unresponsiveness was selected as the method of choice to frustrate the user because its effectiveness was demonstrated by Scheirer et al. (2002). Following the freezing event, the application would either unfreeze with a “Please wait” message (control condition) or unfreeze with a “Please wait” message and a gradual visual change of the user interface (Adaptive condition). The application would automatically revert to the standard look-and-feel after a fixed time.

The visual perception of colour is able to induce measurable and physiological emotional states in humans. Of a number of examples, Kuller et al. (2009) found significant changes in physiological arousal as measured by EEG (brain waves) and EKG (heart rate) when participants were placed inside rooms of different colour, concluding that use of good colour design can be a practical way to improve the overall mood and well-being of people.

The specific colour selection for the adaptive interface was based on existing literature on the relationships between human emotion and colour. First, it has been shown that the colours blue and green (used extensively in the adaptive look-and-feel) are quieting and calming, as opposed to stimulating and arousing (Elliot et al., 2007). Aside from the induction of emotional states, there is some limited evidence of the effects of colour on the rates of physiological recovery. Ali (1972) found that the cortical response to a constant blue light stimulus, as measured by EEG, recovers with less delay than that of a constant red stimulus. A change in colour scheme is also a rather practical and easy-to-implement solution in commercial devices and involves no major (e.g. functionality or layout) changes to the user interface. This was another reason why it was selected as the driving force behind the adaptive user interface.

Another transition element designed into the adaptive user interface was further rounding of the corners of buttons. Rounded corners have been shown to be easier to visually process (Troncoso et al., 2009), and they avoid the negative emotions (e.g. avoidance) caused by visually perceiving sharp edges (PR Newswire, 2010). Furthermore, larger touch targets are easier to tap on with fingers, leading to a more user-friendly design.

The purpose of the freezing event was to cause frustration in the participants. The time limit and incentives were provided to amplify the sense of urgency and therefore the frustration felt once the application became frozen or unresponsive.

In order to avoid participant suspicion of the true purpose of the experiment (e.g. frustrating the participant on purpose), deception was used and approved by the Office of Research Ethics at the University of Waterloo. Recruitment posters and emails, as well as pre-study consent forms and information package made no mention of frustration and disguised the study as simply testing how participants respond to a map game developed by the experimenters.

## 3.3 Methodology

The experimental protocol is outlined in the following section in further detail.

### 3.3.1 Participants

A total of forty undergraduate and graduate students between the ages of 19 to 36 from the University of Waterloo were recruited through e-mail and posters for this study. The participants were equally split into one of two experimental conditions. All participants were asked to complete an initial questionnaire, detailed in the Procedure section. The purpose of the questionnaire was to collect information on participant gender, age, neuroticism or emotional stability, and computer experience. Both groups were gender-balanced and there were no significant differences between the groups in terms of age, self-reported neuroticism (emotional stability), or experience with tablets, touchscreen devices, and map applications. Although all participants completed the experiment, only twenty-four data sets were required. Extra participants were recruited to compensate for several cases where participant data was discarded:

- Two participants' data were discarded due to equipment malfunction during the experiment.
- Three participants reported suspecting that the frustrating events were presented on purpose, and their data was discarded. All participants were verbally asked after the main task if at any point they suspected the true nature of the experiment.
- Eleven participants showed a relatively flat electrodermal response throughout the entire experiment and were excluded from the main dataset. This was in part due to early limitations in the device's true sensitivity (fixed through a firmware update), as well as individual differences in the amount of sweat gland activation.

In the remaining data set, one participant's data was excluded from the control condition following outlier analysis (see section 4.2). Another participant's data was excluded from the Adaptive condition to reduce the age range of the study to 19 to 30 years old, as well as to equalize the cell sizes for data analysis. The remaining and final data set was composed of twenty-two participants (N=11 in each group) with n=6 males and n=5 females in each of the two experimental groups. All participants were compensated $10 for their time. One participant from each group was compensated an additional $100 for having the best performance in their respective group.

### 3.3.2 Apparatus

The experiment occurred in an experiment room with a desk, two chairs, a desktop computer, a monitor, a Blackberry Playbook tablet computer, and a GSR sensor. The experimenter was situated in the experiment room for the duration of the experiment. The GSR sensing system consisted of an Affectiva Q-

Sensor attached to the bottom of the participant's non-dominant hand via the included Velcro strap (Figure 3).


FIGURE 3: AFFECTIVA Q-SENSOR (FROM DIEGOMENDES.COM)

The GSR was wirelessly paired with a desktop computer in the experiment room via a Bluetooth connection. The desktop computer created a local log file containing all raw data from the GSR sensor and also streamed the data onto a MySQL database hosted locally on the computer itself. This database was accessed by the experimenter web page and also the main map application web page. The experimenter accessed the experimenter web page in order to monitor the real-time readings from the GSR sensor, as well as to manually trigger freezing events (Figure 4).

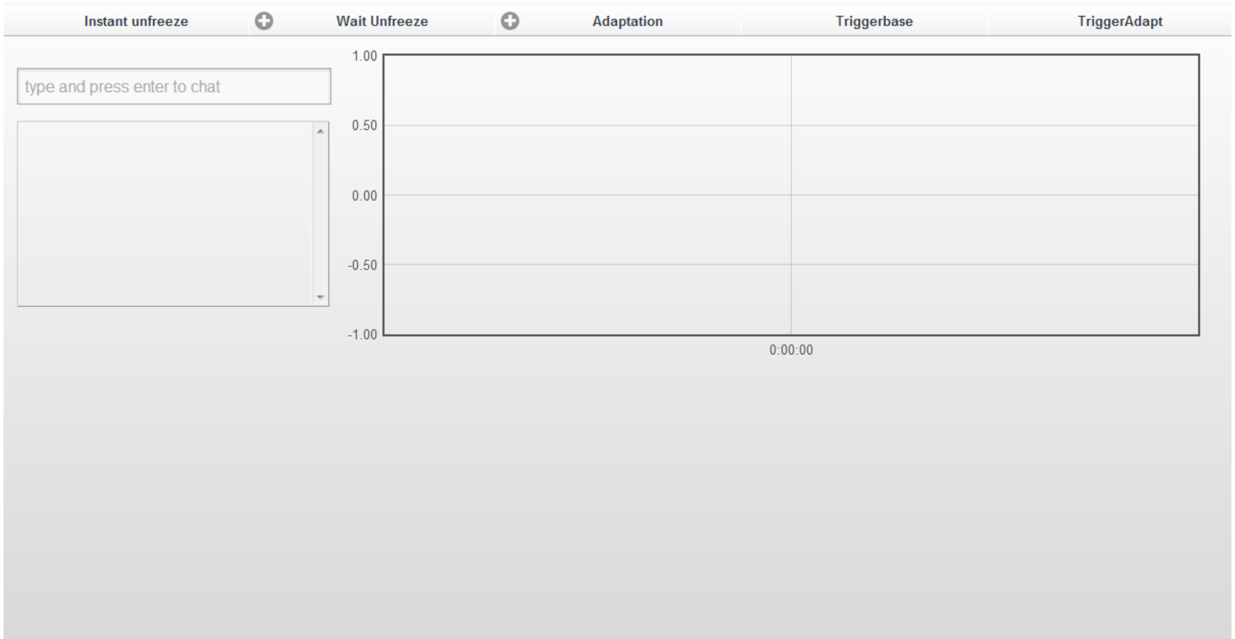


**FIGURE 4: EXPERIMENTER WEB PAGE**

The map application web page was accessed from the Blackberry Playbook, which would be used by the participant to complete the experiment (Figure 5). The application was designed to be familiar to most users by following the general layout and functionality of Google Maps (www.maps.google.com). The desktop version of Google Maps was selected as a basis due to the larger amount of user experience with desktop computers compared to mobile computers. The application was designed to be not as full-featured as commercial map applications of today due to the difficulty in implementation, but also because extra features were not required for the purposes of the experiment. The application was reduced to a simple search text input field, a search button, a "Clear Search Field" button for quick deleting of inputted text, and a "Center Marker" button that allowed users to center the map display onto the search result pin. The text input field was coded to allow instantaneous location suggestions as the user entered a search query.

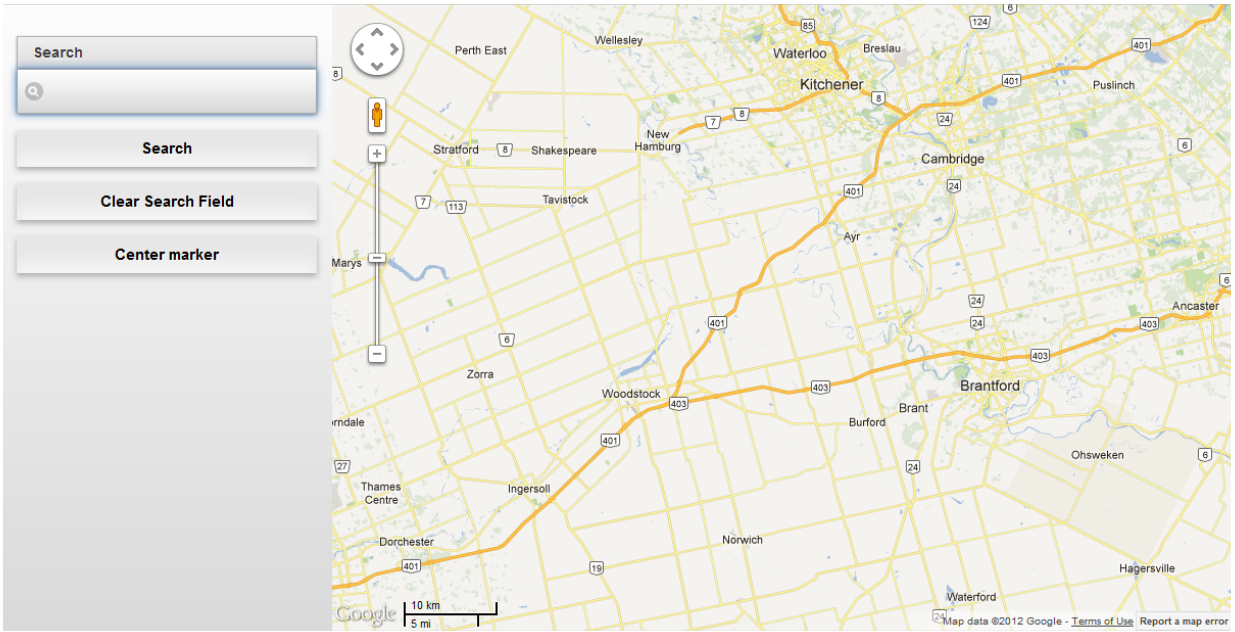


FIGURE 5: DEFAULT APPLICATION

The adaptive look-and-feel of the main application shown above is displayed in Figure 6 below. A green and blue colour scheme was used in this design to induce calmness.

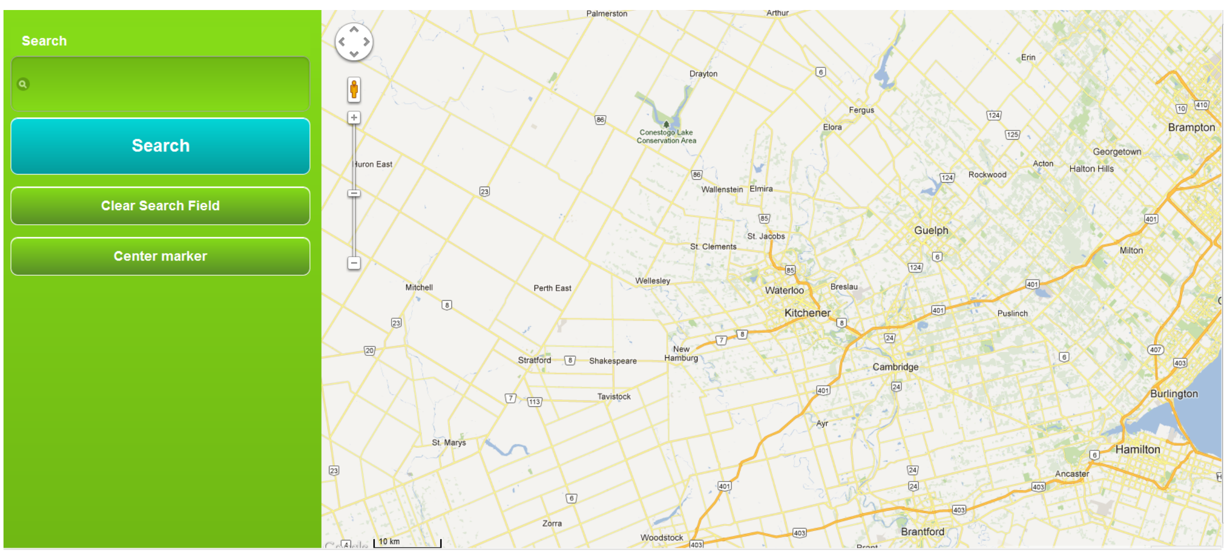


FIGURE 6: ADAPTIVE LOOK-AND-FEEL OF MAIN APPLICATION

The custom map application was developed for the Blackberry Playbook using HTML. It was designed with the ability to transition its look-and-feel between two states, one with the standard mobile look-and-feel using a standard mobile framework (JQuery Mobile) and one with a larger search button and input field, rounded corners, more defined borders for user-interface elements, and cooler colours.

### 3.3.3 Triggers

Participants were allowed 15 minutes to answer as many questions as possible using the map application on the Blackberry Playbook tablet. The specific question sheets can be found in the Appendix.

During this time, the experimenter manually and secretly (using an experimenter dashboard not visible to the participant) triggered a total of five instances where the participant's map application became unresponsive to any touch gestures. These five instances were spread apart roughly evenly in the 15 minutes, with at least two minutes between each trigger. Since participants spent a significant amount of time not interacting with the tablet (e.g. reading or answering questions), the unresponsiveness needed to be triggered manually by the co-located experimenter in order to ensure that they occurred at a time when the participant was using the application.

In both experimental conditions, when unresponsiveness was triggered by the experimenter, the application was frozen until a built-in algorithm detected a significant rise in EDA for up to 10 seconds. This algorithm compared a moving value to the moving average baseline recorded before the stimulus was triggered to determine if a peak has occurred. This algorithm was tuned to be conservative in its selections so it would only catch exaggerated peaks and be less susceptible to noise. As a result, the algorithm was largely unsuccessful and automatic peak detections were uncommon. The 10 second unresponsiveness limit was determined through pilot testing, which revealed that significantly longer delays may have caused participants to disengage from the task, but was long enough to allow for the participant to discover the extent of the unresponsiveness. Also, it allowed the experimenter enough time to silently cancel the triggered event if, at that moment, the participant stopped interacting with the tablet and therefore would not realize that the tablet had become unresponsive.

In the control condition, a "Please wait" pop-up was displayed for one second following the end of the 10-second unresponsiveness window (see Figure 7).

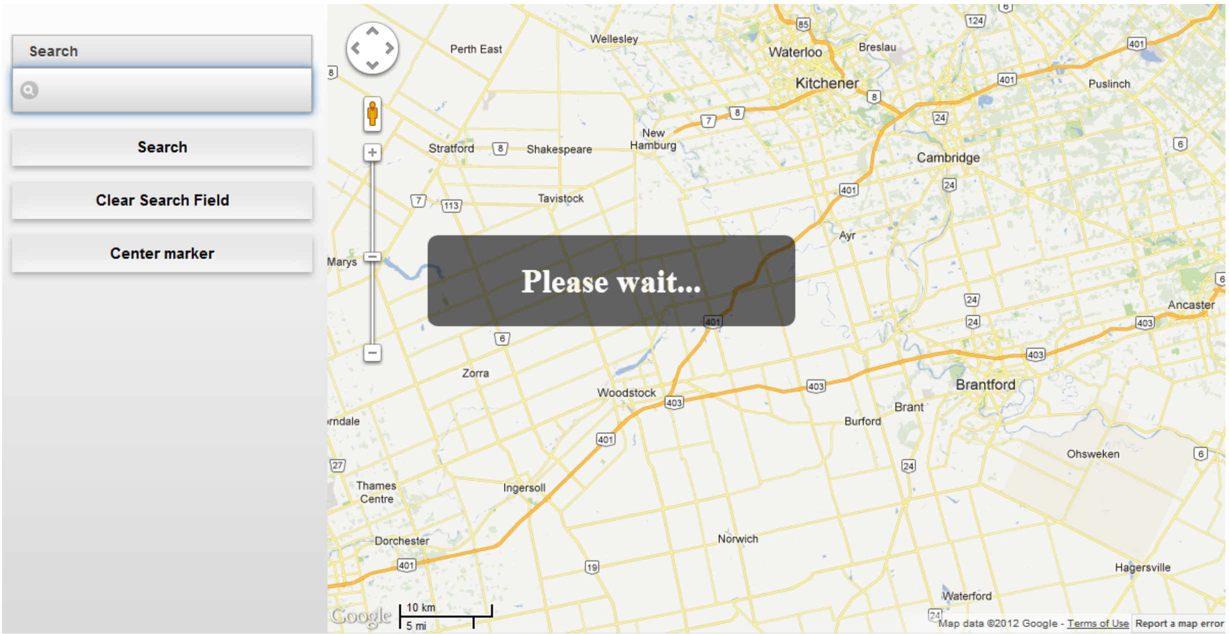


FIGURE 7: CONTROL SCREEN WITH “PLEASE WAIT” POP-UP WHILE RESUMING RESPONSIVENESS

The adaptive condition featured a similar response, but also transitioned, via gradual animation, its look-and-feel while the “Please wait” pop-up was being displayed (see Figure 8). The “Please wait” label was displayed in the adaptive condition in order to control for the effect of providing visual feedback to the user in the control condition.

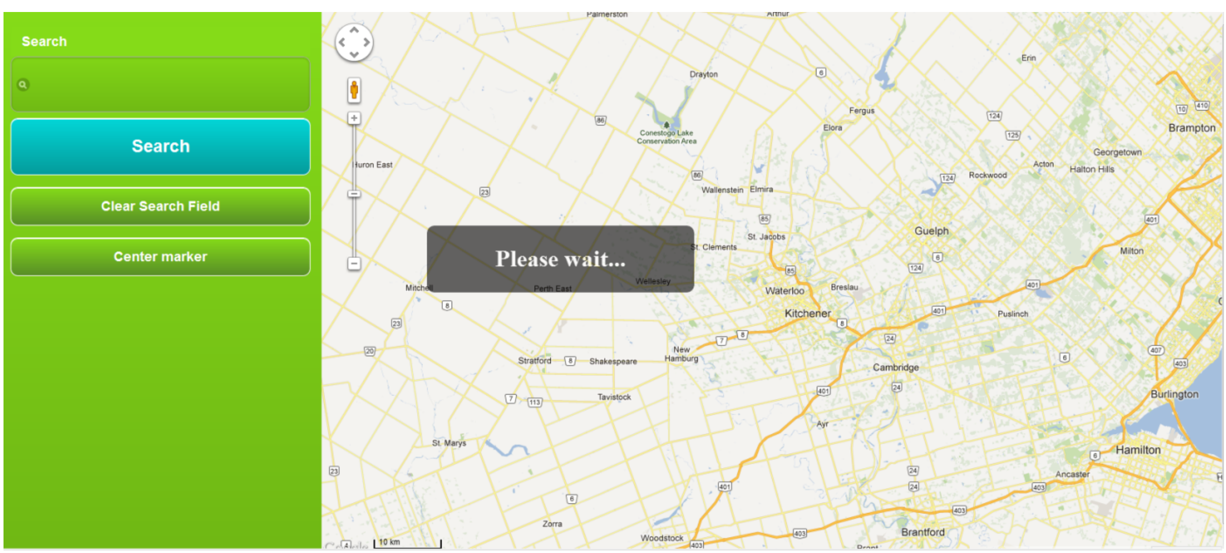


FIGURE 8: ADAPTIVE SCREEN WITH “PLEASE WAIT” POP-UP WHILE RESUMING RESPONSIVENESS

### 3.3.4 PROCEDURE

The experiment consisted of a single 1 hour session. Participants were first given a pre-study information package that excluded the deceptive components of the experiment. Upon providing consent, participants were fitted with the GSR sensor and were asked to complete an online survey consisting of their name, age, computer and mobile device experience, and finally a Mini-IPIP personality questionnaire (Donnellan et al., 2006). This data would be used to compare if neuroticism had a significant impact on the results as well as  to assist in outlier removal if required.

Participants were then given an instruction sheet with steps to familiarize themselves with the customized map application and had the opportunity to ask any questions they had. Participants were shown an instance of an adaptive transition and told that they may or may not experience this transition of look-and-feel throughout the experiment. Following familiarization, participants were given question sheets and a pen. They were instructed to answer as many questions as possible in the 15 minutes of allotted time using the map application on the Blackberry Playbook. Participants were given a chance to have any further questions answered.

To start the main experiment, a 15 minute timer was set up on the desktop monitor and the participant was asked to begin the task.

Five times throughout the experiment, the experimenter manually triggered a 10-second freezing event appropriate for the experimental condition. If participants inquired or complained about the freezing, they were told "it gets stuck sometimes, please continue." This closely followed the protocol used by Scheirer et al. (2002).

Participants were asked to stop once the 15 minute timer alert activated. Each participant was verbally asked about their experience with an initial open-ended question "how did you find the experiment?" If a participant did not include the behaviour or unresponsiveness of the application in their response, the experimenter followed up with "how did you find the application itself?" Further, if participants simply described a freezing or unresponsiveness event, they were asked "how did that make you feel?" This section of the session was recorded with participant consent to capture the audio responses of the participants.

All participants were then given a post-study information package and a second consent form involving the deception used in the study. They were debriefed on the true nature of the experiment and were given the opportunity to not include their data in the experiment.

# 4. DATA ANALYSIS

Along with the raw time-stamped EDA data logged by the skin conductance sensor, participant age, computer experience, personality survey data, and the timestamps of triggers (control and adaptive) were also recorded.

Two major corrections were made to the data prior to analysis. First, participant responses to triggers were analyzed and kept depending on their validity (discussed below). Second, outlier analysis was performed on the remaining data.

## 4.1 TRIGGER SELECTION

Any participant dataset with at least one skin conductance response went through the following selection process.

First, the onset of any SCR needed to be within a 10-second window of the onset of the frustrating trigger, although literature states that the onset of an SCR must be within 3 seconds of the stimulus onset (Andreassi, 2006). The reasoning was as follows:

1. Pilot testing revealed that due to the manual nature of the experiment (e.g. manually triggering the unresponsiveness), most participants did not discover the stimulus immediately.
2. Since unresponsiveness was used as the frustrating stimulus in this experiment, a certain length of discovery (e.g. constantly trying to interact with the system) was required to build up sufficient frustration and cause a measurable SCR. Due to individual differences that may result from age, experience, and personality, a larger analysis window was selected to accommodate for late responders.

Another correction criterion for skin response data was that the end of a triggered event (e.g. the moment the application began to respond) needed to occur sometime after the initial rise of the skin conductance response and before its peak. Otherwise, it could not be reasonably inferred that the end of the unresponsive event (e.g. the adaptation) would have any effect on the following recovery half-time. An example of a valid skin conductance response from the experiment is displayed Figure 9. Note that in this case, the algorithm was able to successfully detect a rise in EDA and automatically triggered the application to unfreeze. Also note that the SCR of this participant occurred with a shorter delay than the typical 1-2 second delay suggested by literature, although it was within the acceptable range.

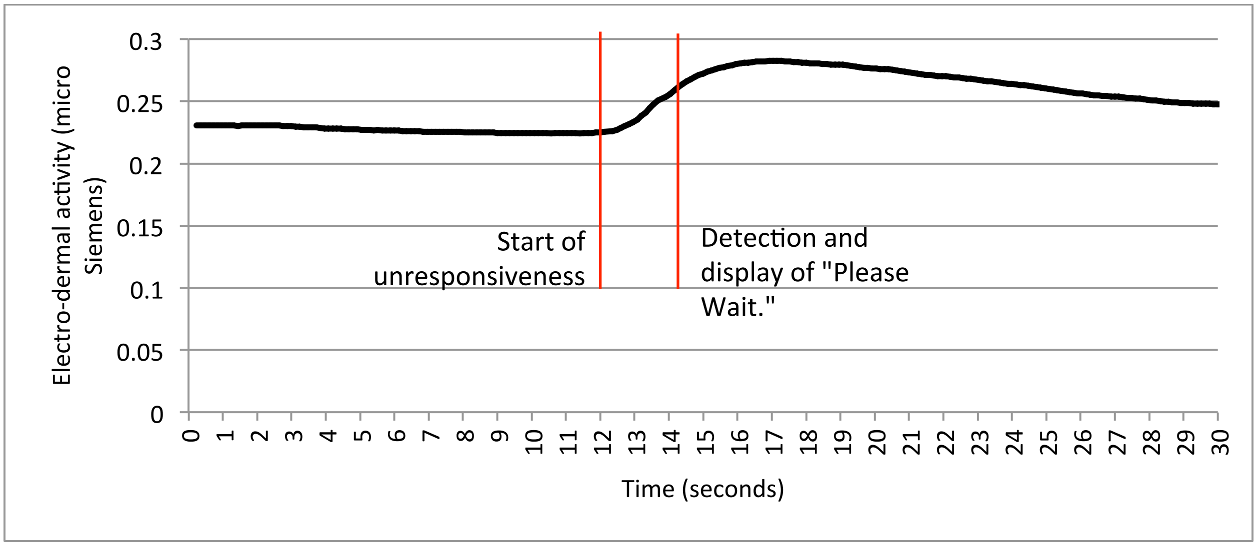


**FIGURE 9: SAMPLE SCR TO A TRIGGERED FRUSTRATING STIMULUS IN ADAPTIVE CONDITION. NOTE THAT THE DETECTION ALGORITHM ACTIVATED IN THIS CASE.**

## 4.2 Outlier Removal

A graphic method was first used to visualize the data and spot potential outliers across each condition. This is shown in Figure 10.

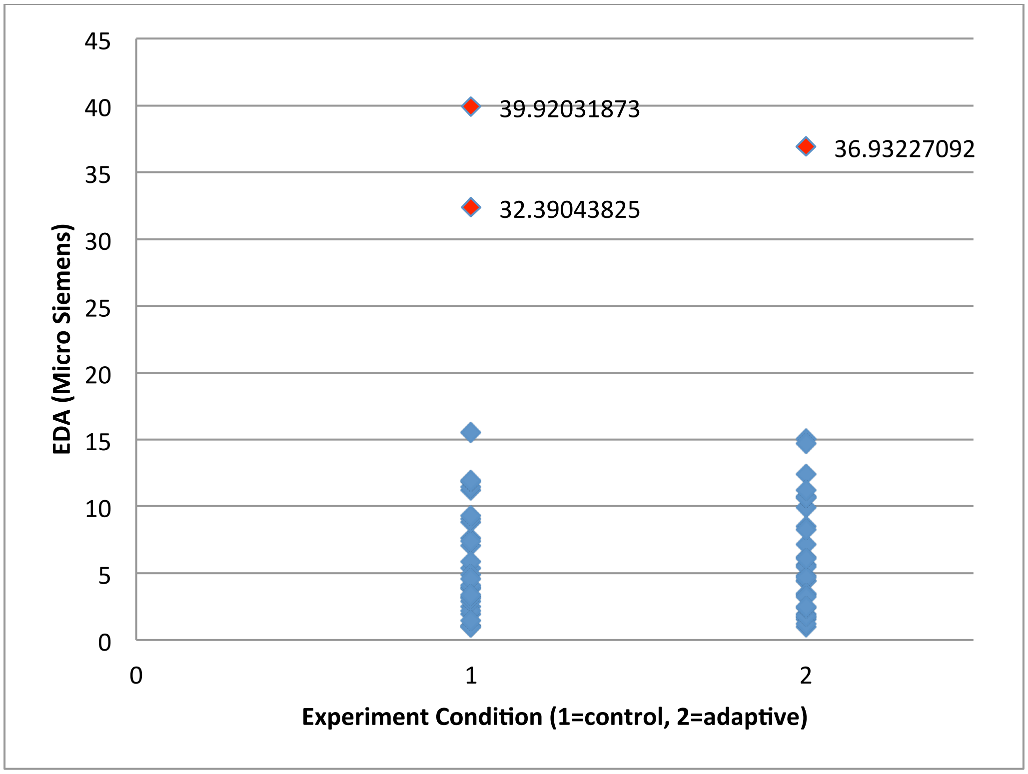


FIGURE 10: CROSS-CONDITION SCATTER OF RECOVERY HALF-TIME DATA. RED POINTS SHOW POTENTIAL OUTLIER DATA.

The Outlier Labeling Rule (Hoaglin et al., 1986; Hoaglin and Iglewicz, 1987) was used to determine statistical outliers. Three data points were identified as outliers (i.e. above the upper bound) and they corresponded to the three points of interest in Figure 10. The remaining data were once again analyzed to confirm no further outliers.

## 4.3 Statistical Analysis

Due to a varying degree of normality in the groups to be analyzed, as well as a relatively small number of samples in most groups that were compared, the Mann-Whitney U test was selected as the preferred method to statistically compare means using SPSS. A Log transform was applied to recovery half-time data in order to de-skew and normalize as much as possible, despite normality not being a necessary assumption of the Mann-Whitney U test.

All error bars in the following graphs represent standard error.

# 5. RESULTS

The observations and data gathered from the pre-study questionnaires, participant responses to frustrating stimuli, their respective recovery responses, and participant feedback are detailed in this chapter.

## 5.1 QUESTIONNAIRES

No significant effect of neuroticism was found on recovery half-time in either of the experimental conditions (shown in Figure 11). No statistically significant differences were found for neuroticism across conditions for any of the following tests (Mann–Whitney U = 56.5, n1 = n2 = 11, P = 0.807 two-tailed).

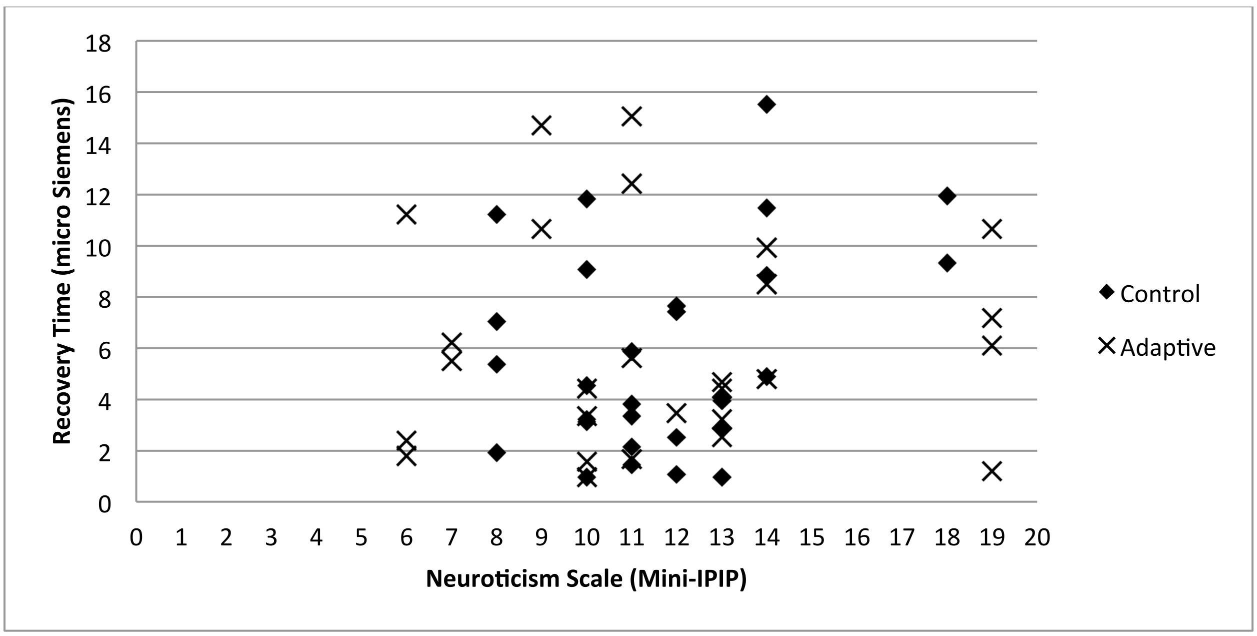


FIGURE 11: COMPARISON OF RECOVERY TIME TO NEUROTICISM

Similarly, there were no statistically significant effects in participant touchscreen, tablet, or map application experience across experimental conditions (Mann–Whitney U = {53.5, 32.5, 44.5}, n1 = 11, n2 = 10, P = {0.934, 0.389, 0.475} two-tailed)

## 5.2 FRUSTRATION

Based on the selected criteria, frustration was successfully induced in 80% of participants in the control condition and 76% of those in the adaptive condition with no statistically significant effects (Mann–Whitney $U = 54.5$, $n1 = n2 = 11$, $P = 0.747$ two-tailed). This is shown in Figure 12.

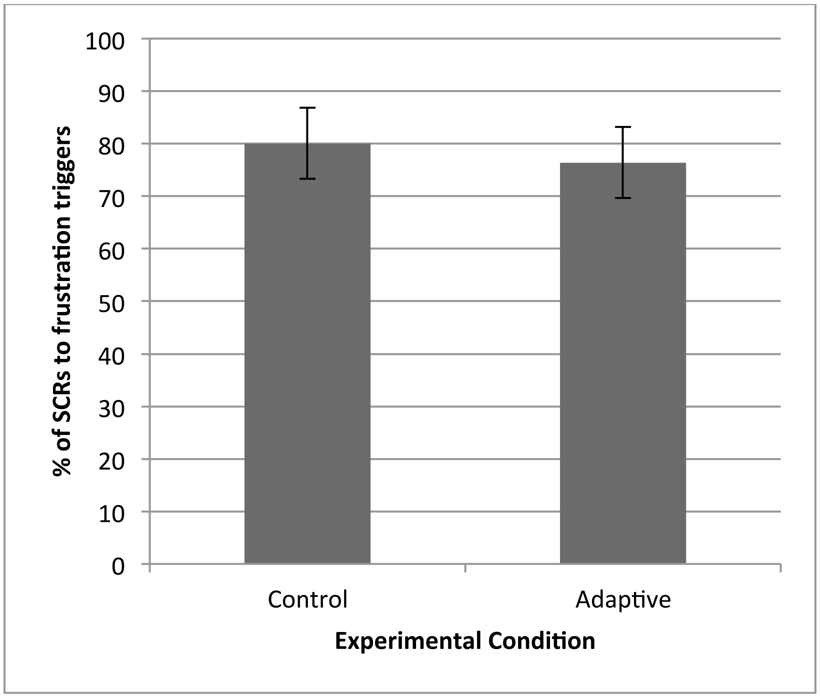


FIGURE 12: PERCENTAGE OF SKIN CONDUCTANCE RESPONSES TO FRUSTRATING STIMULI

Some examples of skin conductance responses from the experiment are shown in Figure 13, Figure 14, Figure 15, and Figure 16 below, with the start and end points of triggered delays (unresponsiveness) labeled. Note that the delays by the application vary up to ten seconds in duration depending on the success of the built-in algorithm to detect a sudden rise in participant's arousal response to the frustrating stimuli.

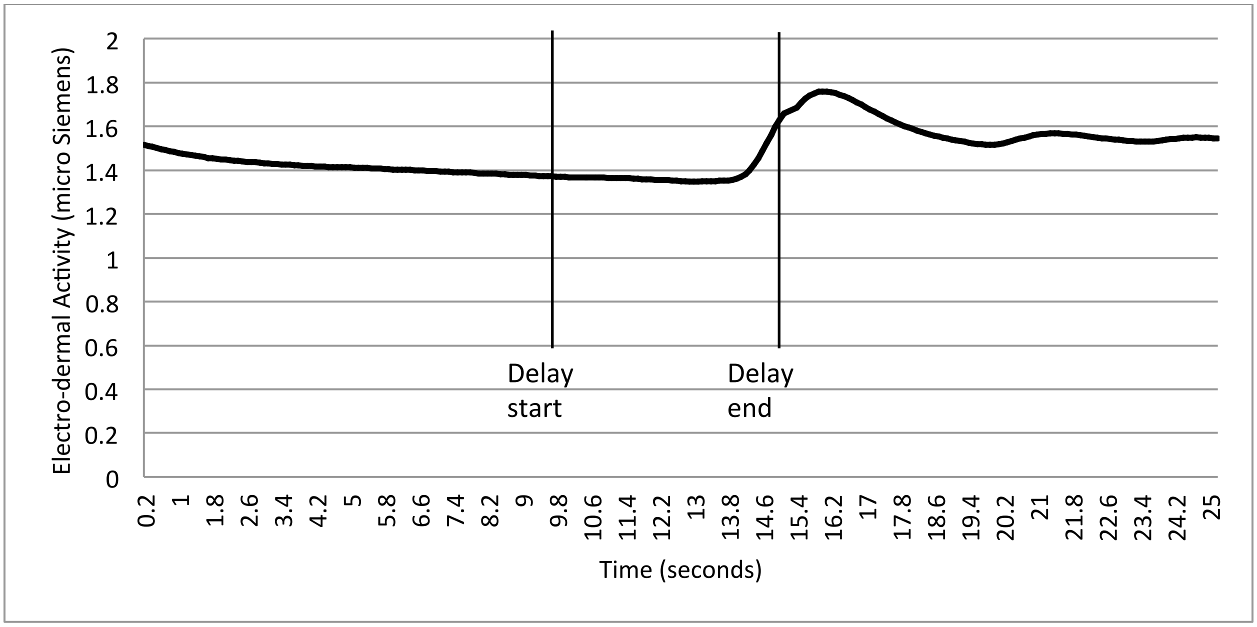


**FIGURE 13: SAMPLE SCR TO A TRIGGERED FRUSTRATING STIMLUS (3 OUT OF 5) IN THE CONTROL CONDITION**

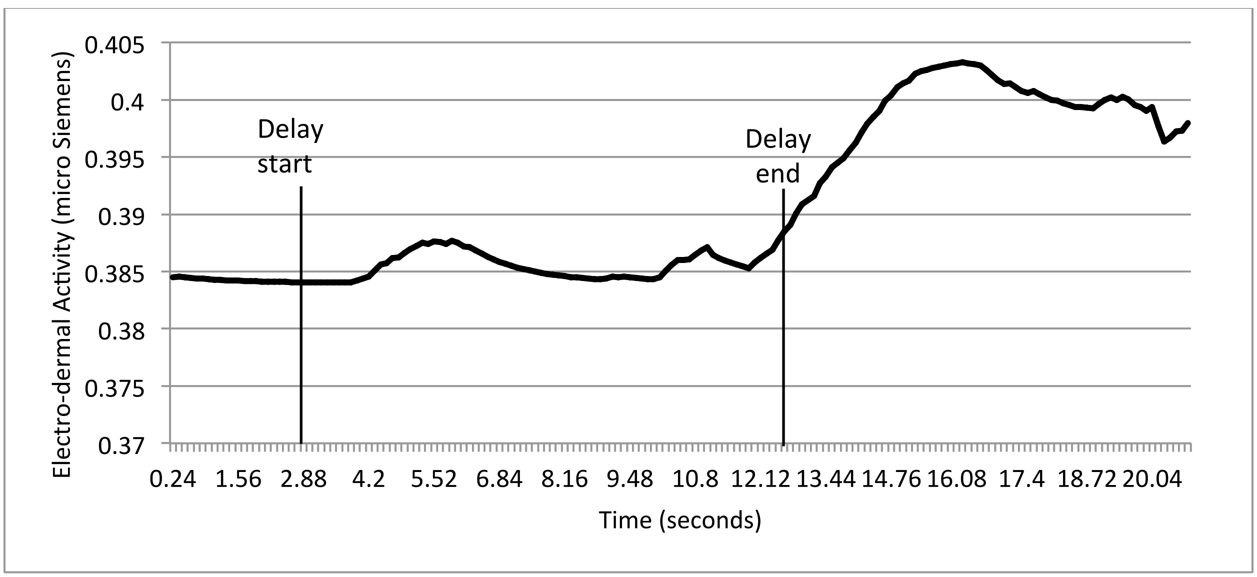


**FIGURE 14: SAMPLE SCR TO A TRIGGERED FRUSTRATING STIMLUS (2 OUT OF 5) IN THE CONTROL CONDITION**

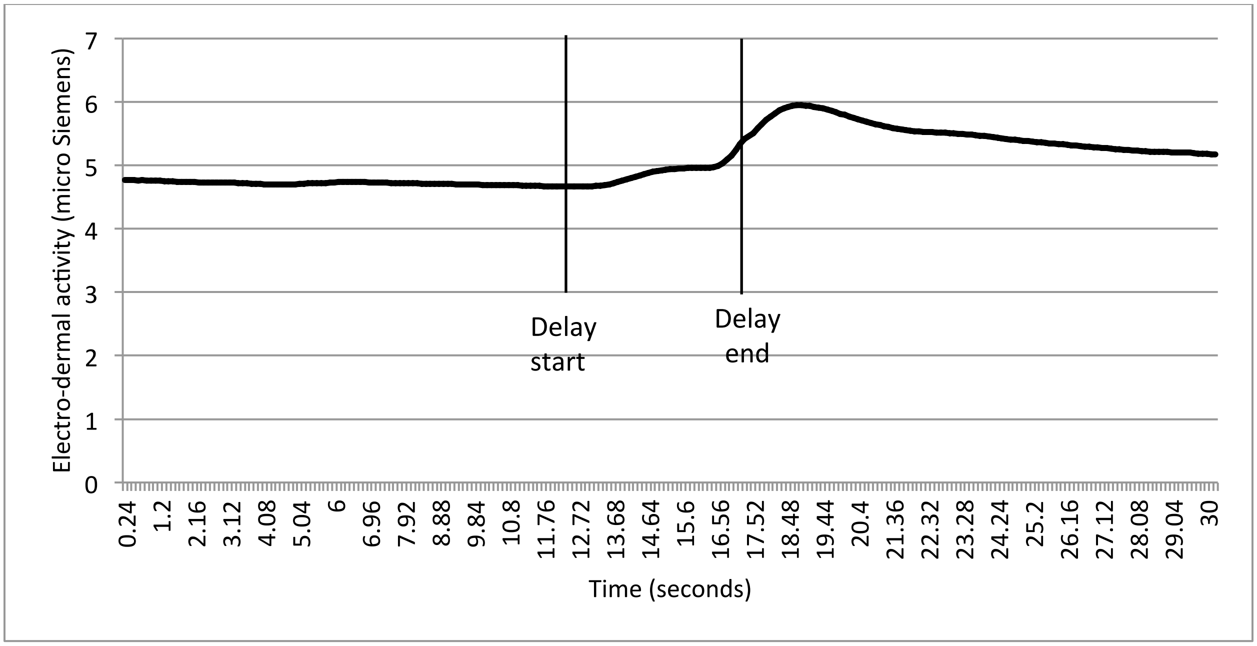


**FIGURE 15: SAMPLE SCR TO A TRIGGERED FRUSTRATING STIMLUS (4 OUT OF 5) IN THE CONTROL CONDITION**

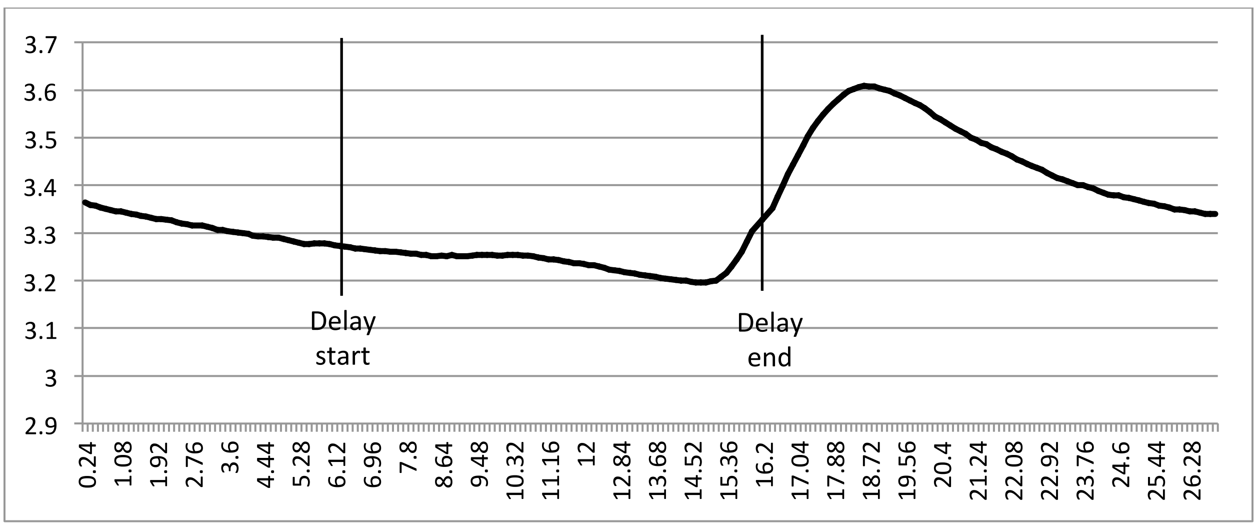


**FIGURE 16: SAMPLE SCR TO A TRIGGERED FRUSTRATING STIMLUS (2 OUT OF 5) IN THE ADAPTIVE CONDITION**

## 5.3 RECOVERY

Figure 17 compares all selected recovery half-time responses to frustrating stimuli across both experimental conditions. Mean recovery half-time for participants in the control condition was 5 seconds compared to 6.2 seconds for participants in the adaptive condition, however this was not statistically significant (Mann–Whitney U = 355, n1 = 29, n2 = 28, P = 0.423 two-tailed).

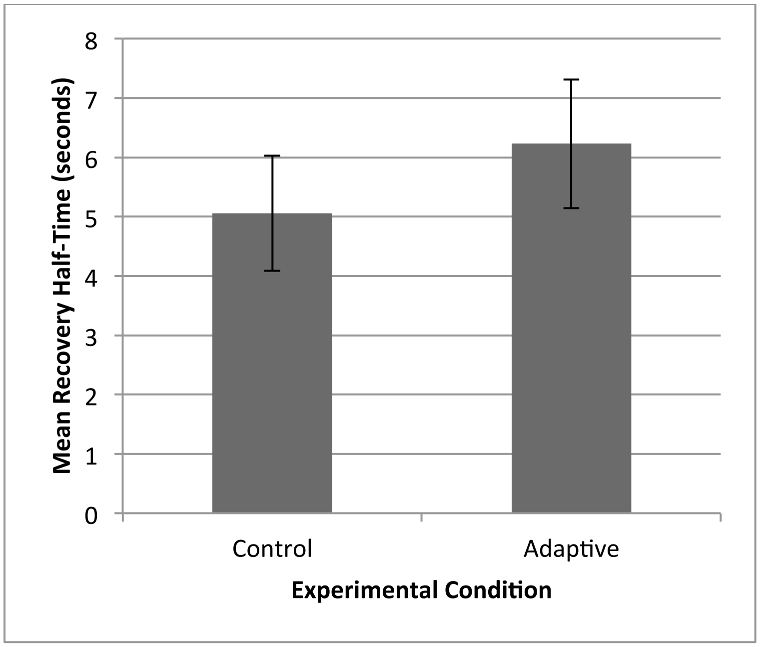


FIGURE 17: MEAN RECOVERY HALF-TIME OF SKIN CONDUCTANCE RESPONSES

However, when recovery half time was examined by trial, a difference was found in the recovery half time for the first trial. Figure 18 looks at the recovery half-time of participants throughout the progression of frustrating stimuli in the order that they were presented. The median Log transforms of recovery times for the initial trial in the control and adaptive conditions were 1.37 and 2.36, respectively; the distributions in the two groups differed significantly (Mann–Whitney U = 2, n1 = n2 = 5, $P < 0.05$ two-tailed). None of the other four trials showed a statistically significant difference between the two conditions (Mann–Whitney U = {14, 16, 11.5, 14}, n1 = {6, 5, 7, 6}, n2 = {5, 7, 5, 5}, P = {0.931, 0.876, 0.364, 0.892} two-tailed). Although a general increasing trend in mean SCR was observed for the Control group and a decreasing trend for that the Adaptive group, no conclusions can be made on this due to high variance. A future study with a larger data set may lower this variance and uncover evidence for any trends that may exist in mean SCR data across trials.

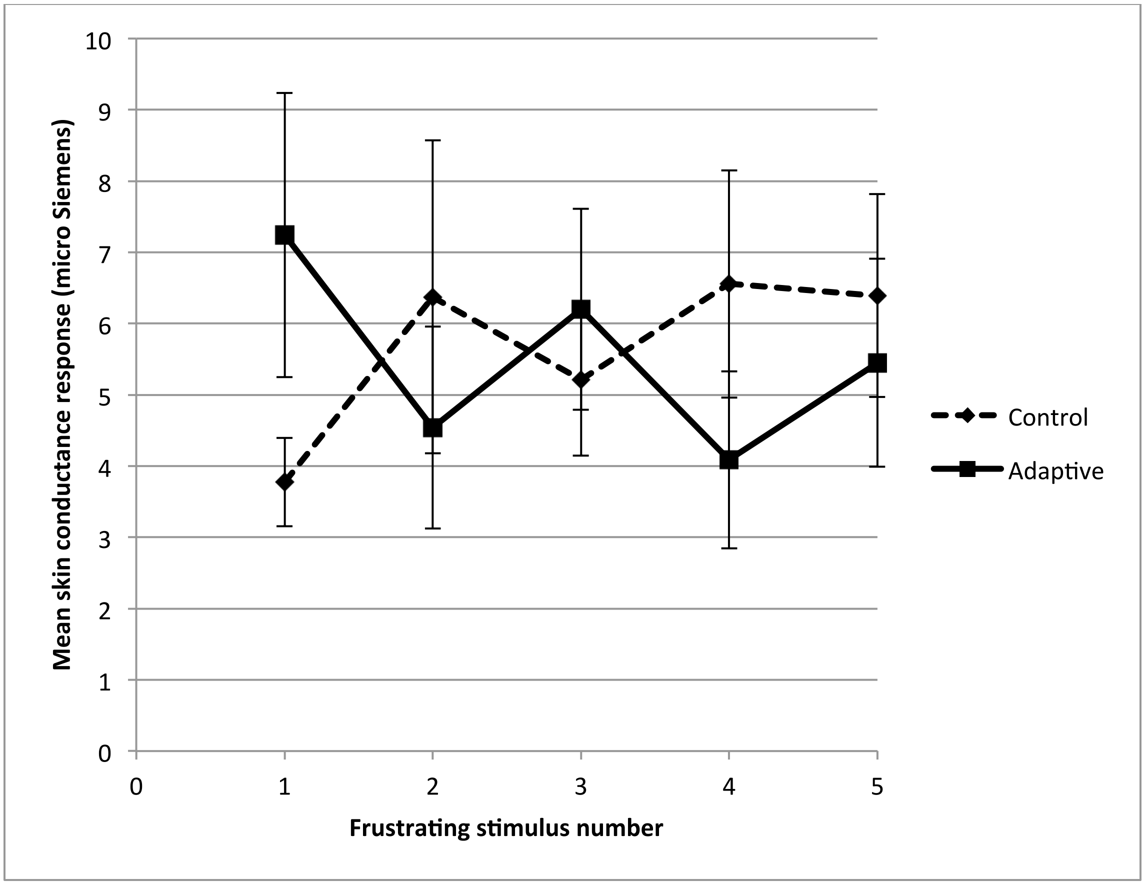


FIGURE 18: PROGRESSION OF SCR TO FRUSTRATING STIMULI ACROSS EXPERIMENTAL CONDITIONS

In order explore comparisons of recovery half-times between males and females; a cross-gender comparison was made in Figure 19. Females exhibited longer mean recovery times of M=6.2 seconds compared to M=4.5 seconds for males, but a Mann-Whitney U test comparison of means based on a normalized Log transform of recovery half-times revealed no statistical significance (Mann–Whitney U = 284, n1 = 21, n2 = 36, P = 0.123 two-tailed).

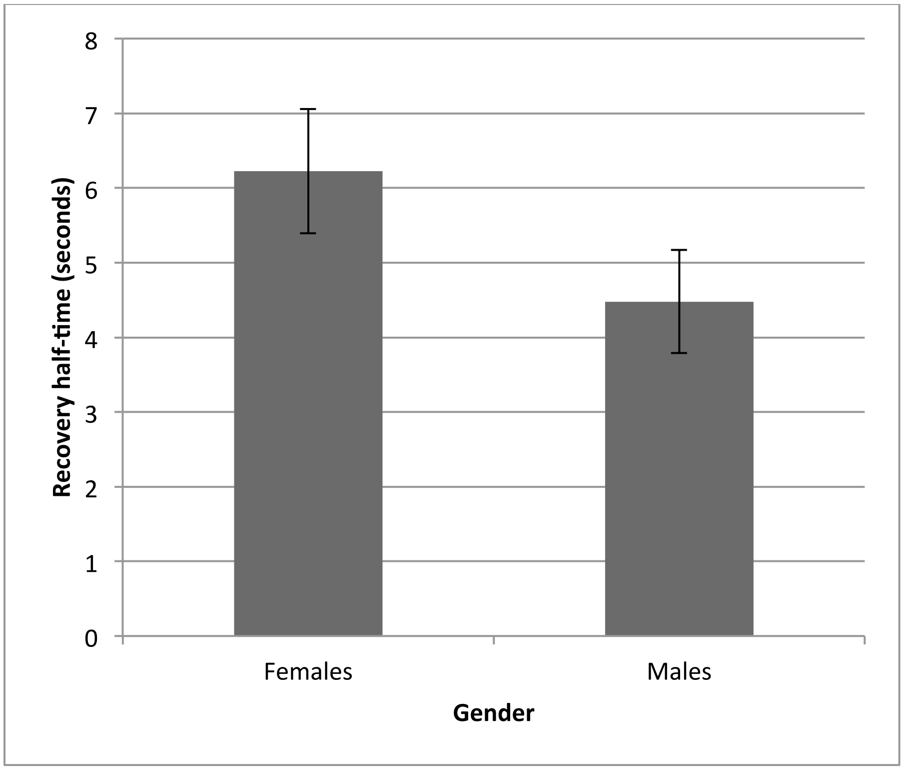


FIGURE 19: CROSS-GENDER COMPARISON OF RECOVERY HALF-TIME MEANS OF ALL TRIALS

## 5.4 PARTICIPANT FEEDBACK

When asked how they felt about the task immediately following their sessions, nearly all participants regardless of condition verbally described their experience as having been frustrating or annoying:

P1: "annoyed a bit, of course", and

*P2: "the screen freezing was annoying".*

However, some participants in the adaptive condition had more detailed responses to the same question regarding the screen freezing, such as:

P3: *"it was stuck for some moments before it changed colour…I tried to zoom or move, I couldn't do it for several seconds",* and

*P4: "whenever the green comes up the…the app stalls and it says 'please wait' so I guess like, cuts off your time from searching…"*, and

*P5: "when it would change colour, it would stop working."*

It should be noted that reference to "green" implies the green and blue colour theme of the adaptive scheme.

# 6. Discussion

The findings on frustration (section 5.2) satisfied the initial hypothesis that frustration would be induced in the majority of trials and represented by a skin conductance response peak when the application was triggered to be unresponsive. The lack of a significant difference between the two experimental conditions was also expected, as the frustrating stimuli did not differ across trials or conditions. This finding also confirms the effectiveness of delays or unresponsiveness in timed and goal-oriented tasks as demonstrated by Scheirer et al. (2002) study.

The observation on mean recovery half-time comparison across condition for all trials did not support the second hypothesis, which stated that the recovery halftime for the skin conductance response peak to the frustrating stimulus would be significantly shorter for participants in the adaptive condition than those in the control condition, implying faster recovery from the frustrated state. Despite this lack of significance, it was also observed that the very first trial exhibited significantly longer recovery times in the adaptive condition. Meanwhile, no significant differences were discovered in any of the subsequent trials. This leads to two important discussions regarding participant habituation and specifics regarding the visual design of the adapted interface state.

First of all, the diminishing effect between the first and subsequent trials may have been due to a learning effect that led participants to anticipate recurring stimuli. Considering that almost all trials were identical in nature and repetitive, participants may have predicted the behaviour of the application for all subsequent trials. In other words, they may have deduced that any subsequent unresponsiveness would be followed by a “please wait” pop-up, an adaptation of the look-and-feel (if they were in the adaptive condition), and resuming responsiveness of the application. In fact, participant feedback appeared to point to this same conclusion – participants in the adaptive condition associated the adaptive look-and-feel to the unresponsiveness of the application. It appeared that associating these two events led to a perceived causal link between the two – in other words, the adaptive event caused the application to stop responding. It should be noted that this discussion refers to associative learning and not cognitive habituation (e.g. lower-level perceptual habituation), a type of non-associative learning. Two possible solutions to break this perception may be to implement a reliable frustration-detection algorithm and to resume the application's responsiveness prior to adaptation. The latter recommendation may break the causal association between the visual feedback of adaptation and the unresponsiveness that it is meant to address. The challenge in this solution would be that a delayed adaptation may occur too late for it to have a measurable effect on the recovery half-time.

The other discussion revolves around the adaptive condition causing significantly longer recovery half-times from induced frustration. The results contradicted the initial hypothesis that a more emotionally calming design with a better user experience would counter the negative emotional aspects of the frustrated user. A plausible explanation may be that despite the adaptive condition exhibiting no change in functionality and having been introduced during initial training, its visual changes were not "relatively small changes at one time" according to adaptive computing guidelines by Duric et al., 2002. Specifically, the issue may have been the greater perceived contrast or chroma in the adaptive screen compared to that of the control, however maintaining the control contrast in the adaptive screen would have resulted in a much less visually noticeable adaptive screen.

The issue of chroma was especially suspect in the results obtained from this experiment. There may have been significant effects resulting from the specific selection of high chroma green and blue colours in the design of the adaptive display. For readers unfamiliar with chroma, it is synonymous with colourfulness, or the perceived intensity of a specific colour. Following the background and research phase of this project, a study published in the *Frontiers of Computer Science* by Wang & Ding (2012) measured the interactions of self-reported arousal and valence emotion data to the lightness and chroma aspects of various colours. They found that a significant difference in self-reported arousal exists between achromatic and chromatic stimuli. They also found that a direct linear relationship between chroma and arousal, as well as chroma and emotional valence. The implications of these findings for this study were significant because the highly chromatic look-and-feel of the adaptive user interface design was a stark contrast to the achromatic colour scheme of the greyish control user interface. This is regardless of the 1-second visual transition (e.g. non-instantaneous) period between the two user interface states. If an increase in perceived colour intensity truly translated to increased physiological arousal, then it could potential affect physiological recovery time as well. Overall, the findings indicate that the specific visual implementation of an adaptive interface used in this experiment was counter-effective. Due to the positive effect of colour on both valence and arousal, the associative learning effect, and a lack of a valence measure, it was inconclusive whether the adaptive design was additionally frustrating (positive arousal, negative valence).

Finally, there is also the issue of measuring physiological recovery using a mere galvanic skin response sensor. Studies that look at physiological reactions to colour have used galvanic skin response as well, but also brain waves (EEG) and heart rate (EKG or ECG). The valence dimension of emotion was not considered in this study due to its difficulty of implementation and required apparatus, but has been used by similar studies as both physiological and self-reported measures. Perhaps more reliable inferences could have been made with the use of these measures.

# 7. Conclusions

The following sections will summarize the findings from this project and draw lessons for future studies and commercial applications of this study.

## 7.1 Summary of Findings

The results confirmed the first hypothesis that frustration would be induced in the majority of trials and represented by a skin conductance response peak when the application was triggered to be unresponsive. This confirmed that using unresponsiveness for durations of up to 10 seconds in a goal-oriented task and a performance incentive is successful in inducing frustration in participants.

Contrary to the second hypothesis, an adaptive user interface consisting of a visually more calm and easy-to-use elements was initially detrimental to the recovery of participants from their frustrated states. This effect was only observed in the first of five trials for each participant, potentially leading to habituation or prediction by participants.

Despite these findings, it should be noted that this was to a large extent an exploratory study aimed at extracting lessons for future studies into affective adaptive computing. The benefits of adaptive and affective user interfaces have been clearly laid out in existing literature. It is only a matter of implementing a suitable adaptive affective display design. Overall, this study led to a series of specific guidelines for designing and implementing future adaptive emotional user interfaces.

## 7.2 Recommendations for Future Work

The following recommendations for future research and commercial devices are made based on the lessons learned from this experiment:

**Investigate the use of chromatic colours with lower perceived colour intensity in the design of adaptive user interfaces.** The findings of this experiment coupled with those of existing literature appear to point towards the design of much more subtle visual adaptations in affective user interfaces. The lack of subtlety in the current experiment was in part due to a large chromatic difference between the affective and default look-and-feel of the application. This, however, does not change the positive effects of the colours green and blue on inducing calm emotions. Thus, future studies that investigate countering high-arousal emotions such as frustration should consider the use of the same colours, except with a much lower chromatic difference between control and adaptive states.

**A distinct dissociation should be designed to exist between frustrating stimuli in a system and that system's adaptive response to said stimuli.** In other words, an adaptation should not present itself as a direct cause of system unresponsiveness. Rather the adaptation should be purposefully delayed and as gradual as possible. Another possible solution may be to introduce dissociative stimuli (e.g. non-adaptive unresponsiveness events) in an attempt to counter the associative learning effect.

**Explore another primary measure of physiological recovery other than the skin conductance recovery half-time.** Heart rate sensing via EKG or the measurement of alpha brain waves via EEG may be a relatively versatile and simple solution to implement. If multiple physiological response sensors are used, they should be combined with appropriate predictive models (e.g. Hidden Markov, Gaussian Prediction) to make more accurate inferences of user affect. It should also be noted that this study only made use of the arousal dimension of emotion. Emotional valence detection would prove extremely useful in gauging user affect.

**A real-time peak prediction algorithm should be implemented**. This algorithm should be used by an adaptive system to automatically transition to an adaptive state if a spike in the difference between current and baseline user arousal has occurred following a potentially frustrating event.

**Forced delays on goal-oriented tasks are strong inductors of frustration** and are relatively versatile and easy to implement for future experiments. Future experiments in adaptive affective computing should consider this simple but effective technique.

# REFERENCES


Ali, M. R. (1972). Pattern of EEG recovery under photic stimulation by light of different colors. *Electroencephalography and Clinical Neurophysiology, 33*(3), 332-335.

Anderson, a K., & Phelps, E. a. (2001). Lesions of the human amygdala impair enhanced perception of emotionally salient events. *Nature*, *411*(6835), 305-9.

Andreassi, J. L. (2006). Psychophysiology: Human Behavior and Physiological Response. Lawrence Erlbaum.

Akgun, M., Akilli, G. K., & Cagiltay, K. (2011). Bringing Affect to Human Computer Interaction. D. Gökcay, & G. Yildirim,(Eds.). *Affective Computing and Interaction: Psychological, Cognitive and Neuroscientific Perspectives*, 308-324.

Banbury, S., Gauthier, M., Scipione, A., and Hou, M. (2005). Intelligent adaptive systems: literature-research of design guidance for intelligent adaptive automation and interfaces. *Defence R&D Canada – Toronto report*.

Bassili, J. N. (1979). Emotion recognition: the role of facial movement and the relative importance of upper and lower areas of the face. *Journal of personality and social psychology, 37*(11), 2049.

Benyon, D., & Murray, D. (1993). Adaptive systems: from intelligent tutoring to autonomous agents. *Knowledge-Based Systems, 6*(4), 197–219.

Caggiano, D. M., & Parasuraman, R. (2004). The role of memory representation in the vigilance decrement. *Psychonomic Bulletin & Review*, *11*, 932-7.

Ceaparu, I., Lazar, J., & Bessiere, K. (2004). Determining causes and severity of end-user frustration. *International Journal of Human-Computer Interaction, 17*(3), 333-356.

Chu, Y., and Rouse, W. B. (1979). Adaptive allocation of decision making responsibility between human and computer in multi-task situations. *IEEE Transactions of Systems, Man and Cybernetics, 9,* 769-778.

Damasio, a R. (1996). The somatic marker hypothesis and the possible functions of the prefrontal cortex. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, *351*(1346), 1413-20.

Donnellan, M. B., Oswald, F. L., Baird, B. M., & Lucas, R. E. (2006). The Mini-IPIP Scales : Tiny-Yet-Effective Measures of the Big Five Factors of Personality. *Psychological Assessment, 18*(2), 192–203.

Duric, Z., Gray, W. D., Heishman, R., Li, F., Rosenfeld, A., Schoelles, M. J., & Wechsler, H. (2002). Integrating perceptual and cognitive modeling for adaptive and intelligent human-computer interaction. *Proceedings of the IEEE, 90*(7), 1272-1289.

Elliot, A. J., Maier, M. a, Moller, A. C., Friedman, R., & Meinhardt, J. (2007). Color and psychological functioning: the effect of red on performance attainment. *Journal of experimental psychology: general*, *136*(1), 154–68.

Erdem, A., & Karaismailoglu, S. (2011). Neurophysiology of Emotions. *Affective Computing and Interaction: Psychological, Cognitive and Neuroscientific Perspectives* (pp. 1-24). Hershey.

Figner, B., & Murphy, R. O. (2011). Using skin conductance in judgment and decision making research. *A handbook of process tracing methods for decision research,* 163-184.

Gardner, A., & Essa, I. (1997). Prosody analysis for speaker affect determination. *In Proceedings of Perceptual User Interfaces Workshop.*

Fernandez, R. (1998). Stochastic modeling of physiological signals with hidden markov models: A step toward frustration detection in human-computer interfaces (Master's thesis, Massachusetts Institute of Technology, Dept. of Electrical Engineering and Computer Science).

Findlater, L., & Mcgrenere, J. (2004). A Comparison of Static, Adaptive, and Adaptable Menus, *Proceedings of the ACM Conference on Human Factors in Computing (CHI 2004) ,* 6(1), 89–96.

Findlater, L., & McGrenere, J. (2008). Comprehensive user evaluation of adaptive graphical user interfaces. *Usable AI,* 10–13.

Freedy, A., Madni, A., and Samet, M. (1985). Adaptive user models: Methodology and applications in man-computer systems. In W. B. Rouse (Ed.) *Advances in man-machine systems research.* JAI Press: Greenwich.

Gajos, K. Z., Czerwinski, M., Tan, D. S., & Weld, D. S. (2006). Exploring the design space for adaptive graphical user interfaces. *Proceedings of the working conference on Advanced visual interfaces - AVI '06*, 201.

Geddes, N.D., and Shalin, V.L. (1997). Intelligent decision aiding for aviation. *Technical Report (No. ISSN 1402-7585).* Prepared for Linkoping Institute of Technology, Linkpoping, Sweden.

Gökçay, D. (2011). Emotional Axes: Psychology, Psychophysiology and Neuroanatomical Correlates. *Affective Computing and Interaction: Psychological, Cognitive and Neuroscientific Perspectives. IGI Global.*

Gross, J. J. (2003). Emotion regulation: Affective, cognitive, and social consequences. *Psychophysiology, 39*(3), 281-291.

Hitchcock, E. M., Warm, J. S., Matthews, G., Dember, W. N., Shear, P. K., Tripp, L. D., Mayleben, D. W., & Parasuraman, R. (2003). Automation cueing modulates cerebral blood flow and vigilance in a simulated air traffic control task. *Theoretical Issues in Ergonomics Science*, *4*, 89-112.

Hoaglin, D.C., Iglewicz, B., and Tukey, J.W. (1986). Performance of some resistant rules for outlier labeling, *Journal of American Statistical Association, 81*, 991-999.

Hoaglin, D. C., and Iglewicz, B. (1987), Fine tuning some resistant rules for outlier labeling, *Journal of American Statistical Association, 82*, 1147-1149.

Janes, C. L. (1982). Electrodermal recovery and stimulus significance. *Psychophysiology*, *19*(2), 129-135.

Joubert, T., Sallé, S.E., Champigneux, G., Grau, J.Y., Sassus, P., and Le Doeuff, H. (1995). The Copilote Electronique project: First lessons as explanatory development starts. In *Proceedings of the 3rd international workshop on human-computer teamwork (Human-Electronic Crew: Can we trust the team?).*

Kapoor, A., Burleson, W., & Picard, R. W. (2007). Automatic prediction of frustration. *International Journal of Human-Computer Studies*, *65*(8), 724–736.

Klein, J., Moon, Y., & Picard, R. W. (2002). This computer responds to user frustration: Theory, design, and results. *Interacting with computers, 14*(2), 119-140.

Krijn, M., Emmelkamp, P. M. G., Biemond, R., de Wilde de Ligny, C., Schuemie, M. J., & van der Mast, C. a P. G. (2004). Treatment of acrophobia in virtual reality: the role of immersion and presence. *Behaviour research and therapy*, *42*(2), 229-39.

Küller, R., Mikellides, B., & Janssens, J. (2009). Color, arousal, and performance—A comparison of three experiments. *Color Research & Application, 34*(2), 141-152.

Lee, H., Choi, Y. S., & Kim, Y. J. (2011). An adaptive user interface based on spatiotemporal structure learning. *Communications Magazine, IEEE*, *49*(6), 118–124.

LiKamWa, R., Liu, Y., Lane, N. D., & Zhong, L. (2011). Can Your Smartphone Infer Your Mood ? *PhoneSense Workshop*.

Lisetti, C. L., & Nasoz, F. (2004). Using Noninvasive Wearable Computers to Recognize Human Emotions from Physiological Signals. *EURASIP Journal on Advances in Signal Processing*, *2004*(11), 1672-1687.

McGrenere, J., Baecker, R. M., & Booth, K. S. (2002). An evaluation of a multiple interface design solution for bloated software. *Proceedings of the SIGCHI conference on Human factors in computing systems Changing our world, changing ourselves - CHI '02*, 164.

Matthews, G. (2001). Individual differences in energetic arousal and sustained attention: A dual-task study. *Personality and Individual Differences, 31,* 575-589.

Matthews, G., Davies, D. R., & Holley, P. J. (1990). Arousal and visual sustained attention: The role of resource availability. *Personality and Individual Differences, 11,* 1159-1173.

Morris, N. M., and Rouse, W. B. (1986). Adaptive aiding for human computer control: Experimental studies of dynamic task allocation. *Technical report (No. AFAMRL-TR-86-0005).* Wright Patterson Air Force Base, OH.

Morrison, J. G., Gluckman, J. P., & Deaton, J. E. (1992). Human performance in complex task environments: a basis for the application of adaptive automation. *Technical report (No. NAWCADWAR-92032-60).*

Nasoz, F., Lisetti, C. L., Alvarez, K., & Finkelstein, N. (2003). Emotion recognition from physiological signals for user modeling of affect. *Proceedings of the 3rd Workshop on Affective and Attitude User Modelling, Pittsburgh, PA, USA.*

Nielsen, J. (1998). Improving the Dreaded 404 Error Message. Retrieved from http://www.useit.com/alertbox/404_improvement.html

Picard, R.W., Vyzas, E., & Healey, J. (2001). Toward machine emotional intelligence: Analysis of affective physiological state. *IEEE Transactions of Pattern Analysis and Machine Intelligence, 23(10)*, 1175–1191.

Picard , R.W. (2001), Building HAL: Computers that Sense, Recognize, and Respond to Human Emotion. *Society of Photo-Optical Instrumentation Engineers. Human Vision and Electronic Imaging VI*, Part of IS&T/SPIE Photonics West 2001.

Picard, R. W., & Klein, J. (2002). Computers that recognise and respond to user emotion: theoretical and practical implications. *Interacting with Computers*, *14*(2), 141-169.

Picard, R. W. (1995). Affective Computing. Cambridge, MA: MIT Press.

Phelps, E. A. (2006). Emotion and cognition: insights from studies of the human amygdala. *Annual Review of Psychology, 57*, 27-53.

Phelps, E. A., O'Connor, K. J., Gatenby, J. C., Gore, J. C., Grillon, C., & Davis, M. (2001). Activation of the left amygdala to a cognitive representation of fear. *Nature neuroscience, 4*(4), 437-441.

Prendinger, H., Mayer, S., Mori, J., & Ishizuka, M. (2003). Persona effect revisited: Using bio-signals to measure and reflect the impact of character-based interfaces. *In Intelligent Virtual Agents* (pp. 283-291). Springer Berlin/Heidelberg.

Prinzel, L.J., Freeman, F.G., Scerbo, M.W., Mikulka, P.J. and Pope, A.T. (1999). A closed loop system for examining psychophysiological measures for adaptive task allocation. *The International Journal of Aviation Psychology, 10*(4), 393–410.

Prinzel, LJ., Parasuraman, R., Freeman, F.G., Scerbo, M, Mikulka., P.J., and Pope, A. (2003). Three experiments examining the use of electroencephalogram, event-related potentials, and heart-rate variability for real- time human-centered adaptive automation design. *Technical Report (No. NASA/TP-2003-212442).*

PR Newswire. (2010). Brain Gap: NeuroFocus Study Reveals What Went Wrong With the Gap's New Brand Logo [Press Release]. Retrieved from http://www.prnewswire.com/news-releases/brain-gap-neurofocus-study-reveals-what-went-wrong-with-the-gaps-new-brand-logo-105165954.html

Richards, J.M. & Gross, J.J. (2000). Emotion Regulation and Memory: The Cognitive Costs of Keeping One's Cool. *Journal of Personality and Social Psychology, 79 (3)*, 410-424.

Russel, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology, 39*(6), 1161–1178.

Scheirer, J., Fernandez, R., Klein, J., & Picard, R. W. (2002). Frustrating the user on purpose: a step toward building an affective computer. *Interacting with Computers*, *14*, 93-118.

St John, M., Kobus, D. A., & Morrison, J. G. (2003). DARPA augmented cognition technical integration experiment (TIE). *DARPA Technical Report TR-1905.*

Troncoso, X. G., Macknik, S. L., & Martinez-Conde, S. (2009). Corner salience varies linearly with corner angle during flicker-augmented contrast: a general principle of corner perception based on Vasarely's artworks. *Spatial vision*, *22*(3), 211–24.

Tzeng, J. (2004). Toward a more civilized design: Studying the effects of computers that apologize. *International Journal of Human-Computer Studies*, *61*, 319–345.

Tzeng, J. (2006). Matching users' diverse social scripts with resonating humanized features to create a polite interface. *International Journal of Human-Computer studies*, *64*, 1230–1242.

Van Velsen, L., Van Der Geest, T., Klaassen, R., & Steehouder, M. (2008). User-centered evaluation of adaptive and adaptable systems: a literature review. *The Knowledge Engineering Review, 23*(03), 261–281.

Vascellaro, J. E. (2012). Apple's Siri Gains Traction–For Some Things. *The Wall Street Journal*. Retrieved from http://blogs.wsj.com/digits/2012/03/26/apple%E2%80%99s-siri-gains-traction-for-some-things/

Veletsianos, G. (2007). Cognitive and Affective Benefits of an Animated Pedagogical Agent: Considering Contextual Relevance and Aesthetics. *Journal of Educational Computing Research, 36*(4), 373-377.

Vincent, J. (2005). Emotional attachment to mobile phones: an extraordinary relationship. *Mobile World*, 93-104.

Wang, S., & Ding, R. (2012). A qualitative and quantitative study of color emotion using valence-arousal. *Frontiers of computer science in China, 6*(4), 469-476.

Wilson, G.F. (2002). Adaptive aiding implemented by psychophysiologically determined operator functional state. *In Proceedings of RTO Human Factors and Medicine Panel (HFM) Symposium held in Warsaw, Poland, 7-9 October 2002*.

Yacoob, Y., & Davis, L. S. (1996). Recognizing human facial expressions from long image sequences using optical flow. *Pattern Analysis and Machine Intelligence, IEEE Transactions on, 18*(6), 636-642.

Young, M. P., Scanneil, J. W., Burns, G. A., & Blakemore, C. (1994). Analysis of connectivity: neural systems in the cerebral cortex. *Reviews in the Neurosciences, 5*(3), 227-250.

Young, M.J. and Eggleston, R.G. (2002). Work-centered decision support. In *Proceedings of RTO Human Factors and Medicine Panel (HFM)* Symposium held in Warsaw, Poland, 7-9 October 2002.

# Appendix A: Experimental Materials

**Title of Project: Physiological Response to a Mobile Map Interface**

**Student Investigator:** ***Behzad Aghaei, baghaei@uwaterloo.ca***

**Faculty Supervisor:** **Catherine Burns, Catherine.burns@uwaterloo.ca**

**Summary of the Project:**

This study is a part of my master's project. The goal of the project is to gain a deeper insight into what makes usable and efficient mobile map software by studying physiological reactions to its use. More specifically, we will use data gathered by video recording, galvanic skin response measurement, and questionnaires regarding personality, workload, and computer experience. We have designed and developed a prototype map application for the Blackberry Playbook, which we are planning to study in our project.

**Procedure:**

Your participation in this study is voluntary. Participation involves performing a search for landmarks on our custom map application based on a set of directions and clues. While performing the searches you will have a galvanic skin response sensor attached to your wrist with a velcro strap. You will be video recorded at this time.

You will receive an introduction to the study, the physiological Galvanic Skin Response sensor, and the custom map application on the Blackberry Playbook, and then you will be asked to

- Complete a short standardized personality questionnaire (Mini-IPIP)
- Take part in a brief training session with map application interface.
- Take part in a timed activity where you will attempt to identify as many landmarks on a provided question sheet as possible.
- Complete a standardized questionnaire of your workload during the task (NASA-TLX) and a custom questionnaire about your general computer/software experience
- Participate in reviewing parts of the videos we recorded of you to help us understand your experience better

Your session will take approximately 1 hour.

During parts of the session a researcher will take notes regarding your responses to the video review. Your computer-based interactions will also be captured and stored in a computer log file. You may decline to respond to questions if you wish. You may withdraw your participation at any time without penalty.

**Confidentiality and Data Security:**

All information you provide is considered completely confidential. Your name will not appear in any publication resulting from this study; however, with your permission anonymous quotations may be used. In these cases participants will be referred to as 'a participant'. Data collected during this study will be retained indefinitely in locked cabinets or on password protected desktop computers in a secure location. Electronic data will not include personal identifying information such as names or student ID.

Your video/audio data, captured from the video recording, will be destroyed up to one year after the completion of the study.

**Remuneration for Your Participation:**

You will receive remuneration for your participation in this study, for a total of $10 if you complete the session. If you choose to withdraw your participation from the study prior to study completion, you will be remunerated for $5.

The highest performing participant in each experimental condition of this study will receive an additional $100. You will be notified which condition you were assigned to following the experiment.

The amount received is taxable. It is your responsibility to report the amount received for income tax purposes.

**Risks and Benefits:**

The risks are no greater than what you experience in every day life. There are no direct benefits to you from participation. However, the results of this research may contribute to the knowledge base of Human Systems Engineering research and to lead to the development of commercial mobile map applications.

**Research Ethics Clearance:**

I would like to assure you that this study has been reviewed and received ethics clearance through the Office of Research Ethics at the University of Waterloo. However, the final decision about participation is yours. Should you have comments or concerns resulting from your participation in this study, please contact Dr. Maureen Nummelin, the Director, Office of Research Ethics, at 1-519-888-4567, Ext. 36005 or maureen.nummelin@uwaterloo.ca.

Thank you for your assistance in this project.

## CONSENT FORM

By signing this consent form, you are not waiving your legal rights or releasing the investigator(s) or involved institution(s) from their legal and professional responsibilities.

______________________________________________________________________________

**Project:** ***Evaluating a Mobile Map Interface***

I have read the information presented in the information letter about a study being conducted ***by Behzad Aghaei*** of the Department of ***Systems Design Engineering***, under the supervision of Professor ***Catherine Burns.*** I have had the opportunity to ask any questions related to this study, to receive satisfactory answers to my questions, and any additional details I wanted.

I am aware that I may allow excerpts from the conversational data collected for this study to be included in scientific presentations and/or publications, with the understanding that any quotations will be anonymous.

I am aware that I may withdraw my consent for any of the above statements or withdraw my study participation at any time without penalty by advising the researcher.

This project has been reviewed by, and received ethics clearance through, the Office of Research Ethics at the University of Waterloo. I was informed that if I have any comments or concerns resulting from my participation in this study, I may contact Dr. Maureen Nummelin, the Director, Office of Research Ethics, at 1-519-888-4567, Ext. 36005 or maureen.nummelin@uwaterloo.ca..

| | **Please Circle One** | **Please initial Your Choice** |
|---|---|---|
| With full knowledge of all foregoing, I agree, of my own free will to participate in this study. | YES NO | ________ |
| I agree to let my conversation during the study be directly quoted, anonymously, in presentations of research results | YES NO | ________ |

Participant Name: ___________________________________________
(Please print)

Participant Signature: ________________________________________

Witness Name: ______________________________________________
(Please print)

Witness Signature: ___________________________________________

Date: ___________________________________

**There are phrases describing people's behaviors. Please use the rating scale below to describe how accurately each statement describes you. Describe yourself as you generally are now, not as you wish to be in the future. Describe yourself as you honestly see yourself, in relation to other people you know of the same sex as you are, and roughly your same age. So that you can describe yourself in an honest manner, your responses will be kept in absolute confidence. Please read each statement carefully, and then fill in the bubble that corresponds to the number on the scale.** *

| | 1=Very Inaccurate | 2=Moderately Inaccurate | 3=Neither Inaccurate nor Accurate | 4=Moderately Accurate | 5=Very Accurate |
|---|---|---|---|---|---|
| 1. Am the life of the party * | ◯ | ◯ | ◯ | ◯ | ◯ |
| 2. Sympathize with others' feelings * | ◯ | ◯ | ◯ | ◯ | ◯ |
| 3. Get chores done right away * | ◯ | ◯ | ◯ | ◯ | ◯ |
| 4. Have frequent mood swings * | ◯ | ◯ | ◯ | ◯ | ◯ |
| 5. Have a vivid imagination * | ◯ | ◯ | ◯ | ◯ | ◯ |
| 6. Don't talk a lot * | ◯ | ◯ | ◯ | ◯ | ◯ |
| 7. Am not interested in other people's problems * | ◯ | ◯ | ◯ | ◯ | ◯ |
| 8. Often forget to put things back in their proper place * | ◯ | ◯ | ◯ | ◯ | ◯ |
| 9. Am relaxed most of the time * | ◯ | ◯ | ◯ | ◯ | ◯ |
| 10. Am not interested in abstract ideas * | ◯ | ◯ | ◯ | ◯ | ◯ |
| | 1=Very Inaccurate | 2=Moderately Inaccurate | 3=Neither Inaccurate nor Accurate | 4=Moderately Accurate | 5=Very Accurate |
| 11. Talk to a lot of different people at parties * | ◯ | ◯ | ◯ | ◯ | ◯ |
| 12. Feel others' emotions * | ◯ | ◯ | ◯ | ◯ | ◯ |
| 13. Like order * | ◯ | ◯ | ◯ | ◯ | ◯ |
| 14. Get upset easily * | ◯ | ◯ | ◯ | ◯ | ◯ |
| 15. Have difficulty understanding abstract ideas * | ◯ | ◯ | ◯ | ◯ | ◯ |
| 16. Keep in the background * | ◯ | ◯ | ◯ | ◯ | ◯ |
| 17. Am not really interested in others * | ◯ | ◯ | ◯ | ◯ | ◯ |
| 18. Make a mess of things * | ◯ | ◯ | ◯ | ◯ | ◯ |
| 19. Seldom feel blue * | ◯ | ◯ | ◯ | ◯ | ◯ |
| 20. Do not have a good imagination * | ◯ | ◯ | ◯ | ◯ | ◯ |

Back Submit

## Computer Experience Questionnaire

Please answer the questions below in terms of months or years.

How much experience do you have with computers?

How much experience do you have with touchscreen devices?

How much experience do you have with tablet computers (e.g. Blackberry Playbook or iPad)?

How much experience do you have with map software such as Google Maps or Mapquest?

## Post-Study Information Page

**Title of Project:** ***Exploring A Mobile Device Interface That Adapts Its Look and Functionality To Suit User State As Measured By a Biosensor***

**Student Investigator:** ***Behzad Aghaei, baghaei@uwaterloo.ca***
**Faculty Supervisor:** ***Prof. Catherine Burns, Systems Design Engineering, c4burns@uwaterloo.ca***

We appreciate your participation in our study and thank you for spending the time helping us with our research. The purpose of this post-study information page is to provide additional details about the study you just participated in.

You were originally told that this study was designed to investigate your physiological response to the use of our map software. While we were indeed interested in studying your physiological response to our software, the study objectives were more involved from what we originally explained to you. The physiological response of interest was your skin conductance (electro-dermal activity) that was measured using the Galvanic Skin Response (GSR) sensor you wore. These readings are closely related to your state of physiological arousal, which will help us infer how frustrated (or not) you were feeling at specific points of the experiment.

In this study, we are actually interested in how we can address the frustration caused by everyday software through the use of visually 'pleasant' or 'calming' adaptations to the user interface of that software. This study is important because past research has outlined the many benefits of adaptive and emotional user interfaces (e.g., Picard & Vizas, 2001; Lavie & Meyer, 2010). Deception has been used in previous research to induce frustration (Scheirer et al., 2002). Examining the effects of affective adaptation to address the emotional state of users can pave the way for commercialized affective adaptive user interfaces in our every day appliances.

It was necessary to have participants think that they were simply participating in a typical user test study because any prior knowledge or suspicion of intentional frustration may have reduced the effect of the frustrating elements (which we built into our software) on your physiology.

Some people in this study were randomized into our baseline/control group that experienced frustrating elements from the software**,** and others were randomized into our adaptive condition, which in addition to the frustrating elements also experienced adaptation of the software's user interface following frustration.

As stated in the pre-study consent form, the highest performing participant in each experimental condition of this study will receive an additional $100. If you are the highest performing participant in your group, you will be notified via e-mail following the completion of the study.

The reason we needed to use deception in this study was because we needed participants' behaviour and attitudes to be as natural as possible. To summarize

1. The purpose of this study was to examine the effectiveness of user interface adaptation on user frustration with a software application.
2. Freezing and unresponsiveness in the software application were pre-scripted and intentional events with the purpose of inducing frustration in our participants.

We apologize again for not providing you with complete and accurate information about the purpose, objectives and procedures for this study, but we hope you understand why this was necessary.

Since this study involves some aspects that you were not told about before starting, it is very important that you not discuss your experiences with any other students who potentially could be in this study, until the end of the term. If people come into the study knowing about our specific predictions, as you can imagine, it could influence their results, and the data we collect would not be usable. Also please do not make this post-study information document available to other students.

The information you provided will be kept confidential by not associating your name with your responses. All of the data will be summarized and no individual could be identified from these summarized results. The data will be stored with all identifying or potentially identifying information removed. Electronic data, with no personal identifiers, will be stored indefinitely on a password-protected computer in a secured laboratory.

Since some elements of the study are different from what was originally explained, there is another consent form for you to read and complete, if you are willing to allow us to use the information you have provided. This consent is also a record that the full purpose of the study was explained to you.

This study was reviewed and received ethics clearance through the Office of Research Ethics at the University of Waterloo. Should you have comments or concerns resulting from your participation in this study, please contact Dr. Maureen Nummelin, the Director, Office of Research Ethics, at 1-519-888-4567, Ext. 36005 or maureen.nummelin@uwaterloo.ca..

We really appreciate your participation and hope that this has been an interesting experience for you.

( ) YES ( ) NO I read the Post-Study Information page.

## POST-DEBRIEFING CONSENT FORM

**Title of Project:** ***Exploring A Mobile Device Interface That Adapts Its Look and Functionality To Suit User State As Measured By a Biosensor***

**Faculty Supervisor:** ***Prof. Catherine Burns, Systems Design Engineering, c4burns@uwaterloo.ca***

**Student Investigator:** ***Behzad Aghaei, baghaei@uwaterloo.ca***

During the debriefing session, I learned that it was necessary for the researchers to disguise the real purpose of this study. I realize that this was necessary since having full information about the actual purpose of the study might have influenced the way in which I responded to the tasks and this would have invalidated the results. Thus, to ensure that this did not happen, some of the details about the purpose of the study initially were not provided (or were provided in a manner that slightly misrepresented the real purpose of the study). However, I have now received a complete verbal and written explanation as to the actual purpose of the study and have had an opportunity to ask any questions about this and to receive acceptable answers to my questions.

I have been asked to give permission for the researchers to use my data (or information I provided) in their study, and agree to this request. I am aware that I may withdraw this consent by notifying the Faculty Supervisor of this decision.

I am aware this study has been reviewed and received ethics clearance through the Office of Research Ethics and I may contact the Director, Office of Research Ethics at 519-888-4567 Ext. 36005 if I have any concerns or comments resulting from my involvement in this study.

**Participant's Name:** ____________________________________________________

**Participant's Signature:** ________________________________________________

**Date:** ________________________________

**Witness' Name:** ________________________________________________________

**Witness' Signature:** ____________________________________________________